On the Use of a Hybrid Approach to Contrast Endmember Induction Algorithms

Miguel A. Veganzones Carmen Hernández

Computational Intelligence Group
Basque Country University

The 5th International Conference on Hybrid Artificial Intelligence Systems (HAIS'10), San Sebastián 2010

Overview

1. Endmember Induction Algorithms (EIAs)
2. Proposed methodology
3. Experimentation

Hyperspectral images

- Optical passive sensors.
- VNIR (Visible/Near InfraRed), SWIR (Short-Wave InfraRed).
- High number of bands: >100.
- High spectral resolution.
- Contiguity: regular spectrum measures.
- Applications: Earth Observation, industrial quality processes.

Hyperspectral cube

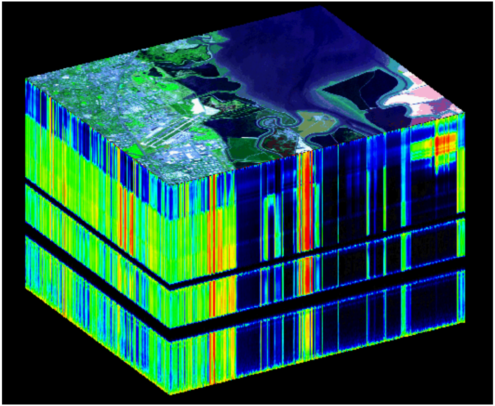
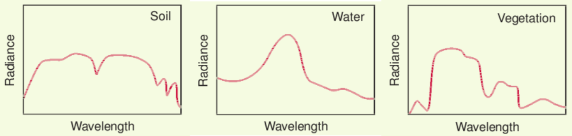


Figure: Image from the JPL's Airborne Visible/Infrared Imaging Spectrometer flying up to 20.000 meters over Moffett Field, California.

Endmembers

- Spectral signatures of objects on a given scale, resolution and wavelengths.
- Patterns of emission and absorption (shape).



Spectral libraries

- Collection of spectral signatures from different materials obtained by in-lab spectroscopy techniques.
- Some public libraries:
  - United States Geological Survey (USGS) Spectroscopy Lab.
  - Advance Spaceborne Thermal Emission and Reflection Radiometer (ASTER).
- Manual user guided process.
- A priori information.

Endmember Induction Algorithms (EIAs)

- Try to induce the spectral signatures of the materials from the image itself.
- No a priori information.
- Automatic unsupervised methodology.
- Approaches: geometrics, lattice computing, heuristics.

EIAs Validation

- Provide an objective meassure of the goodness of the EIAs.
- Directly: hability to detect known materials.
  - Groundtruth knowledge: hard and error prone.
- Indirectly: by application performance.
  - Often there is no need (discriminative features) or possibility (mixtures, noise) to induce the real spectral sinatures.
  - Classification accuracy: lack of groundtruth, few classes, poor conclusions ...

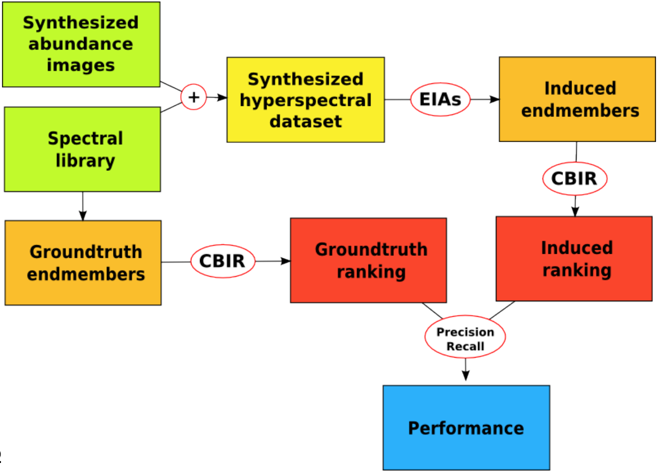
Overview

1. Endmember Induction Algorithms (EIAs)

2. Proposed methodology

3. Experimentation

Schema


Synthesized abundance images
+
Spectral library
Synthesized hyperspectral dataset
EIAs
Induced endmembers
CBIR
Groundtruth endmembers
CBIR
Groundtruth ranking
Induced ranking
Precision Recall
Performance

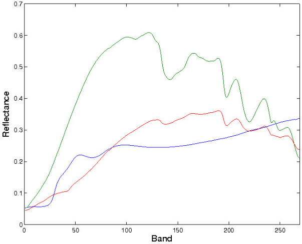
o

Methodology

1. Build a dataset of synthetic hyperspectral images.
2. Ground truth ranking:
    1. Compute the dissimilarities between each pair of images in the dataset on the basis of the image ground truth endmembers.
    2. For each image, rank all the images in the dataset with respect to their ground truth dissimilarity.
3. Induced ranking:
    1. For eah image, compute its endmembers using the selected EIA.
    2. Compute the dissimilarities between each pair of images in the dataset on the basis of the image induced endmembers.
    3. For each image, rank all the images in the dataset with respect to their induced dissimilarity.
4. Compare the rankings obtained by the use of ground truth and induced endmembers.

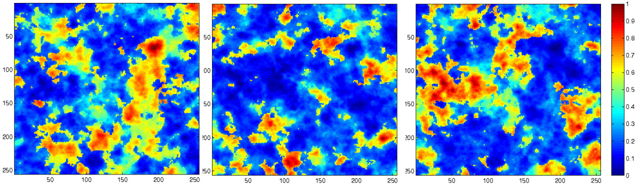
Synthetic Hyperspectral Image Module

- The synthetic hyperspectral images are generated as linear mixtures of a set of spectra with synthesized abundance images.



Left: three spectral signatures (groundtruth endmembers) selected from the spectral library.

Below: three generated synthetic abundance images.



CBIR module

- A CBIR system is based on the definition of a dissimilarity measure between the images.
- Given two hyperspetral images, H_ξ, H_γ, we compute the following distance matrix:

$$D_{\xi,\gamma} = \left[d_{i,j}; i = 1, \ldots, p_\xi; j = 1, \ldots, p_\gamma\right]$$

where $d_{i,j}$ is any defined distance between the two endmembers, e_i^ξ, e_j^γ.

- Then, the dissimilarity is given by:

$$S\left(H_\xi, H_\gamma\right) = \left(\|m_r\| + \|m_c\|\right)\left(\left|p_\xi - p_\gamma\right| + 1\right)$$

where m_r and m_c are the vectors built of the minimal values of the distance matrix $D_{\xi,\gamma}$ by rows and columns respectively.

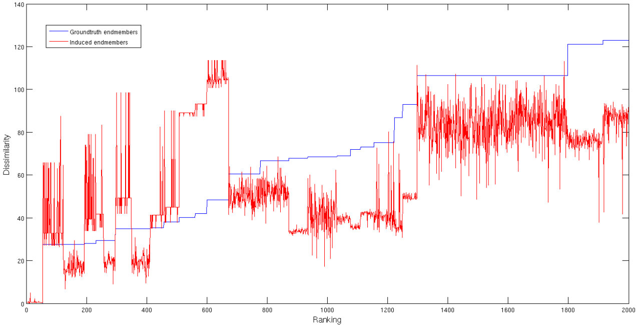
Performance measures

- From a CBIR point of view the goal is to retrieve the k more similar images from the dataset given a query image.
- Use of precision and recall measures:

$$\text{precision}_k = \frac{|R \cap T|}{|T|} \qquad \text{recall}_k = \frac{|R \cap T|}{|R|}$$

  where T is the set of returned images and R is the set of images relevant to the query of size k.
- R and T are defined in base to the groundtruth ranking for each query of size k.

Similarities comparison



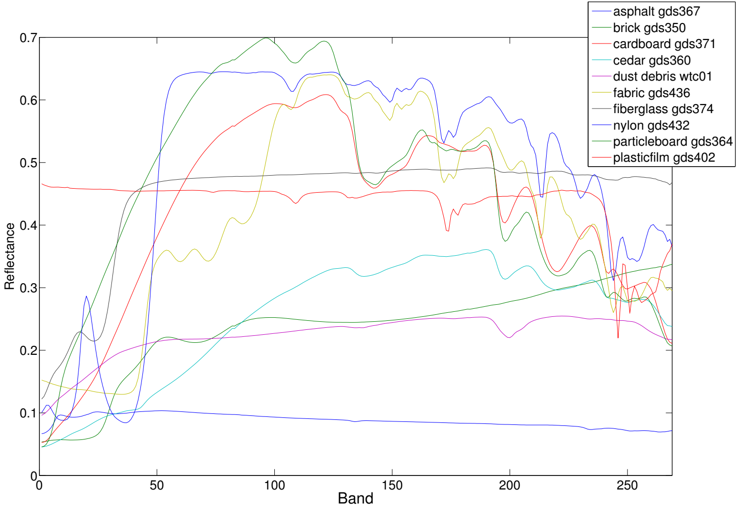
Overview

1. Endmember Induction Algorithms (EIAs)
2. Proposed methodology
3. Experimentation

Datasets

- We have synthesized a total of 6000 hyperspectral images divided in three datasets of 2000 images each.
- Each dataset is defined by the number of endmembers in the repository of groundtruth endmembers selected from USGS library: 5, 10 or 20 endmembers.
  - This represents an increasing diversity in the materials presented in the dataset.
- The procedure to generate an image is the following:
  - Select the number of endmembers presented on the image (2 to 5).
  - Select randomly the endmembers from the repository.
  - Generate the synthetic abundance images as Gaussian Random Fields with Mattern correlation function.
  - Compute the linear combination of the abundance images and the endmembers.

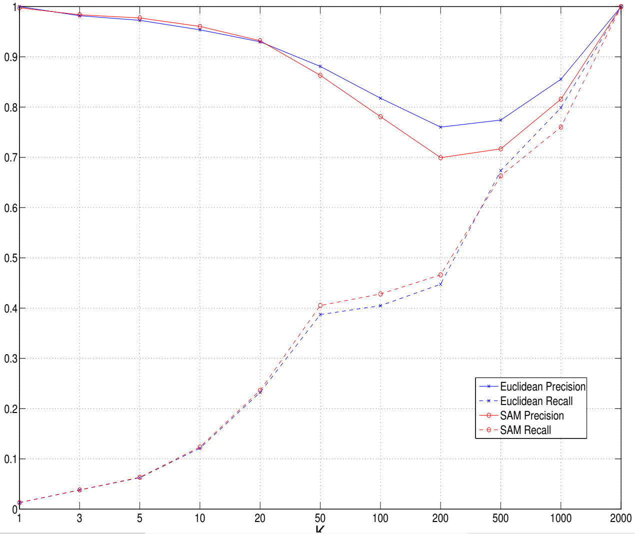
10-Dataset endmembers



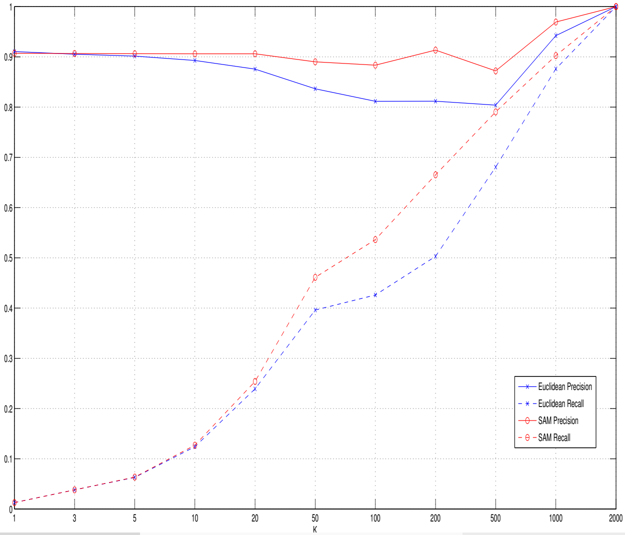
Endmember Induction Algorithms

- We have tested the performance of two EIAs:
  - N-FINDER: it works by inflating a simplex inside the data, beginning with a random set of pixels.
  - Endmembers Induction Heuristic Algorithm (EIHA): based on the equivalence between Strong Lattice Independence and Affine Independence.
- Metric: we have used the Euclidean distance and the Spectral Angle Distance.

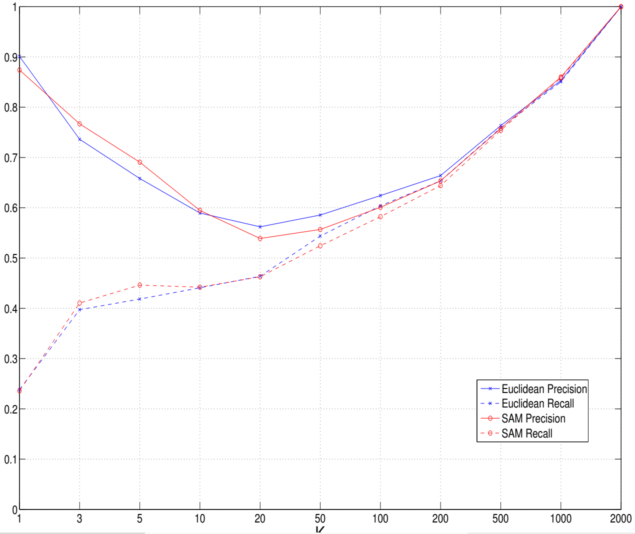
Experimental results: 5-dataset - EIHA



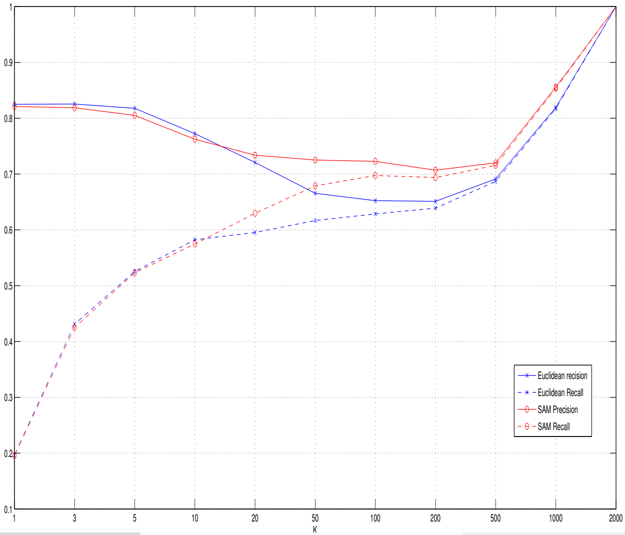
Experimental results: 5-dataset - N-FINDER



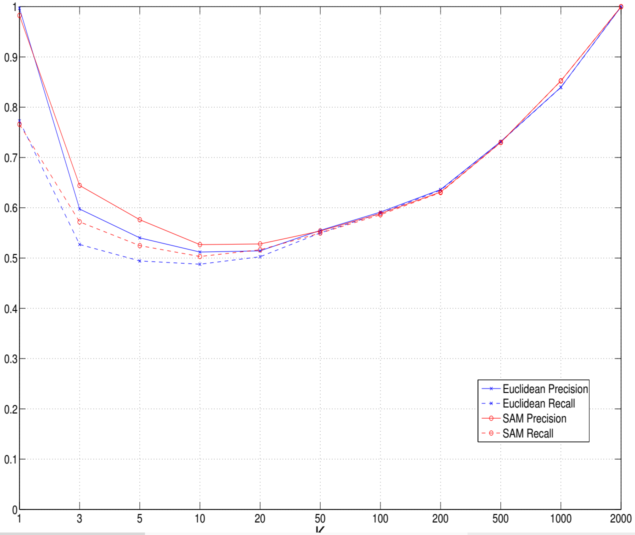
Experimental results: 10-dataset - EIHA



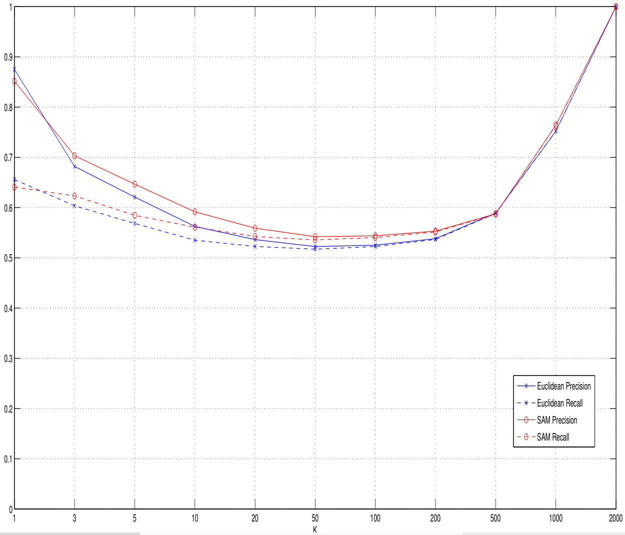
Experimental results: 10-dataset - N-FINDER



Experimental results: 20-dataset - EIHA



Experimental results: 20-dataset - N-FINDER



Conclusions

- We propose a hybrid approach for the evaluation and comparison of Endmember Induction Algorithms.
- We propose the use of CBIR based performance measures based on the spectral information of the images.
- The performance of the two tested algorithms are similar.
    - Focus on other aspects: computational cost, induced endmembers meaning, ...
- Future: more experimental research.
    - Try more algorithms.
    - Pixel purity of abundance images.

Thanks

Thank you very much for your attention.

- *Contact*:
  - Miguel Angel Veganzones
  - Computational Intelligence Group
  - Universidad del Pais Vasco (UPV/EHU)
  - E-mail: **miguelangel.veganzones@ehu.es**
  - Web: **http://www.ehu.es/computationalintelligence**