Multivariate data analysis

Application to hyperspectral imagery

Mireille GUILLAUME
Institut Fresnel/Ecole Centrale de Marseille
Sao Carlos French-Brazilian summer school 2010, 5-11 september

1. Data and Issues
Data
Issues
2. Principal component analysis and Independent component analysis
3. Metrics and clustering
4. Object detection
General framework
Presentation of recent works
5. Unmixing
General framework
Presentation of recent works
MDMD-NMF algorithm
Discussion
Results
Conclusion

1. Data and Issues
Data
Issues
2. Principal component analysis and Independent component analysis
3. Metrics and clustering
4. Object detection
General framework
Presentation of recent works
5. Unmixing
General framework
Presentation of recent works
MDMD-NMF algorithm
Discussion
Results
Conclusion

1. Data and Issues
Data
Issues
2. Principal component analysis and Independent component analysis
3. Metrics and clustering
4. Object detection
General framework
Presentation of recent works
5. Unmixing
General framework
Presentation of recent works
MDMD-NMF algorithm
Discussion
Results
Conclusion

1. Data and Issues
Data
Issues
2. Principal component analysis and Independent component analysis
3. Metrics and clustering
4. Object detection
General framework
Presentation of recent works
5. Unmixing
General framework
Presentation of recent works
MDMD-NMF algorithm
Discussion
Results
Conclusion

1. Data and Issues
Data
Issues
2. Principal component analysis and Independent component analysis
3. Metrics and clustering
4. Object detection
General framework
Presentation of recent works
5. Unmixing
General framework
Presentation of recent works
MDMD-NMF algorithm
Discussion
Results
Conclusion

Table of contents

1. Data and Issues
   - Data
   - Issues

2. Principal component analysis and Independent component analysis

3. Metrics and clustering

4. Object detection
   - General framework
   - Presentation of recent works

5. Unmixing
   - General framework
   - Presentation of recent works
     - MDMD-NMF algorithm
     - Discussion
     - Results
     - Conclusion

Data

Multivariate data : many observations of variate characters for objects representative of a population.

objects : individuals
characters : variables

→ statistics and analysis of the population

EXAMPLES

- For each person of a population, observe size, weight, eyes color, hair color, number of children, age, etc...
- For an acoustic signal, observe time, frequency

Data

EXAMPLES

- For each person of a population, observe size, weight, eyes color, hair color, number of children, age, etc...
- For an acoustic signal, observe time, frequency
- For each pixel of an image, observe the wavelengths
  pixel = object
  wavelength = character

The data **D** : usual representation

	$\mathbf{d_1}$	$\mathbf{d_2}$	..	$\mathbf{d_l}$	$\mathbf{d_L}$
1				d_{1l}	.
2					.
.				.	.
i				d_{il}	.
N	.	.	.	.	d_{NL}

$\mathbf{d_l}$: character or variable l (for example age or size)
i : index of the observed object or individual

Data

EXAMPLES

- For each person of a population, observe size, weight, eyes color, hair color, number of children, age, etc...
- For an acoustic signal, observe time, frequency
- For each pixel of an image, observe the wavelengths
  pixel = object
  wavelength = character

The data **D** : usual representation

	$\mathbf{d_1}$	$\mathbf{d_2}$	..	$\mathbf{d_l}$	$\mathbf{d_L}$
1				d_{1l}	.
2					.
.				.	.
i				d_{il}	.
N	.	.	.	.	d_{NL}

$\mathbf{d_l}$: character or variable l (for example age or size)
i : index of the observed object or individual

Data

EXAMPLES

- For each person of a population, observe size, weight, eyes color, hair color, number of children, age, etc...
- For an acoustic signal, observe time, frequency
- For each pixel of an image, observe the wavelengths
  pixel $=$ object
  wavelength $=$ character

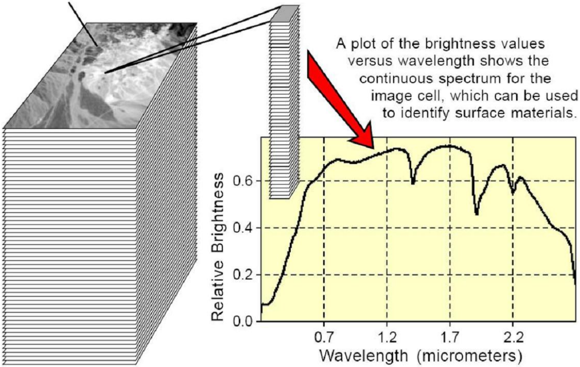
What can we do with that ?

Analysis of a population

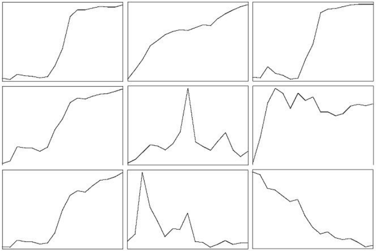
- Which characters are relevant (how many) ?
  $\longrightarrow$ Dimension reduction
- How to compare two different objects ? $\longrightarrow$ Metrics
- Can we group some objects together ?
  $\longrightarrow$ Classification
- Can we detect an object in the population ? $\longrightarrow$ Detection

Hyperspectral imaging

- Images collection
- Spectra collection



Hyperspectral data

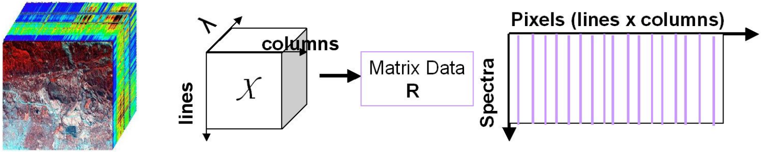


Collection of spectra

Collection of images

Hyperspectral data

Usual representation



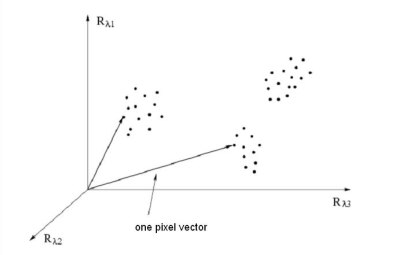
Cube

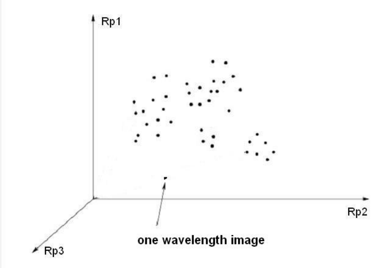
Matrix **R**: a pixel is a vector ($\mathbf{R}' = \mathbf{D}$)

pixel = object or individual
wavelength = variable or character

Hyperspectral data

Space representation





objects space
vector pixels

variables space
vector images

Multivariate statistical models

Let **x** be a random vector representing the pixels vectors,

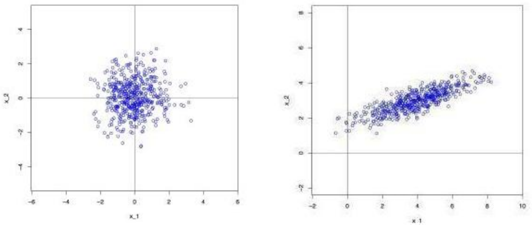
$\mu = E[\mathbf{x}]$

Gaussian density law

$$f_G(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}} \frac{1}{|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\right)$$

Σ : covariance matrix of the data, $\Sigma = E[\mathbf{x}\mathbf{x}']$

500 observations i.i.d. $\mathcal{N}_2(0, I_2)$: $\mathcal{N}_2(\mu, \Sigma)$, $\mu = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$ $\Sigma = 0.4 \times \begin{pmatrix} 5 & 3 \\ 3 & 2.25 \end{pmatrix}$



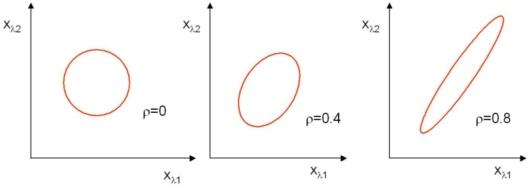
2-D scatterplot

Multivariate statistical models

Gaussian density law

Constant density levels are ellipsoids

$$(\mathbf{x}-\mu)\Sigma^{-1}(\mathbf{x}-\mu) = cte \qquad (1)$$



Example : 2-D Normal distribution

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} ; \rho = \frac{cov(x_1,x_2)}{\sigma_1\sigma_2}$$

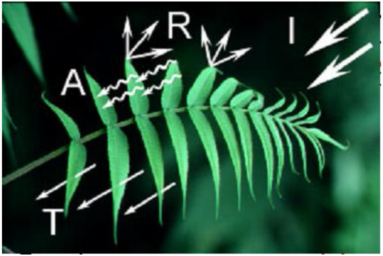
linear correlation coefficient

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

$$|\Sigma| = det\Sigma = (\sigma_1\sigma_2)^2(1-\rho^2)$$

Motivation

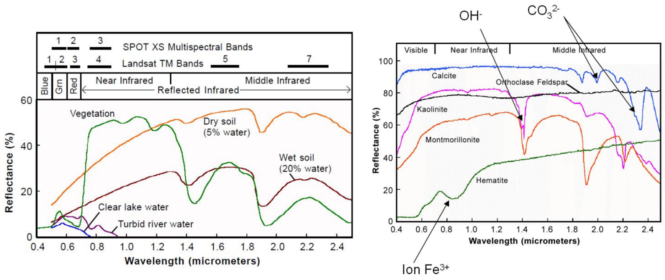
- The reflectance spectrum or the emitting spectrum is representative of the observed material → object identification



spectral signatures of variate materials
spectral libraries : ASTER : http://speclib.jpl.nasa.gov
USGS : http://speclab.cr.usgs.gov/spectral.lib04/spectra-lib04.html

Motivation

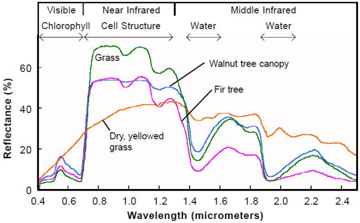
- The reflectance spectrum or the emitting spectrum is representative of the observed material → object identification

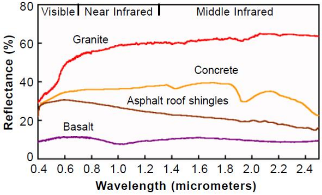


spectral signatures of variate materials
spectral libraries : ASTER : http://speclib.jpl.nasa.gov
USGS : http://speclab.cr.usgs.gov/spectral.lib04/spectra-lib04.html

Motivation

- The reflectance spectrum or the emitting spectrum is representative of the observed material → object identification



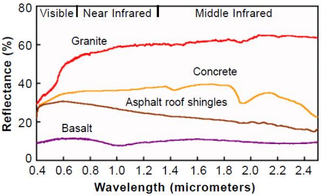


Sample spectra from the ASTER Spectral Library. ASTER will be one of the instruments on the planned EOS AM-1 satellite, and will record image data in 14 channels from the visible through thermal infrared wavelength regions as part of NASA's Earth Science Enterprise program.

spectral signatures of variate materials
spectral libraries : ASTER : http://speclib.jpl.nasa.gov
USGS : http://speclab.cr.usgs.gov/spectral.lib04/spectra-lib04.html

Motivation

- The reflectance spectrum or the emitting spectrum is representative of the observed material → object identification

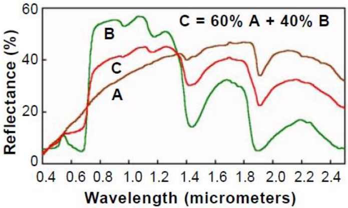


Sample spectra from the ASTER Spectral Library. ASTER will be one of the instruments on the planned EOS AM-1 satellite, and will record image data in 14 channels from the visible through thermal infrared wavelength regions as part of NASA's Earth Science Enterprise program.

spectral signatures of variate materials

Motivation

- The reflectance spectrum or the emitting spectrum is representative of the observed material $\rightarrow$ object identification
- $\rightarrow$ Quantitative estimation of the material abundances in each pixel (sub-pixel)



Example of a composite spectrum (C) that is a linear mixture of two spectra: A (dry soil) and B (green vegetation).

Linear mixing model

$\mathbf{X} = \mathbf{AS}$ or $\quad \mathbf{R} = \mathbf{X} + \mathbf{N}$

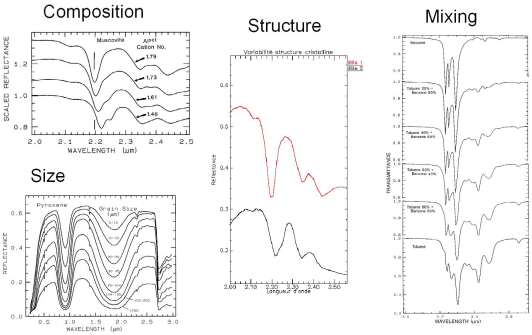
A pure materials "endmembers" matrix

S abundances fractions for each pixel and each endmember

$\mathbf{x}_i = \sum_j a_{ji} \mathbf{s}_j$

Difficult points

→ Spectral variability



→ Physical data models for corrections

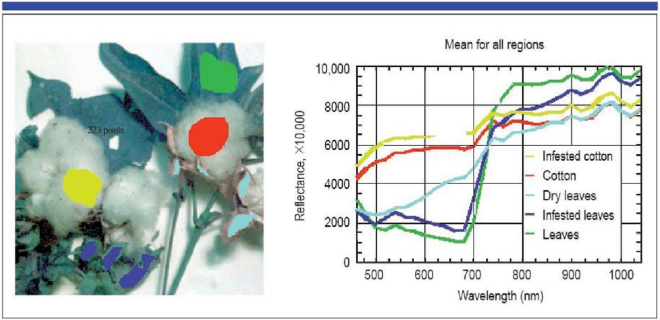
Applications

Many applications :

- Military (detection),
- Agriculture (ecosystems)
- Geoscience
- Industrial (survey, mines),
- Astronomy

Applications

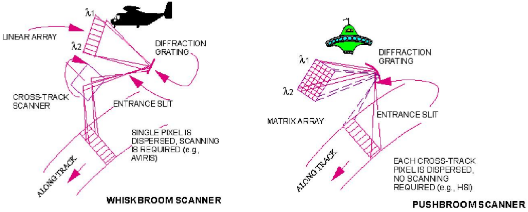
EXAMPLE



Analysis of infested and healthy cotton plants reveals significant spectral differences, thus allowing automated inspection.

Applications

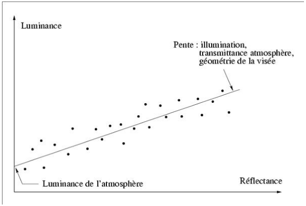
For launched ground monitoring applications such as teledetection



- Radiance spectra observed
- Needs to correct solar illumination
- Needs to correct atmospherical absorption

$Reflectance(\lambda) = \frac{L_o(\lambda}{L_{sol}(\lambda)T(\lambda)cos\theta} - \frac{L_{atm}(\lambda}{L_{sol}(\lambda)T(\lambda)cos\theta}$ L_o observed luminance, L_{sol}andL_{atm} resp. solar and atmospheric luminance, $T(\lambda)$ atmospherical transmittance, θ angle illumination.

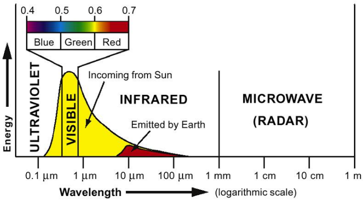
Applications

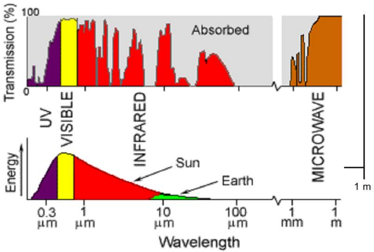


- Radiance spectra observed
- Needs to correct solar illumination
- Needs to correct atmospherical absorption

$Reflectance(\lambda) = \frac{L_o(\lambda}{L_{sol}(\lambda)T(\lambda)cos\theta} - \frac{L_{atm}(\lambda}{L_{sol}(\lambda)T(\lambda)cos\theta}$ L_o observed luminance, L_{sol}andL_{atm} resp. solar and atmospheric luminance, $T(\lambda)$ atmospherical transmittance, θ angle illumination.

Spectral window





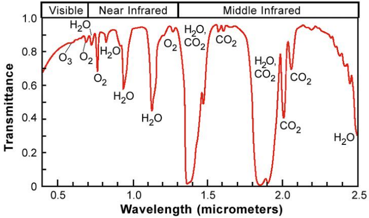
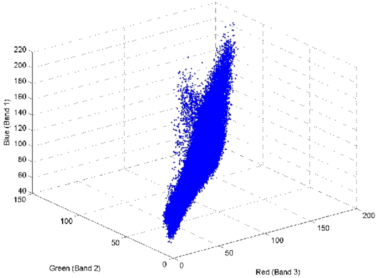


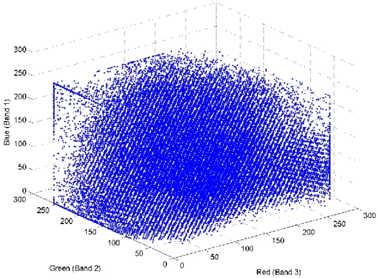
Table of contents

1. Data and Issues
  - Data
  - Issues
2. Principal component analysis and Independent component analysis
3. Metrics and clustering
4. Object detection
  - General framework
  - Presentation of recent works
5. Unmixing
  - General framework
  - Presentation of recent works
    - MDMD-NMF algorithm
    - Discussion
    - Results
    - Conclusion

PCA

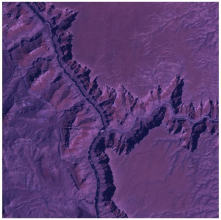
Principal components: projections of the data on its mains directions in the multidimensional space.
Main direction = direction for which the variance of the projected data is maximum
maximum energy or maximum inerty

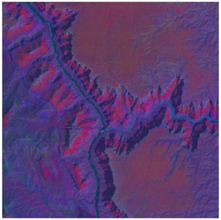


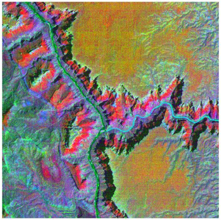


PCA

Correlated image and uncorrelated images







PCA

Center the data $\rightarrow \mathbf{X}$

Problem

First component : find a unit vector $\mathbf{u}$ such as $\mathbf{X}'\mathbf{u}$ has maximum dispersion (maximum energy)
Following components : find orthogonal unit vectors such as $\mathbf{X}'\mathbf{u}$ has maximum dispersion

Can be solved with Lagrangian multipliers formulation.

Solution

- First component: maximize $\mathcal{L} = \mathbf{u}_1'\mathbf{X}\mathbf{X}'\mathbf{u}_1 - \lambda(\mathbf{u}_1'\mathbf{u}_1 - 1)$
- Following components: maximize $\mathcal{L} = \mathbf{u}_2'\mathbf{X}\mathbf{X}'\mathbf{u}_2 - \lambda(\mathbf{u}_2'\mathbf{u}_2 - 1) - \delta\mathbf{u}_2'\mathbf{u}_1$
- Solution: $\mathbf{X}\mathbf{X}'\mathbf{U} = \Lambda\mathbf{U}$, $\quad \mathbf{U} = \begin{pmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_L \end{pmatrix}$

Principal directions: eigenvectors of $\mathbf{X}\mathbf{X}'$

PCA

Eigenvectors and eigenvalues values equation

$\mathbf{XX'U} = \Lambda\mathbf{U}$

$\mathbf{XX'} \approx E[\mathbf{xx'}] = \Sigma$ covariance matrix of the random vector $\mathbf{x}$

PCA and KL

PCA $\approx$ Karhunen-Loëve transform (statistical point of vue)

Principal components :

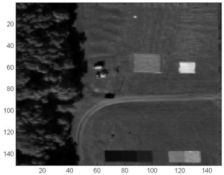
$$\mathbf{y} = \mathbf{U'x}$$

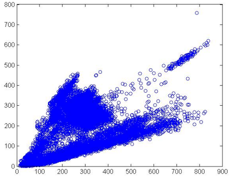
Property : $\mathbf{y}$ is uncorrelated, mean=0

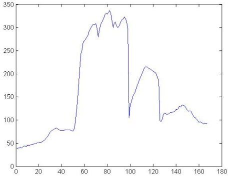
The transform $\mathbf{z} = \Lambda^{-1/2}\mathbf{U'}$ whitens the data.

Example

PCA of a HYDICE scene. HYDICE: sensor with $1m - 60cm$ resolution, 220 spectral bands and spectral resolution of 10 nm

a 

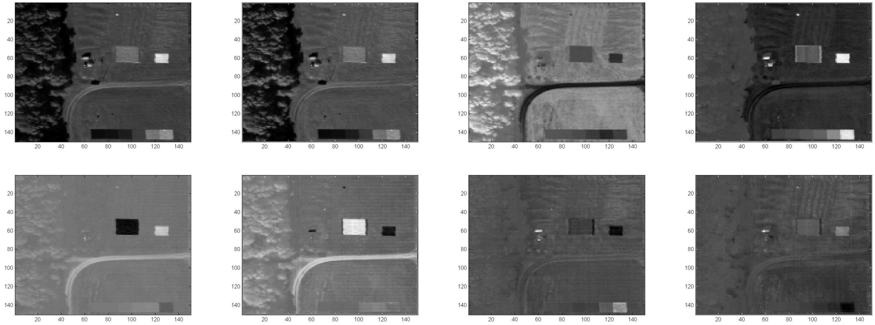
b 

c 

Mean of the data cube (a), scatterplot of the data cube for wavelengths number 60 and 120 (b), and mean spectrum (c)

Example

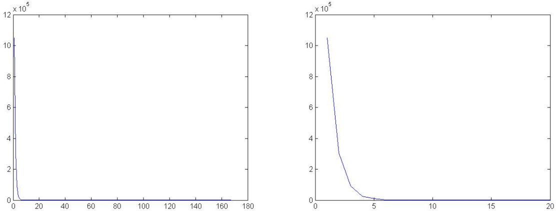
First components of a HYDICE scene.



Mean of the data cube and 7 first PCA components (components 1,3,5,7 are negative)

Example

Eigenvalues of the same HYDICE scene

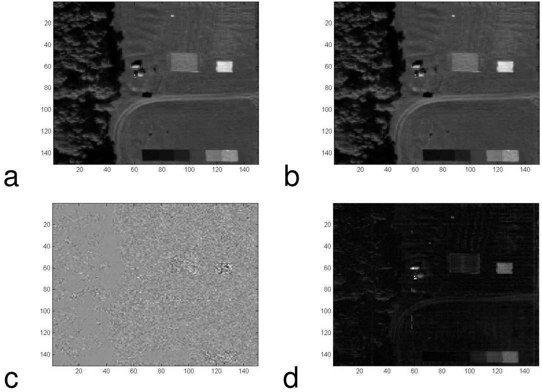


Eigenvalues of the corresponding eigenvectors in HYDICE scene

What reduced dimension can we choose ?

Make the dimension reduction

Example



Mean value of original data (a), mean value of reconstructed data with only 5 components (b), direct mean error reconstruction (c) and sqrt of quadratic error reconstruction (d).

Why reduce the dimension ?

Hughes phenomenon : the curse of dimensionality

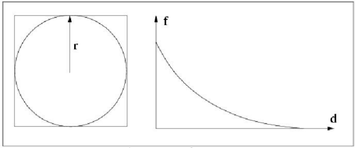
- Complexity : total number of possible different numerical values of the data   hyperspectral image with 220 spectral bands $C = (2^{16})^{2}00 \approx 10^{1060}$, for one image $C = 2^{16} = 256^{2}$
- Needs the observation of $10^{530} \times 10^{530}$ to be able to "fill" the space
- Usually, the observed image is sparse in the multidimensional space
- Not adequate for the estimation of parameters (not enough observations !)
- Some other phenomenon : the volume concentrates in the "shell" of the distributions

Why reduce the dimension ?

Hughes phenomenon : the curse of dimensionality

Fraction of the hypercube volume
Containing an hypersphere

$$f = Vs(r) / Vc(r) = \pi^{D/2} / (D\, 2^{D-1}\, \Gamma(D/2))$$

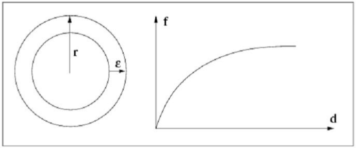


$$\mathrm{Lim}_{D\to\infty} f = 0$$

$\Rightarrow$ Volume concentrates
In the corners when D $\uparrow$

Fraction of the shell volume of an
hypersphere

$$f = Vc(r, \varepsilon) / Vs(r) = 1 - (1 - \varepsilon / r)^D$$



$$\mathrm{Lim}_{D\to\infty} f = 1 \quad \forall \varepsilon > 0$$

$\Rightarrow$ Volume concentrates
in the shell when D $\uparrow$

Conclusion

??????

Independent Component Analysis

Instead of UNCORRELATED, INDEPENDENT components

Uncorrelated

$E[C_i C_j] = 0$

Independent

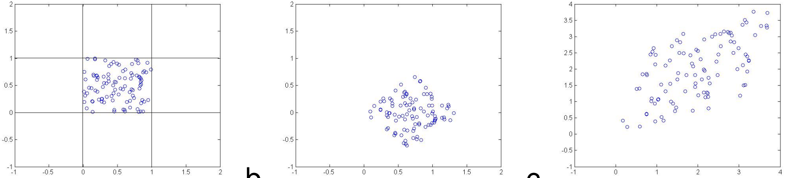
$P(C_i, C_j) = P(C_i)P(C_j)$

$$\textit{Independent} \Rightarrow \textit{Uncorrelated}$$

For Gaussian random variables only

$$\textit{Independent} \Leftrightarrow \textit{Uncorrelated}$$

Independent and uncorrelated



a b c

a: observations of uncorrelated and independent uniform variables $f(x_1|x_2) = f(x_1)$
b: observations of uncorrelated and dependent uniform variables $f(x_1|x_2) \neq f(x_1)$
c: observations of correlated uniform variables

Independent Component Analysis

Model

$\mathbf{x} = \mathbf{A}\mathbf{s}$

A mixing matrix, **s** sources

Hypothesis

- The components of **s** (sources) are mutually independent
- The columns vectors of **A** are linearly independent
- At most one component is Gaussian

Independent Component Analysis

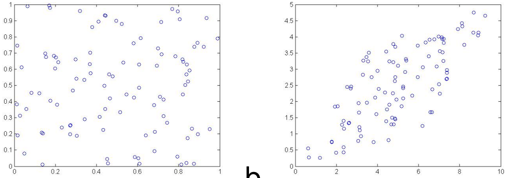
Ambiguities and Problems

- The solution is defined up to a multiplicative constant (undetermined energy) $x_j = \sum_i \left(\frac{a_{ij}}{\alpha_i}\right)(s_i \alpha_i)$
- The variance is fixed to one for each component $E[s_i^2] = 1$
- The IC's cannot be ordered by decreasing energy
- The solution is defined up to a permutation matrix $\mathbf{x} = \mathbf{A}\mathbf{P}^{-1}\mathbf{P}\mathbf{s}$

Independent component analysis

EXAMPLE

a b



Independent uniform source data $\mathbf{s}_i$ (a) and (observed $\mathbf{x}_i$) mixed data with matrix $\mathbf{A} = \begin{pmatrix} 4 & 6 \\ 4 & 1 \end{pmatrix}$ (b)

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}$$

Goal : find the matrix $\mathbf{A}$ and the sources $\mathbf{s}_i$ from the observed $\mathbf{x}_i$

Knowledge : independence of the sources $\rightarrow$ criterion

Independent Component Analysis

Measures of independence

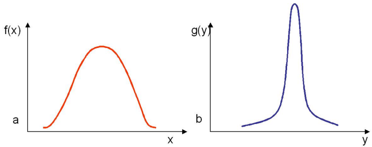
- CLT $\rightarrow$ each linear combination of two independent r.v. is "more Gaussian" than the r.v. themselves
- search for a matrix $\mathbf{W}' \approx \mathbf{A}^{\dagger}$ such as $\tilde{\mathbf{s}} = \mathbf{W}'\mathbf{x}$
- *INDEPENDENCE* $\leftrightarrow$ NON GAUSSIAN distribution for $\tilde{\mathbf{s}}$

Each column $\mathbf{w}_l$ of $\mathbf{W}$ gives an independent direction
The vector $\mathbf{x}$ is projected onto the directions $\mathbf{w}_l$, $l = 1 \ldots L$

Independent Component Analysis

Measure of Nongaussianity with Negentropy

- Gaussian variable has the largest Entropy among all r.v. with same variance
  Entropy $H(x) = -\sum_i p(x_i) \log p(x_i)$
  Differential Entropy
  $H(x) = -\int p_x(\xi) \log p_x(\xi) d\xi$
- Negentropy
  $J(x) = H(x_{Gauss}) - H(x))$
- $\rightarrow$ maximize the Negentropy

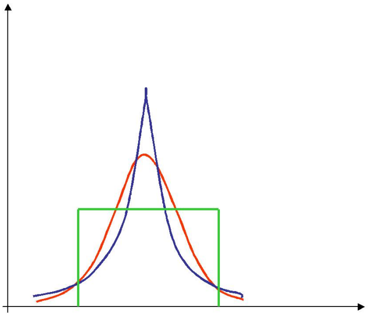


Large entropy r.v. density (a) and small entropy r.v. density (b)

Independent Component Analysis

Measure of Nongaussianity with Kurtosis

- Gaussian variable has null Kurtosis
  Kurtosis $kurt(x) = E[x^4] - 3(E[x^2])^2$
- Positive Kurtosis : supergaussian variable (ex Laplacian)
- Negative Kurtosis : subgaussian variable (ex uniform)
- $\rightarrow$ maximize $|kurt(x)|$



Gaussian, supergaussian (Laplace) and subgaussian (uniform) distributions

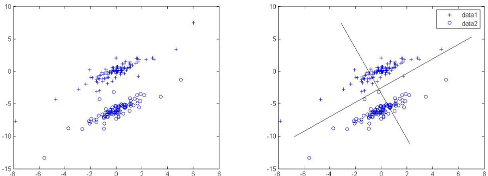
Independent Component Analysis

FastIca algorithm with deflation approach

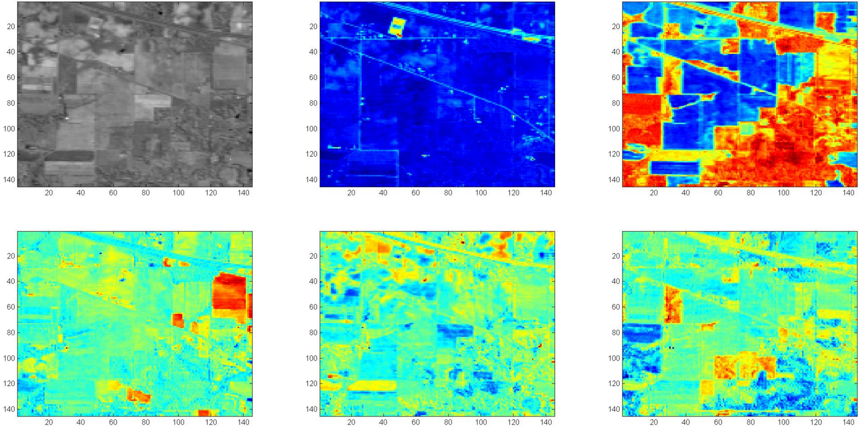
Centering and whitening the data(uncorrelated and white data)
Find iteratively directions vectors **w** such as the data projection $\mathbf{w}'\mathbf{x}$ maximizes the Nongaussianity (maximizes the kurtosis)

- ICA: find the most independent (most interesting) directions
- PCA : find the principal (most energy) directions

EXAMPLE



IC’s of Indian site AVIRIS data



Mean image and five IC’s after dimension reduction with PCA

Independent Component Analysis

Application : anomalies detection with ICA (*A. Huck, ICIP 2008*)

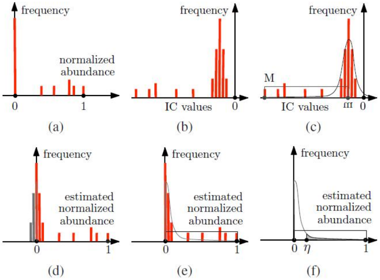


Fig. 1. Histogram schemes (a) of a rare endmember abundance map and (b) of the associated IC histogram; (c) scheme of the IC histogram probabilistic model, (d) histogram scheme of the estimated normalized abundance map, (e) scheme of the normalized abundance map probabilistic model and (f) scheme of the threshold computation;

Algorithm 1

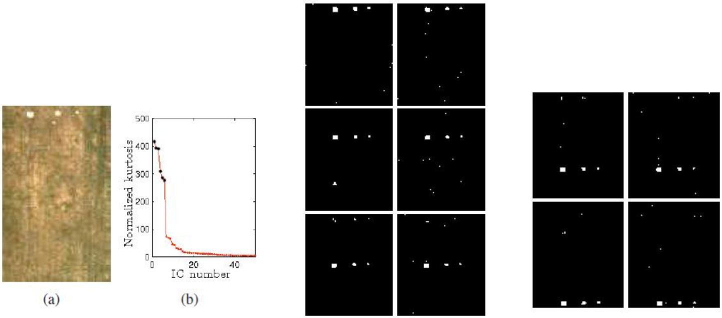
Inputs: HSI and P_{fa};
FastICA analysis;
Select ICs with high normalized kurtosis and process each one as follows:

1. Model the IC histogram with the mixture of a normal and a uniform pdf. An EM algorithm is used to estimate their parameters (Fig.1(c));
2. Estimate the normalized abundance map from the IC by the mean of a piecewise linear transform depending on the parameters of the estimated pdfs (Fig.1(d));
3. Model the estimated abundance map with the mixture of a half-normal and a uniform pdf whose parameters depend on the parameters estimated in step 1 (Fig.1(e));
4. Finally, compute the threshold. The detection mask is obtained thresholding the estimated normalized abundance map (Fig.1(f));

Ouputs: Anomaly detection masks.

Independent Component Analysis

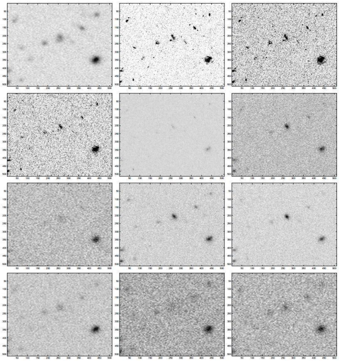
Independent components of HYDICE scene estimated by the algorithm



Usefull in finding ”anomalies”, which have ”peaky” distributions.

Independent Component Analysis

Application : segmentation of a 12-component astronomical image after PCA and after ICA (*F. Flitti, GRETSI 2003*)



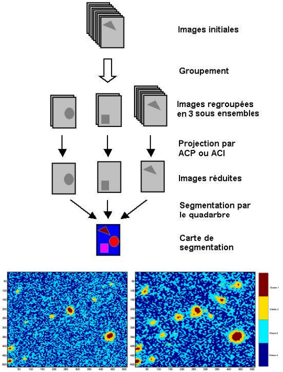


Table of contents

1. Data and Issues
   - Data
   - Issues

2. Principal component analysis and Independent component analysis

3. Metrics and clustering

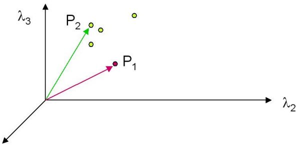
4. Object detection
   - General framework
   - Presentation of recent works

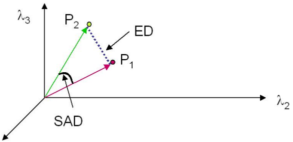
5. Unmixing
   - General framework
   - Presentation of recent works
      - MDMD-NMF algorithm
      - Discussion
      - Results
      - Conclusion

Metrics

How to compare two pixels vectors ?
Algebraic distances

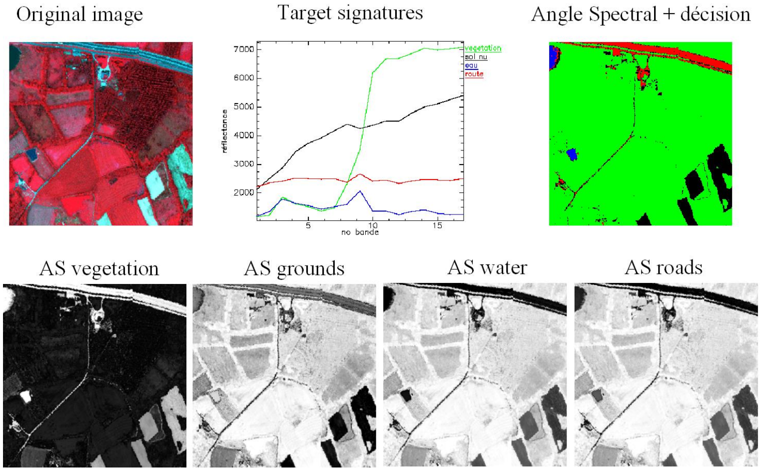
- Euclidian L_2 distance $ED(\mathbf{x}_1, \mathbf{x}_2) = \sqrt{\sum_l (x_{1,l} - x_{2,l})^2}$
- L_1 distance $CBD(\mathbf{x}_1, \mathbf{x}_2) = \sum_l |x_{1,l} - x_{2,l}|$
- Spectral angle $SAD = \cos^{-1}\left(\frac{\mathbf{x}_1 \mathbf{x}_2}{\|x_1\| \|x_2\|}\right)$





Metrics

Spectral angle



Metrics

How to compare two pixels vectors ?
Statistical distances

- Mahalanobis distance
  $MD(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}_1 - \mathbf{x}_2)'\Sigma^{-1}(\mathbf{x}_1 - \mathbf{x}_2))$
- Spectral information divergence
  $SID(\mathbf{x}_1, \mathbf{x}_2) = \sum_l p_{1,l} log\frac{p_{1,l}}{p_{2,l}} + p_{2,l} log\frac{p_{2,l}}{p_{1,l}}$ with
  $p_{j,l} = \frac{x_{j,l}}{\sum_l x_{j,l}}$ and Kullback Liebler pseudo-distance
  $KLD(\mathbf{x}_1, \mathbf{x}_2) = p_{1,l} \ log\frac{p_{1,l}}{p_{2,l}}$

SID and SAD often give very similar results.

Metrics

Statistical non parametric proposed distance

- Kendall's TAU
- Rank correlation coefficient (non linear correlation)

Let (x_1, x_2) and $(\tilde{x}_1, \tilde{x}_2)$ be two realizations of (X_1, X_2). The two observed vectors are said to be concordant if $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) > 0$, and discordant if $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) < 0$. Kendall's τ coefficient is defined as :

$$\tau(X_1, X_2) = P\left[(X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0\right] - P\left[(X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0\right] \quad (2)$$

$(\tilde{X}_1, \tilde{X}_2)$ being a couple of continuous random variables, independent of (X_1, X_2) and following the same probability law.
$\tau(X_1, X_2) =$ probability of concordance $-$ probability of discordance of the random variables X_1 and X_2.

Metrics

Statistical non parametric proposed distance

- Kendall's TAU
- Rank correlation coefficient (non linear correlation)

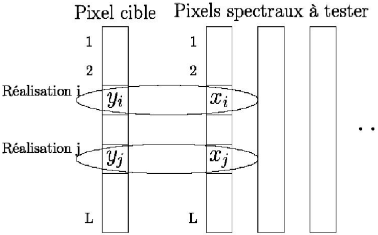
Empiric estimator from N observations $\{x_{l1}\}_{l=1...N}$ of X_1 and N observations $\{x_{k2}\}_{k=1...N}$ of X_2 :

$$\hat{\tau} = \frac{2}{N(N-1)} \sum_{l=1}^{N-1} \sum_{k=l+1}^{N} sign\left[(x_{l1} - x_{k1})(x_{l2} - x_{k2})\right]. \qquad (2)$$

Metrics

Statistical non parametric proposed concordance measure

- Kendall's TAU
- Rank correlation coefficient (non linear correlation)



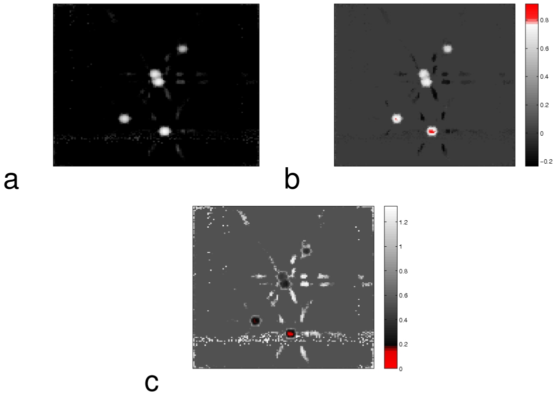
Metrics

How to compare two pixels vectors ?
Geometric distance

- Orthogonal projection distance
  $OPD(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1'\mathbf{P}_2\mathbf{x}_1 + \mathbf{x}_2'\mathbf{P}_1\mathbf{x}_2,$ with
  $\mathbf{P}_j = \mathbf{I}_L - \mathbf{x}_j(\mathbf{x}_j'\mathbf{x}_j)^{-1}\mathbf{x}_j'$

Comparison of SAD and TAU

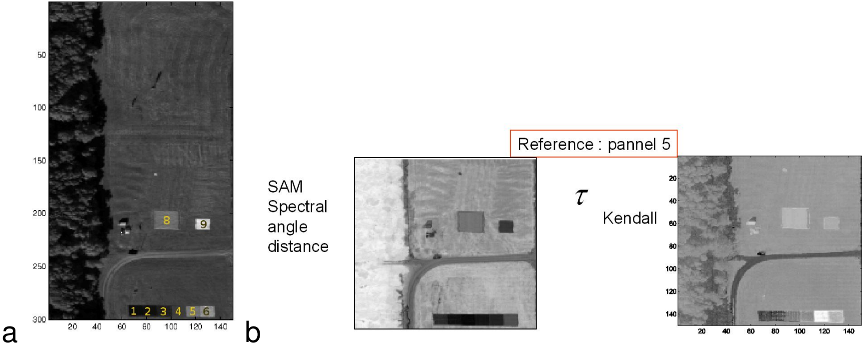
EXAMPLE reconstructed grism image



a: Original image, b: distance map with TAU, c: distance map with SAD

Comparison of SAD and TAU

EXAMPLE HYDICE with pannels

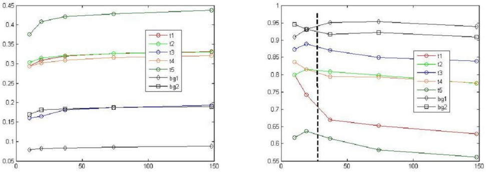


a: Original image, b: distance map with TAU and SAD

Comparison of SAD and TAU

EXAMPLE HYDICE with pannels

Comparison SAD - tau as a function of the number of spectral bands



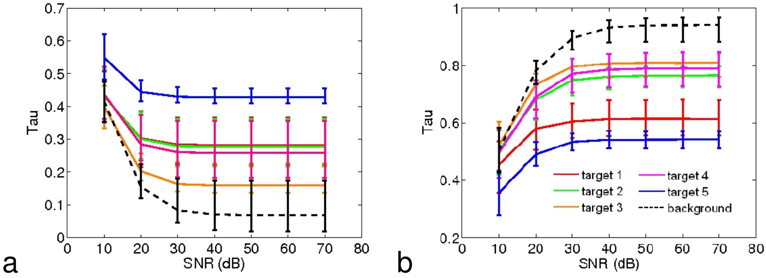
C Spectral angle between targets and background

Concordance between targets and backgroung

c: comparison as a function of spectral bands number

Comparison of SAD and TAU

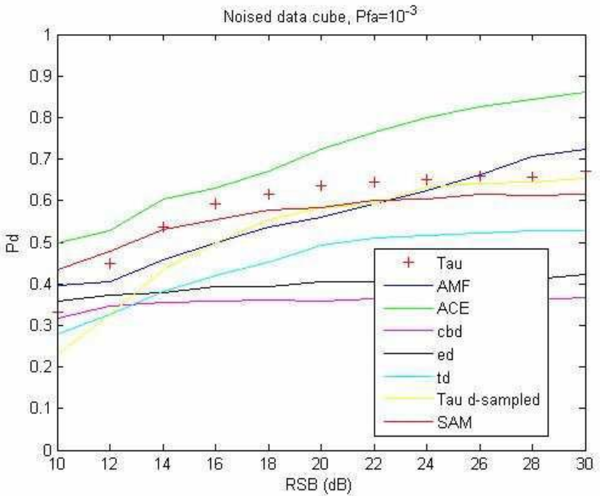
EXAMPLE discriminant power



a: SAD, b: TAU

Comparison of many distances

EXAMPLE sensitivity to noise



Application to unsupervised classification

Clustering : group the data into homogeneous classes, without knowing the classes spectra signatures.
EXAMPLE K-means
If K clusters, K inertia centers

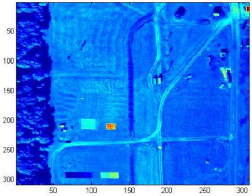
Inertia

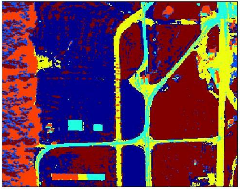
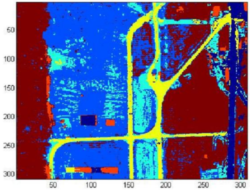
Inertia or dispersion of a set of objects $\mathbf{x}_n$ with an inertia center G_k: $I_k = \frac{1}{N}\sum_n d(\mathbf{x}_n, G_k)$

Intra-class inertia in the total observation $I_C = \sum_{k=1}^{K} I_k P_k$; P_k is the weight of the class k
Inter-class inertia in the total observation
$I_O = \sum_{k=1}^{K} P_k d(G_k, G)$; P_k is the weight of the class k
$I = I_O + I_C = cst \quad \rightarrow$ Maximize I_O, minimize I_C

Application to classification

EXAMPLE K-means

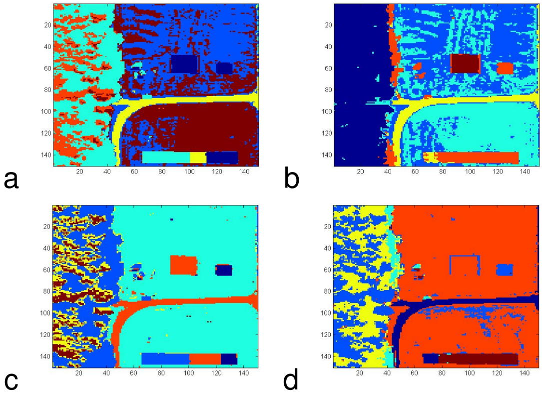
a 

b 
c 

a: image in false colors
b: K-means with ED
c: K-means with TAU distance

Application to classification

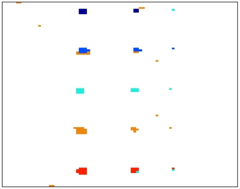
EXAMPLE K-means, 6 classes

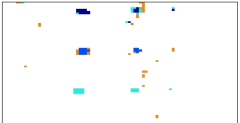


a: K-means with CBD, b: K-means with SAD, c: K-means with ED, d: K-means with correlation

Application to classification

EXAMPLE K-means with TAU (top), and K-means with SAD (bottom), 6 classes





Application to classification

EXAMPLE Other recent methods

a b c

a: image in false colors

b: Maximum Likelihood classification; c: Kernel SVM (support vector machine) classification

Table of contents

1. Data and Issues
    Data
    Issues
2. Principal component analysis and Independent component analysis
3. Metrics and clustering
4. Object detection
    General framework
    Presentation of recent works
5. Unmixing
    General framework
    Presentation of recent works
        MDMD-NMF algorithm
        Discussion
        Results
        Conclusion

Problem

Supervised detection

The target spectrum is known
Goal : detect all the pixels containing this target
$\rightarrow$ find the ”closer” pixels (metric)
spectral variability $\rightarrow$ statistical models of the target and of the background (non-target)

Unsupervised detection

The target signature is not known
Goal : find ”anomalous” pixels : different from the background
spectral variability $\rightarrow$ statistical models of the background

Supervised detection

Many algorithms, based on the Likelihood ratio test or on subspace projection.
The final detector depends on

- The model for spectral variability
- The unknown parameters of the statistical model
- Considering or not full pixel detection or sub-pixel detection
- In the case of sub-pixel detection, the mixing model

Example of matched filter

Two hypothesis

H_0: background, $\mathbf{x} \sim \mathcal{N}(\mu_b, \Sigma_b)$
H_1: target, $\mathbf{x} \sim \mathcal{N}(\mu_t, \Sigma_t)$

Likelihood ratio test

$$\log[\Lambda] = \log\left[\frac{P(\mathbf{x}|H_1)}{P(\mathbf{x}|H_0)}\right] \begin{array}{c} > \\ < \end{array} \eta$$

Matched filter with $\Sigma_t = \Sigma_b$

$D_{MF}(\mathbf{x}) = y = (\mu_t - \mu_b)'\Sigma^{-1}\mathbf{x}$ detection map

Example of matched filter

Two hypothesis

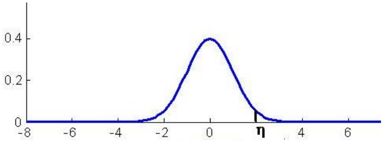
H_0: background, $\mathbf{x} \sim \mathcal{N}(\mu_b, \Sigma_b)$
H_1: target, $\mathbf{x} \sim \mathcal{N}(\mu_t, \Sigma_t)$

Matched filter with $\Sigma_t = \Sigma_b$

$D_{MF}(\mathbf{x}) = y = (\mu_t - \mu_b)'\Sigma^{-1}\mathbf{x}$ detection map

CFAR property

y follows a gaussian law $\rightarrow$ False alarm probability
$P_{FA} = \int_{\eta}^{\infty} p(y|H_0)dy$



Detection filters with parameters estimation

Example of Adaptive matched filter

Two hypothesis

H_0: background, $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \Sigma)$
H_1: target, $\mathbf{x} \sim \mathcal{N}(\mu_t, \Sigma)$
$\mu_t = b\mathbf{s}$, b unknown

Generalized Likelihood Ratio Test

$$\log[\Lambda] = \log\left[\frac{\max_b P(\mathbf{x}; b|H_1)}{P(\mathbf{x}|H_0)}\right] \begin{matrix} > \\ < \end{matrix} \eta$$

Adaptive Matched filter

$D_{AMF}(\mathbf{x}) = y = \frac{(\mathbf{s}'\Sigma^{-1}\mathbf{x})^2}{\mathbf{s}'\Sigma^{-1}\mathbf{s}}$

Unsupervised detection

Anomaly detection

Training data : $\{\mathbf{y}_j, j = 1..N\}$ on background data, parameters θ_0 under H_0
test data :$\{\mathbf{x}_i\}$ with parameters θ_0 or θ_1 under H_0 or H_1

Two hypothesis

H_0: background, $\mathbf{x} \sim \mathcal{N}(\mu, \Sigma)$
H_1: target, $\mathbf{x} \sim \mathcal{N}(\mathbf{s}, \Sigma)$
$\mathbf{s}$, Σ unknown

RX anomaly detector

$$D_{RX}(\mathbf{x}) = (\mathbf{x} - \hat{\mu})'\hat{\Sigma}^{-1}(\mathbf{x} - \hat{\mu})$$

Mahalanobis distance between tested $\mathbf{x}$ and estimated background mean vector μ

Anomalous component pursuit

Anomalies detection using projection pursuit with FastICA deflation algorithm.

Data modeling

$$
\begin{array}{lll}
\mathrm{H}_0: & \mathbf{r}_\zeta = \mathbf{b}_\zeta^d & \sim \mathcal{N}(\mathbf{m}, \Sigma) \\
\mathrm{H}_1: & \mathbf{r}_\zeta = \mathbf{b}_\zeta^d + (\mathbf{t}^d - \mathbf{m}) & \sim \mathcal{N}(\mathbf{t}^d, \Sigma)
\end{array}
$$

After whitening

$$
\begin{array}{lll}
\mathrm{H}_0: & \mathbf{z}_\zeta = \mathbf{b}_\zeta & \sim \mathcal{N}(\mathbf{0}_{L,1} \\
\mathrm{H}_1: & \mathbf{z}_\zeta = \mathbf{b}_\zeta + \mathbf{t} &
\end{array}
$$

$$\mathbf{t} = \mathbf{D}^{-1/2}\mathbf{U}^T \left(\mathbf{t}^d - \mathbf{m}\right).$$

Anomalous component pursuit

Anomalies detection using projection pursuit with FastICA deflation algorithm.

Data modeling

$$\mathbf{r}_\zeta = \mathbf{b}_\zeta^d + \beta_\zeta \left(\mathbf{t}^d - \mathbf{m}\right) \quad i.e. \quad \mathbf{z}_\zeta = \mathbf{b}_\zeta + \beta_\zeta \mathbf{t} \ , \qquad (3)$$

where β_ζ follows a Bernoulli distribution of parameter p:

$$\beta_\zeta \sim \mathcal{B}(p) \text{ , and p is small} \qquad (4)$$

We search for a projector $\mathbf{w}$

$$\mathbf{s} = \mathbf{w}^T \mathbf{Z} \qquad (5)$$

Anomalous component pursuit

Anomalies detection using projection pursuit with FastICA deflation algorithm.

Assumption

A_2: $\mathbf{w}$ is parallel to $\mathbf{t}$ $\Rightarrow$
$\mathbf{w}$ locally maximizes the kurtosis.

Finally, the model can be extended to many anomaly classes as follows :

$$\mathbf{r}_\zeta = \mathbf{b}^d_\zeta + \sum_{j=1}^{J_a} \beta_{\zeta,j}(\mathbf{t}^d_j - \mathbf{m}) \quad i.e. \quad \mathbf{z}_\zeta = \mathbf{b}_\zeta + \sum_{j=1}^{J_a} \beta_{\zeta,j}\mathbf{t}_j \ , \qquad (6)$$

Anomalous component pursuit

Algorithm : Estimation of one projector with FastICA
Choose an initial normalized vector $\mathbf{w}$;
Until convergence do:

1. $\mathbf{w} \leftarrow \mathrm{E}_{\mathbf{z}}\left[\mathbf{z}g(\mathbf{w}^T\mathbf{z})\right] - \mathrm{E}_{\mathbf{z}}\left[g'(\mathbf{w}^T\mathbf{z})\right]$;
2. $\mathbf{w} \leftarrow \mathbf{P}_{\mathbf{W}}^{\perp}\mathbf{w}$
3. $\mathbf{w} \leftarrow \mathbf{w}/\|\mathbf{w}\|_2$

The initialization is made with *RX* algorithm, and we choose the most anomalous pixel detected with RX.

Anomalous component pursuit

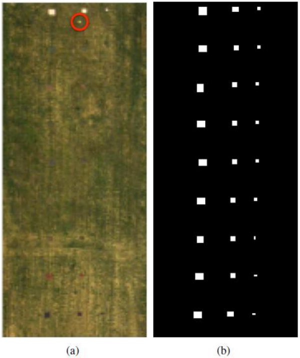


Fig. 1. (a) Visual representation of the analyzed scene and (b) ground truth detection mask. The target pixels are white, background pixels are black. Most panels are hard to perceive in (a). The circled anomaly in (a) is not included in the ground truth mask (b)

Anomalous component pursuit

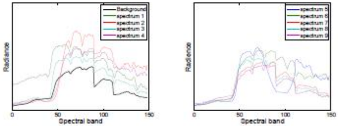


Fig. 2. Average background spectrum and 9 panel spectra.

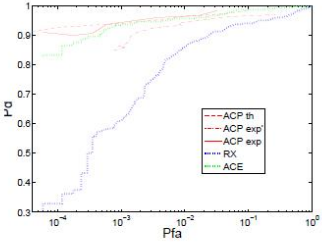
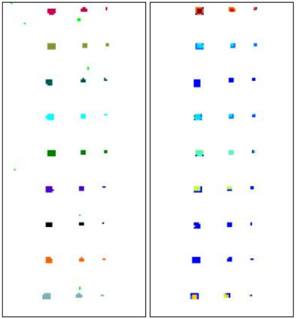


Fig. 3. Detection ROC curves of ACP, ACE and RX. ACP and RX are unsupervised detectors whereas ACE is supervised. These curves have been obtained from the HYDICE dataset presented in Fig.1.

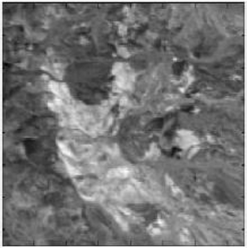
Anomalous component pursuit



(a) (b)

Fig. 6. Target discrimination. (a) ACP "discrimination map" for $\kappa = 10^{-6}$. 10 spectral classes are found, one of them corresponds to the rock visible in Fig.1(a), under the first row second column target. The other 9 correspond to the 9 panel materials. (b) Semi-supervised segmentation of the "true" targets only: the target panels presented in Fig.1(b) are segmented with the K-means algorithm, with random initial conditions, SAD measure and 9 classes.

Anomalous component pursuit



Spectral mean of the selected cuprite data scene

Anomalous component pursuit

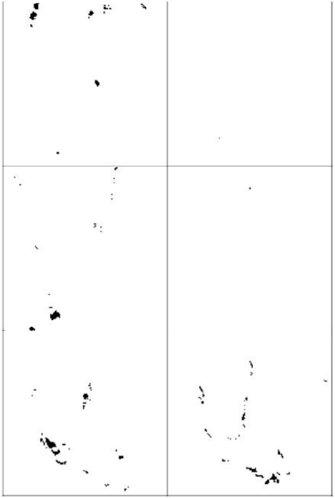
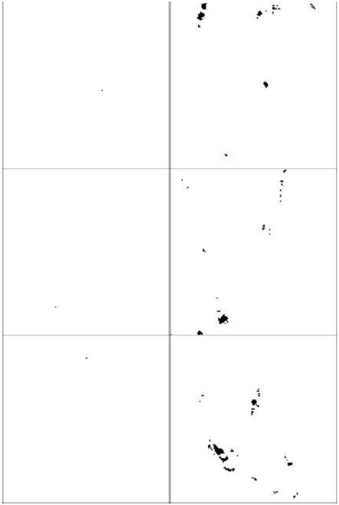


Fig. 9. First six anomaly detection masks of cuprite data. Here the anomalies are black and the background is white

Anomalous component pursuit



ig. 10. First six anomaly detection masks of cuprite data with one anomaly
ıserted in coordinates (90,80)

Anomalous component pursuit

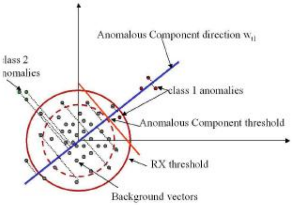


Fig. 11. Schematic representation of ACP process

Table of contents

1. Data and Issues
  - Data
  - Issues

2. Principal component analysis and Independent component analysis

3. Metrics and clustering

4. Object detection
  - General framework
  - Presentation of recent works

5. Unmixing
  - General framework
  - Presentation of recent works
    - MDMD-NMF algorithm
    - Discussion
    - Results
    - Conclusion

Problem

UNMIXING

- Geometrical approach
- Statistical approach
- Non-negative matrix factorization : algebraic approach

Problem

UNMIXING

- Geometrical approach
- Statistical approach
- Non-negative matrix factorization : algebraic approach

Problem

UNMIXING

- Geometrical approach
- Statistical approach
- Non-negative matrix factorization : algebraic approach

Why NMF ?

Linear mixing model:

$$\mathbf{R} = \mathbf{X} + \mathbf{N} \tag{7}$$

$$\mathbf{X} = \mathbf{A}\mathbf{S} \tag{8}$$

A: reflectances of endmembers
$\rightarrow C_1$: non-negative
S: abundances of endmembers
$\rightarrow C_2$: non-negative
$\rightarrow C_3$: sum-to-one

Problem

UNMIXING

- Geometrical approach
- Statistical approach
- Non-negative matrix factorization : algebraic approach

Why not NMF ?

Ill posed problem:

- Is the solution unique ?
- Which criterion to be optimized ?
- Which algorithm (convexity) ?

Basic formulation of NMF

Find two matrices $\hat{\mathbf{A}}$ and $\hat{\mathbf{S}}$ such as:

$$\mathbf{X} \simeq \hat{\mathbf{A}}\hat{\mathbf{S}} \tag{7}$$

Minimize the reconstruction quadratic error (RQE):

$$\mathrm{R}QE(\mathbf{A}, \mathbf{S}) = \|\mathbf{R} - \mathbf{A}\mathbf{S}\|_F^2 \tag{8}$$

$\rightarrow$ Ensures C_1 and C_2 (needs normalization to enforce C_3)
$\rightarrow$ Does not ensure unicity of solution: needs regularization
$\rightarrow$ Not convex for **A** and **S** simultaneously

Advanced formulations of NMF

The objective function RQE is regularized :
$f(\mathbf{A},\mathbf{S}) = \mathrm{R}QE(\mathbf{A},\mathbf{S}) + \lambda_A f_A(\mathbf{A}) + \lambda_S f_S(\mathbf{S})$

SOME EXAMPLES

- Sum-to-one constraint : STU-NMF
  $f_{STU}(\mathbf{A},\mathbf{S}) = \mathrm{R}QE(\mathbf{A},\mathbf{S}) + \delta . STU(\mathbf{S})$,
  $STU(\mathbf{S}) = \sum_{i=pixels} \left(\sum_{j=endmembers} s_{ij} - 1 \right)^2$
- Minimum-volume constraint : MVC-NMF
  $f_{MVC-NMF} = f_{STU}(\mathbf{A},\mathbf{S}) + \lambda_J . J(\mathbf{A})$, $J(\mathbf{A}) = \frac{(J-1)!}{2} V^2 [\mathbf{A}]$
- Minimum spectral dispersion constraint : MD-NMF (Minidisco)
  $f_{MD-NMF} = f_{STU}(\mathbf{A},\mathbf{S}) + \lambda_A . D_A(\mathbf{A})$, $D_A(\mathbf{A}) = L \sum_{j=endmembers} \sigma^2_{A_j}$

Advanced formulations of NMF

The objective function RQE is regularized :
$f(\mathbf{A},\mathbf{S}) = \mathrm{R}QE(\mathbf{A},\mathbf{S}) + \lambda_A f_A(\mathbf{A}) + \lambda_S f_S(\mathbf{S})$

SOME EXAMPLES

- Sum-to-one constraint : STU-NMF
  $f_{STU}(\mathbf{A},\mathbf{S}) = \mathrm{R}QE(\mathbf{A},\mathbf{S}) + \delta.STU(\mathbf{S})$,
  $STU(\mathbf{S}) = \sum_{i=pixels}\left(\sum_{j=endmembers} s_{ij} - 1\right)^2$
- Minimum-volume constraint : MVC-NMF
  $f_{MVC-NMF} = f_{STU}(\mathbf{A},\mathbf{S}) + \lambda_J.J(\mathbf{A})$, $J(\mathbf{A}) = \frac{(J-1)!}{2} V^2\,[\mathbf{A}]$
- Minimum spectral dispersion constraint : MD-NMF (Minidisco)
  $f_{MD-NMF} = f_{STU}(\mathbf{A},\mathbf{S}) + \lambda_A.D_A(\mathbf{A})$, $D_A(\mathbf{A}) = L\sum_{j=endmembers}\hat{\sigma}^2_{A_j}$

Advanced formulations of NMF

The objective function RQE is regularized :
$f(\mathbf{A},\mathbf{S}) = RQE(\mathbf{A},\mathbf{S}) + \lambda_A f_A(\mathbf{A}) + \lambda_S f_S(\mathbf{S})$

SOME EXAMPLES

- Sum-to-one constraint : STU-NMF
  $f_{STU}(\mathbf{A},\mathbf{S}) = RQE(\mathbf{A},\mathbf{S}) + \delta . STU(\mathbf{S})$,
  $STU(\mathbf{S}) = \sum_{i=pixels}\left(\sum_{j=endmembers} s_{ij} - 1\right)^2$
- Minimum-volume constraint : MVC-NMF
  $f_{MVC-NMF} = f_{STU}(\mathbf{A},\mathbf{S}) + \lambda_J . J(\mathbf{A})$, $J(\mathbf{A}) = \frac{(J-1)!}{2} V^2 [\mathbf{A}]$
- Minimum spectral dispersion constraint : MD-NMF (Minidisco)
  $f_{MD-NMF} = f_{STU}(\mathbf{A},\mathbf{S}) + \lambda_A . D_A(\mathbf{A})$, $D_A(\mathbf{A}) = L \sum_{j=endmembers} \hat{\sigma}^2_{A_j}$

Advanced formulations of NMF

The objective function RQE is regularized :
$f(\mathbf{A},\mathbf{S}) = \mathrm{R}QE(\mathbf{A},\mathbf{S}) + \lambda_A f_A(\mathbf{A}) + \lambda_S f_S(\mathbf{S})$

SOME EXAMPLES

- Sum-to-one constraint : STU-NMF
  $f_{STU}(\mathbf{A},\mathbf{S}) = \mathrm{R}QE(\mathbf{A},\mathbf{S}) + \delta.STU(\mathbf{S})$,
  $STU(\mathbf{S}) = \sum_{i=pixels} \left(\sum_{j=endmembers} s_{ij} - 1 \right)^2$
- Minimum-volume constraint : MVC-NMF
  $f_{MVC-NMF} = f_{STU}(\mathbf{A},\mathbf{S}) + \lambda_J.J(\mathbf{A})$, $\quad J(\mathbf{A}) = \frac{(J-1)!}{2} V^2 [\mathbf{A}]$
- Minimum spectral dispersion constraint : MD-NMF (Minidisco)
  $f_{MD-NMF} = f_{STU}(\mathbf{A},\mathbf{S}) + \lambda_A.D_A(\mathbf{A})$, $\quad D_A(\mathbf{A}) = L \sum_{j=endmembers} \hat{\sigma}^2_{A_j}$

MDMD-NMF

MOTIVATION

Consider abundance sparsity

- Most of the time, an observed pixel contains only few mixed endmembers among the J contained in **A**
- The abundances should be either small ($< \frac{1}{J}$) or large ($> \frac{1}{J}$)

MDMD-NMF

MOTIVATION

Consider abundance sparsity

- Most of the time, an observed pixel contains only few mixed endmembers among the J contained in **A**
- The abundances should be either small ($< \frac{1}{J}$) or large ($> \frac{1}{J}$)
- Maximum dispersion of the abundances : obtain abundances the most far from $\frac{1}{J}$ under STU constraint

MDMD-NMF

FORMULATION OF MDMD-NMF

Criterion

Minimize the regularized function

$$f_{MDMD-NMF} = f_{STU}(\mathbf{A}, \mathbf{S}) + \lambda_A . D_A(\mathbf{A}) + \lambda_S . D_S(\mathbf{S}) \qquad (9)$$

with

$$D_S(\mathbf{S}) = -J \sum_{i=pixels} \hat{\sigma}^2_{S_i} = -\left\| \mathbf{S} - \frac{1}{J} \mathbf{I}_{JI} \right\|_F^2 \qquad (10)$$

In which $\hat{\sigma}^2_{S_i}$ is the dispersion of the abundances for the pixel i

MDMD-NMF

IMPLEMENTATION

Algorithm

- Alternate Gradient
- Multiscale Armijo/Lin based technique for μ_S and μ_A

$$\mathbf{S} \leftarrow \mathbf{S} - \mu_S \left(\bar{\mathbf{A}}^T (\bar{\mathbf{A}}\mathbf{S} - \bar{\mathbf{X}}) - \lambda_S (\mathbf{S} - \frac{1}{J}\mathbf{1}_{JI}) \right) \quad (11)$$

$$\mathbf{A} \leftarrow \mathbf{A} - \mu_A \left((\mathbf{A}\mathbf{S} - \mathbf{X})\mathbf{S}^T + \lambda_A (\mathbf{A} - \frac{1}{L}\mathbf{1}_{L,L}\mathbf{A}) \right) \quad (12)$$

- $\bar{\mathbf{X}}$ and $\bar{\mathbf{A}}$ include the sum-to-one constraint:

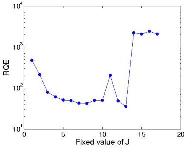
$$\bar{\mathbf{X}} = \begin{bmatrix} \mathbf{X} \\ \delta \cdot \mathbf{1}_{1I} \end{bmatrix} \quad \bar{\mathbf{A}} = \begin{bmatrix} \mathbf{A} \\ \delta \cdot \mathbf{1}_{1J} \end{bmatrix} \quad (13)$$

MDMD-NMF

IMPLEMENTATION

Parameters

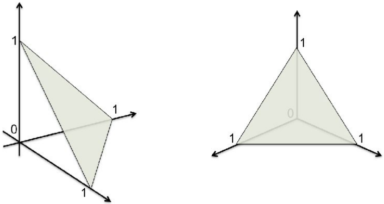
- Initialization with VCA
- Regularization parameters: $\delta = 1$, $\lambda_A = 0.01$, $\lambda_S = 0.01 \cdot J$
- Stop criterion: on the objective function RQE instead of $f_{MDMD-NMF} \rightarrow \mathrm{RQE}^{k-100} < \min_{k'=0,\ldots,99} \mathrm{RQE}^{k-k'}$
  Let RQE locally increase in order to avoid local minima
- Estimation of J: find the best RQE



RQE as a function of *J* for real data

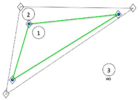
Relationship with minimum volume methods

C_3: the abundance vectors are localized in a J-simplex $\mathcal{S}_s$, which vertices are on the axes of the associated base.

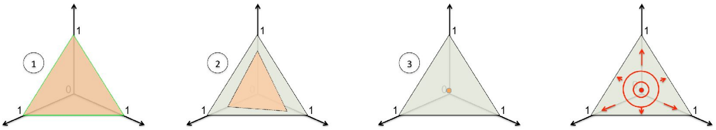


In the abundances' space, possible location of the abundance vectors for $J = 3$

Relationship with minimum volume



Data enclosing simplex



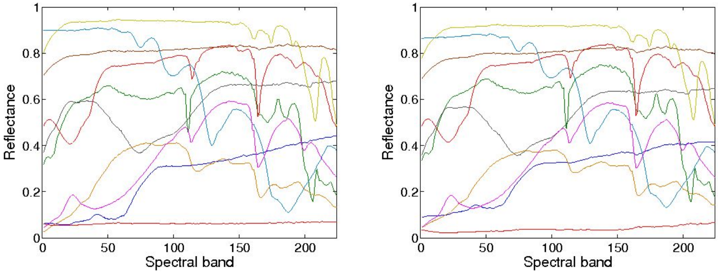
Abundance vectors location corresponding to (1) minimum volume simplex with pure pixels, (2) no pure pixel, (3) all abundances equal to $\frac{1}{J}$, and constraint on the dispersion of the abundances

Results for simulated data

Data generation

- J endmembers randomly extracted from USGS library
- Abundances generated according to a Dirichlet density law
- Selection of *I* abundance vectors with maximum value equal to a fixed threshold ξ
- ι is the ratio of null abundances

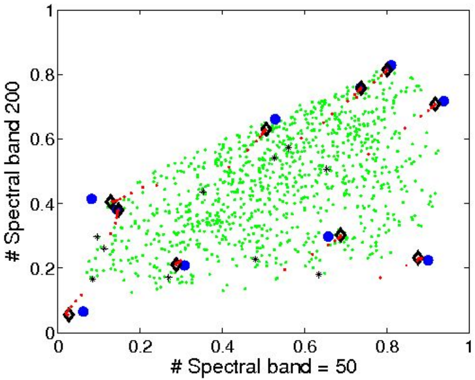
Results for simulated data



MDMD-NMF results for $J = 10$, $\xi = 0.7$, $\iota = 0.8$, I=1000, SNR=40 dB.

True and estimated spectra

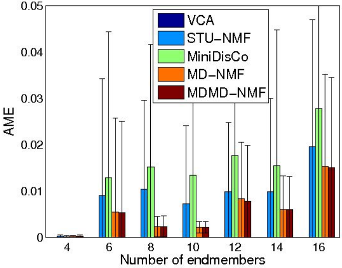
Results for simulated data



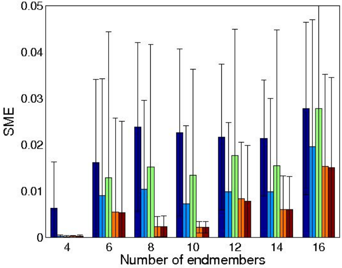
MDMD-NMF results for $J = 10$, $\xi = 0.7$, $\iota = 0.8$, I=1000, SNR=40 dB. Scatterplot

Results for simulated data

As the number of endmembers J varies
Mean over 20 runs



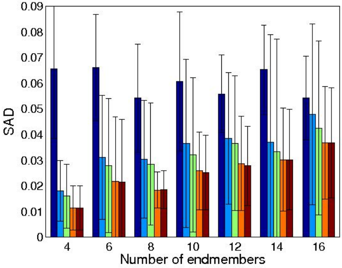
AME: abundance mean error

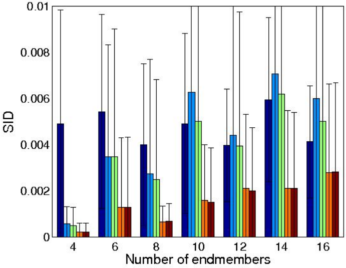


SME: spectral mean error

Results for simulated data

As the number of endmembers *J* varies

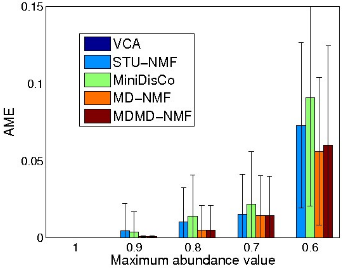




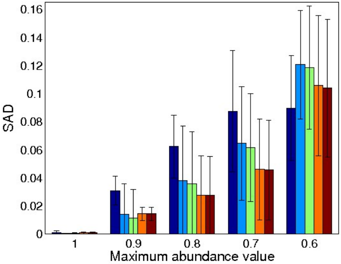
SAD: spectral angle distance SID: spectral information divergence

Results for simulated data

As the maximum abundance varies

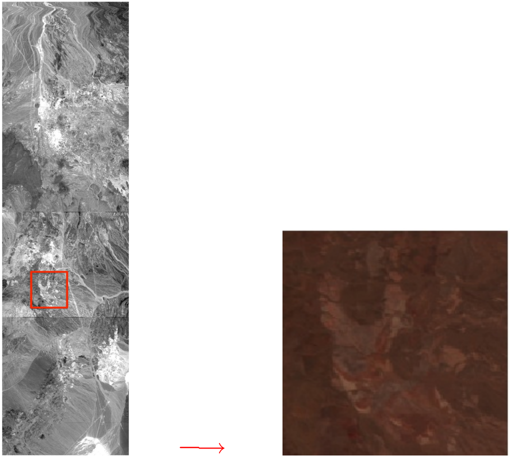


AME: abundance mean error



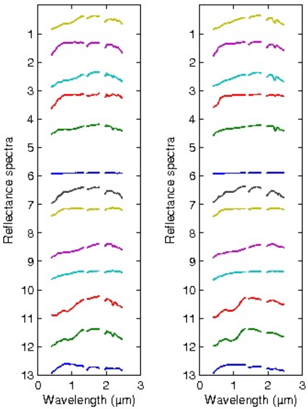
SAD: spectral angle distance

Analysis of Cuprite data



Ground truth available for the whole data

Results for MDMD-NMF

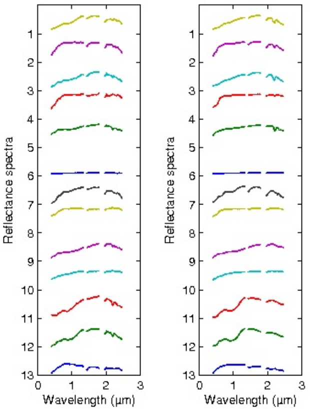


	USGS references	SAD
1	Kaolin/Smect KLF508 85%K	4.8°
2	Kaolin/Smect KLF511 12%K	2.6°
3	Kaolin/Smect KLF508 85%K	4.2°
4	Perthite HS415.3B	1.8°
5	Muscovite IL107	5.2°
6	Brookite HS443.2B	5.3°
7	Nontronite SWa-1.a	3.6°
8	Microcline HS151.3B	4.3°
9	Corrensite CorWa-1	2.9°
10	Quartz GDS74 Sand Ottawa	2.9°
11	Goethite WS220	8.6°
12	Goethite WS219 (limonite)	5.6°
13	Dry_Long_Grass AV87 − 2	5.1°
	Mean SAD	4.4°

Estimated spectra and associated ones in USGS

7 to 8 distinct references found

Results for MDMD-NMF

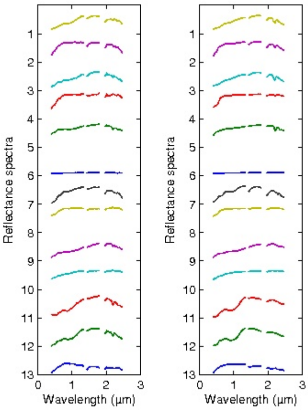


Estimated spectra and associated ones in USGS

	USGS references	SAD
1	Kaolin/Smect KLF508 85%K	4.8°
2	Kaolin/Smect KLF511 12%K	2.6°
3	Kaolin/Smect KLF508 85%K	4.2°
4	Perthite HS415.3B	1.8°
5	Muscovite IL107	5.2°
6	Brookite HS443.2B	5.3°
7	Nontronite SWa-1.a	3.6°
8	Microcline HS151.3B	4.3°
9	Corrensite CorWa-1	2.9°
10	Quartz GDS74 Sand Ottawa	2.9°
11	Goethite WS220	8.6°
12	Goethite WS219 (limonite)	5.6°
13	Dry_Long_Grass AV87 – 2	5.1°
	Mean SAD	4.4°

7 to 8 distinct references found

Results for MDMD-NMF

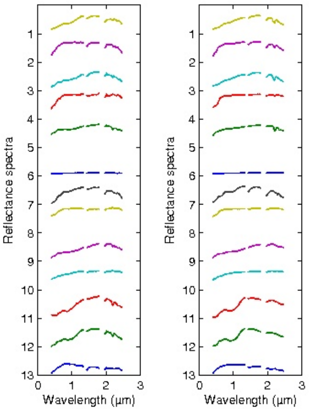


Estimated spectra and associated ones in USGS

	USGS references	SAD
1	Kaolin/Smect KLF508 85%K	4.8°
2	Kaolin/Smect KLF511 12%K	2.6°
3	Kaolin/Smect KLF508 85%K	4.2°
4	Perthite HS415.3B	1.8°
5	Muscovite IL107	5.2°
6	Brookite HS443.2B	5.3°
7	Nontronite SWa-1.a	3.6°
8	Microcline HS151.3B	4.3°
9	Corrensite CorWa-1	2.9°
10	Quartz GDS74 Sand Ottawa	2.9°
11	Goethite WS220	8.6°
12	Goethite WS219 (limonite)	5.6°
13	Dry_Long_Grass AV87 – 2	5.1°
	Mean SAD	4.4°

7 to 8 distinct references found

Results for MDMD-NMF

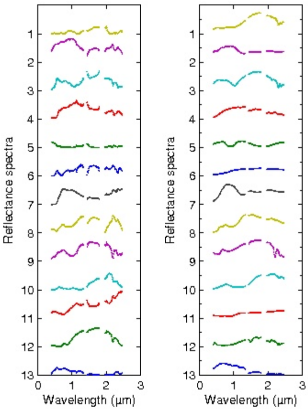


	USGS references	SAD
1	Kaolin/Smect KLF508 85%K	4.8°
2	Kaolin/Smect KLF511 12%K	2.6°
3	Kaolin/Smect KLF508 85%K	4.2°
4	Perthite HS415.3B	1.8°
5	Muscovite IL107	5.2°
6	Brookite HS443.2B	5.3°
7	Nontronite SWa-1.a	3.6°
8	Microcline HS151.3B	4.3°
9	Corrensite CorWa-1	2.9°
10	Quartz GDS74 Sand Ottawa	2.9°
11	Goethite WS220	8.6°
12	Goethite WS219 (limonite)	5.6°
13	Dry_Long_Grass AV87 – 2	5.1°
	Mean SAD	4.4°

Estimated spectra and associated ones in USGS

7 to 8 distinct references found

Results for MVC-NMF

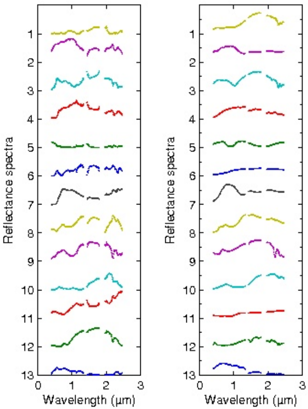


Estimated spectra and associated ones in USGS

	USGS references	SAD
1	Smaragdite HS290.3B	33.7°
2	Almandine WS475	7.1°
3	Actinolite HS315.4B	12.3°
4	Tumbleweed ANP92 − 2C Dry	9.3°
5	Hypersthene PYX02.h > 250u	42.4°
6	Desert_Varnish GDS141	15.0°
7	Spessartine HS112.3B	15.5°
8	Goethite WS219 (limonite)	15.5°
9	Lepidolite NMNH105541	10.5°
10	Siderite HS271.3B	19.3°
11	Hematite GDS69.d 30 − 45um	10.3°
12	Azurite WS316	16.5
13	Opal WS732	16.1°
	Mean SAD	17.2°

3 to 4 distinct references found

Results for MVC-NMF

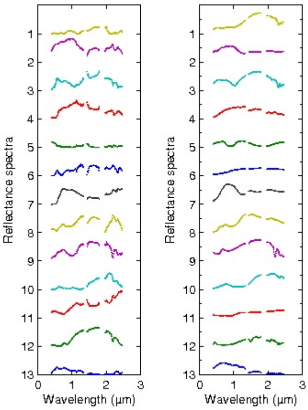


Estimated spectra and associated ones in USGS

	USGS references	SAD
1	Smaragdite HS290.3B	33.7°
2	Almandine WS475	7.1°
3	Actinolite HS315.4B	12.3°
4	Tumbleweed ANP92 – 2C Dry	9.3°
5	Hypersthene PYX02.h > 250u	42.4°
6	Desert_Varnish GDS141	15.0°
7	Spessartine HS112.3B	15.5°
8	Goethite WS219 (limonite)	15.5°
9	Lepidolite NMNH105541	10.5°
10	Siderite HS271.3B	19.3°
11	Hematite GDS69.d 30 – 45um	10.3°
12	Azurite WS316	16.5
13	Opal WS732	16.1°
	Mean SAD	17.2°

3 to 4 distinct references found

Results for MVC-NMF

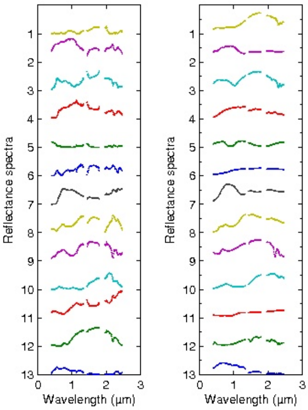


Estimated spectra and associated ones in USGS

	USGS references	SAD
1	Smaragdite HS290.3B	33.7°
2	Almandine WS475	7.1°
3	Actinolite HS315.4B	12.3°
4	Tumbleweed ANP92 – 2C Dry	9.3°
5	Hypersthene PYX02.h > 250u	42.4°
6	Desert_Varnish GDS141	15.0°
7	Spessartine HS112.3B	15.5°
8	Goethite WS219 (limonite)	15.5°
9	Lepidolite NMNH105541	10.5°
10	Siderite HS271.3B	19.3°
11	Hematite GDS69.d 30 – 45um	10.3°
12	Azurite WS316	16.5
13	Opal WS732	16.1°
	Mean SAD	17.2°

3 to 4 distinct references found

Results for MVC-NMF

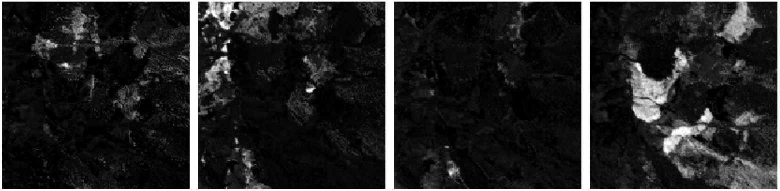


Estimated spectra and associated ones in USGS

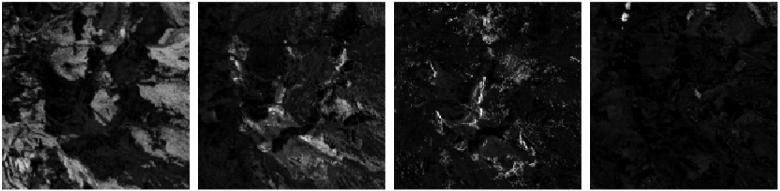
	USGS references	SAD
1	Smaragdite HS290.3B	33.7°
2	Almandine WS475	7.1°
3	Actinolite HS315.4B	12.3°
4	Tumbleweed ANP92 – 2C Dry	9.3°
5	Hypersthene PYX02.h > 250u	42.4°
6	Desert_Varnish GDS141	15.0°
7	Spessartine HS112.3B	15.5°
8	Goethite WS219 (limonite)	15.5°
9	Lepidolite NMNH105541	10.5°
10	Siderite HS271.3B	19.3°
11	Hematite GDS69.d 30 – 45um	10.3°
12	Azurite WS316	16.5
13	Opal WS732	16.1°
	Mean SAD	17.2°

3 to 4 distinct references found

Abundances maps obtained with MDMD-NMF

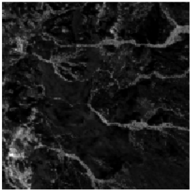


$\xi_1 = 0.97$ $\xi_2 = 0.84$ $\xi_3 = 0.76$ $\xi_4 = 0.90$

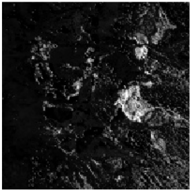


$\xi_5 = 0.83$ $\xi_6 = 1.00$ $\xi_7 = 0.99$ $\xi_8 = 0.79$

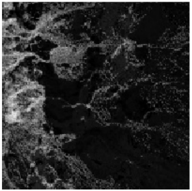
Abundances maps obtained with MDMD-NMF



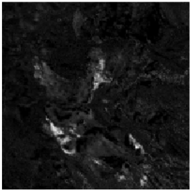
$\xi_9 = 0.72$



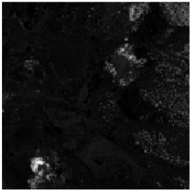
$\xi_{10} = 0.91$



$\xi_{11} = 0.81$



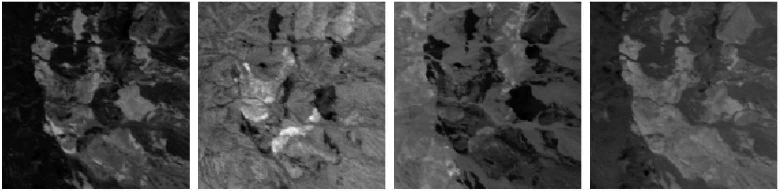
$\xi_{12} = 0.96$



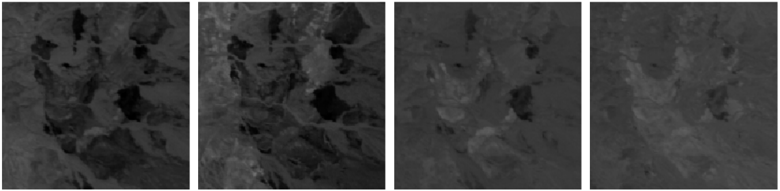
$\xi_{13} = 0.96$

Figure: Abundance maps given by MDMD-NMF and ξ_j for each *endmember j*.

Abundances maps obtained with VCA algorithm



$\xi_1 = 0.33$ $\xi_2 = 0.17$ $\xi_3 = 0.14$ $\xi_4 = 0.19$



$\xi_5 = 0.15$ $\xi_6 = 0.13$ $\xi_7 = 0.13$ $\xi_8 = 0.16$

Abundances maps obtained with VCA algorithm

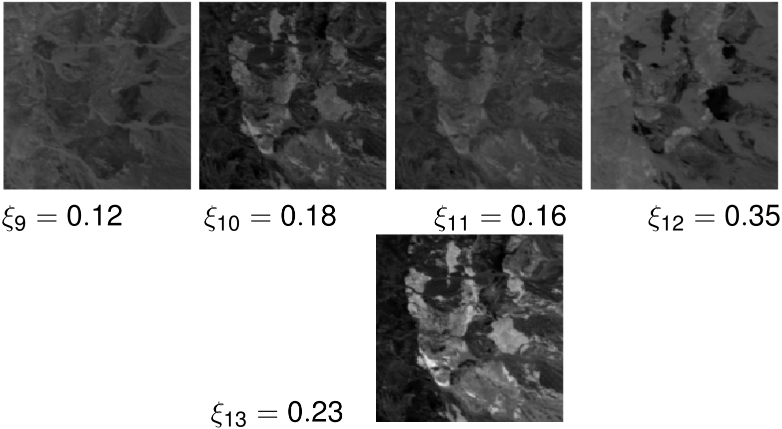


Figure: Abundance maps obtained with VCA and ξ_j for each *endmember j*.

Conclusion

- MDMD-NMF performs well on synthetic data
  - Number of endmembers
  - Robust for $\xi \geq 0.7$
  - No numerical instabilities
- Real data:

Conclusion

- MDMD-NMF performs well on synthetic data
  - Number of endmembers
  - Robust for $\xi \geq 0.7$
  - No numerical instabilities
- Real data:

Conclusion

- MDMD-NMF performs well on synthetic data
  - Number of endmembers
  - Robust for $\xi \geq 0.7$
  - No numerical instabilities
- Real data:

Conclusion

- MDMD-NMF performs well on synthetic data
  - Number of endmembers
  - Robust for $\xi \geq 0.7$
  - No numerical instabilities
- Real data:
  - Good separability of the spectra
  - Good identification power
  - Estimated abundances in the domain of validity

Conclusion

- MDMD-NMF performs well on synthetic data
  - Number of endmembers
  - Robust for $\xi \geq 0.7$
  - No numerical instabilities
- Real data:
  - Good separability of the spectra
  - Good identification power
  - Estimated abundances in the domain of validity

Conclusion

- MDMD-NMF performs well on synthetic data
  - Number of endmembers
  - Robust for $\xi \geq 0.7$
  - No numerical instabilities
- Real data:
  - Good separability of the spectra
  - Good identification power
  - Estimated abundances in the domain of validity

Conclusion

- MDMD-NMF performs well on synthetic data
  - Number of endmembers
  - Robust for $\xi \geq 0.7$
  - No numerical instabilities
- Real data:
  - Good separability of the spectra
  - Good identification power
  - Estimated abundances in the domain of validity

Conclusion

- MDMD-NMF performs well on synthetic data
  - Number of endmembers
  - Robust for $\xi \geq 0.7$
  - No numerical instabilities
- Real data:
  - Good separability of the spectra
  - Good identification power
  - Estimated abundances in the domain of validity

Conclusion

- MDMD-NMF performs well on synthetic data
  - Number of endmembers
  - Robust for $\xi \geq 0.7$
  - No numerical instabilities
- Real data:
  - Good separability of the spectra
  - Good identification power
  - Estimated abundances in the domain of validity