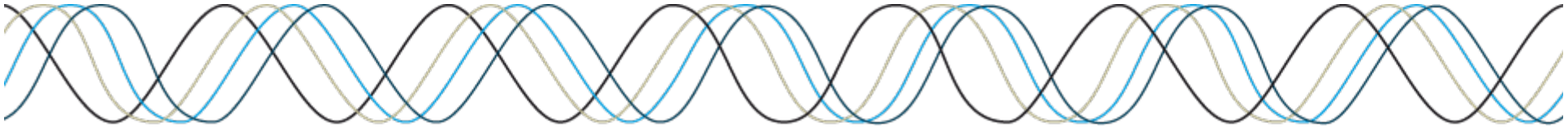
Spectral unmixing



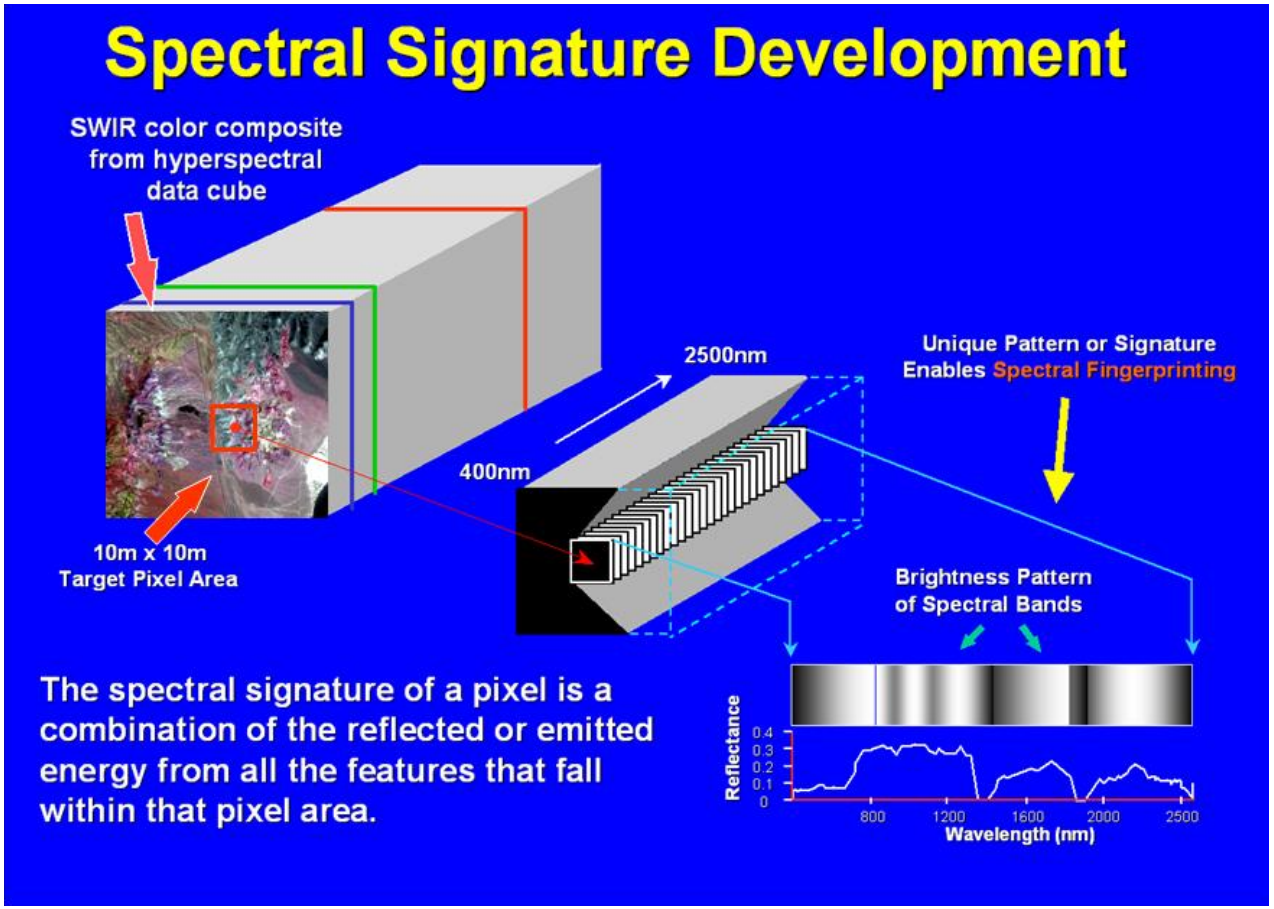
gipsa-lab

Grenoble | images | parole | signal | automatique | laboratoire

Outlines

1. Introduction to sub-pixel analysis
2. Linear mixture model
3. Nonlinear mixture model
4. Bi-linear mixture model

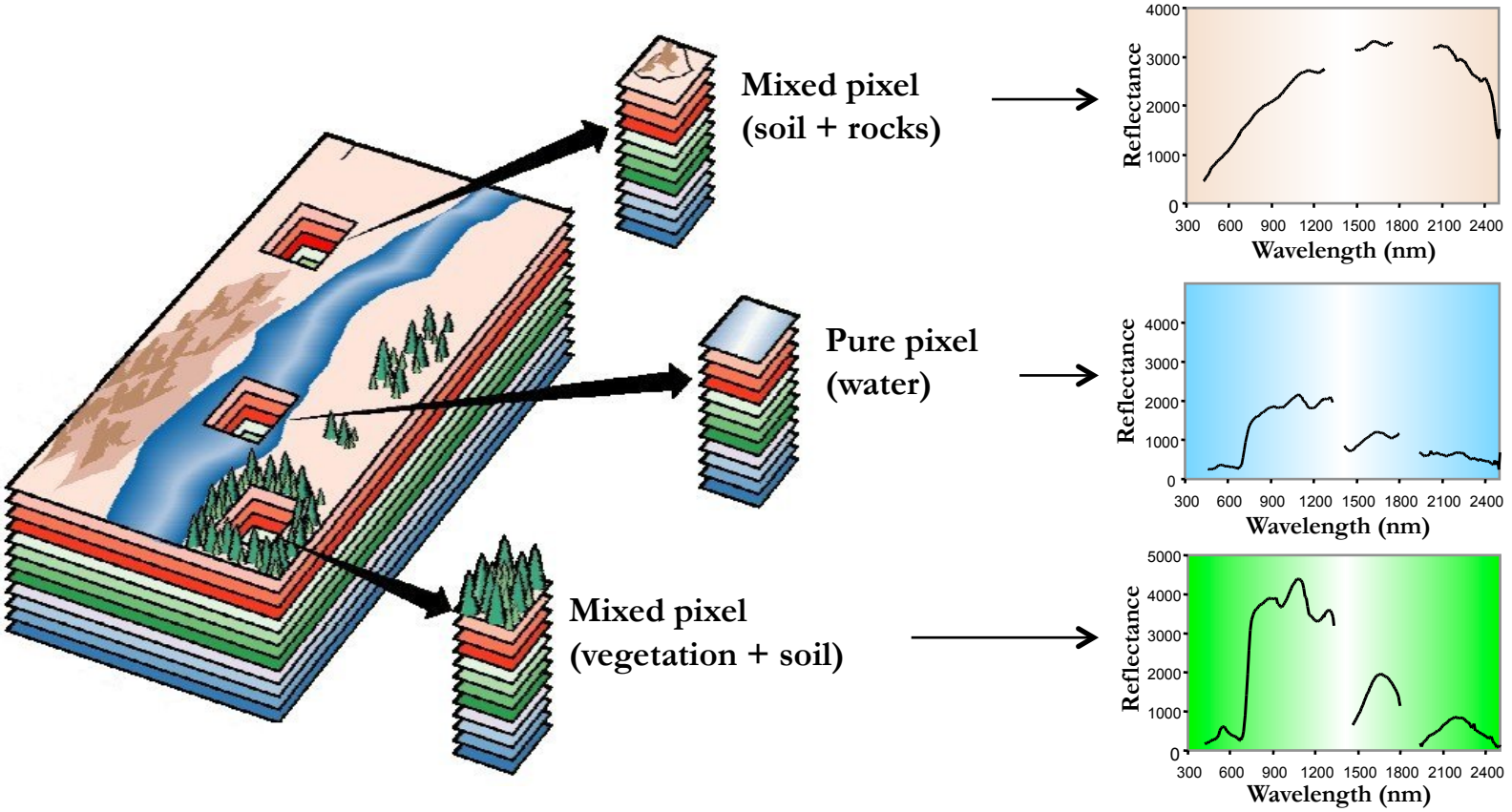
Spectral mixture



Some particularities of hyperspectral data are not to be found in other types of image data:

- Mixed pixels
- Sub-pixel targets

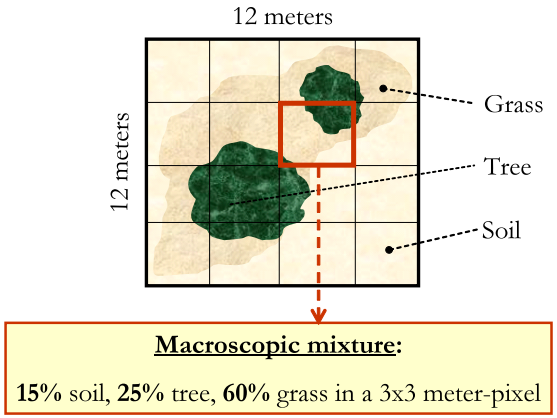
Spectral mixture



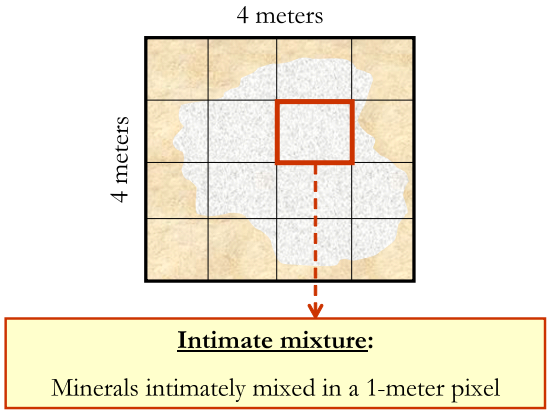
Mixed pixel:
The signal detected by a sensor into a single pixel is frequently a combination of numerous disparate signals.

Spectral mixture

Usually one of the reason why mixed pixels exist depends on one of two reasons:



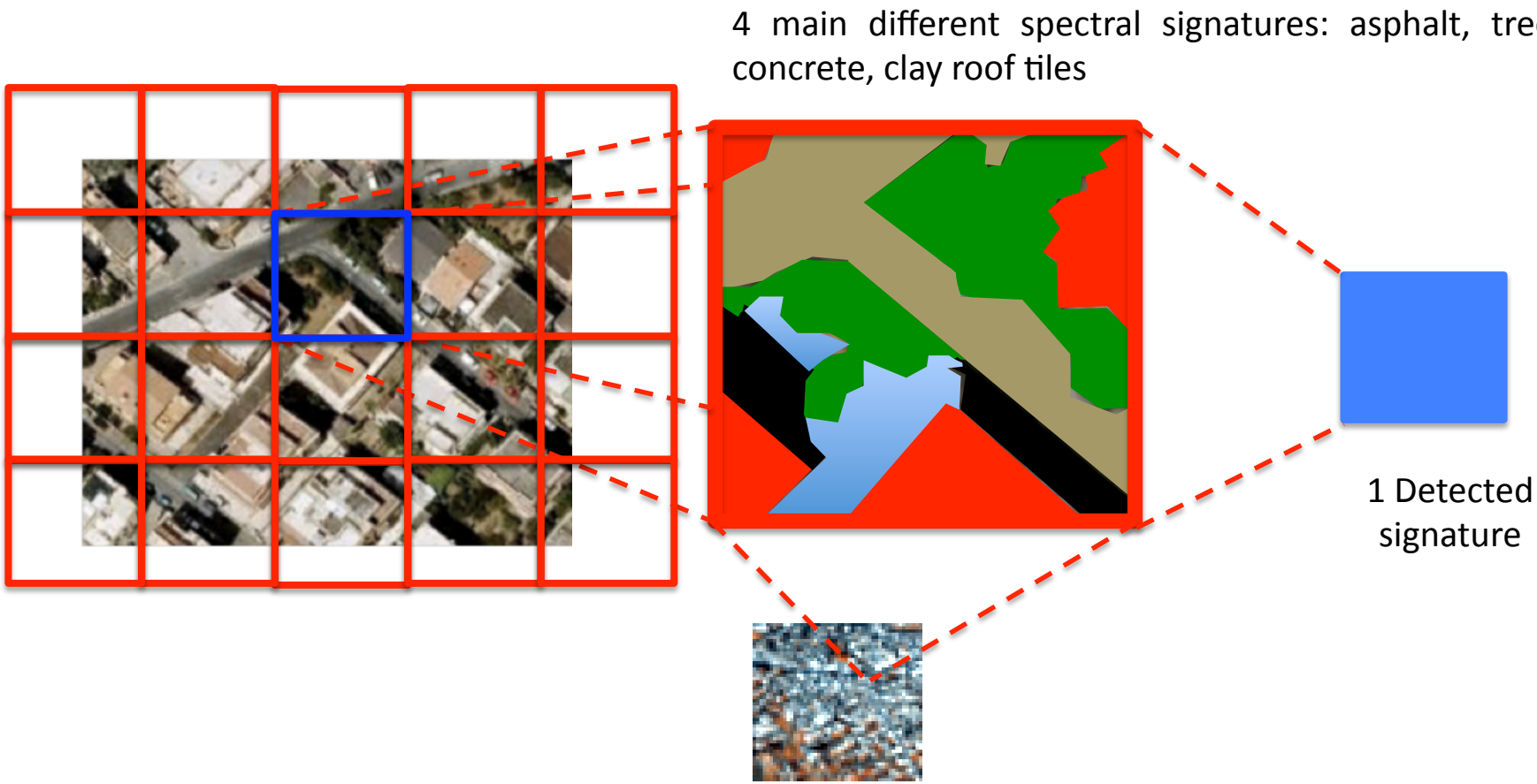
If the resolution is low enough that disparate materials can jointly occupy a single pixel, the resulting spectral measurement will be the composite of the individual spectra.



mixed pixels can result when distinct materials are combined into a homogeneous mixture.

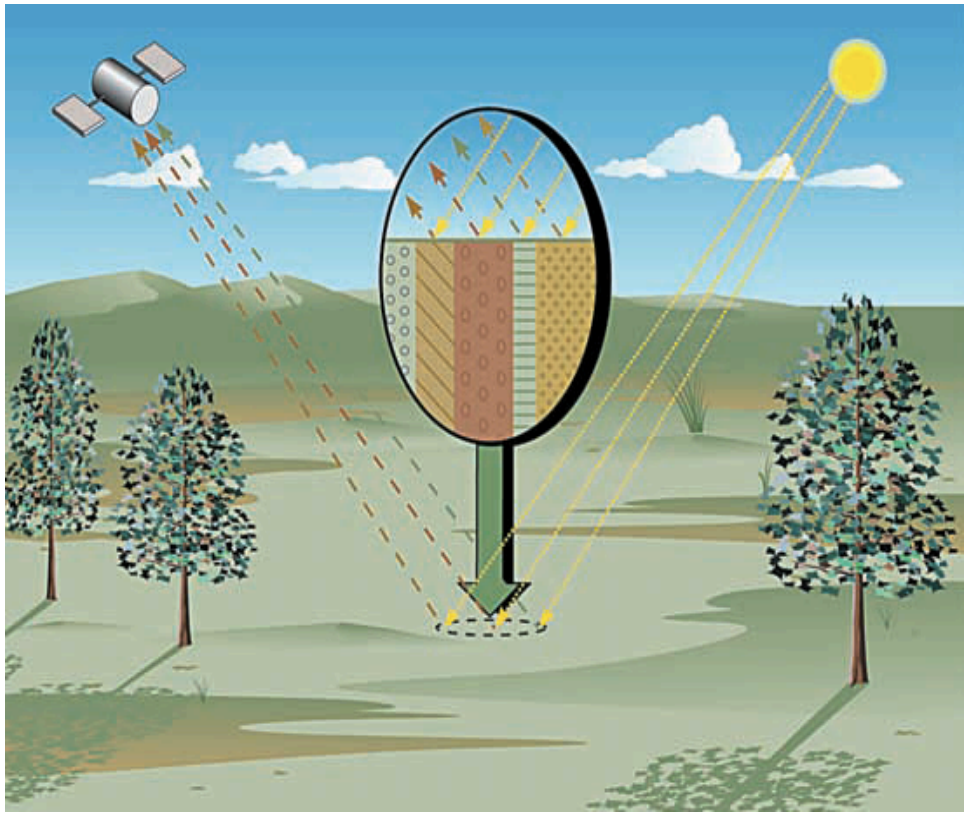
Spectral mixture

In a hyperspectral image, each pixel, depending on the spatial resolution of the sensor, cover a certain area of the ground. The detected signal is a combination of the signals produced by the different types covered by the pixel



Linear spectral mixture

If the endmembers in a pixel appear in spatially segregated patterns similar to a square checkerboard, these systematics are basically linear. In this case the spectrum of a mixed pixel is a linear combination of the endmember spectra weighted by the fractional area coverage of each endmember in a pixel.

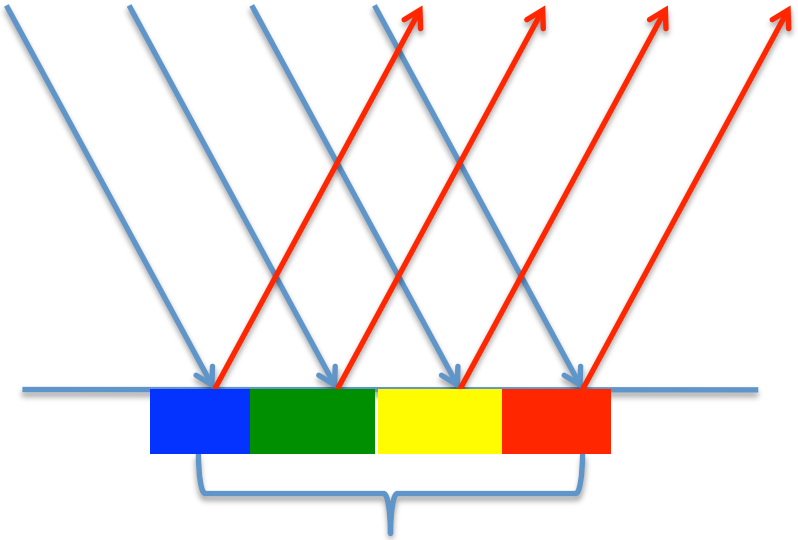


$$x = \sum_{i=1}^{M} a_i s_i + w$$

Linear mixture

Linear spectral mixture

$$E_0 \qquad E_0 e^{-\alpha_1} \quad \cdots \quad E_0 e^{-\alpha_N} \qquad \Rightarrow \qquad E_S = E_0 \sum_{k=1}^{N} e^{-\alpha_k}$$

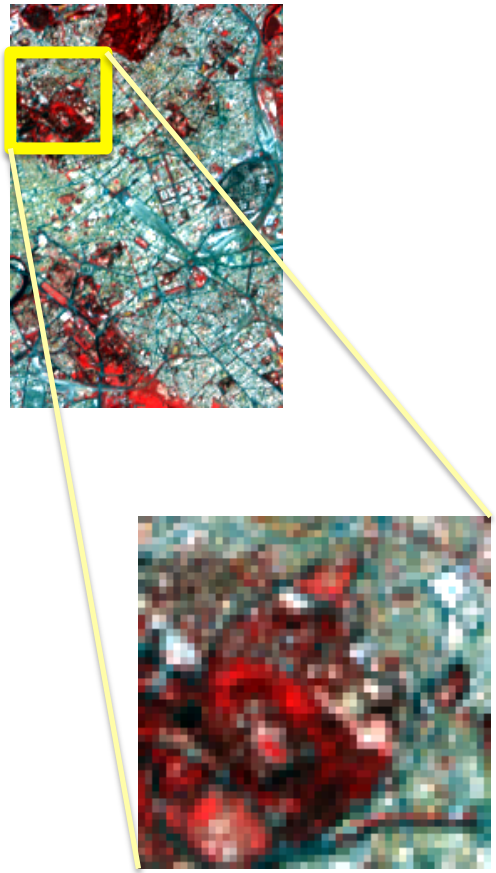


Target pixel

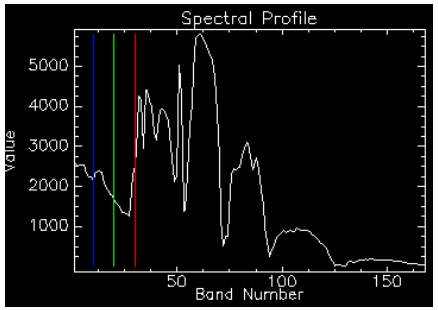
If the total surface area is considered to be divided proportionally between the different materials in the within the pixel then the reflected radiation will convey the characteristics of the associated media with the same proportions.

Linear spectral mixture

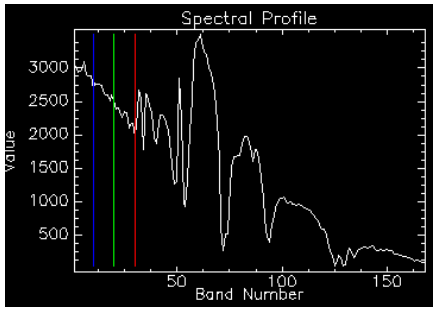
Linear mixture



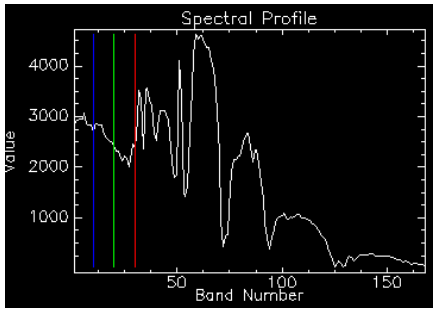
Vegetation (60%)



Buildings (40%)



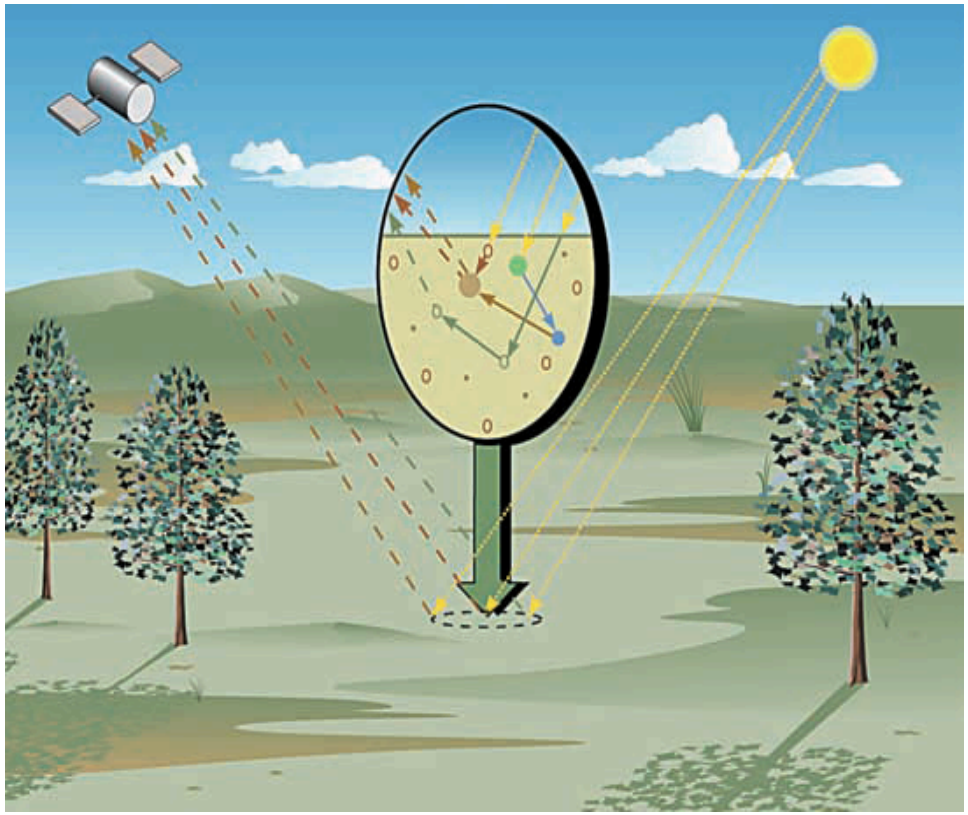
Mixed pixel
0.6*veg+0.4*build



There exists a linear relationship between the fractional abundance of the substances comprising the area being imaged and the spectra in the reflected radiation.

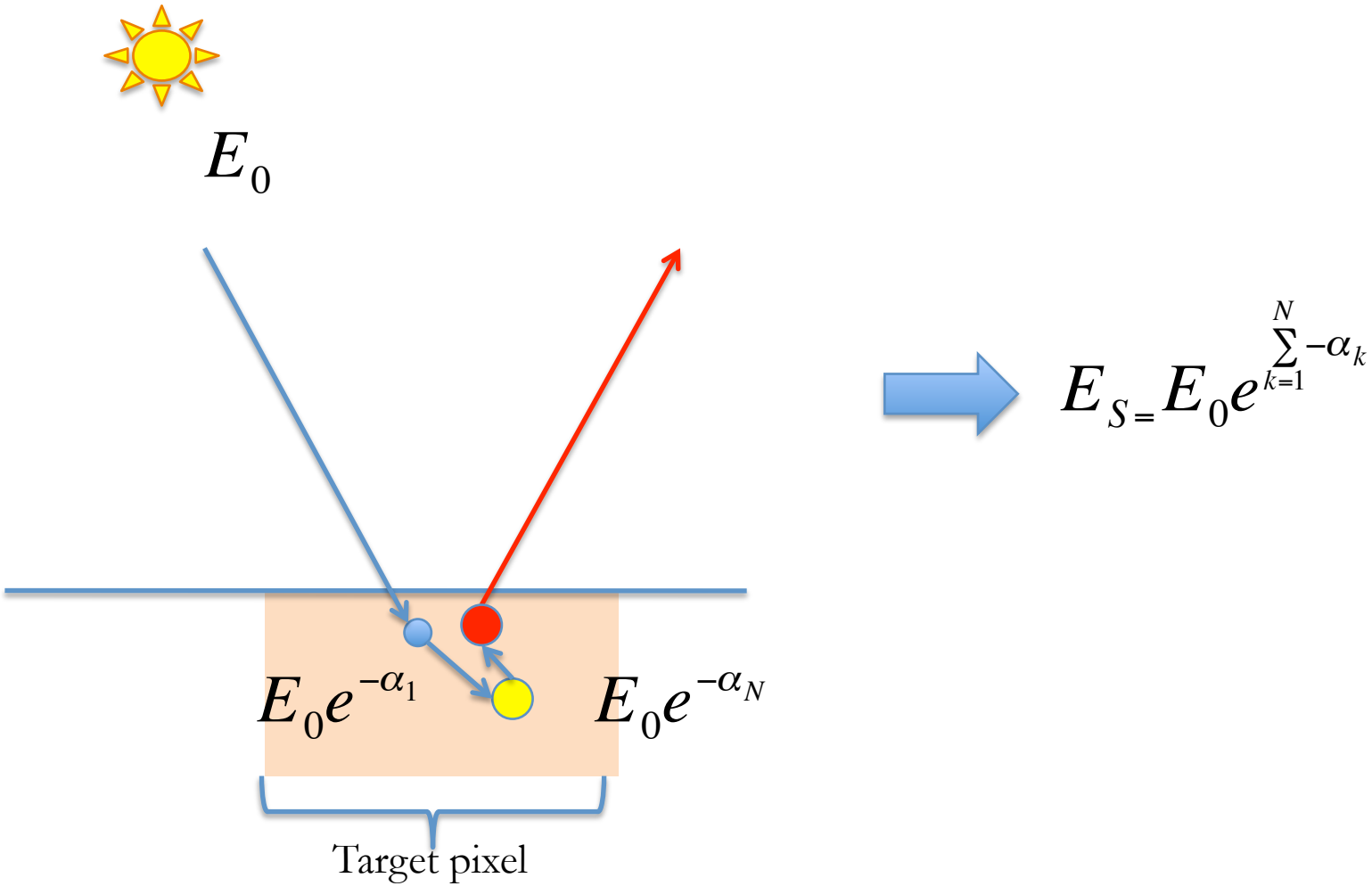
Non linear spectral unmixing

If the components of interest in a pixel are in an intimate association, like sand grains of different composition in a beach deposit, light typically interacts with more than one component as it is multiply scattered, and the mixing systematics between these different components are nonlinear.

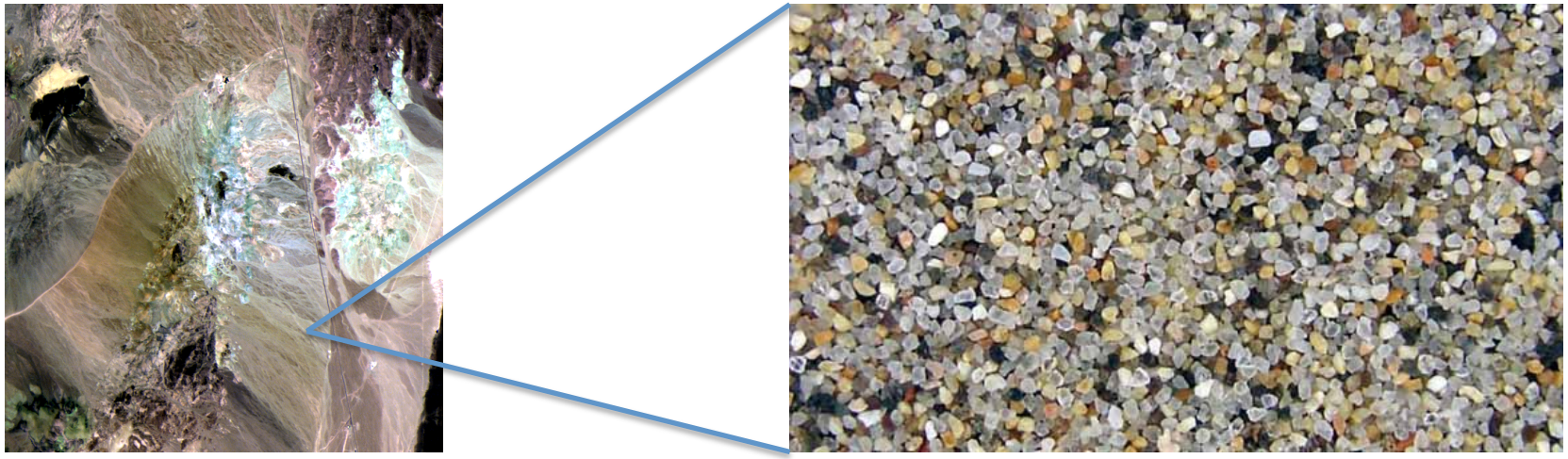


Nonlinear mixture

Nonlinear spectral mixture

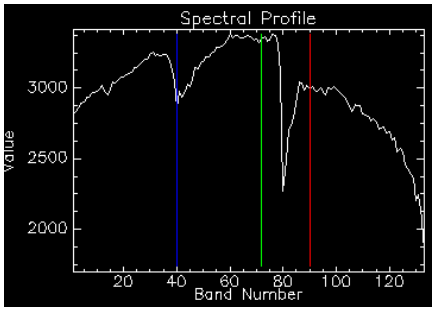


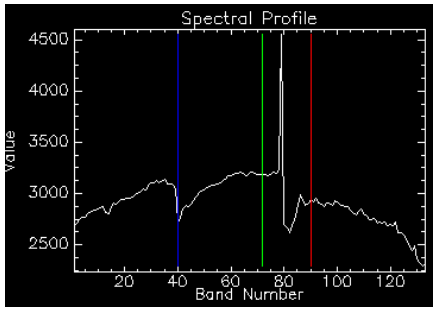
Non linear spectral unmixing

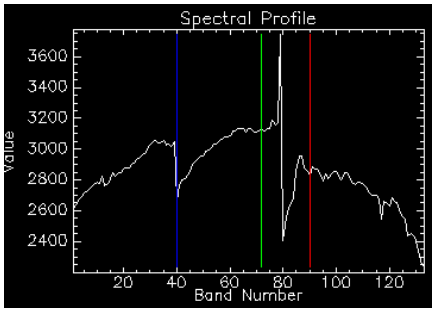


Nonlinear mixture

Calcite **Kaolinite**



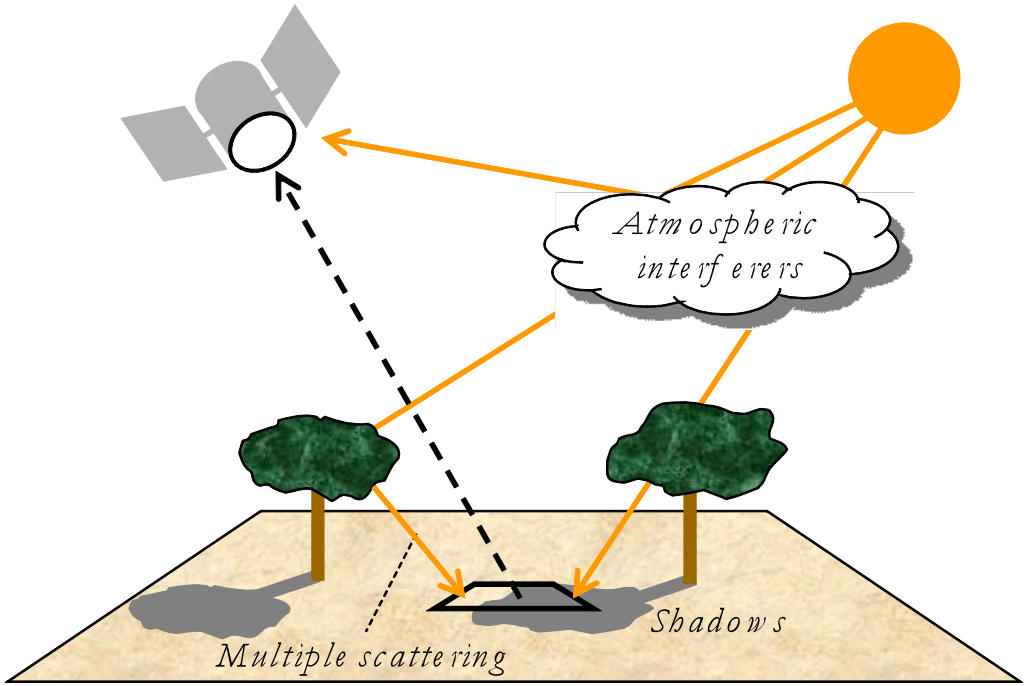




Nonlinear spectral mixture

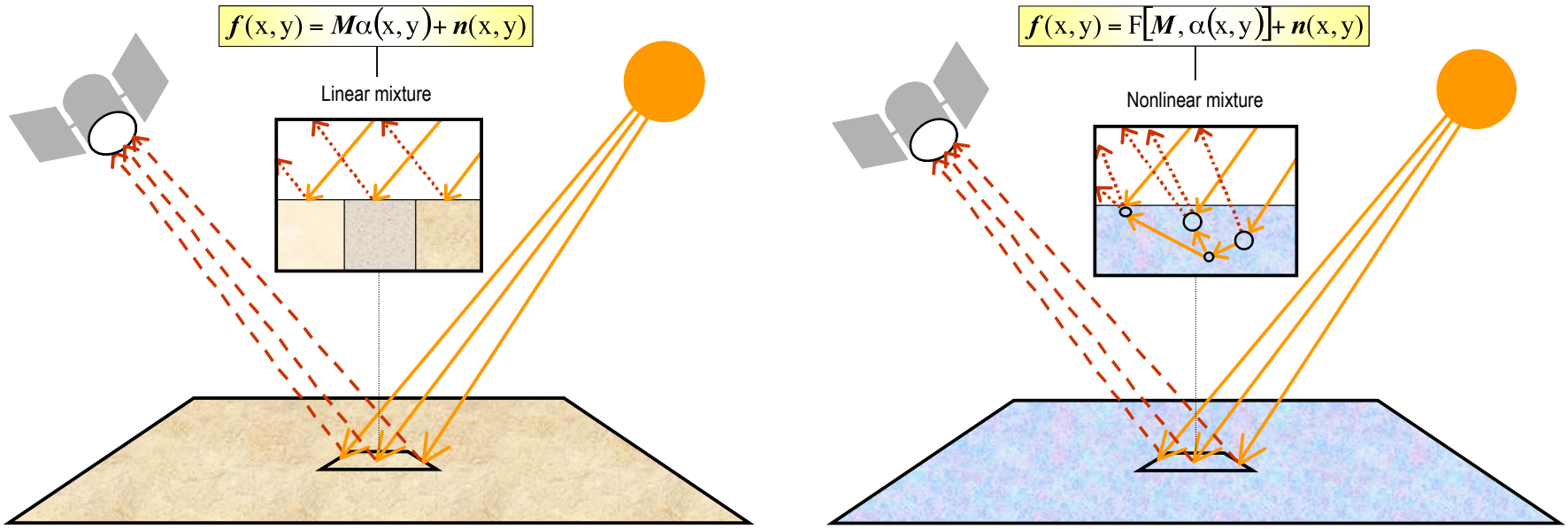
Such nonlinear effects have been recognized in spectra of:

- particulate mineral mixtures
- aerosols and atmospheric particles
- vegetation and canopy.



Linear vs nonlinear mixing model

- *Linear mixture model*
  - ✓ *Assumes that endmember substances are sitting side-by-side within the FOV.*
- *Nonlinear mixture model*
  - ✓ *Assumes that endmember components are randomly distributed throughout the FOV.*
  - ✓ *Multiple scattering effects.*



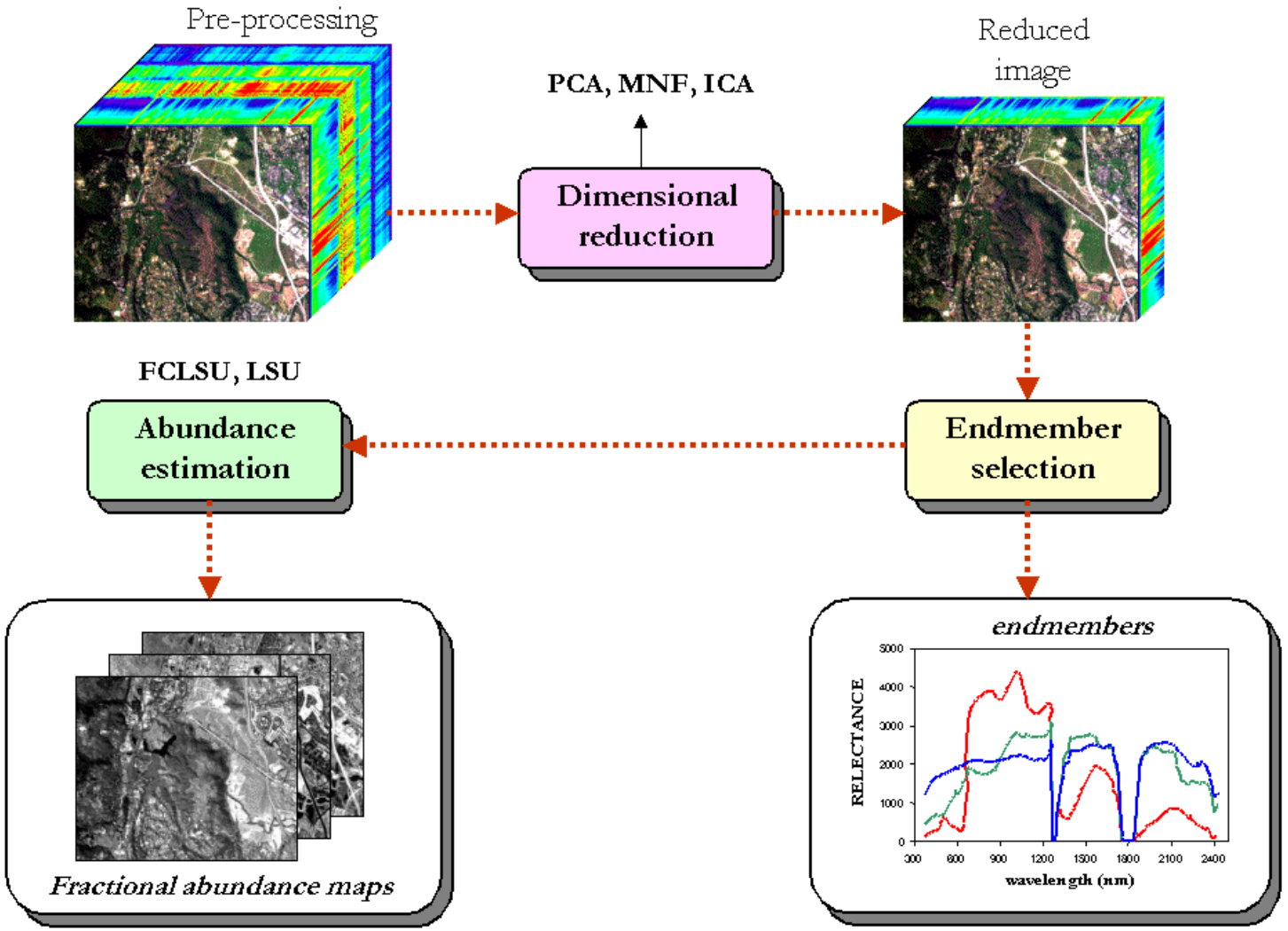
spectral unmixing

The basic premise of mixture modeling is that within a given scene, the surface is dominated by a small number of distinct materials that have relatively constant spectral properties.
These distinct substances (e.g., water, grass, mineral types) are called *endmembers*, and the fractions in which they appear in a mixed pixel are called *fractional abundances*.

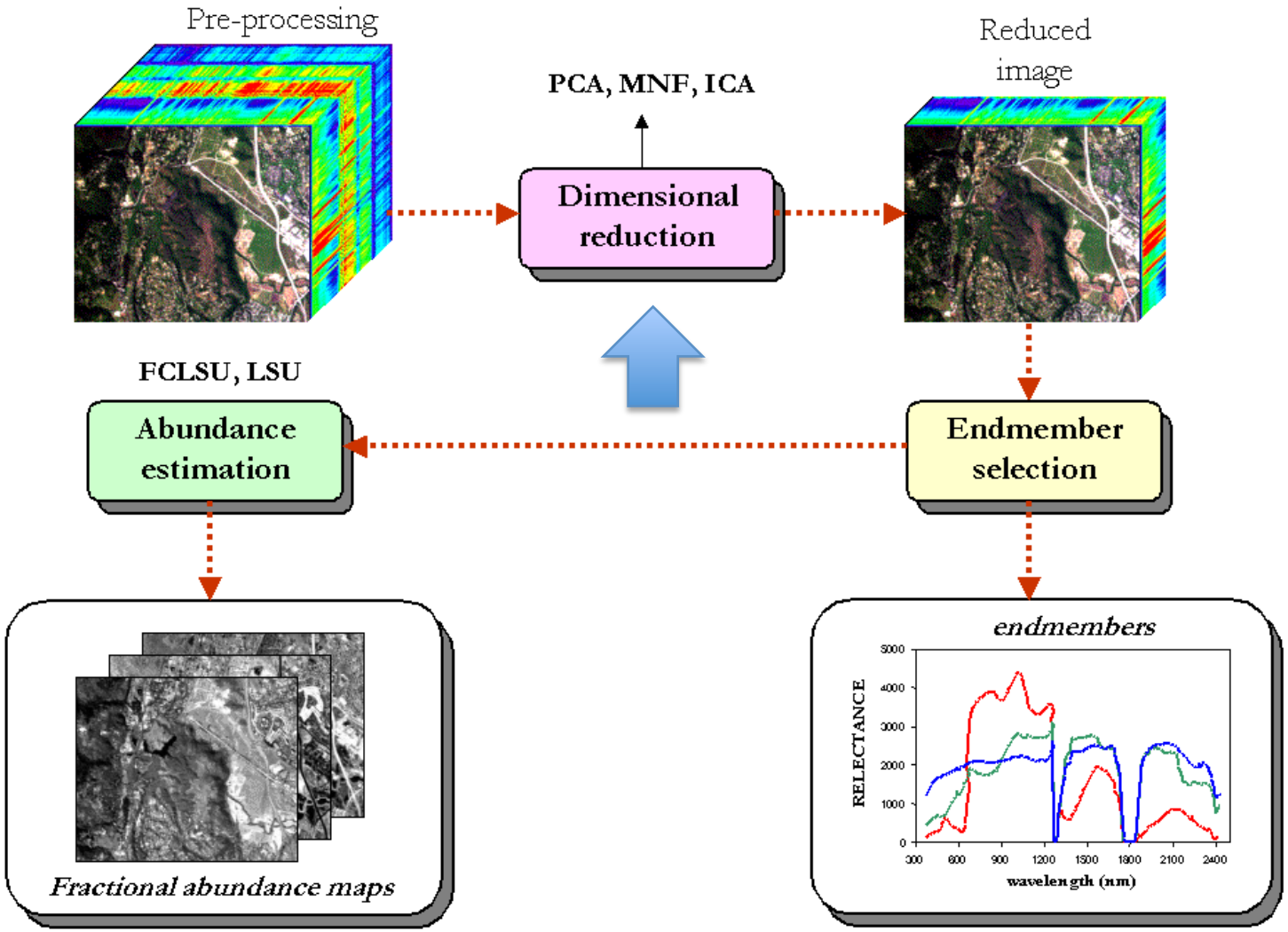
- The goal is to find the endmembers that can be used to "unmix" other mixed pixels

- A crucial issue is how to find spectral endmembers.

spectral unmixing scheme



Dimensionality reduction



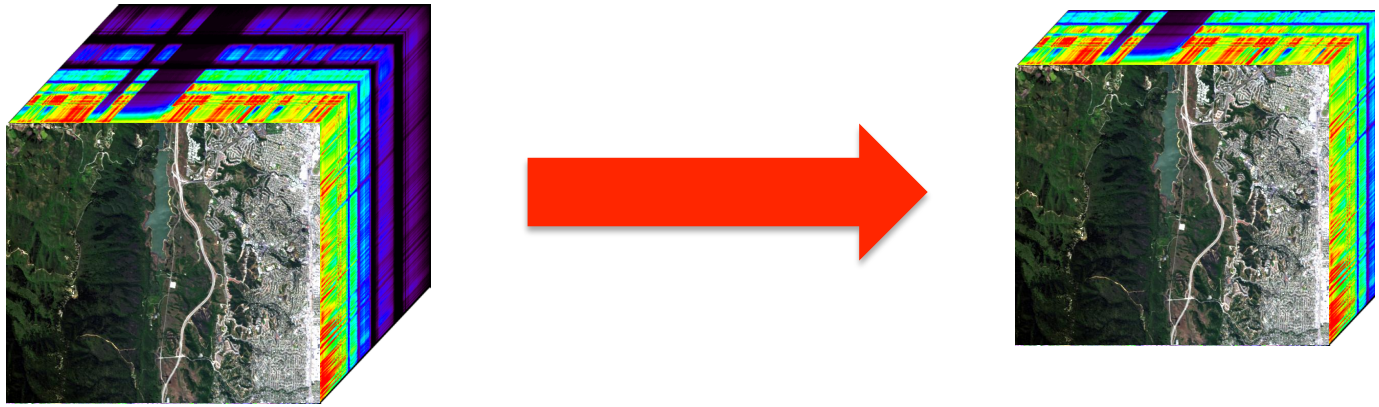
Dimensionality reduction

A dataset composed of hundreds of high correlated narrowband channels may cause problems in the:

•processing phase (complexity)

•inversion phase (Hughes phenomenon).

dimensionality reduction may become a key parameter to obtain a good performance



Dimensionality reduction

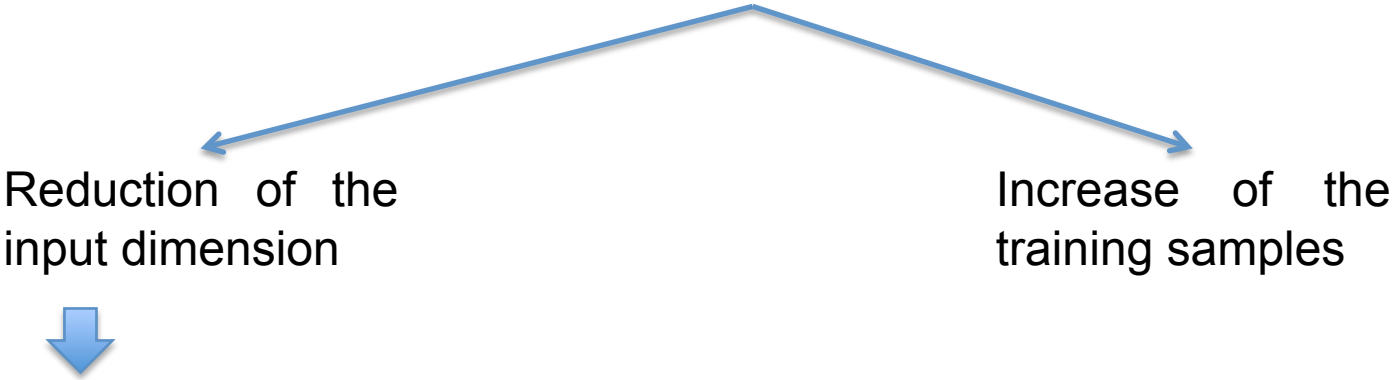
The dimensionality reduction pre-processing is extremely useful to enhance the computing performances. It depends on the intrinsic information retained by the HSI and can strongly influence the result of the unmixing process.

- Determination of intrinsic dimensionality (ID) for remotely sensed imagery is a very challenging problem. *ID is defined as the minimum number of parameters required to account for the observed properties of the data*. The true dimensionality of multivariate data is difficult to determine in practice due to the unknown ground truth.

- *Virtual Dimensionality (VD) is defined as the minimum number of spectrally distinctive signal sources*. It can be used as the reference for the number of endmembers. In general, VD is much larger than ID.

Dimensionality reduction – Hughes effect

Hughes demonstrated that having a fixed number of training samples, the ability of a classification algorithm to correctly predict the effective result, decreases as the input dimension increase.



Intrinsic dimension

Any low-dimensional data space can trivially be turned into a higher dimensional space by adding redundant or randomized dimensions, from this assumption many high-dimensional data sets can be reduced to lower dimensional data without significant information loss.

Hughes, G.F., 1968. "On the mean accuracy of statistical pattern recognizers", IEEE Transactions on Information Theory, IT-14:55-63.

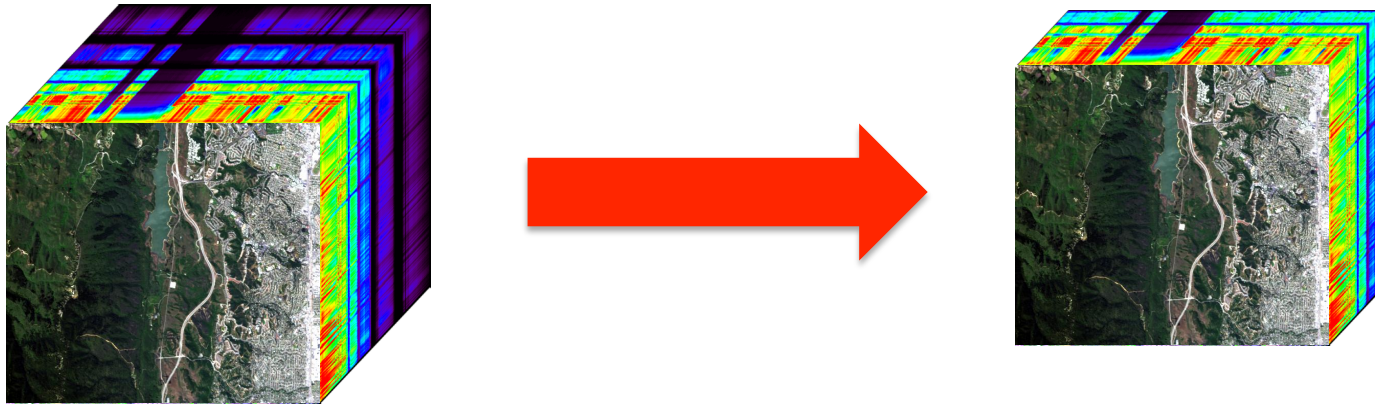
Feature selection/extraction

Feature reduction, or dimensionality reduction, is the process of reducing the number of features of the considered data.

Dimensionality reduction can be divided into:

- **"feature selection" algorithms:** Selection of a sub-optimal subset of the original set of features

- **"feature extraction" algorithms:** Projection of the original feature space into a lower dimensional subspace that preserves most of the information



Feature selection

Feature selection algorithms perform a search through the space of feature subsets, and, as a consequence, must address four basic issues affecting the nature of the search:

1. **Starting point.** Selecting a point in the feature subset space from which to begin the search can affect the direction of the search. One option is to begin with no features and successively add attributes. In this case, the search is said to proceed forward through the search space. Conversely, the search can begin with all features and successively remove them. In this case, the search proceeds backward through the search space. Another alternative is to begin somewhere in the middle and move outwards from this point.

2. **Search organization.** An exhaustive search of the feature subspace is prohibitive for all but a small number of features. With N initial features there exist 2^N possible subsets.

3. **Evaluation strategy.** A criterion to estimate the final accuracy of the feature subsets.

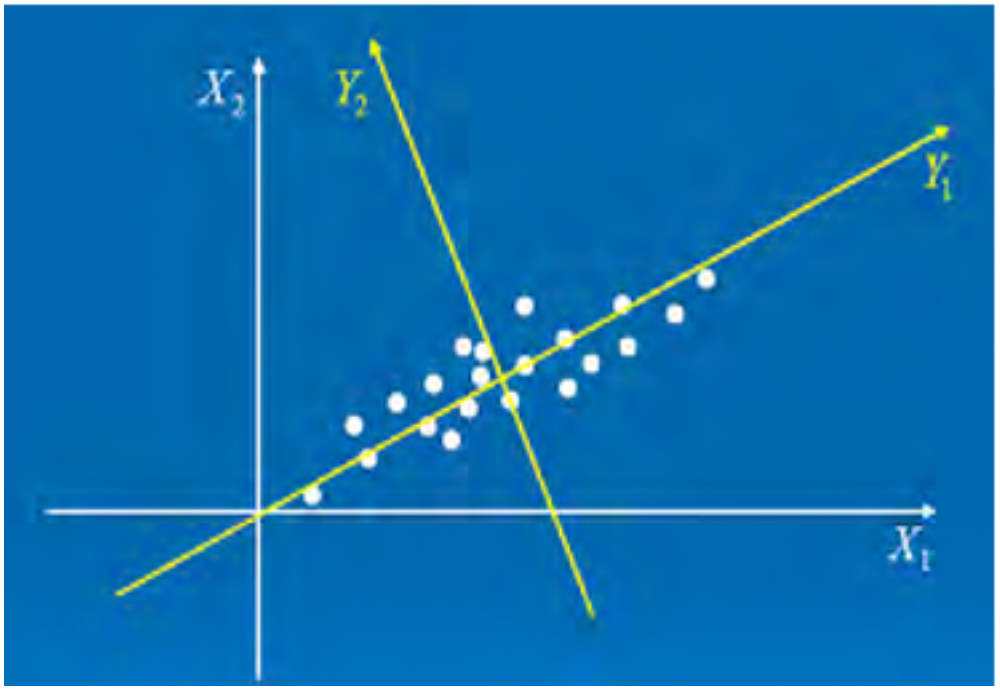
4. **Stopping criterion.** A feature selector must decide when to stop searching through the space of feature subsets. Depending on the evaluation strategy, a feature selector might stop adding or removing features when none of the alternatives improves upon the merit of a current feature subset. Alternatively, the algorithm might continue to revise the feature subset as long as the merit does not degrade.

PCA

Non-parametric approach for relevant information extraction from a redundant+noisy data

Eliminate the redundance by means of autocorrelation

Geometrically speacking, PCA change the reference system to enhance the data structure



PCs have the following properties:

•Are linear combinations of the original variables

•Are mutually not correlated

MNF

MNF is a linear transformation that consists of two separate PCA rotations and a noise whitening step:

•Use the noise covariance matrix to decorrelate and rescale the noise in the data (noise whitening). This results in transformed data in which the noise has unit variance and no band-to-band correlations.

•Perform a standard Principal Components transformation of the noise-whitened data.

The MNF transform, like the PCA transform, is an eigenvector procedure, but based on the covariance structure of the noise in the image data set.

Goal of the MNF transform is to select component in a way that maximizes the signal-to-noise ratio (rather than the information content).

Independent Component analysis

ICA is a quite powerful technique to separate a multivariate signal into additive subcomponents supposing the mutual statistical independence of the non-Gaussian source signals.

The goal is to find a linear representation of non-gaussian data so that the components are statistically independent, or as independent as possible.

Such a representation seems to capture the essential structure of the data in many applications, including feature extraction and signal separation.

$$x_j = a_{j1}s_1 + a_{j2}s_2 + \ldots + a_{jn}s_n$$

Independent Component analysis – The cocktail party problem

Imagine to be in a room where two people are speaking simultaneously. You have two microphones, which you hold in different locations. The microphones give you two recorded time signals, which we could denote by **x1(t)** and **x2(t)**, with **x1** and **x2** the amplitudes, and **t** the time index. Each of these recorded signals is a weighted sum of the speech signals emitted by the two speakers, which we denote by **s1(t)** and **s2(t) .** We could express this as a linear equation:

$$x_1(t) = a_{11}s_1 + a_{12}s_2$$

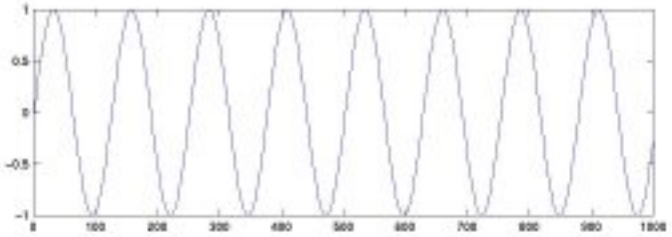
$$x_2(t) = a_{21}s_1 + a_{22}s_2$$

It would be very useful if you could now estimate the two original speech signals **s1(t)** and **s2(t)** , using only the recorded signals **x1(t)** and **x2(t)**.

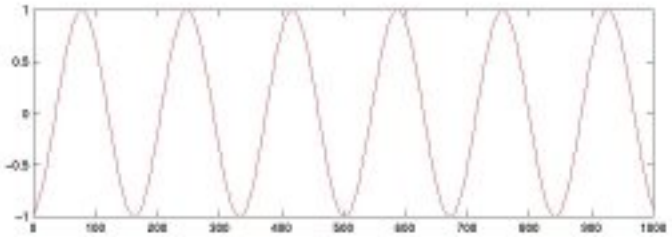
If we knew the parameters ***a***, we could solve the linear equation by classical methods. If you don't know the ***a***, the problem is considerably more difficult.

Independent Component analysis – The cocktail party problem

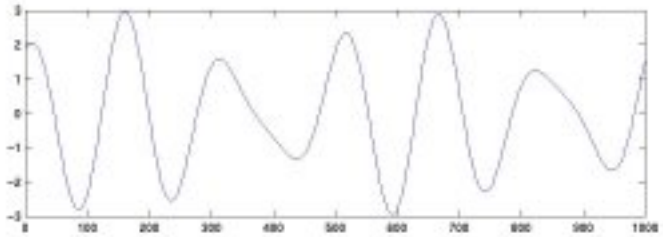
s1



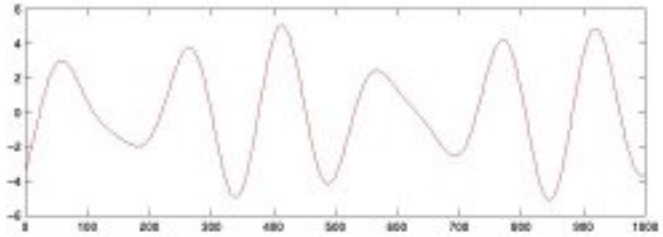
s2



x1(t)



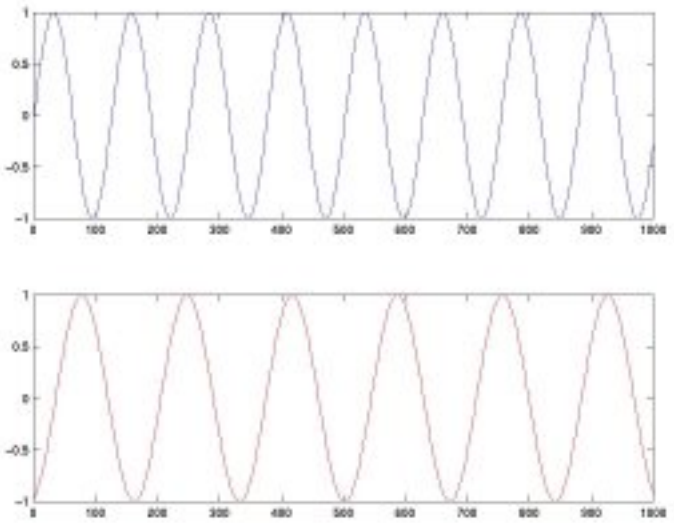
x2(t)



Independent Component analysis

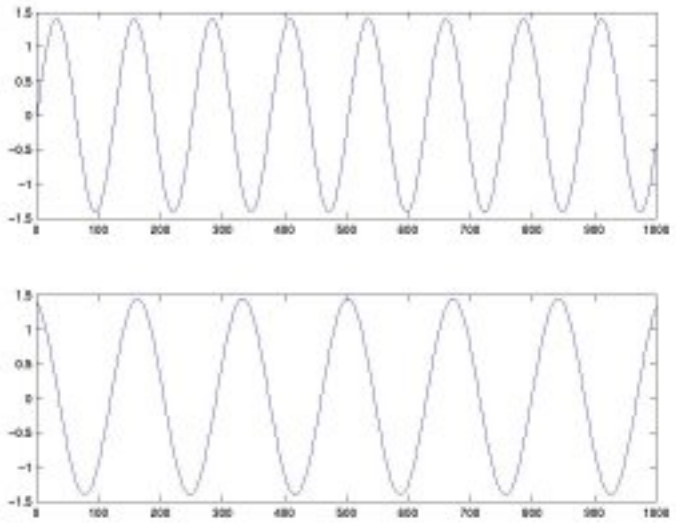
We can use some information on the statistical properties of the signals ***s(t)*** to estimate the a. To do this it is enough to assume that s1(t) and s2(t) , at each time instant **t**, are statistically independent.

Original



Differences in amplitude

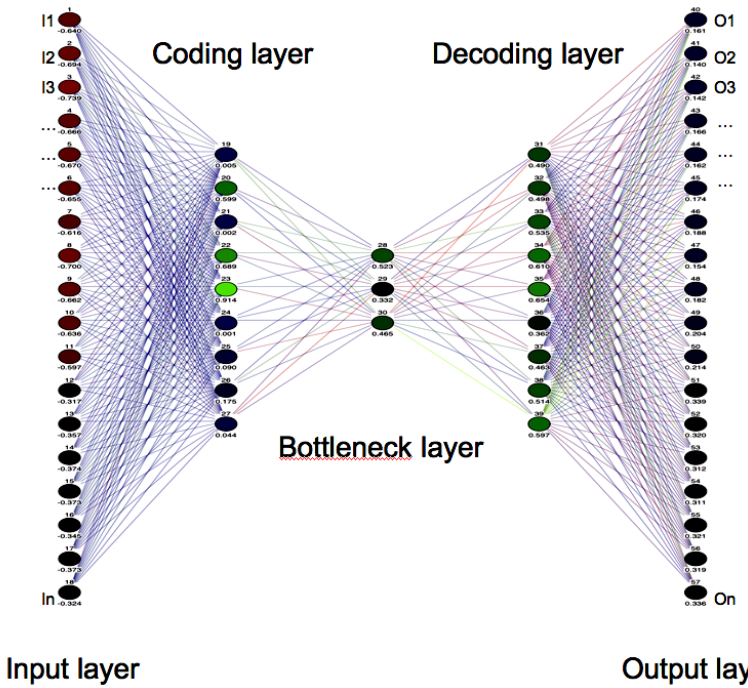
Reconstructed

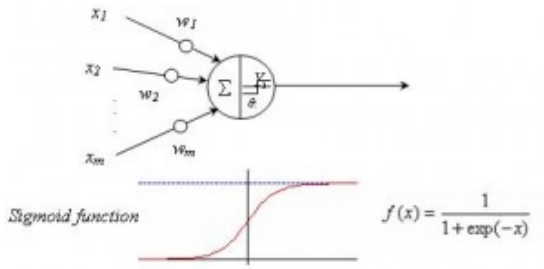


ICA algorithms return signals that are very close to the original source signals

Nonlinear Principal Components Analysis

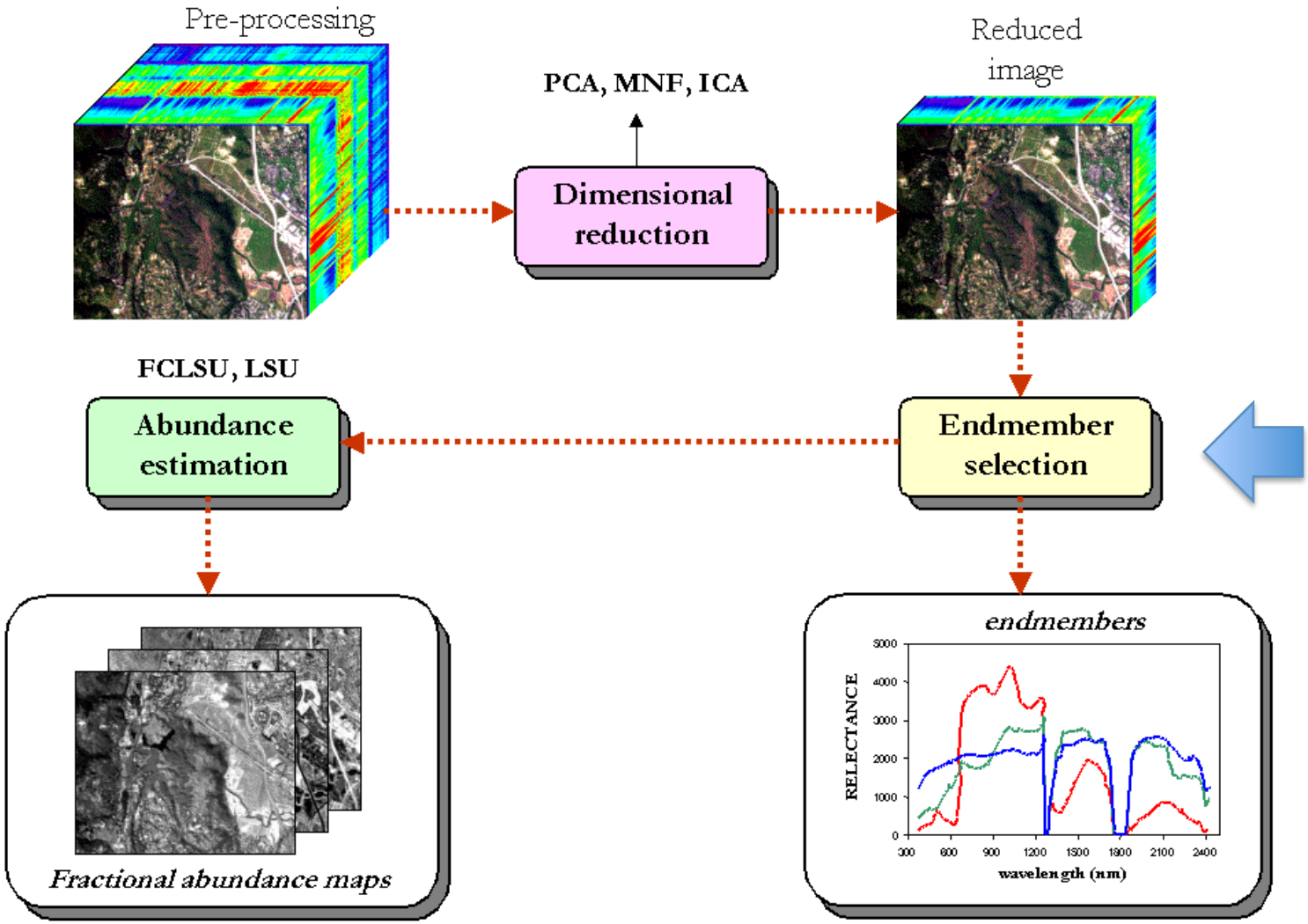
Nonlinear generalization of the standard PCA, performed by Autoassociative Neural Networks (AANN)





Multi-layer neural networks (NN) of a conventional type, featuring feed-forward connections and sigmoidal nodal transfer functions, trained by back-propagation or similar algorithms.

Endmember extraction



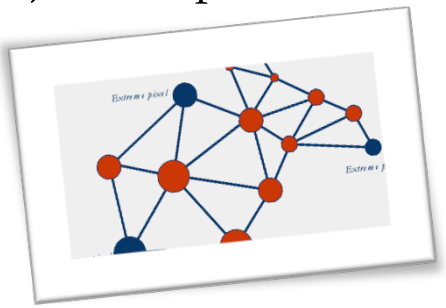
Endmember extraction

Endmember extraction algorithms: simultaneous versus sequential

Different approaches depending on the actual **implementation**:

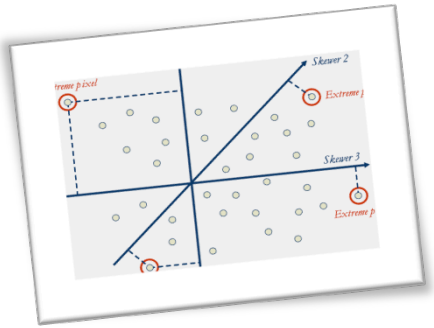
Simultaneous algorithms: they assume that the number of endmembers *p* is known a priori. For each value of *p*, they recalculate all the endmembers (they do not take advantage of previously calculated *p*-1 endmembers). Example: N-FINDR.

N-FINDR algorithm



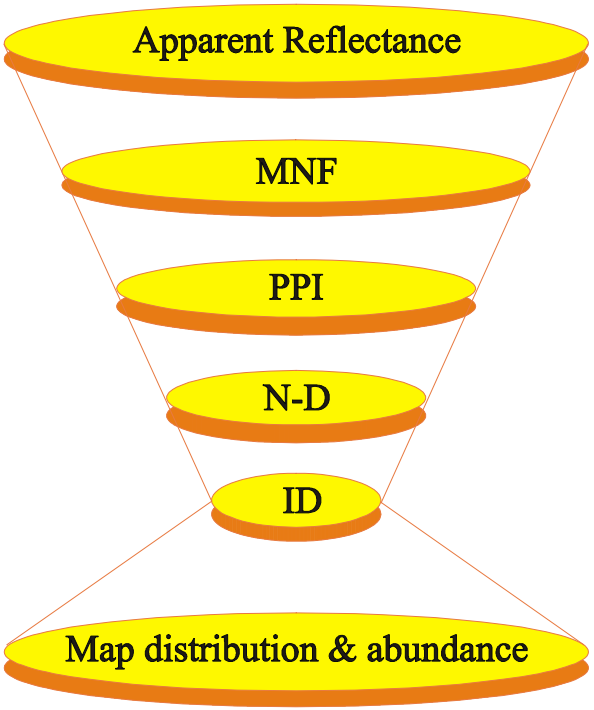
Sequential algorithms: they produce a set of endmembers in sequential order, i.e., if we are searching for *p* endmembers, a subset of previously calculated *p*-1 endmembers would always be in the final set. Examples: PPI, OSP, VCA.

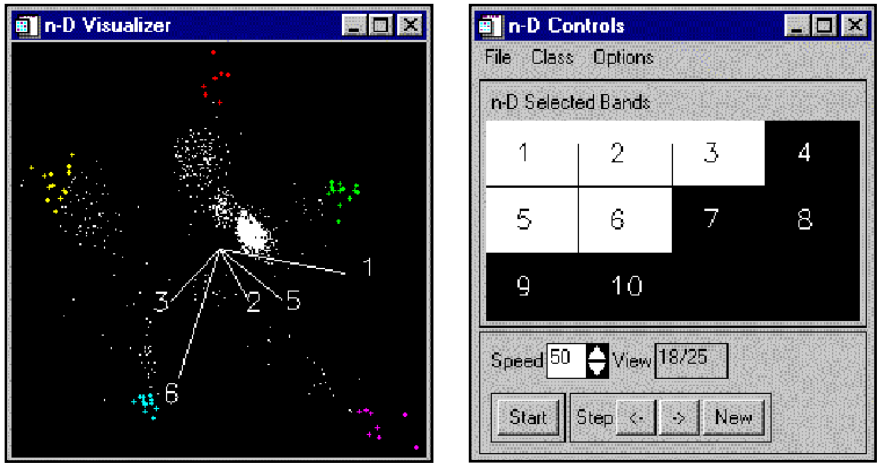
Pixel Purity Index (PPI)



Pixel Purity Index (PPI)

- Parameters: *k* (number of skewers) and *t* (cut-off threshold)
- It is not an iterative algorithm (no convergence criteria)
- No estimation of number of endmembers *a priori*
- Manual intervention required to select a final set of endmembers

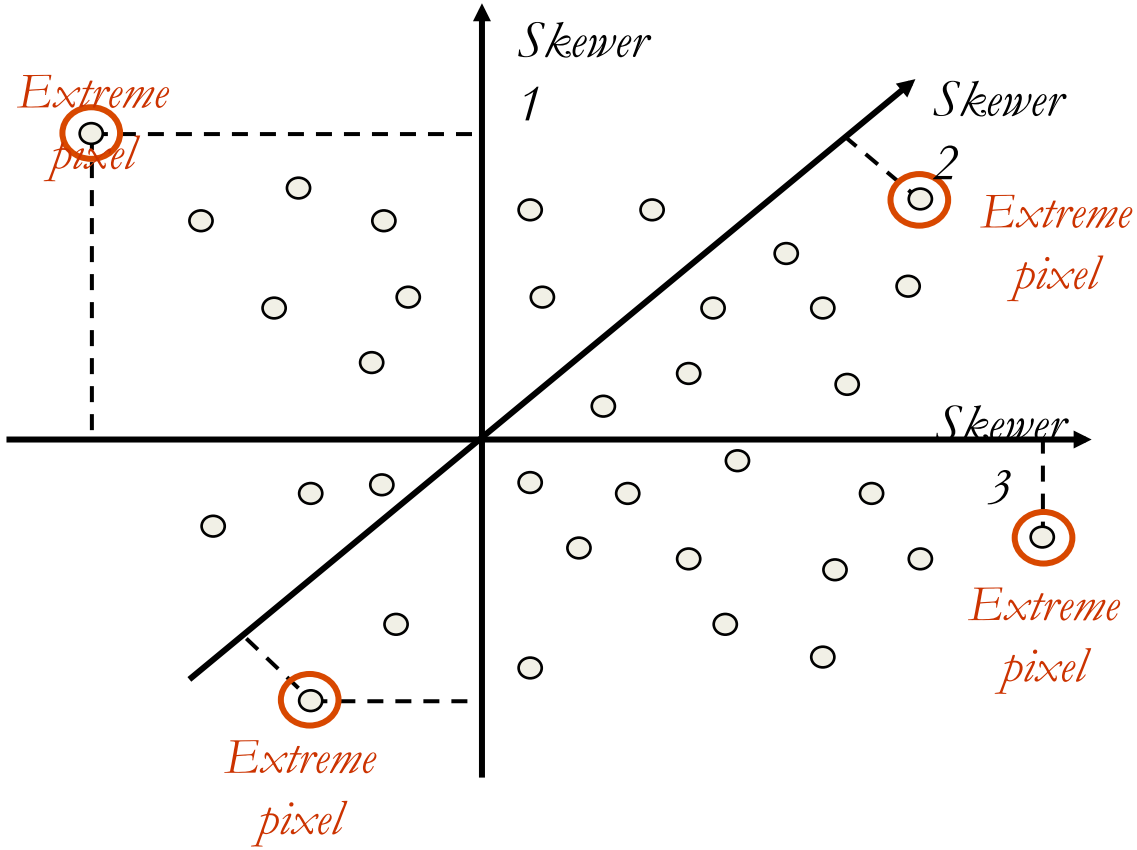




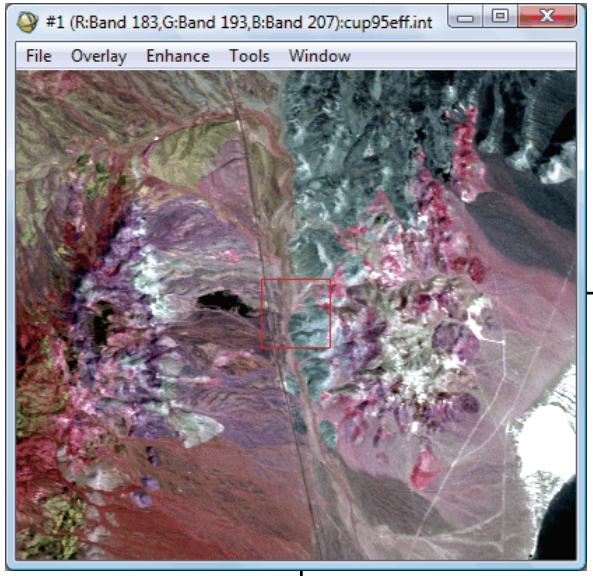
ENVI's N-Dimensional visualization tool

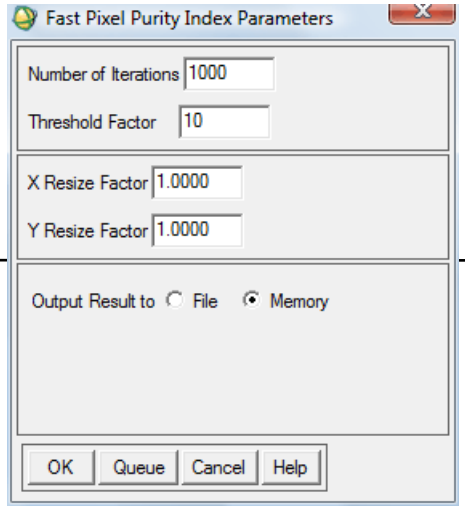
Pixel Purity Index (PPI)

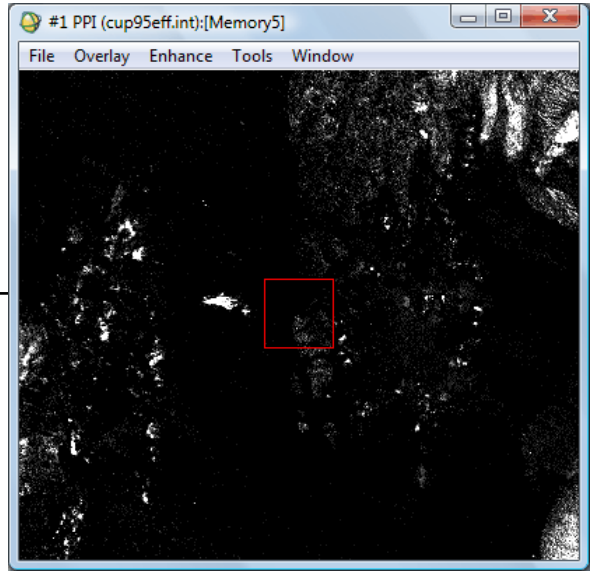
Pixel Purity Index (PPI) - Works by projecting each pixel onto one vector from a set of random vectors spanning the reflectance space. A pixel receives a score when it represent an extremum of all the projections. Pixels with the highest scores are deemed to be spectrally pure.

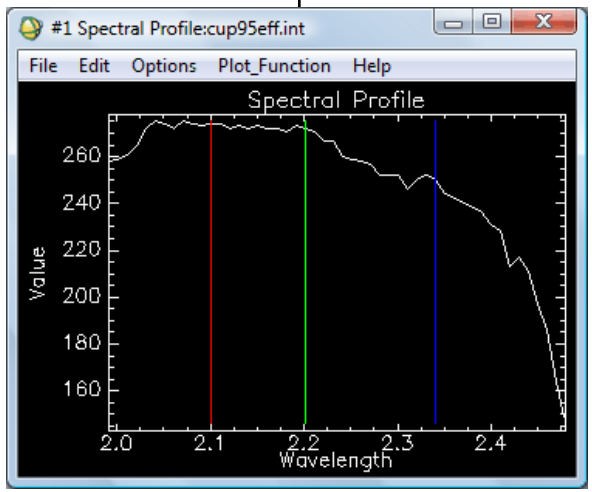


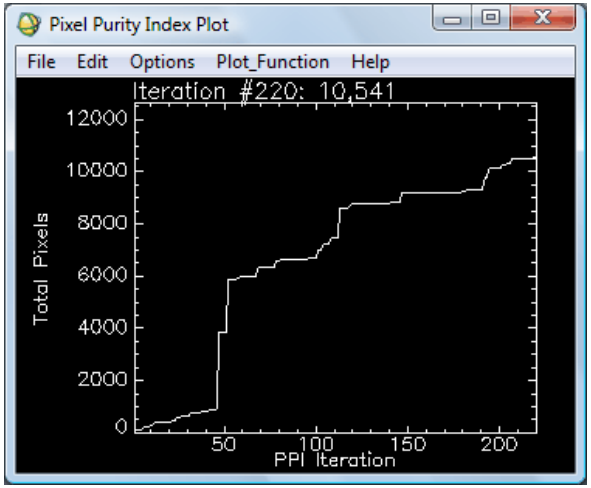
Pixel Purity Index (PPI)

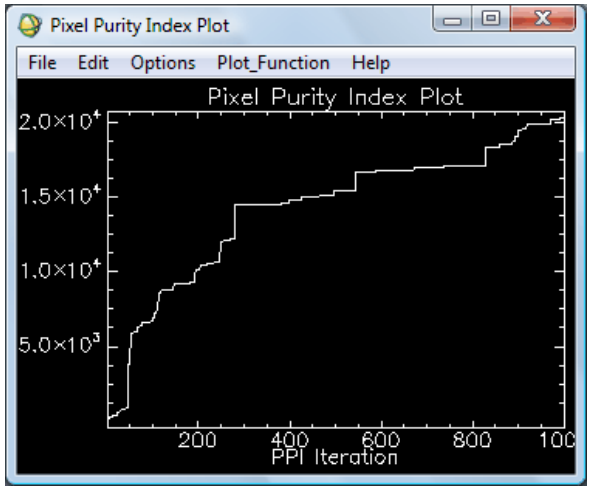










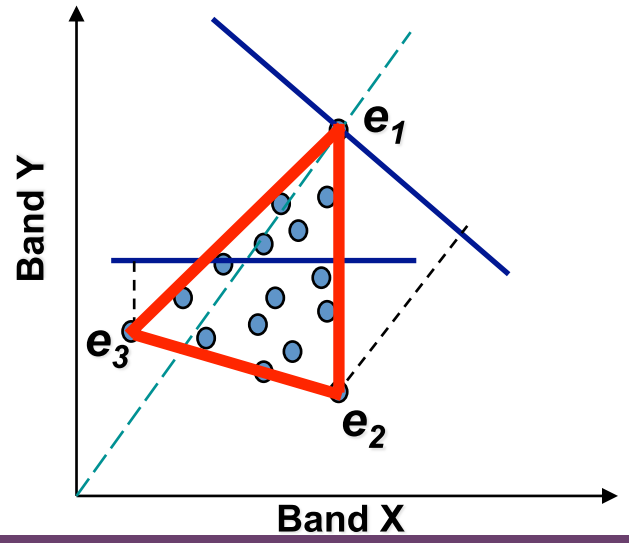


Demo will be performed using ITTVIS Envi 4.5 (http://www.ittvis.com)

Orthogonal Subspace Projection

Spectral-based method developed by Harsanyi and Chang (IEEE-TGRS 1994):

1. Select the brightest pixel from the scene as the initial one
2. Apply an orthogonal subspace projection operator to all the pixels in the original image to identify the most orthogonal pixel to the first one (second *endmember*)
3. Apply the orthogonal subspace projector again to all pixels, looking for the most orthogonal one with regards to the first two (third *endmember*)
4. Repeat the process until a predetermined number of *endmember*s is found



Orthogonal subspace projector:

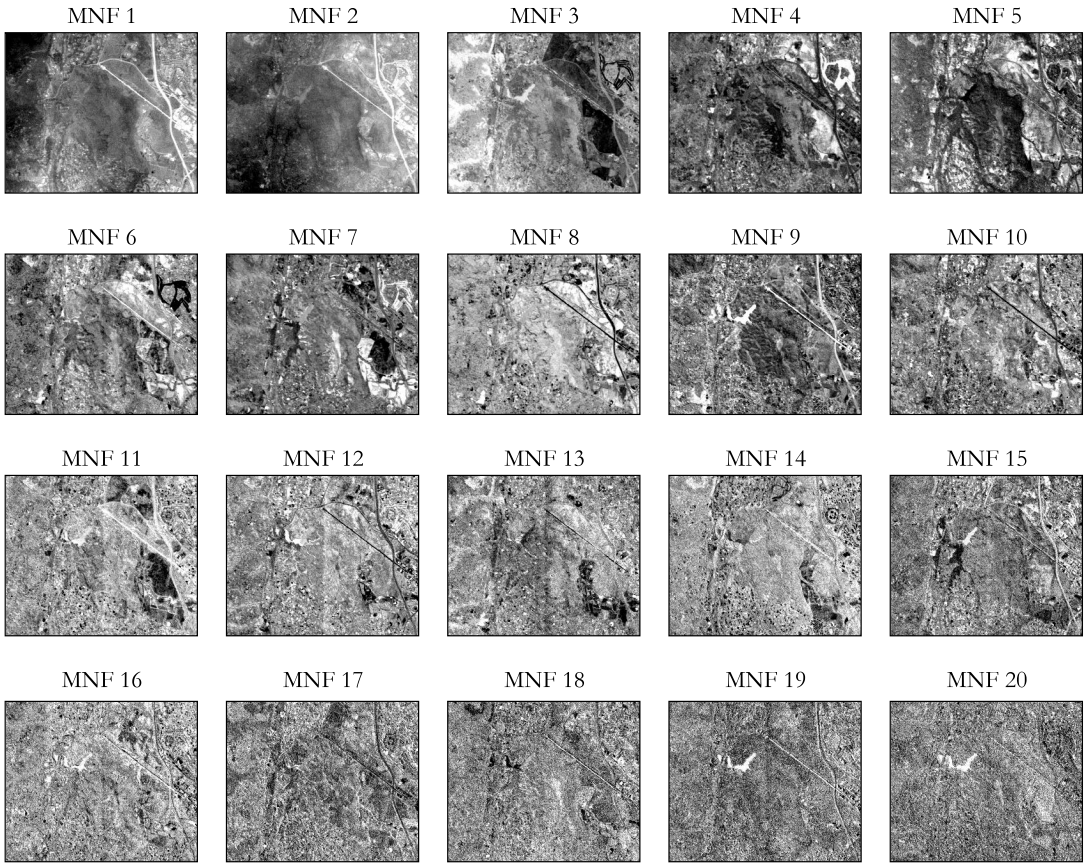
$$P_{\mathbf{U}}^{\perp} = \mathbf{I} - \mathbf{U}\left(\mathbf{U}^{T}\mathbf{U}\right)^{-1}\mathbf{U}^{T}$$

where **I** is the identity matrix and **U** is a matrix in which the newly extracted pixels (*endmembers*) are incorporated

N-FINDR

Spectral-based method developed by Winter (Proceedings of SPIE, 2003):

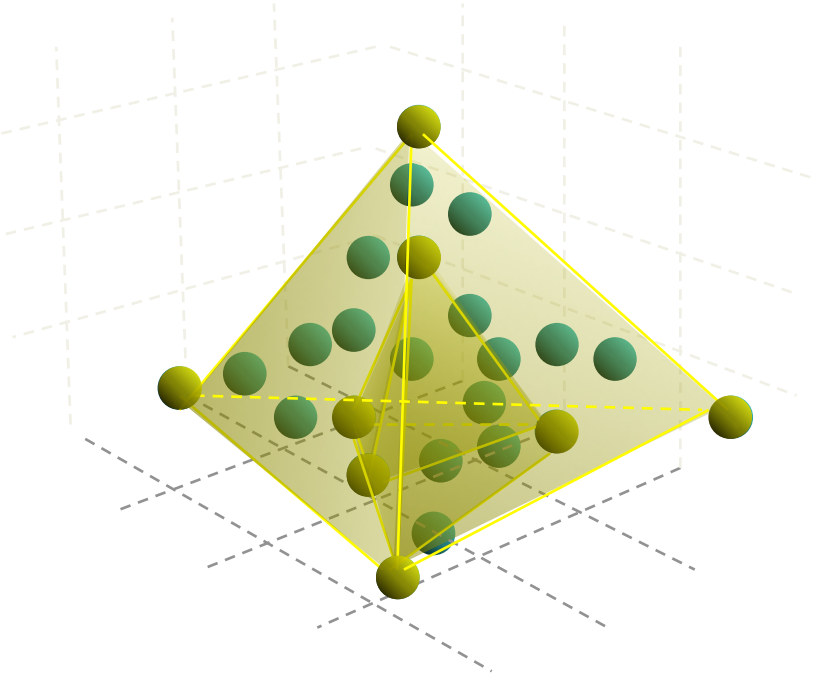
1. Dimensional reduction using MNF (ordering resulting components in terms of signal to noise ratio) or PCA (ordering components in terms of variance)



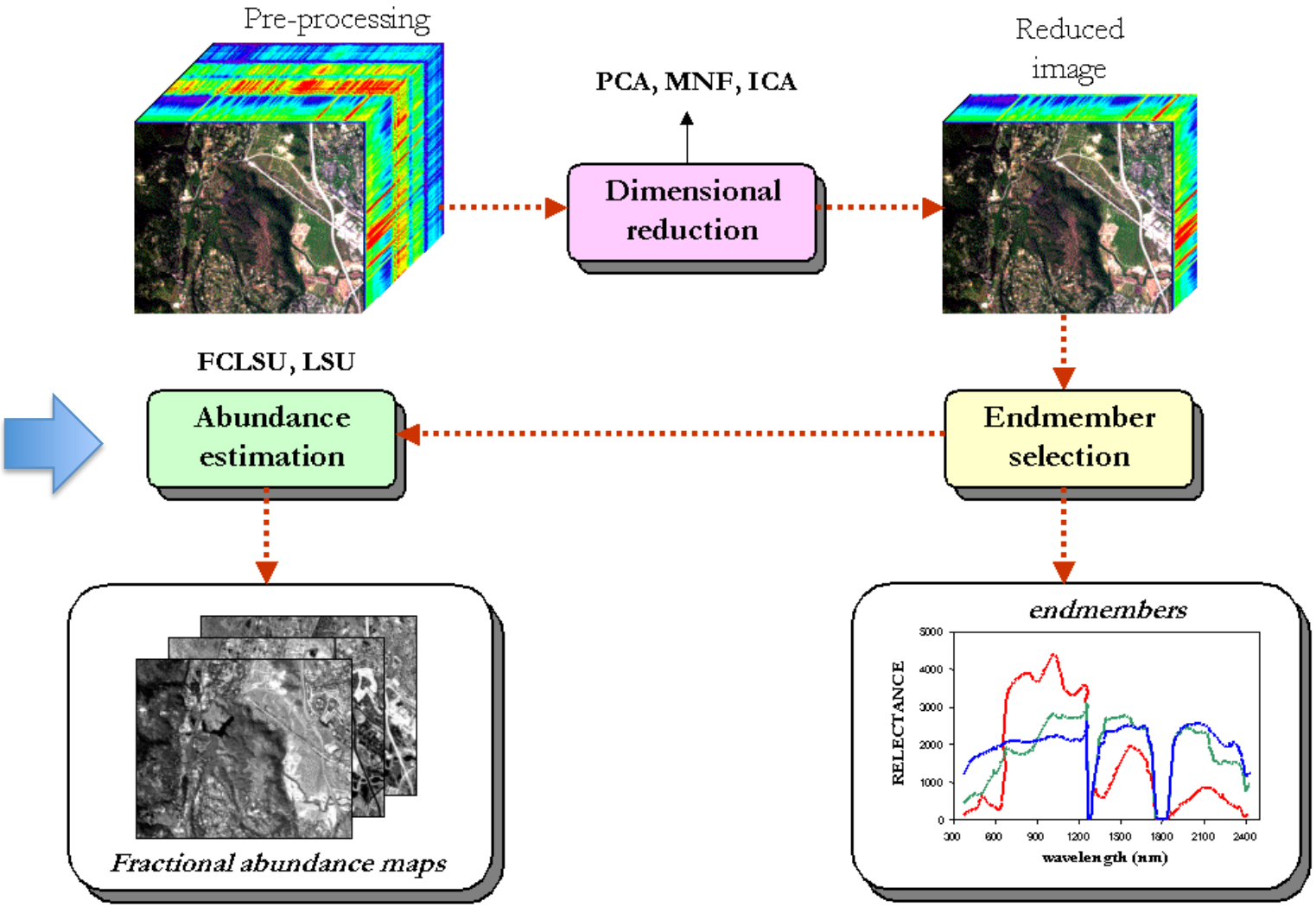
N-FINDR

Spectral-based method developed by Winter (Proceedings of SPIE, 2003):

1. Dimensional reduction using MNF (ordering resulting components in terms of signal to noise ratio) or PCA (ordering components in terms of variance)
2. Random initial selection of *endmembers* according to a predetermined number
3. Test each pixel in each *endmember* position and retain combinations with *highest volume*



Endmember extraction



Outlines

1. Introduction to sub-pixel analysis
2. Linear mixture model
3. Nonlinear mixture model
4. Bi-linear mixture model

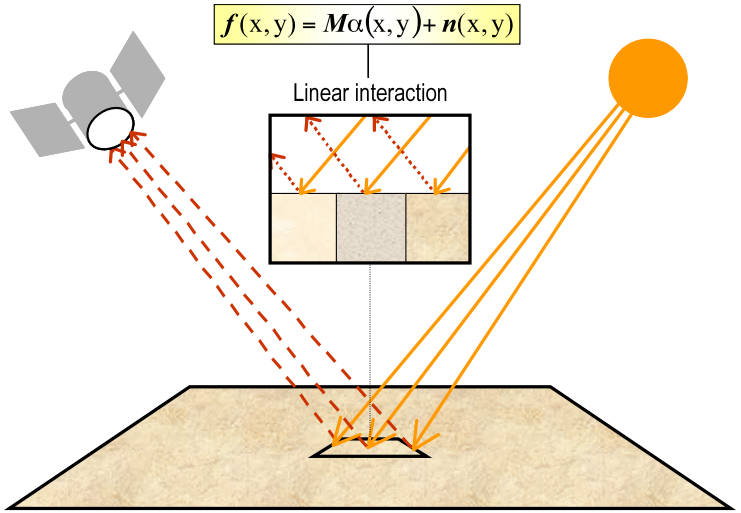
Linear spectral unmixing (LSU)

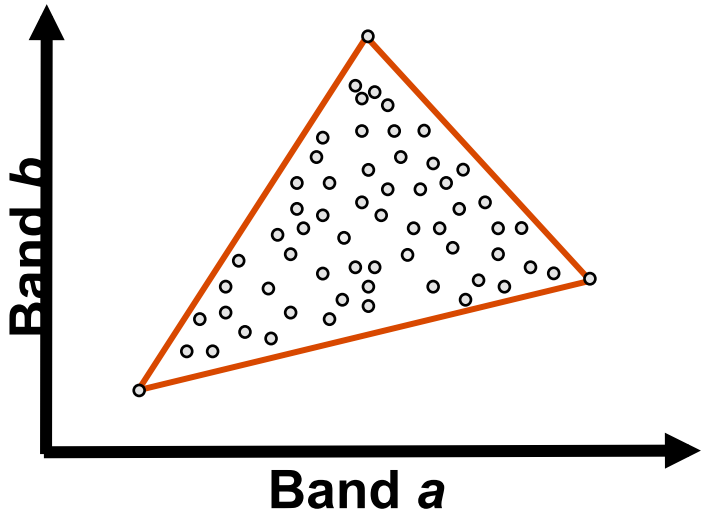
To be extimated

$$x = \sum_{i=1}^{M} a_i s_i + w$$

Known terms

The goal is to find extreme pixel vectors (*endmembers*) that can be used to "**unmix**" other mixed pixels in the data using a *linear mixture model*.





Abundance estimation

Unconstrained least squares linear unmixing

Non-negativity constrained least squares linear unmixing

Fully constrained least squares linear unmixing

Weighted least squares for linear unmixing

Unconstrained abundance estimation

- When all the endmember information (i.e., the number of endmembers and their spectral signatures) are known, abundances can be estimated via the least squares solution.
- If the two abundance constraints are ignored, abundance estimation is to find α such that the pixel reconstruction error:

$$e = \left\| \mathbf{r} - \mathbf{M}\hat{\boldsymbol{\alpha}} \right\|^2$$

is minimized. The least squares (LS) solution is:

$$\hat{\boldsymbol{\alpha}} = \left(\mathbf{M}^T\mathbf{M}\right)^{-1}\mathbf{M}^T\mathbf{r}$$

Non-negative constrained abundance estimation

- If the abundance non-negativity constraint needs to be relaxed, the problem of abundance estimation becomes a constrained optimization problem:

$$\min e = \min f(\boldsymbol{\alpha}) = \mathbf{r}^T\mathbf{r} - 2\mathbf{r}^T\mathbf{M}\boldsymbol{\alpha} + \boldsymbol{\alpha}^T\mathbf{M}^T\mathbf{M}\boldsymbol{\alpha}$$

$$\text{Subject to}: 0 \le \alpha_i \le 1, \text{for} 1 \le i \le p$$

- This optimization problem with inequality constraints can be solved by quadratic programming since the objective function is a quadratic function.

Fully constrained abundance estimation

- If both abundance non-negativity and sum-to-one constraints need to be relaxed, the constrained optimization problem becomes:

$$\min e = \min f(\boldsymbol{\alpha}) = \mathbf{r}^T\mathbf{r} - 2\mathbf{r}^T\mathbf{M}\boldsymbol{\alpha} + \boldsymbol{\alpha}^T\mathbf{M}^T\mathbf{M}\boldsymbol{\alpha}$$

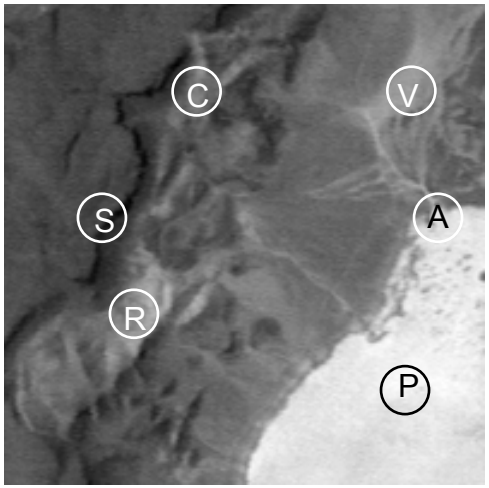
$$\text{Subject to : } \alpha_1 + \alpha_2 + \cdots + \alpha_p = 1$$

$$0 \le \alpha_i \le 1, \text{for} 1 \le i \le p$$

- Actually, the sum-to-one constraint can be easily satisfied by adding a row vector with all elements being one to the endmember matrix **M**, adding an element one to the pixel vector **r**, and solving the resulting least squares problem.

$$\tilde{\mathbf{M}} = \begin{bmatrix} \mathbf{M} \\ \mathbf{1} \end{bmatrix} \qquad \tilde{\mathbf{r}} = \begin{bmatrix} \mathbf{r} \\ 1 \end{bmatrix} \qquad \tilde{\mathbf{r}} = \tilde{\mathbf{M}}\boldsymbol{\alpha} + \mathbf{n}$$

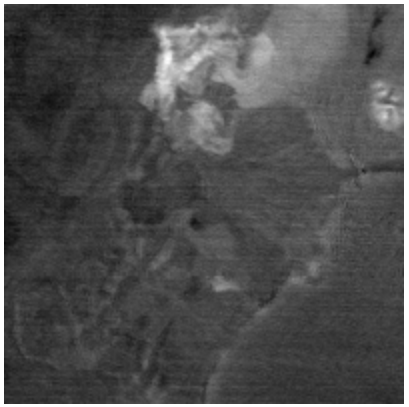
Aviris Lunar Lake experiment



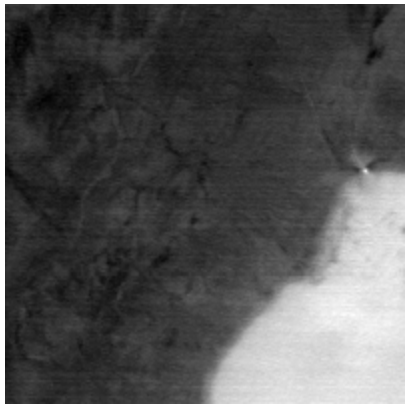
- Five endmembers: Cinder (C), Playa lake (P), Rhyolite (R), Vegetation (V), Shade (S)
- Six endmembers: Cinder (C), Playa lake (P), Rhyolite (R), Vegetation (V), Shade (S), Anomaly (A)

Pixel reconstruction error = $\dfrac{\sum_{i=1}^{N} \left\| \mathbf{r}_i - \hat{\mathbf{r}}_i \right\|^2}{N}$

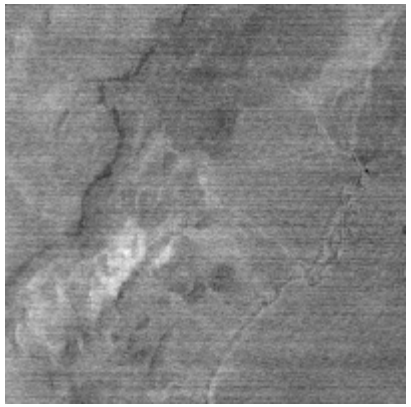
Aviris Lunar Lake experiment



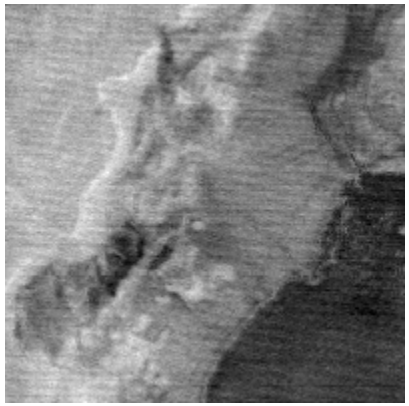
Cinder



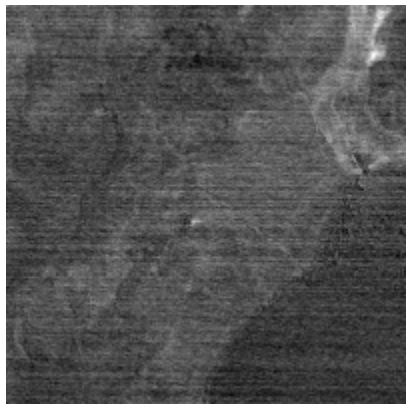
Playa lake



Rhyolite



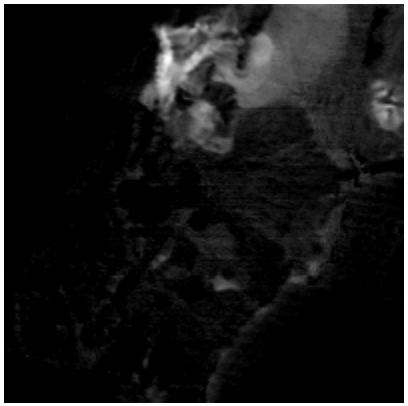
Shade



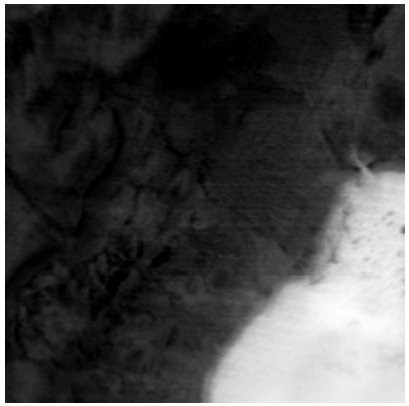
Vegetation

Abundance maps from unconstrained least squares (UCLS) linear unmixing with five endmember (pixel reconstruction error = 8.85×10^3)

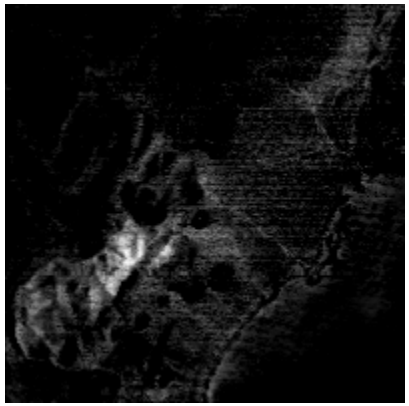
Aviris Lunar Lake experiment



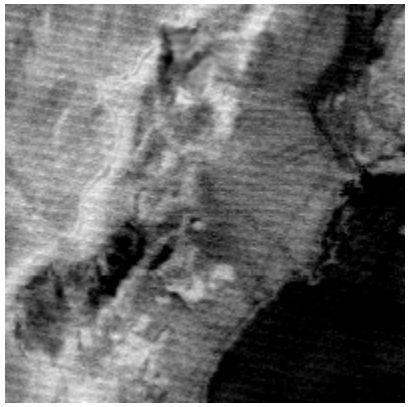
Cinder



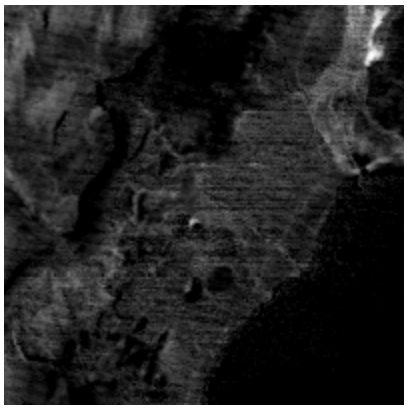
Playa lake



Rhyolite



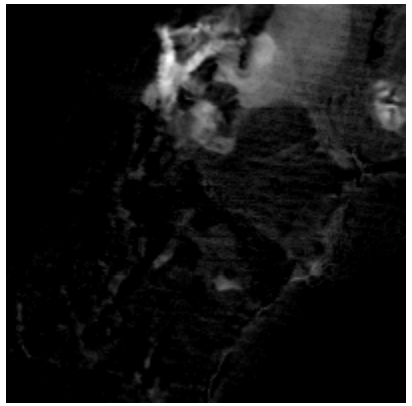
Shade



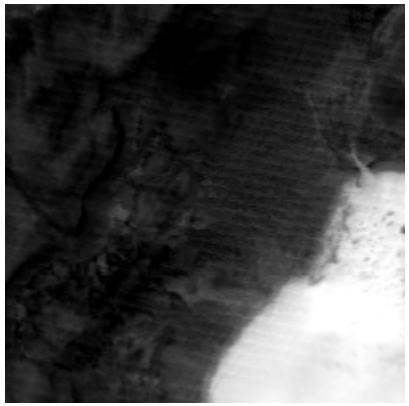
Vegetation

Abundance maps from non-negativity constrained least squares (NCLS) linear unmixing with five endmember (pixel reconstruction error = 1.02×10^4)

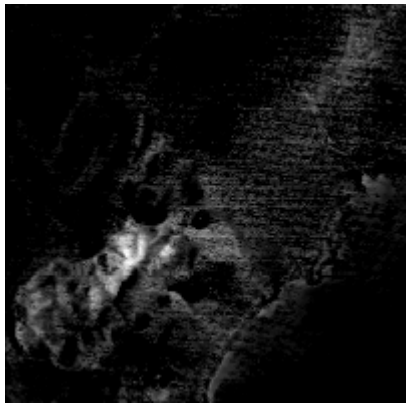
Aviris Lunar Lake experiment



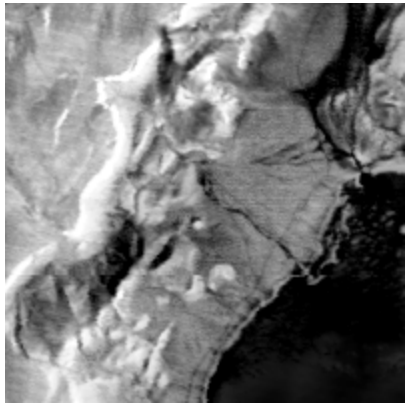
Cinder



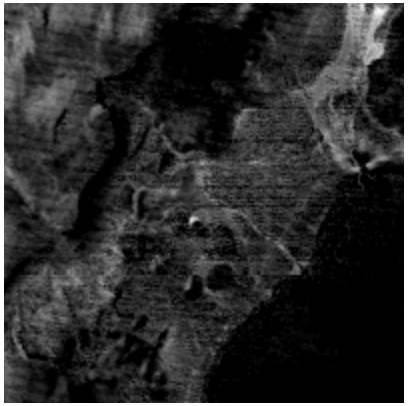
Playa lake



Rhyolite



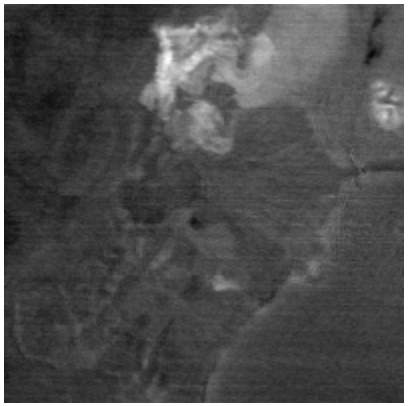
Shade



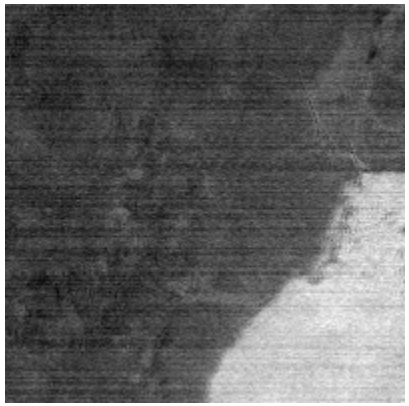
Vegetation

Abundance maps from fully constrained least squares (FCLS) linear unmixing with five endmember (pixel reconstruction error = 1.43×10^4)

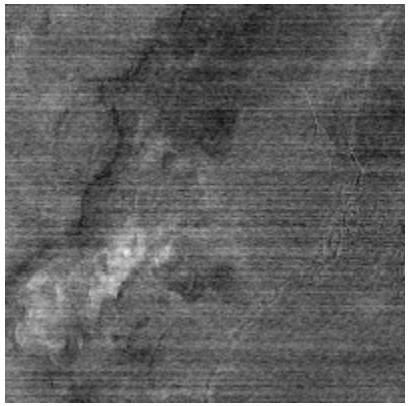
Aviris Lunar Lake experiment



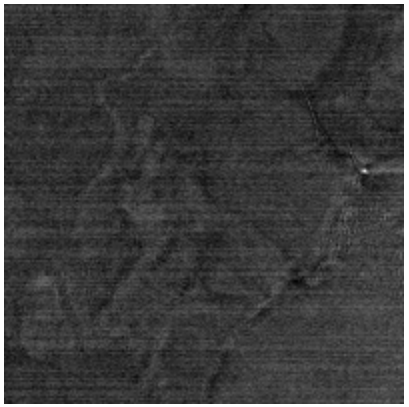
Cinder



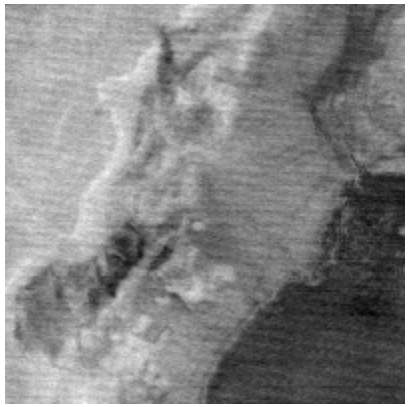
Playa lake



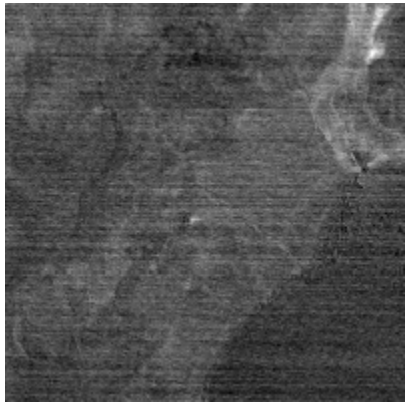
Rhyolite



Anomaly



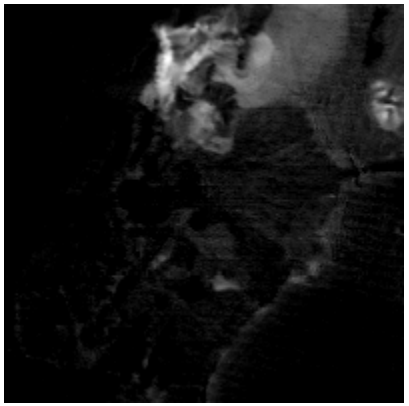
Shade



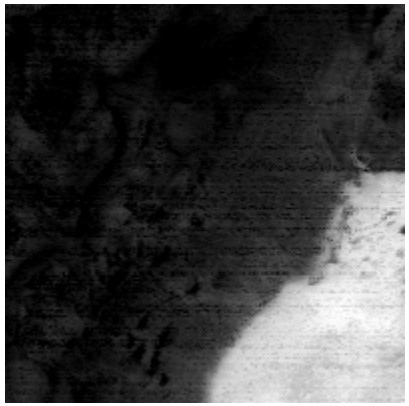
Vegetation

Abundance maps from unconstrained least squares (UCLS) linear unmixing with six endmember (pixel reconstruction error = 8.64×10^3)

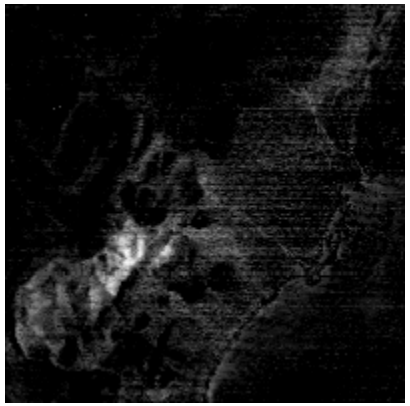
Aviris Lunar Lake experiment



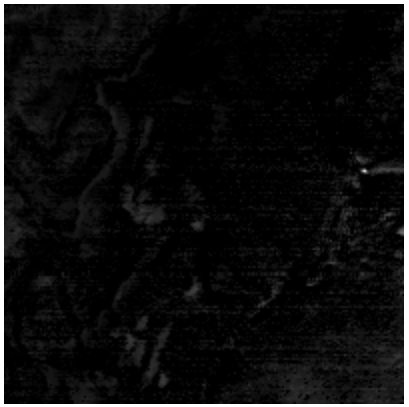
Cinder



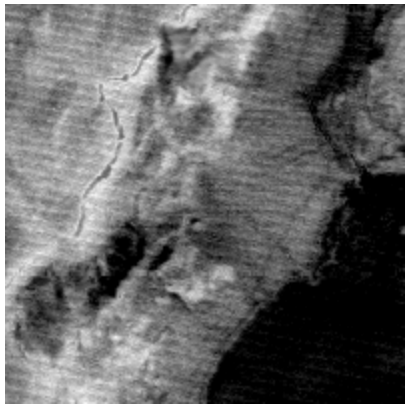
Playa lake



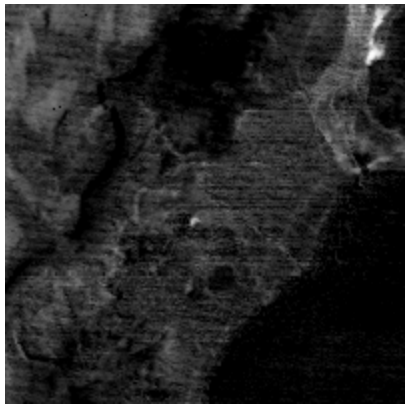
Rhyolite



Anomaly



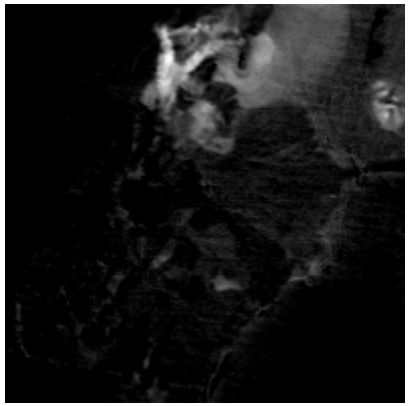
Shade



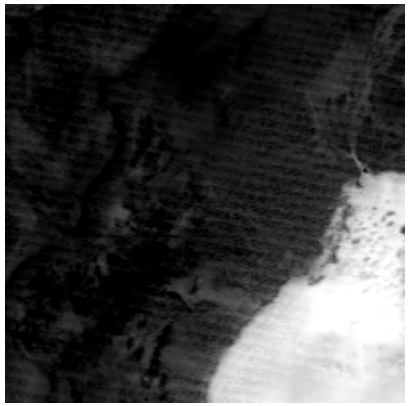
Vegetation

Abundance maps from non-negativity constrained least squares (NCLS) linear unmixing with six endmember (pixel reconstruction error = 9.56×10^3)

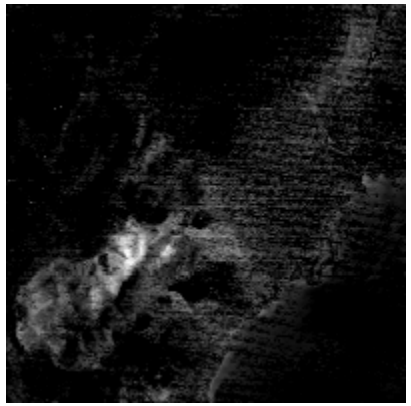
Aviris Lunar Lake experiment



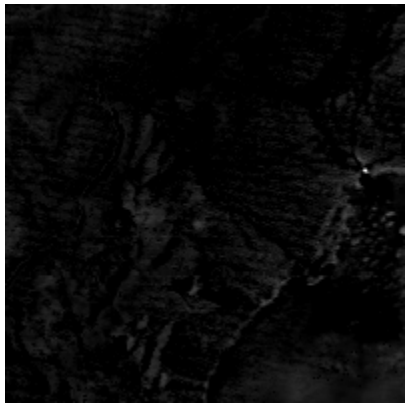
Cinder



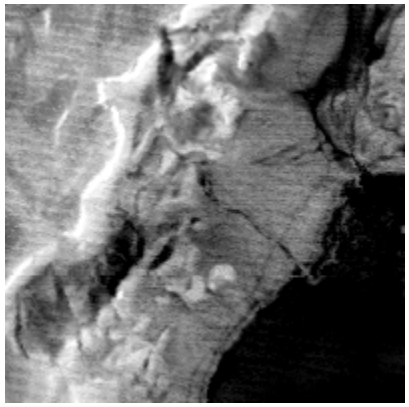
Playa lake



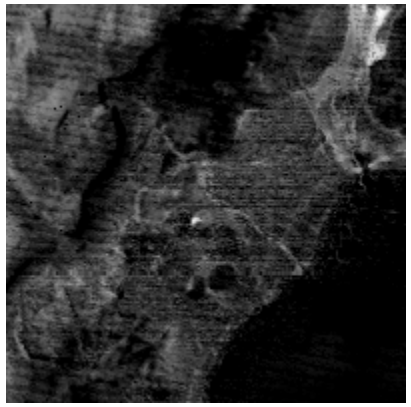
Rhyolite



Anomaly



Shade



Vegetation

Abundance maps from fully constrained least squares (FCLS) linear unmixing with six endmembers (pixel reconstruction error = 1.35×10^4)

Aviris Lunar Lake experiment

Pixel Reconstruction Errors

	UCLS	NCLS	FCLS
Five endmembers	8.85×10^3	1.02×10^4	1.43×10^4
Six endmembers	8.64×10^3	9.56×10^3	1.35×10^4

- Pixel reconstruction error is decreased when using more endmembers.
- Pixel reconstruction error is increased when imposing more constraints.
- Comparison in terms of pixel reconstruction error should be made for comparable cases.
- Evaluation is more reasonable in terms of abundance accuracy. However, it is difficult due to the unavailability of ground truth in practice.

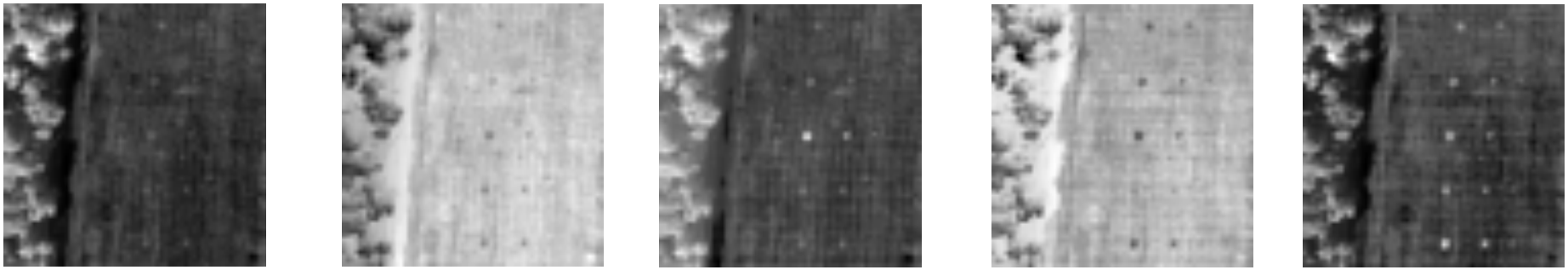
Weighted Least Square

- When only partial endmember information is known, abundances can be estimated via the weighted least squares (WLS) solution. For instance, we may know the foreground endmembers only, but the information of background endmembers is more difficult to be determined.
- The WLS solution for the abundances of the known endmembers can be estimated as:

$$\hat{\boldsymbol{\alpha}} = \left(\mathbf{M}^T \boldsymbol{\Sigma}^{-1} \mathbf{M}\right)^{-1} \mathbf{M}^T \boldsymbol{\Sigma}^{-1} \mathbf{r}$$

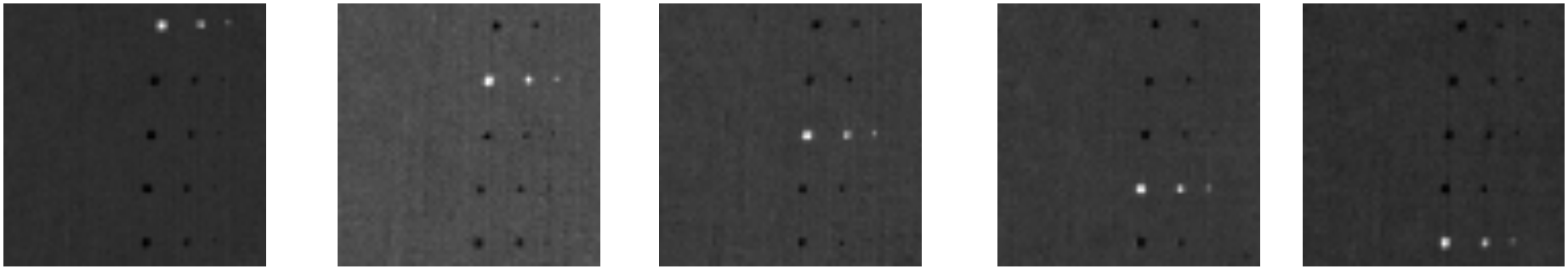
where $\boldsymbol{\Sigma}$ is background covariance matrix.

Weighted Least Square



P1 P2 P3 P4 P5

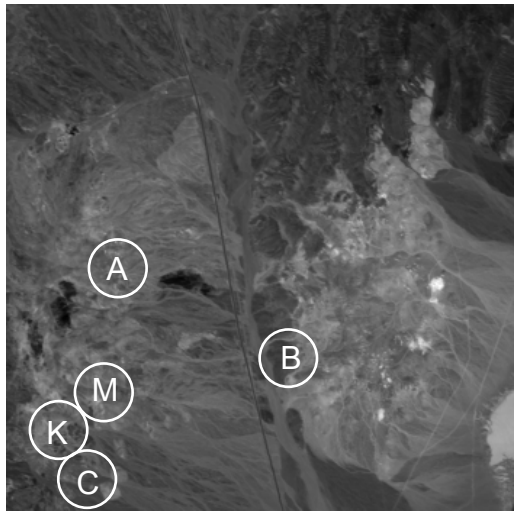
Abundance maps from UCLS linear unmixing when only five panel signatures are known



P1 P2 P3 P4 P5

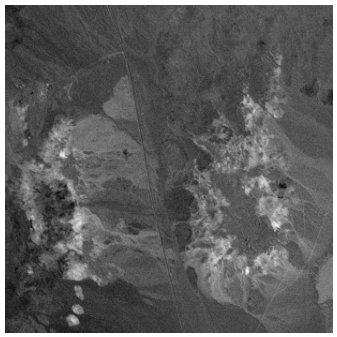
Abundance maps from weighted least squares (WLS) linear unmixing when only five panel signatures are known

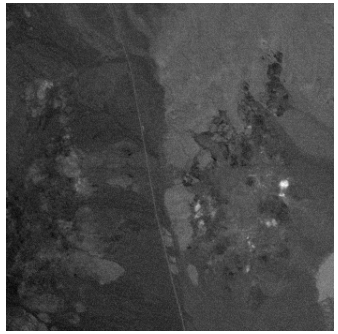
Weighted Least Square

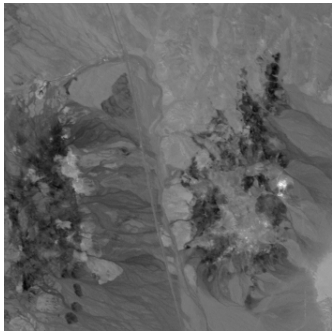


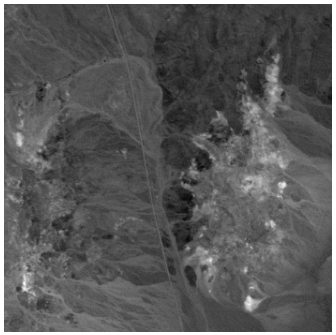
five known endmembers:
alunite (A), buddingtonite (B), calcite (C), kaolinite (K), muscovite (M)

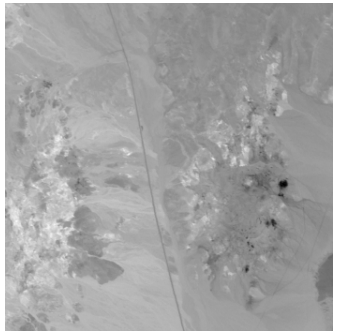
Weighted Least Square


alunite (A)

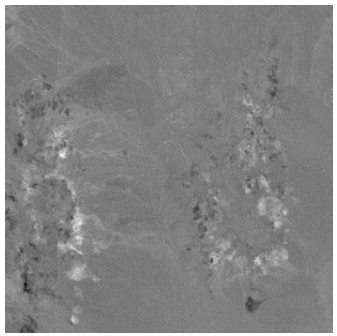

buddingtonite (B)

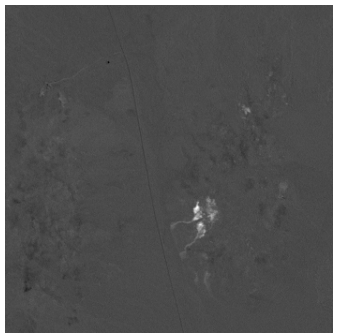

calcite (C)

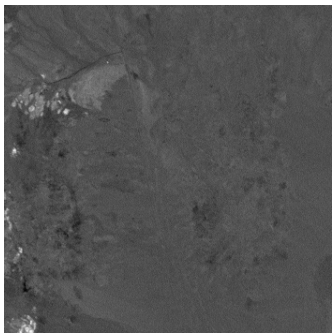

kaolinite (K)

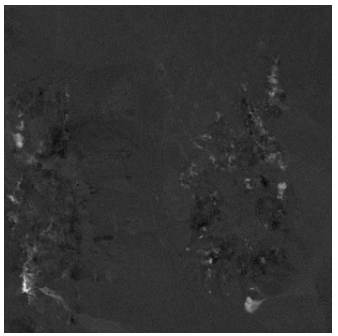

muscovite (M)

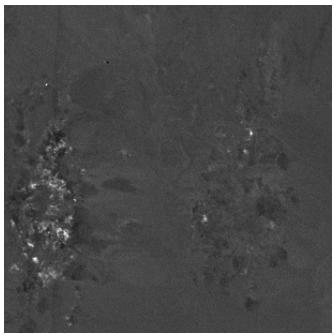
Abundance maps from UCLS linear unmixing when only five mineral signatures are known


alunite (A)


buddingtonite (B)


calcite (C)


kaolinite (K)


muscovite (M)

Abundance maps from weighted least squares (WLS) linear unmixing when only five mineral signatures are known

remarks

- Whether abundance constraints should be imposed or not depends on practical application. It has been argued that if the model is accurate (i.e., the number of endmembers and their signatures), the two constraints should be satisfied automatically.
- The non-negativity constraint is more important than the sum-to-one constraint. Due to noise and spectral variability, reinforcing the sum-to-one constraint may be prone to induce additional estimation error.
- The discussed abundance estimation methods are based on minimum pixel reconstruction error (L2). Abundance estimation is also doable if the error is redefined (L1 error, spectral angle error, etc.), which results in a more complicated constrained optimization problem.
- When endmembers are unknown, endmember signatures should be extracted or estimated first. Some endmember extraction algorithms can provide abundance estimates simultaneously.

remarks

- There exists another group of abundance estimation based on blind source separation, which does not require endmember signatures to be known *a prior*. Widely used matrix factorization-based blind source separation methods include
  - Independent component analysis (ICA)
  - Nonnegative matrix factorization (NMF)

  These methods are more frequently used for unsupervised (soft) classification.

Outlines

1. Introduction to sub-pixel analysis
2. Linear mixture model
3. Nonlinear mixture model
4. Bi-linear mixture model

Nonlinear spectral unmixing

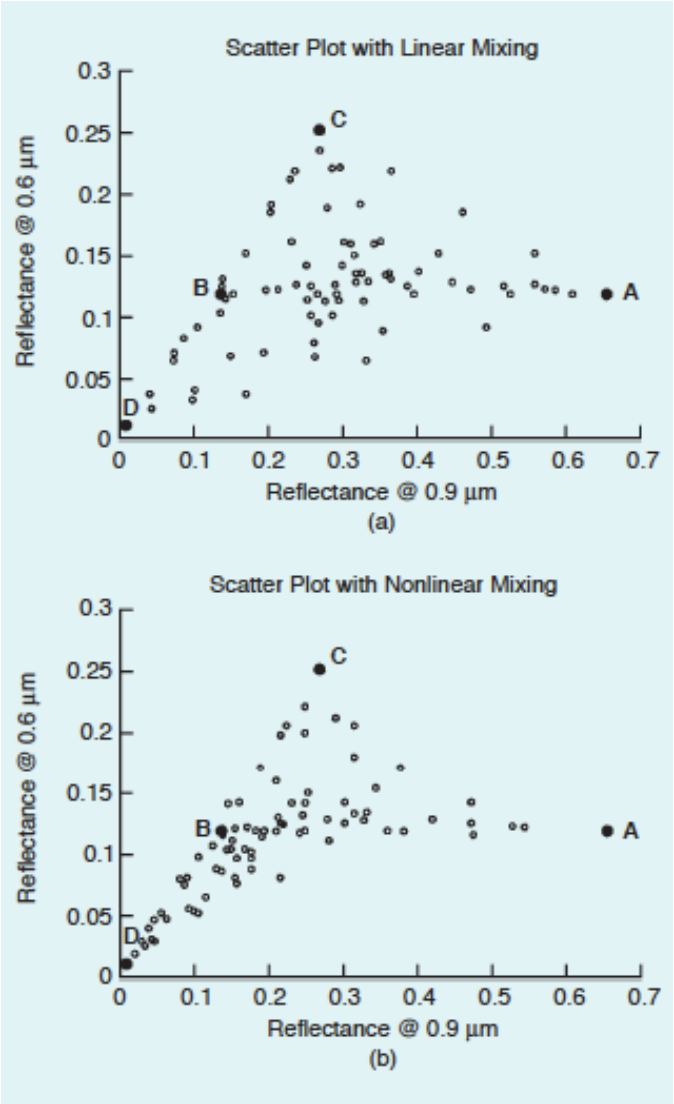
Linear versus nonlinear unmixing revisited

Neural network-based spectral unmixing

Architectures and training algorithms

Experiments using laboratory data

Linear vs nonlinear mixing



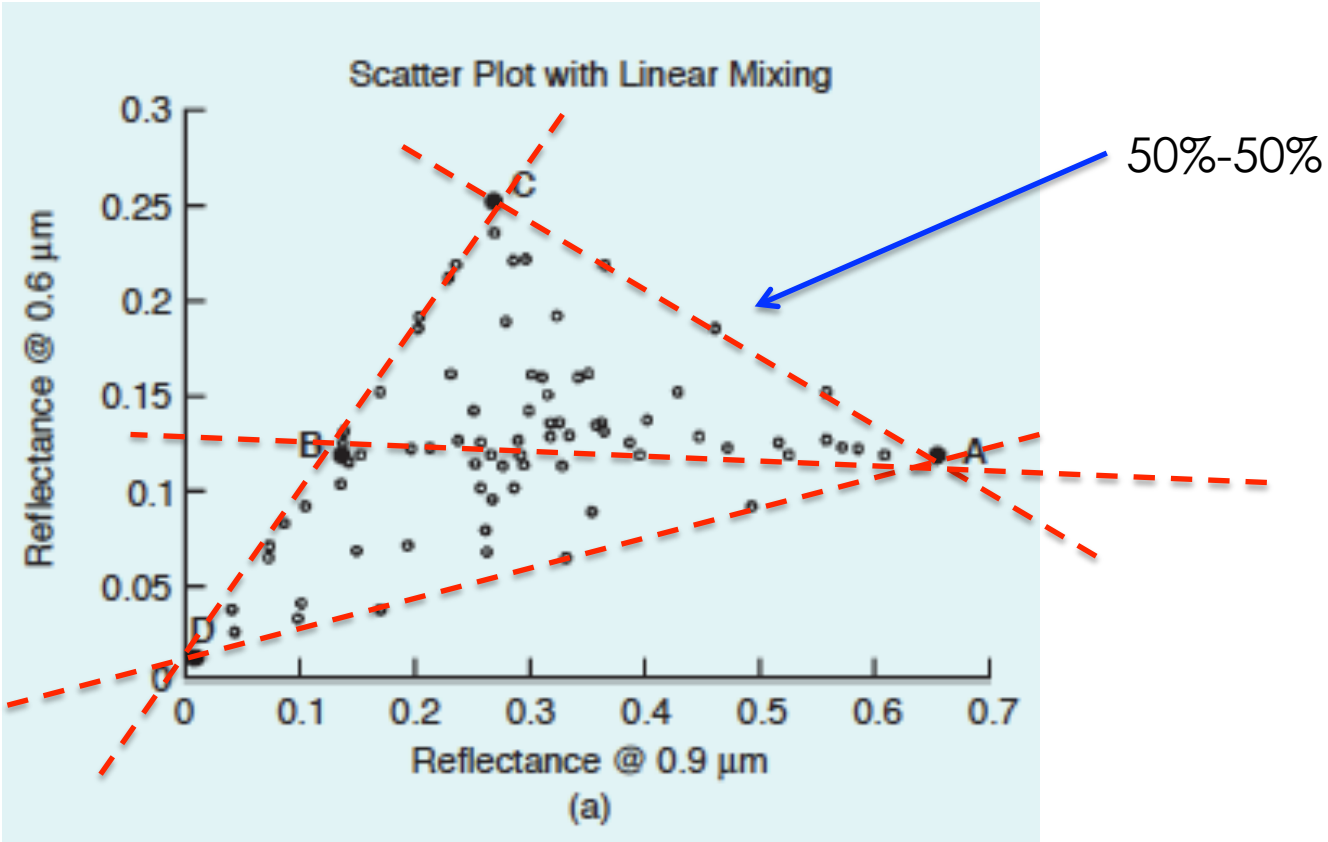
Linear or nonlinear?

Linear approach is useful, but valid for few cases

Nonlinear approaches are very difficult to be implemented

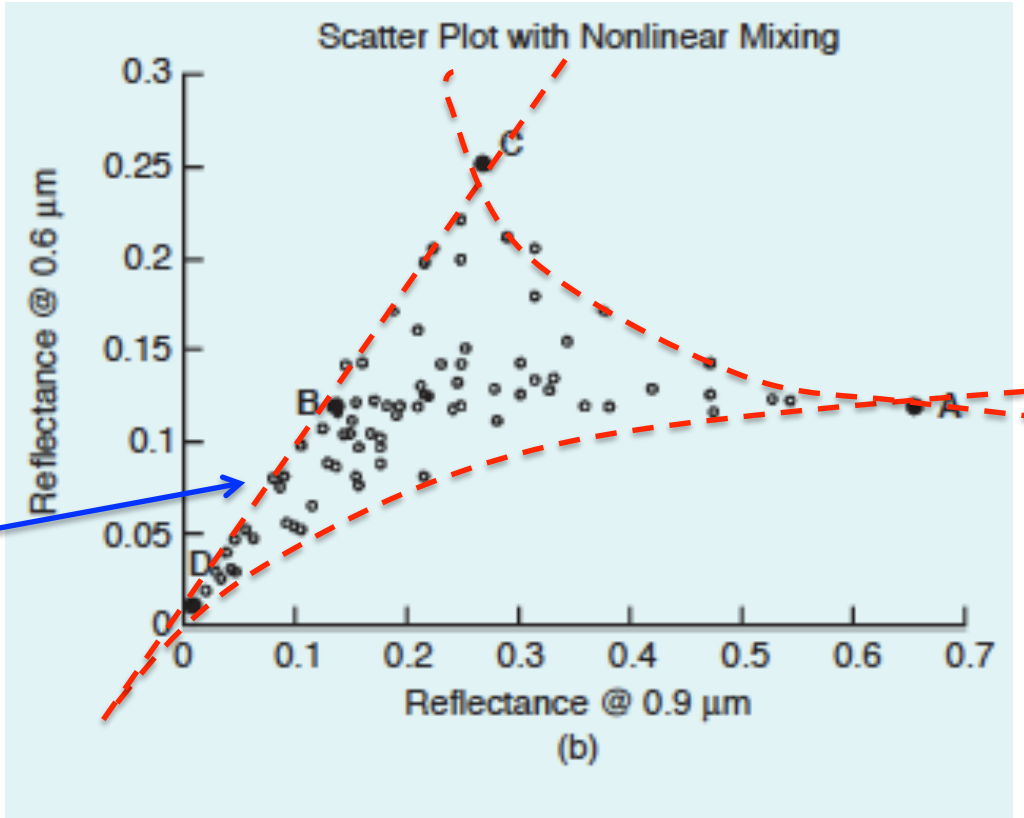
The effects of nonlinear mixing on reflectance spectra can be quite dramatic

Linear vs nonlinear mixing

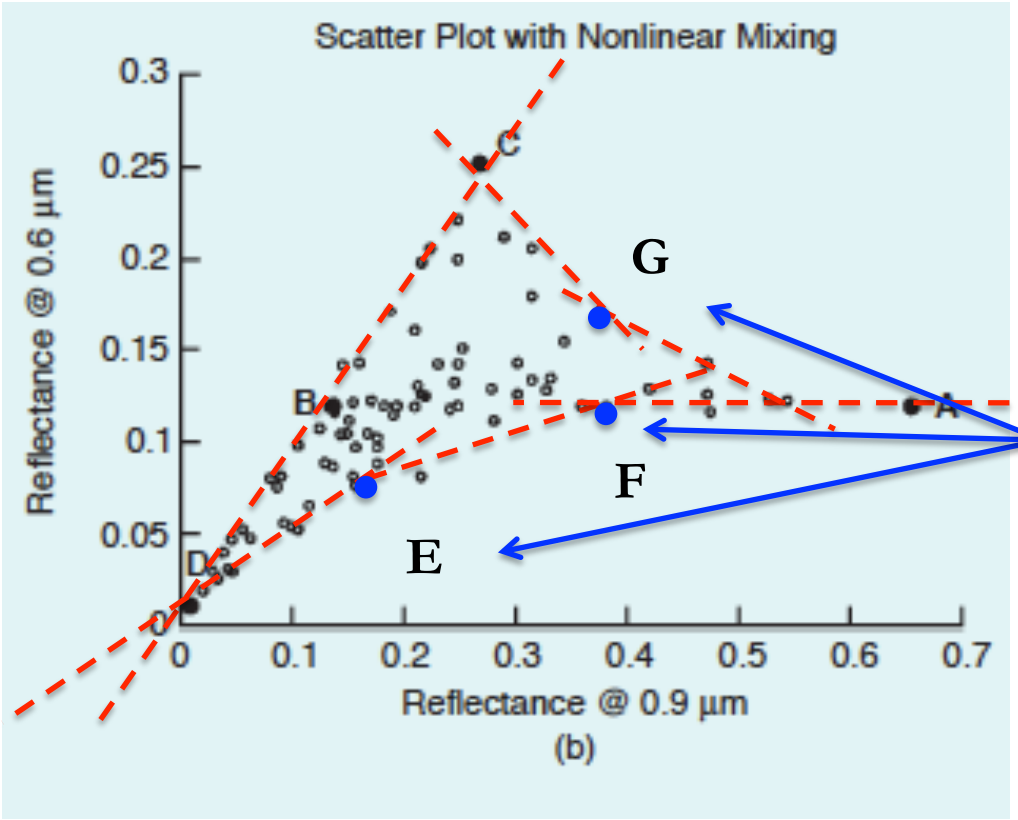


Linear vs nonlinear mixing

Data cloud shifted to low albedo area

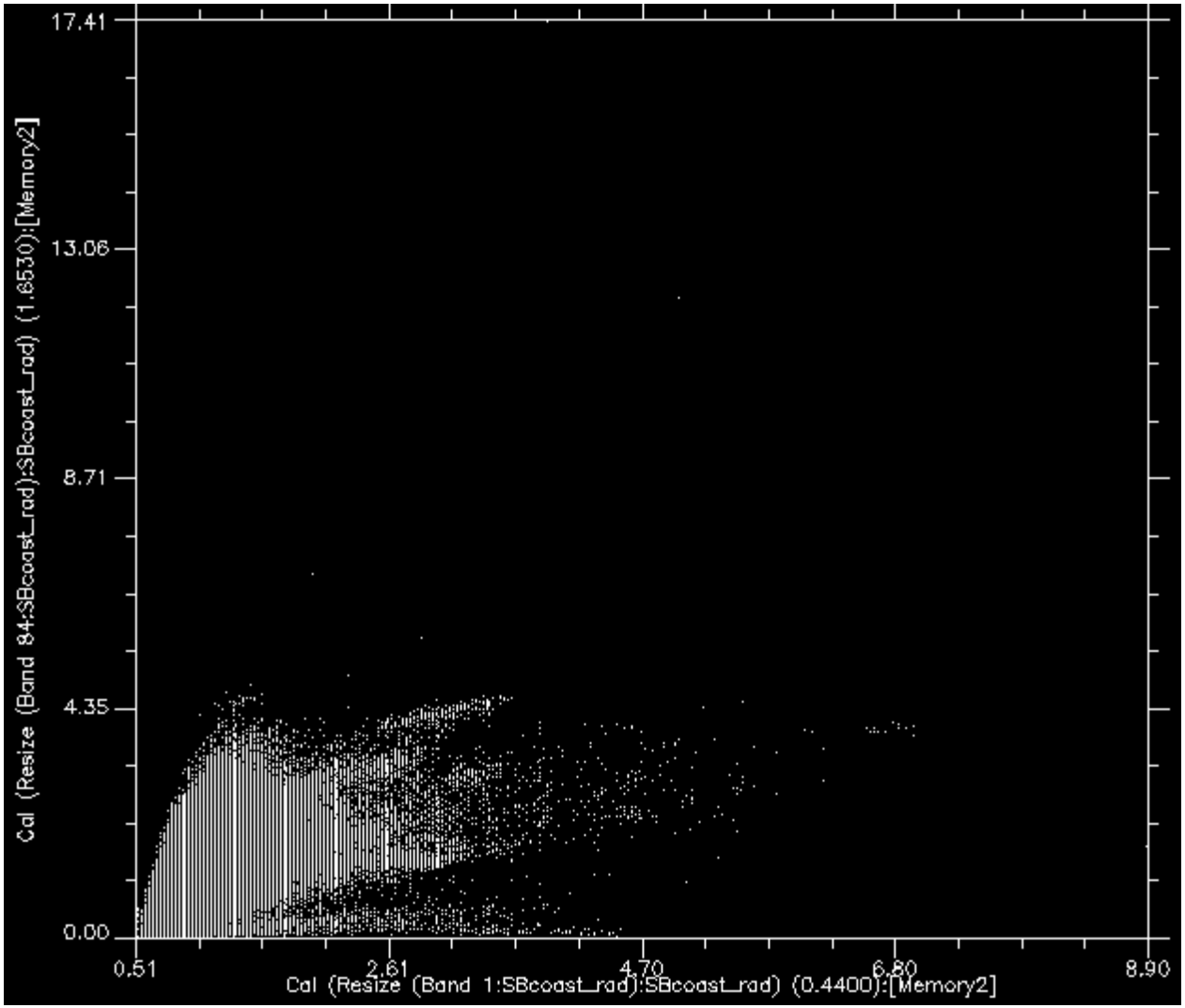


Linear vs nonlinear mixing



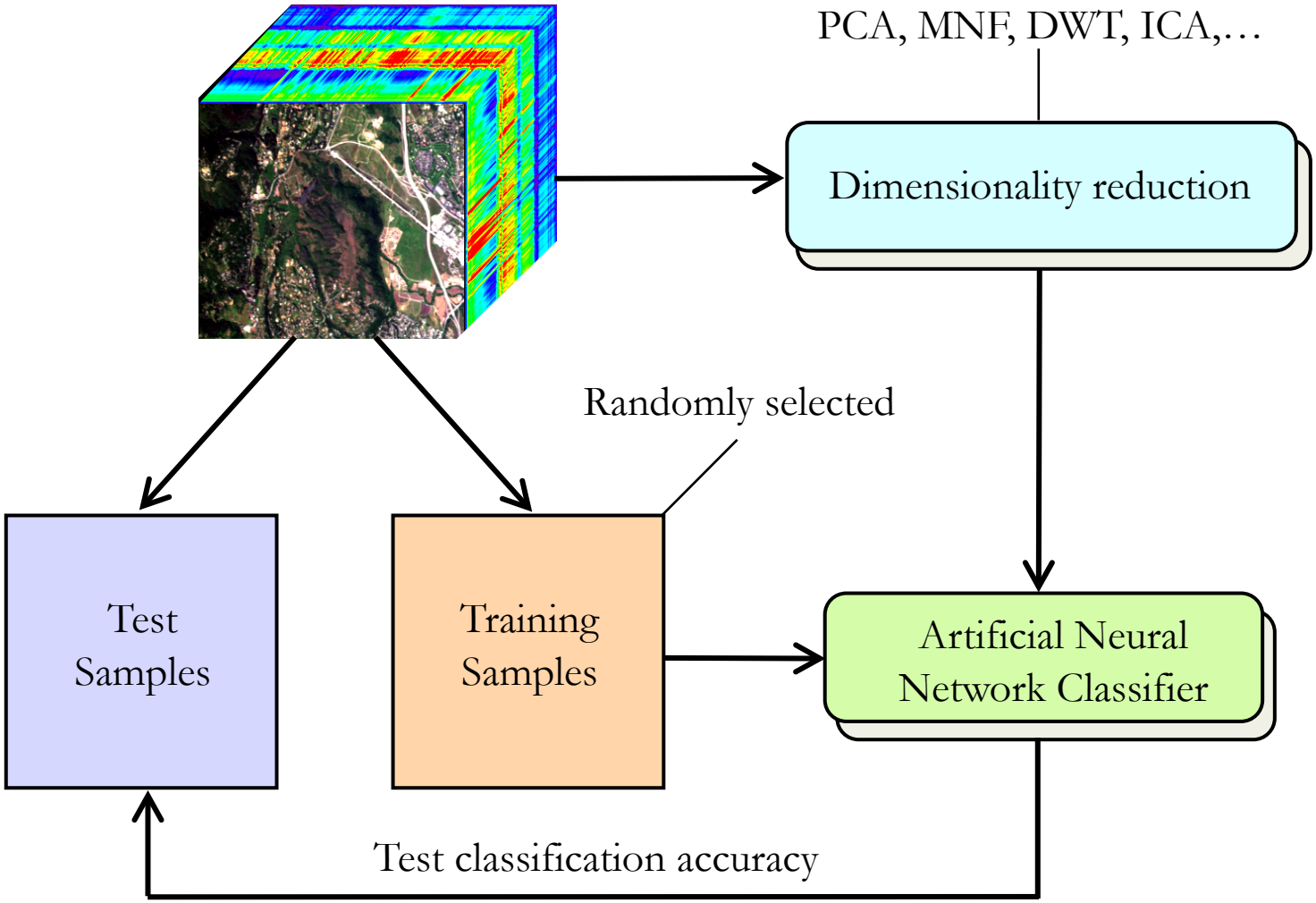
False endmembers!!

Linear vs nonlinear mixing



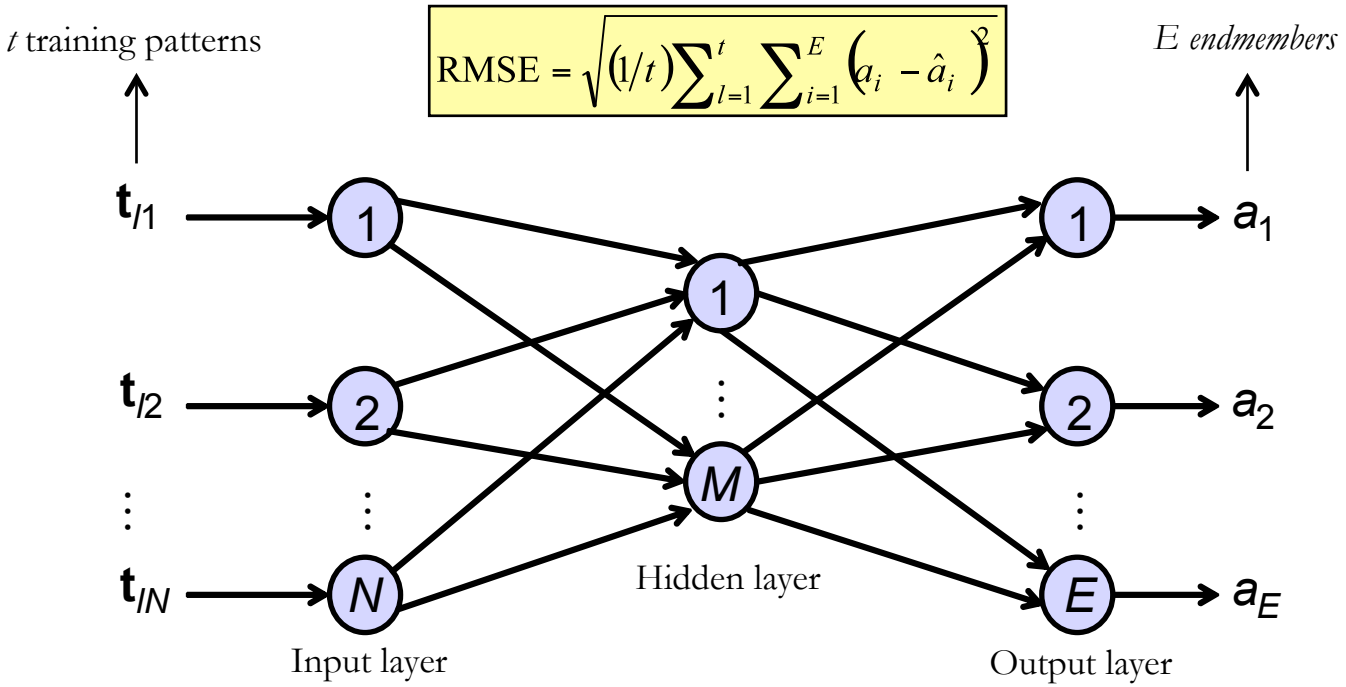
Standard NN based unmixing

- Incorporates a learning process based on known, *randomly selected* training samples.



Standard NN based unmixing

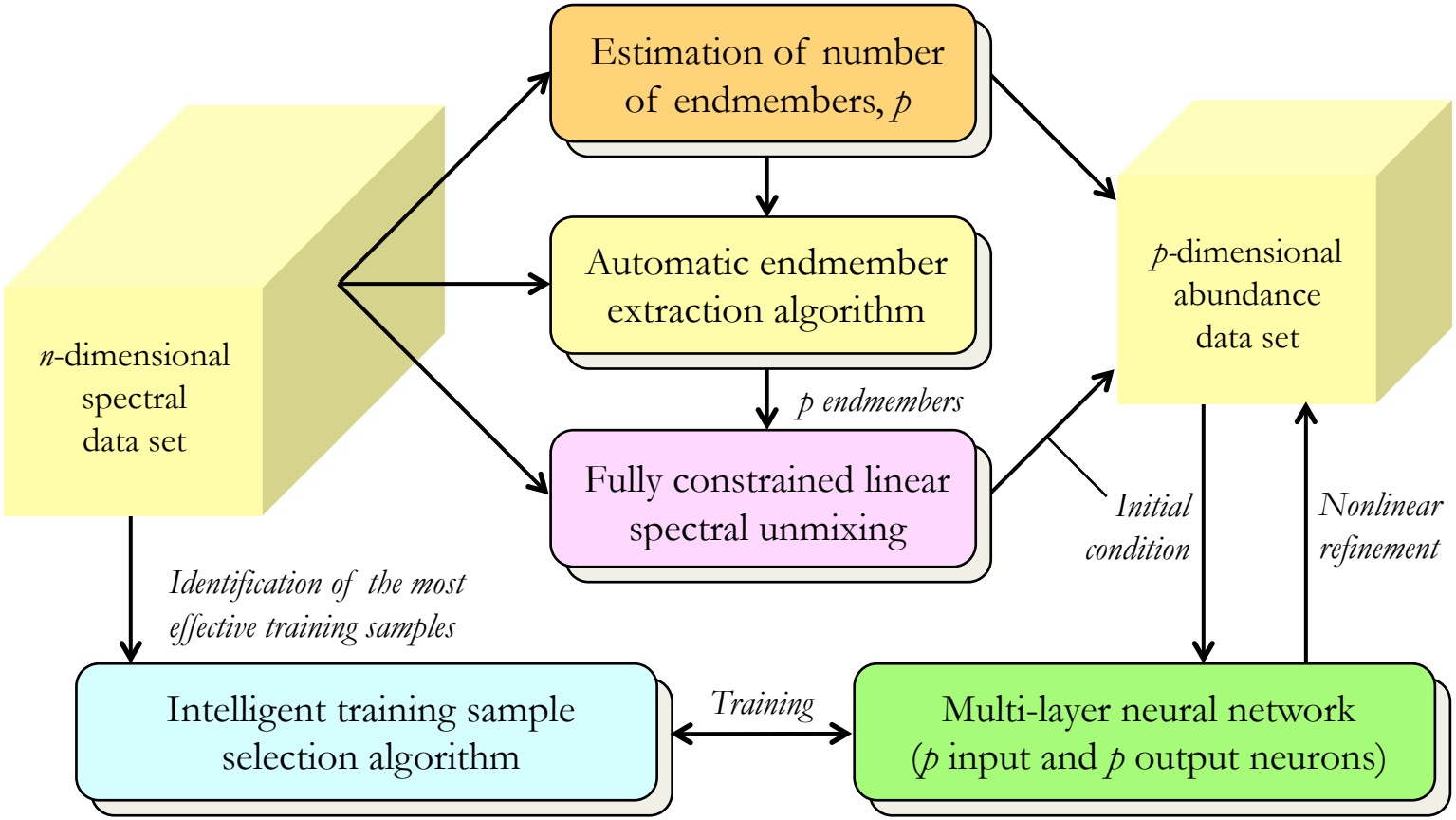
- Back-propagation learning from ***selected training samples***.
- The ***training stage*** is the most crucial: *how many training samples* are required?.



Standard NN based unmixing

Combined linear/nonlinear architecture #1

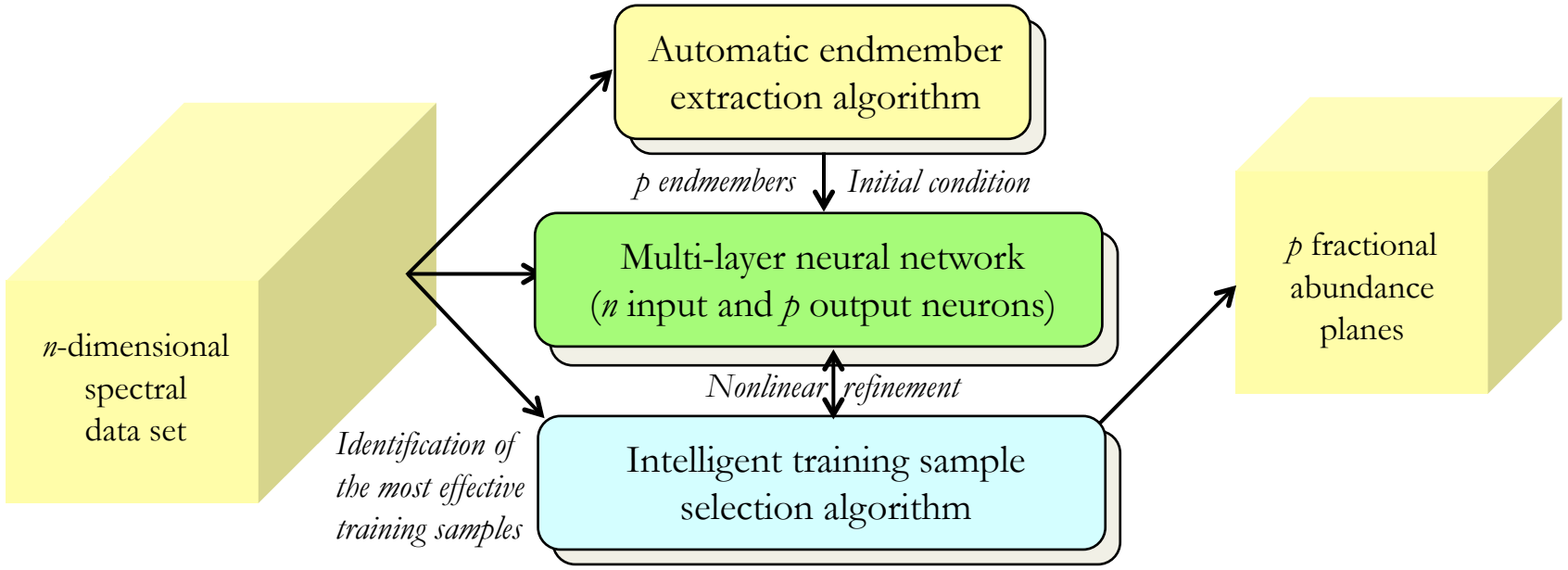
- Nonlinear ***refinement of linear abundance estimations*** (using FCLSU).
- Limitation: this approach does not take into account the *full spectral information*.



Standard NN based unmixing

Combined linear/nonlinear architecture #2

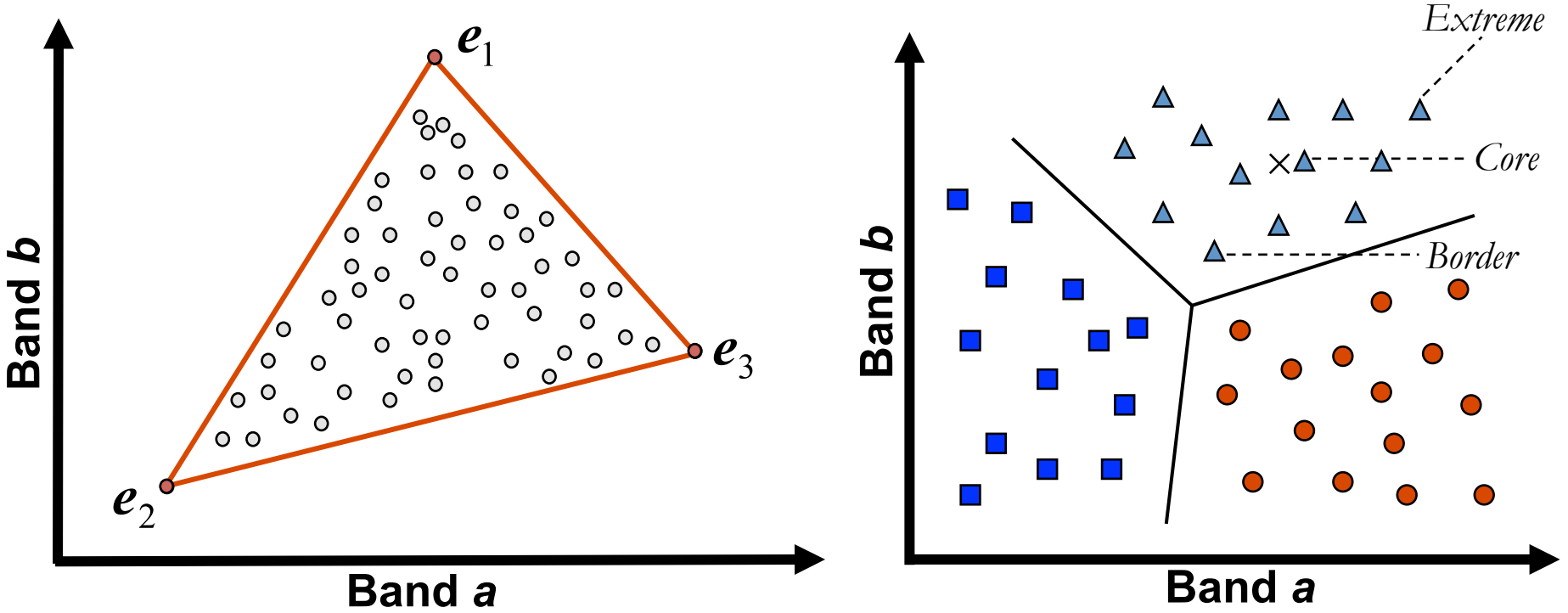
- MLP initially trained using only *p* endmembers provided by an *automatic algorithm*.
- Nonlinear refinement from a *linear* initial weight condition using training samples.
- The *full spectral information* is used throughout the process.
  - ✓ *Choice of endmember finding algorithms for initialization (AMEE used in this work).*
  - ✓ *Choice of intelligent training sample selection algorithms*



Standard NN based unmixing

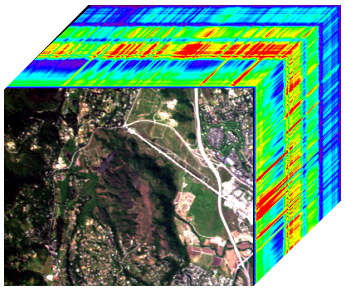
Training of neural network architecture

- Need for algorithms able to *automatically search* for the most useful *training samples*.
- Potential to direct ground-truth data collection to the most useful training sites.
  - ✓ *In the linear mixture model, the most highly pure patterns are of interest.*
  - ✓ *In the nonlinear mixture model, "border" training patterns are required.*



Standard NN based unmixing

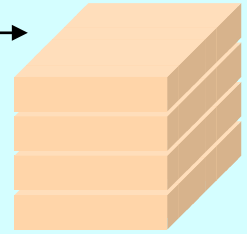
Border-training Selection Alg. (BTA)



Parallel automatic endmember extraction algorithm

p endmembers

1. **Partition** the input data using *spatial-domain decomposition*



2. **Process independently** at each *local* partition:

   2.1. *Spectral screening* to associate each signature $\mathbf{t}_i$ to pure class $\mathbf{e}_j$ (p classes)

   2.2. For each signature, compute the *Mahalanobis distance* from each pure class as:

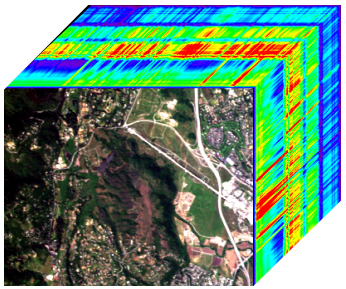
$$MD(\mathbf{t}_i, \mathbf{e}_j) = (\mathbf{t}_i, \boldsymbol{\mu}_j)^T \mathbf{K}_i^{-1} (\mathbf{t}_i, \boldsymbol{\mu}_j)$$

   2.3. Compute "***borderness***" for $\mathbf{t}_i$ as *difference between the two smallest values of* $MD(\mathbf{t}_i, \mathbf{e}_j)$

3. **Gather** results and select a set of samples *in decreasing order according to their borderness score*

Standard NN based unmixing

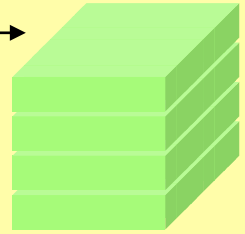
Mixed-signature Selection Alg. (MSA)



Parallel automatic endmember extraction algorithm

p endmembers

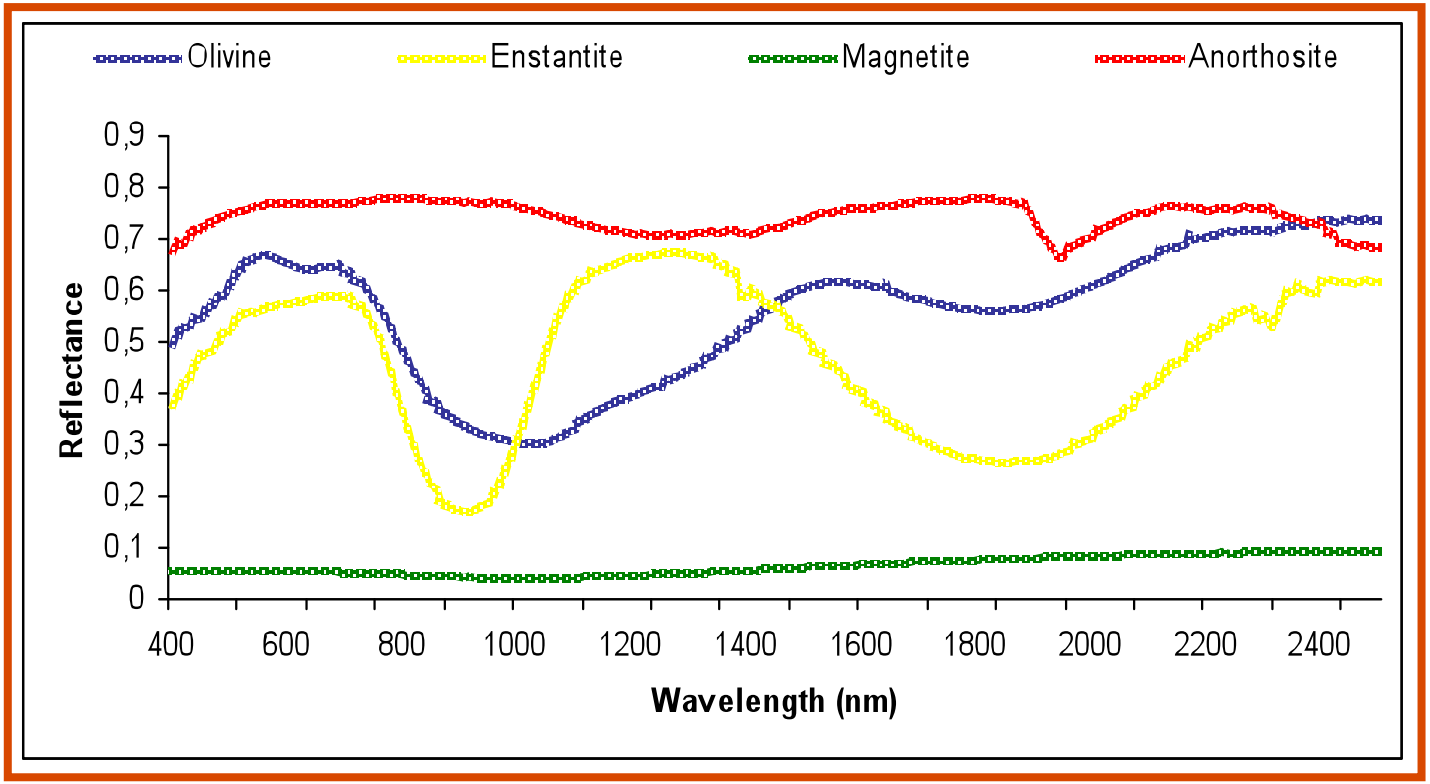
1. **Partition** the input data using *spatial-domain decomposition*



2. **Compute the centroid** using $\boldsymbol{c}_p = (1/p)\sum_{i=1}^{p} \boldsymbol{e}_i$ and broadcast it to workers
3. **Process independently** at each *local* partition:
   - 3.1. Calculate *eccentricity* of each *signature* $\mathbf{t}_i$ as SAD($\mathbf{t}_i$, $\mathbf{c}_p$) an *retain the one with lowest score*
   - 3.2. Calculate *eccentricity* of each *endmeber* ei as SAD($\mathbf{e}_i$, $\mathbf{c}_p$) and *remove the one with lowest score*
   - 3.3. Calculate $\boldsymbol{c}_{p-1} = (1/p-1)\sum_{i=1}^{p-1} \boldsymbol{e}_i$ and iterate from 3.1 *until a set of training samples is obtained*
3. **Gather** results and select the most highly mixed signatures by comparing with $\mathbf{c}_p$

Mustard's database with known proportions

- Database of 26 pure and mixed (binary & ternary) spectra.
  - ✓ *Collected using RELAB, a bidirectional spectrometer.*
  - ✓ *211 spectral bands in the range 0.4 – 2.5 μm.*
  - ✓ *Ground-truth information about fractional abundances is available for each spectra.*



Mustard's database with known proportions

	True abundance fractions			Sample selection order		
Signature	Anorthosite	Enstatite	Olivine	MSA	OSP	Maximin
p_1	100	-	-	20th	1st	1st
p_2	-	100	-	19th	2nd	5th
p_3	-	-	100	18th	3rd	8th
b_1	90	-	10	17th	19th	4th
b_2	75	-	25	13th	8th	7th
b_3	50	-	50	9th	5th	10th
b_4	25	-	75	12th	12th	6th
b_5	10	-	90	16th	9th	13th
b_6	-	90	10	15th	18th	9th
b_7	-	75	25	11th	13th	2nd
b_8	-	50	50	8th	7th	12th
b_9	-	25	75	10th	16th	14th
b_{10}	-	10	90	14th	14th	3rd
c_1	16.16	16.24	67.60	6th	20th	11th
c_2	16.13	67.85	16.02	7th	11th	16th
c_3	67.81	15.99	16.20	5th	6th	19th
c_4	16.05	41.83	42.12	2nd	4th	18th
c_5	41.92	16.11	41.97	4th	10th	17th
c_6	41.83	41.77	16.40	3rd	15th	15th
c_7	33.61	33.03	33.36	1st	17th	20th

Pure signatures

Binary mixtures

Ternary mixtures

Outlines

1. Introduction to sub-pixel analysis
2. Linear mixture model
3. Nonlinear mixture model
4. Bi-linear mixture model

Bi-linear spectral unmixing

bilinear models are defined by spectral components appearing in the widely used LMM but also by bilinear terms corresponding to possible interactions between the different materials of the scene.

$$x = \sum_{i=1}^{M} a_i s_i + w$$

$$x = \sum_{r=1}^{R} a_r s_r + \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \beta_{i,j} m_i \bullet m_j + w$$

product of the *i*th and *j*th spectra

amplitude of the interaction term due to the *i*th and *j*th components

$$m_i \bullet m_j = \begin{pmatrix} m_{1,i} m_{1,j} \\ \cdot \\ \cdot \\ \cdot \\ m_{L,i} m_{L,j} \end{pmatrix}$$

Bi-linear spectral unmixing

bilinear models are defined by spectral components appearing in the widely used LMM but also by bilinear terms corresponding to possible interactions between the different materials of the scene.
Nascimento model:

$$x = \sum_{i=1}^{M} a_i s_i + w$$

$$x = \sum_{r=1}^{R} a_r s_r + \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \beta_{i,j} m_i \bullet m_j + w$$

product of the *i*th and *j*th spectra

$$m_i \bullet m_j = \begin{pmatrix} m_{1,i} m_{1,j} \\ \cdot \\ \cdot \\ \cdot \\ m_{L,i} m_{L,j} \end{pmatrix}$$

amplitude of the interaction term due to the *i*th and *j*th components

$$a_r \geq 0$$

$$\beta_{i,j} \geq 0$$

$$\sum_{r=1}^{R} a_r + \sum_{i=1}^{R-1} \sum_{j=1+1}^{R} \beta_{i,j} = 1$$

Bi-linear spectral unmixing

By considering the interaction $m_i * m_j$ as a new spectral component having fractions $\beta_{i,j}$ the formula can be rewritten as:

$$x = \sum_{p=1}^{R^*} a_p^* m_p^* + w$$

Where:

$$a_p^* = a_r \quad , m_p^* = m_r \qquad p = 1,\ldots,R$$

$$\beta_{i,j} \geq 0$$

$$a_p^* = \beta_{i,j} \; , m_p^* = m_i \bullet m_j \qquad \mathrm{R} + 1 \leq p \leq R^*$$

Bi-linear spectral unmixing

Fan's model:

$$x = \sum_{r=1}^{R} a_r s_r + \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} a_i a_j m_i \bullet m_j + w$$

The Fan's model assumes that the amplitudes of the interactions depends on the component fractions involved in the mixture.

Bi-linear spectral unmixing

Generalized model:

$$x = \sum_{r=1}^{R} a_r s_r + \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \gamma_{i,j} a_i a_j m_i \bullet m_j + w$$

quantifies the interactions terms between the different spectral components.

The generalized model assumes that the contribution of the interaction term $m_i * m_j$ is proportional to the fractions of the involved components with an amplitude $\lambda_{i,j} a_i a_j$

Generalized method is more flexible

Summary

- VD estimation
  - To estimate the number of endmembers to be extracted
- Dimensionality reduction (optional)
  - To improve the performance of endmember extraction
- Endmember extraction
  - To determine endmember signatures, which may or may not be image pixels
- Endmember selection (optional)
  - To optimize the endmember set that is actually used for unmixing
- Abundance estimation (linear or nonlinear)
  - To estimate endmember abundance, where linear and nonlinear unmixing may be integrated.