

# CBIR for hyperspectral images

Miguel A. Veganzones

Grupo Inteligencia Computacional  
Universidad del País Vasco

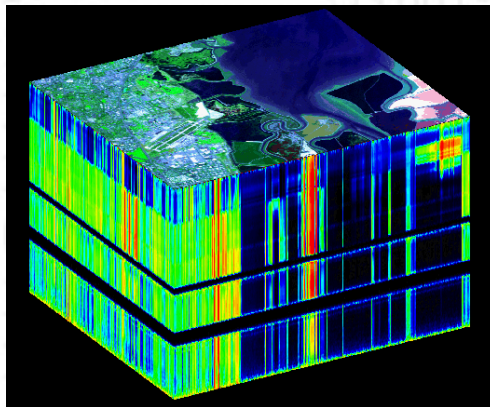
# Outline

- 1 Introduction
  - Hyperspectral images
  - CBIR systems
- 2 Feature Characterization
  - Endmember induction and unmixing
  - Information quantification
- 3 CBIR system for hyperspectral images
  - Queries
  - Retrieval
- 4 Experiment
  - Design
  - Results
  - Conclusions

# Outline

- 1 Introduction
  - Hyperspectral images
  - CBIR systems
- 2 Feature Characterization
  - Endmember induction and unmixing
  - Information quantification
- 3 CBIR system for hyperspectral images
  - Queries
  - Retrieval
- 4 Experiment
  - Design
  - Results
  - Conclusions

# AVIRIS cube

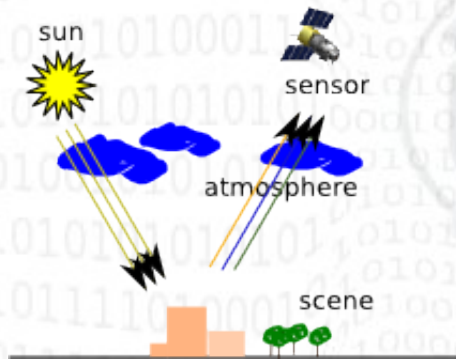


**Figure:** Imagen tomada desde el JPL's Airborne Visible/Infrared Imaging Spectrometer volando a 20.000 metros sobre Moffett Field, California.

# Hiperespectrales VS Multiespectrales

- Número de bandas:
  - Color/Multiespectrales: 3-10 bandas.
  - Hiperespectrales: >100.
- Resolución espectral: longitud de onda/ancho de banda
  - Color/Multiespectrales: orden de 10.
  - Hiperespectrales: orden de 100.
- Contigüidad:
  - Color/Multiespectrales: muestreos irregulares del espectro.
  - Hiperespectrales: muestreos regulares del espectro.

# Sistemas de imagen hiperespectral

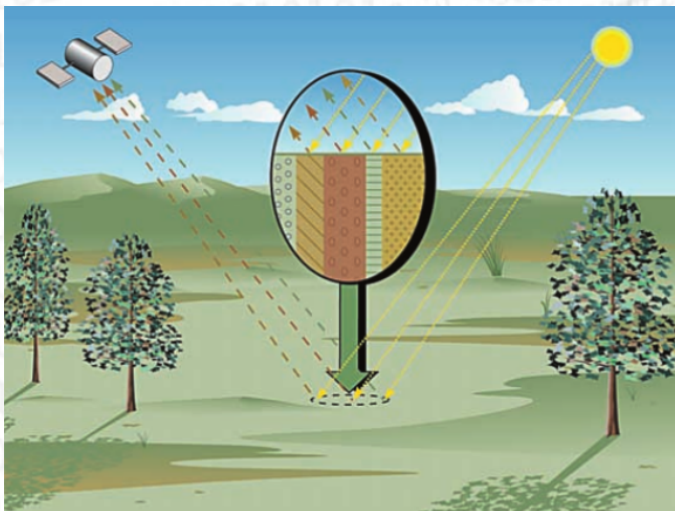


# Información espacial/espectral

- Información espacial:
  - Cada pixel representa un espacio determinado de la escena.
  - Depende de la altitud y apertura del sensor.
- Información espectral:
  - Se obtiene mediante un interferómetro o prisma.
  - Un conversor convierte la radiancia muestreada en cada señal espectral.

# Modelo de mezcla lineal

## Ilustración





# Modelo de mezcla lineal

## Formulación

### LMM

- $H = A \cdot E + \eta$
- $\mathbf{h}(x, y) = a(x, y)_1 \cdot \mathbf{e}_1 + a(x, y)_2 \cdot \mathbf{e}_2 + \dots + a(x, y)_p \cdot \mathbf{e}_p + \eta$

donde:

- $H$  es una imagen hiperespectral con dimensiones espaciales  $m \times n$  y con  $d$  bandas espectrales.
- $A$  es una imagen de abundancias espectrales con dimensiones espaciales  $m \times n$ .
- $E$  es un conjunto de  $p$  firmas espectrales (endmembers) con  $d$  bandas.
- $\eta$  es ruido aditivo.

# Modelo de mezcla lineal

## Restricciones

### LMM

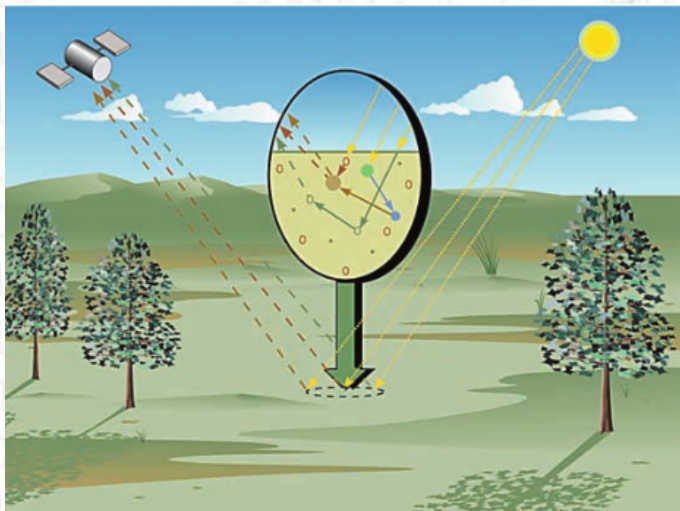
- $H = A \cdot E + \eta$
- $\mathbf{h}(\mathbf{x}, \mathbf{y}) = a(x, y)_1 \cdot \mathbf{e}_1 + a(x, y)_2 \cdot \mathbf{e}_2 + \dots + a(x, y)_p \cdot \mathbf{e}_p + \eta$

sujeito a:

- Abundance Non-negative Constraint (ANC):  $a(x, y)_i \geq 0$
- Abundance Sum-to-one Constraint (ASC):  $\sum_i a(x, y)_i = 1$

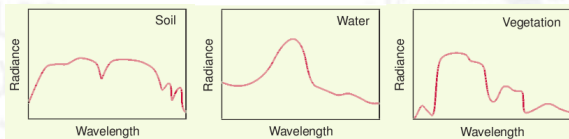
# Modelo de mezcla no lineal

## Ilustración



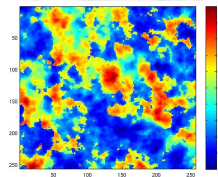
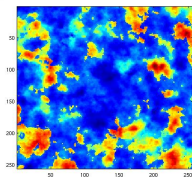
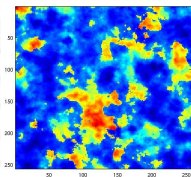
# Endmembers

- Firmas espectrales de distintos materiales a una escala, resolución y frecuencias dadas.
- USGS library: firmas espectrales de multitud de materiales obtenidas mediante técnicas de espectroscopía con microscopios en laboratorio.



# Imágenes de abundancia

- Indican la proporción de cada material en la imagen.
- Información espacial.



# Demezclado (Unmixing)

- Obtener las imágenes de abundancia a partir de la imagen hiperespectral original y un conjunto de firmas espectrales (endmembers).
- Estimación mediante mínimos cuadrados (Least-Squares Estimation).

# Outline

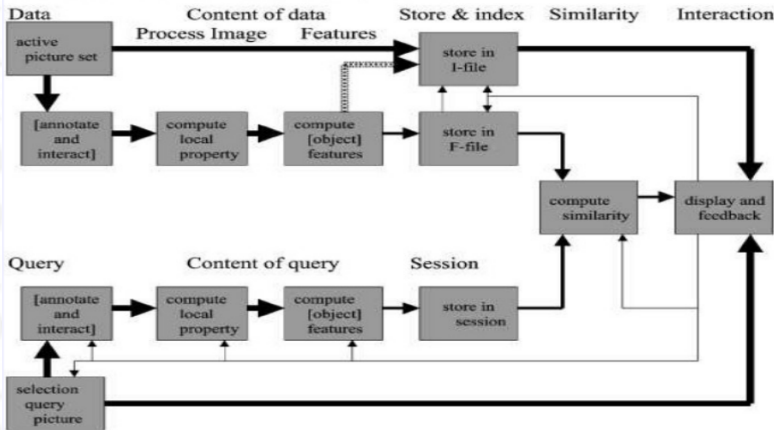
- 1 Introduction
  - Hyperspectral images
  - **CBIR systems**
- 2 Feature Characterization
  - Endmember induction and unmixing
  - Information quantification
- 3 CBIR system for hyperspectral images
  - Queries
  - Retrieval
- 4 Experiment
  - Design
  - Results
  - Conclusions

# Objetivos

- Recuperar información de grandes bases de datos (imágenes).
- Superar las deficiencias de los métodos tradicionales basados en metadatos.
- Usar la información contenida en las imágenes como base para las búsquedas.
- Elaboración de métricas basadas en la caracterización de la información contenida en las imágenes.



# Descripción



\* From "Content-Based Image Retrieval at the end of the early years". W.M.Smeulder et al. IEEE Trans. on Pattern Analysis and Machine Intelligence (2000)

# Retrieval feedback

- Salto semántico: existe una brecha entre la información semántica buscada por el usuario y la caracterización de la información de las imágenes.
- Especialmente importante en dominios amplios (variabilidad del catálogo de imágenes).
- Retrieval feedback: proceso iterativo por el cual el usuario refina la búsqueda en función de los resultados previos (selección de resultados positivos y negativos).

# Outline

- 1 Introduction
  - Hyperspectral images
  - CBIR systems
- 2 Feature Characterization
  - Endmember induction and unmixing
  - Information quantification
- 3 CBIR system for hyperspectral images
  - Queries
  - Retrieval
- 4 Experiment
  - Design
  - Results
  - Conclusions

# Endmember induction

- Induce the set of endmember that form the hyperspectral image.
- It must be an automatic and, desirably, a fast process.
- Different methodologies: geometrical, heuristics, morphological, ...
- More or less, they all follow the linear mixing model.

# Unmixing

- Extract the abundancies of each endmember in the hyperspectral image.
- Different methods depending on the restrictions to the model and the provided information about the endmembers.

# Outline

- 1 Introduction
  - Hyperspectral images
  - CBIR systems
- 2 **Feature Characterization**
  - Endmember induction and unmixing
  - **Information quantification**
- 3 CBIR system for hyperspectral images
  - Queries
  - Retrieval
- 4 Experiment
  - Design
  - Results
  - Conclusions

# Feature vector

- The feature vector describing an hyperspectral image is defined by:
  - A set of endmembers.
  - A probability of occurrence of each endmember.
  - The abundance images of each endmember.
- The number of features for each image is variable: each image has a different number of endmembers.

# Methodologies

- Endmember induction:
  - Virtual dimensionality methods: helps to tune the induction method.
  - Morphological methods: fast and automatic.
- Abundance extraction:
  - Least Squares method.
  - Full-Constrained Least Squares method.
- Modelling occurrence and spatial information:
  - Statistics: mean, variance, kurtosis, ...
  - Markov Random Fields.



# Outline

- 1 Introduction
  - Hyperspectral images
  - CBIR systems
- 2 Feature Characterization
  - Endmember induction and unmixing
  - Information quantification
- 3 CBIR system for hyperspectral images
  - Queries
  - Retrieval
- 4 Experiment
  - Design
  - Results
  - Conclusions

# Spectral queries

- Images containing a set of specific endmembers:

$$Q = \{\cup_{i=1}^n E_i\}.$$

- Images containing a set of specific endmembers and not containing a distinct set of specific endmembers:

$$Q = \{\cup_{i=1}^n E_i\} \wedge \neg \{\cup_{j=1}^m E_j\}, \text{ where } E_i \neq E_j, \forall i, j.$$

# Spectral/Spatial queries

- Images containing a set of specific endmembers whose occurrence is given by a probability distribution function:

$$Q = \{\bigcup_{i=1}^n F(E_i)\}.$$

- Images containing a set of specific endmembers with a determined spatial distribution:  $Q = \{\bigcup_{i=1}^n (E_i \wedge A_i)\}.$

# Outline

- 1 Introduction
  - Hyperspectral images
  - CBIR systems
- 2 Feature Characterization
  - Endmember induction and unmixing
  - Information quantification
- 3 CBIR system for hyperspectral images
  - Queries
  - Retrieval
- 4 Experiment
  - Design
  - Results
  - Conclusions

# System definition

- The user defines the query.
- The system characterizes the set of images: endmembers and abundancies (independently denoted as samples).
- The system compares the query characteristics with the characteristics of the database images and establish a similarity ordered list,  $S$ .
- The system presents to the user the  $k > 0$  first images of  $S$ , denoted by  $S_k$ .
- The user redefines the query by adding relevance information in  $S_k$ .

## Only positive samples

- When the query is defined only as a set of positive samples.
- Model the positive class and retrieve the images with higher probability of being an occurrence of the modelled class.
- Kernel One-Class Support Vector Machine (KOC-SVM) [5].

# Positive and negative samples

- The query is defined as a set of positive and negative samples.
- The positive samples form a well defined class.
- The negative samples form a very heterogeneous group that cannot be modelled as a class but it gives usable information.
- Alternatives:
  - Classic two-classes SVM.
  - One-class SVM /SVDD for both positive/negative classes [5, 6].

# Only negative samples

- The query is defined only as a set of negative samples.
- The desired class cannot be modelled but negative samples can be used to restrict the search.
- One-class SVM / SVDD for negative samples [6].



## With occurrence probability

- The query is defined by a set of positive and/or negative samples and an associated probability distribution function for each.
- Each endmember has associated a probability of occurrence.
- Alternatives:
  - Modifications of the previous methodologies.
  - Weighted kernel density estimations.

## With spatial distribution

- The query is defined by a set of positive / negative samples and abundance images associated to them.
- Each endmember has associated an abundance image.
- Alternatives:
  - Modifications of previous methodologies.
  - Model the spatial information independently (MRF).

# Outline

- 1 Introduction
  - Hyperspectral images
  - CBIR systems
- 2 Feature Characterization
  - Endmember induction and unmixing
  - Information quantification
- 3 CBIR system for hyperspectral images
  - Queries
  - Retrieval
- 4 Experiment
  - Design
  - Results
  - Conclusions

# Outline

- 1 Introduction
  - Hyperspectral images
  - CBIR systems
- 2 Feature Characterization
  - Endmember induction and unmixing
  - Information quantification
- 3 CBIR system for hyperspectral images
  - Queries
  - Retrieval
- 4 Experiment
  - Design
  - Results
  - Conclusions

# Outline

- 1 Introduction
  - Hyperspectral images
  - CBIR systems
- 2 Feature Characterization
  - Endmember induction and unmixing
  - Information quantification
- 3 CBIR system for hyperspectral images
  - Queries
  - Retrieval
- 4 Experiment
  - Design
  - Results
  - Conclusions

## For Further Reading I

-  Hyperspectral Data Exploitation: Theory and Applications. Chein-I Chang. 2007.
-  Hyperspectral Imaging: Techniques for Spectral detection and Classification. Chein-I Chang. 2003.
-  Signal Theory Methods in Multispectral Remote Sensing. David A. Landgrebe. 2003.
-  Remote Sensing: the Image Approach, 2<sup>nd</sup> Edition. John R. Scott. 2007.
-  One-Class SVM for Learning in Image Retrieval. Yunqiang Chen, Xiang Zhou, Thomas S. Huang. Proc. IEEE Int. Conf. on Image Processing, Thessaloniki, Greece. 2001.

## For Further Reading II



Non-Relevance Feedback Document Retrieval Based on One-Class SVM and SVDD. Takashi Onoda, Hiroshi Murata, Seiji Yamada. International Joint Conference on Neural Networks, Vancouver, Canada. 2006.

# Questions?

*Thank you very much for your attention.*

- Contact:
  - Miguel Angel Veganzones
  - Grupo Inteligencia Computacional
  - Universidad del País Vasco - UPV/EHU (Spain)
  - E-mail: miguelangel.veganzones@ehu.es
  - Web page: <http://www.ehu.es/computationalintelligence>