Image analysis of hyperspectral data using mathematical morphology

M. Dalla Mura[1] and M. Fauvel[2]

[1] GIPSA-Lab, Grenoble Institute of Technology - France
[2] Universitï¿½ de Toulouse, INP-ENSAT, UMR 1201 DYNAFOR, France

Tutorial WHISPERS 2014

24-27 June 2014, Lausanne, Switzerland

History

2001 M. Pesaresi & J.A. Benediktsson: **Morphological Profile**

2003 J.A. Benediktsson, M. Pesaresi & K. Arnason: **Morphological Profile**

2005 J.A. Benediktsson, J. A. Palmason & J. Sveinsson: **Extended MP**

2009 M. Fauvel, J. Chanussot & J.A. Benediktsson: **Kernel PCA based EMP**

2010 M. Dalla Mura, J.A. Benediktsson, Waske & L. Bruzzone: **Attribute Profile**

2011 M. Dalla Mura, A. Villa, J.A. Benediktsson & L. Bruzzone: **ICA based AP**

2014 WHISPERS Tutorial!

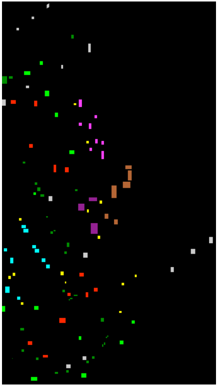
VHR Hyperspectral Images: Urban Area

ROSIS: 1.5 m/pixel - 115 spectral channels - 430 nm, 860 nm.

- Spatial resolution: Lower than 2.5 meter by pixel
- Spectral resolution: More than 100 spectral channels (up to 4096!)

Classification

Original Data



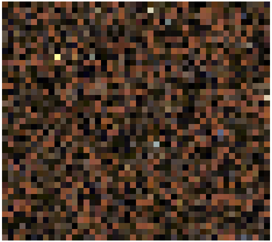
Ground-truth

Thematic Map

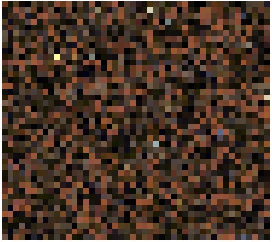
Classification: A pattern recognition approach

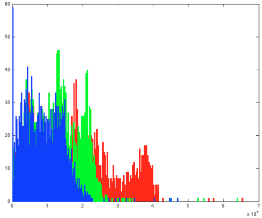
1. Feature extraction: One vector of attributes extracted for every pixels
2. Pattern recognition algorithms: Support Vectors Machines ...

Why spatial information? 1/2

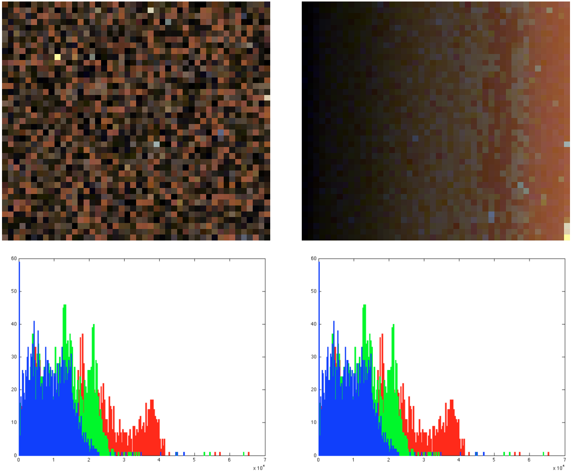


Why spatial information? 1/2

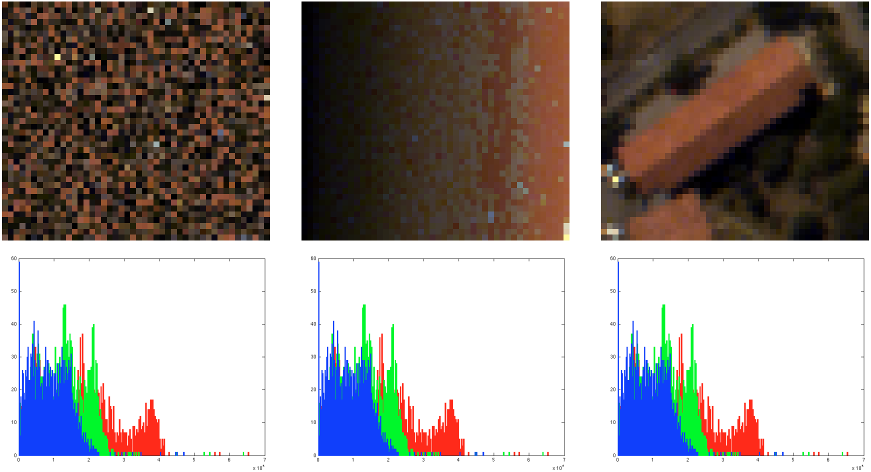




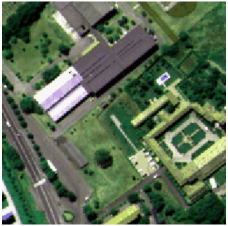
Why spatial information? 1/2



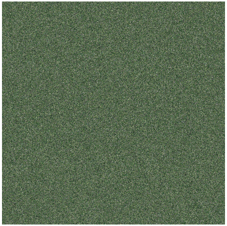
Why spatial information? 1/2



Why spatial information? 2/2



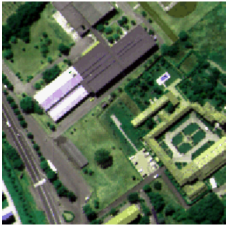
Random permutation
of the spatial position
of the pixels



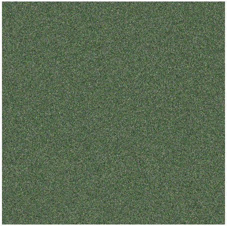
Spectral classification → Same classification !! ← Spectral classification

Need to incorporate information from the spatial domain

Why spatial information? 2/2



Random permutation
of the spatial position
of the pixels



Spectral classification → Same classification !! ← Spectral classification

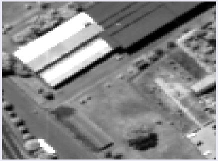
---

Need to incorporate information from the spatial domain

Dilation and Erosion

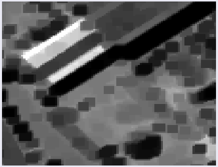
Definition (Image)

$$f : \mathcal{D}_f \subset \mathbb{Z}^d \rightarrow \{0, \ldots, f_{max}\}$$
$$f^c = f_{max} - f$$



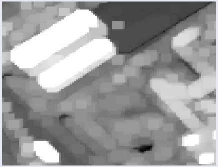
Definition (Erosion)

$$\epsilon_B(f) = \bigwedge_{b \in B} f_{-b}$$

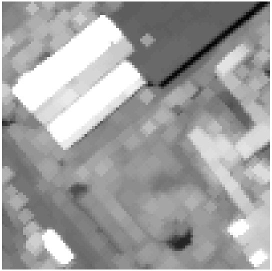


Definition (Dilation)

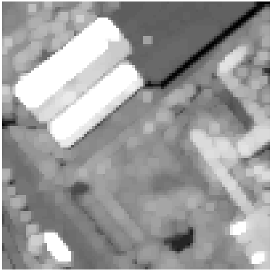
$$\delta_B(f) = \bigvee_{b \in B} f_{-b}$$



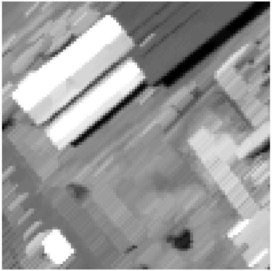
Effect of the structuring element B



0	0	1	0	0
0	1	1	1	0
1	1	1	1	1
0	1	1	1	0
0	0	1	0	0



1	1	1
1	1	1
1	1	1

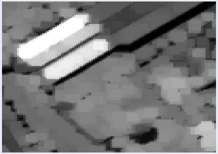


0	0	0	0	1
0	0	0	1	0
0	0	1	0	0
0	1	0	0	0
1	0	0	0	0

Opening and Closing

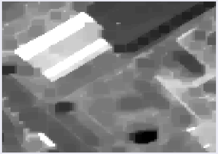
Definition (Opening)

$$\gamma_B(f) = \delta_B \circ \epsilon_B(f)$$



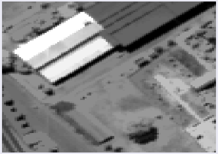
Definition (Closing)

$$\phi_B(f) = \epsilon_B \circ \delta_B(f)$$



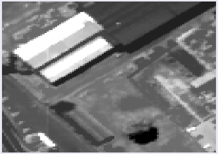
Definition (Opening by reconstruction)

$$\gamma_B^r(f) = Rec_f \circ \epsilon_B(f)$$



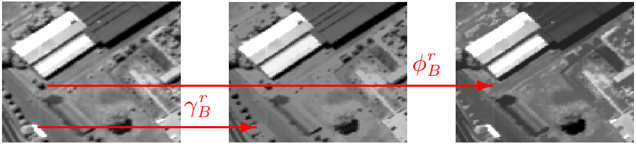
Definition (Closing by reconstruction)

$$\phi_B^r(f) = Rec_f^* \circ \delta_B(f)$$



Opening and Closing Profile

- For a given B, $\gamma_B^r(f)$ (resp. $\phi_B^r(f)$) indicates which clear (dark) objects fit B.



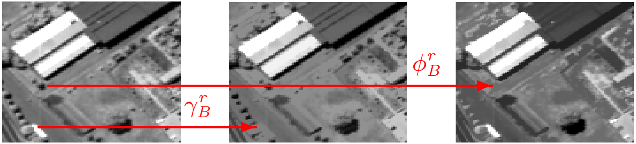
- Applying γ_{B_i} with a set of $\left\{B_i | B_i \subset B_{i+1}, i \in [1, \ldots, n]\right\}$: Opening Profile
- Applying ϕ_{B_i} with a set of $\left\{B_i | B_i \subset B_{i+1}, i \in [1, \ldots, n]\right\}$: Closing Profile

Definition (Granulometry)

A granulometry is a transformation having a size parameter n, and with the following properties: *Anti-extensivity*, *Increasingness* and *Absorption*.

Opening and Closing Profile

- For a given B, $\gamma_B^r(f)$ (resp. $\phi_B^r(f)$) indicates which clear (dark) objects fit B.



- Applying γ_{B_i} with a set of $\left\{B_i | B_i \subset B_{i+1}, i \in [1, \ldots, n]\right\}$: Opening Profile
- Applying ϕ_{B_i} with a set of $\left\{B_i | B_i \subset B_{i+1}, i \in [1, \ldots, n]\right\}$: Closing Profile

Definition (Granulometry)

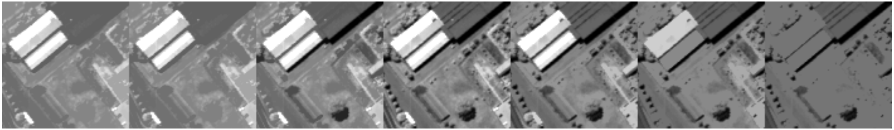
A granulometry is a transformation having a size parameter n, and with the following properties: *Anti-extensivity*, *Increasingness* and *Absorption*.

Morphological Profile

Definition (Morphological profile)

The Morphological Profile of size n is a $(2n+1)$-dimensional vector such as:

$$\mathsf{MP}(\mathbf{x}) = \Big[\mathsf{CP}_n(\mathbf{x}), f(\mathbf{x}), \mathsf{OP}_n(\mathbf{x})\Big].$$



$\mathsf{CP}(\mathbf{x}) \longleftarrow \quad f(\mathbf{x}) \quad \longrightarrow \mathsf{OP}(\mathbf{x})$

For a given pixel $\mathbf{x}$, information include in the $\mathsf{MP}(\mathbf{x})$ are:

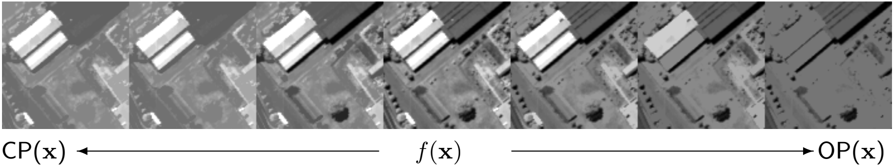
- Contrast: Is the structure to which the pixel belongs to darker or lighter than his surrounding neighbors?
- Size: Is the structure to which the pixel belongs to small or big compared to B_i?

Morphological Profile

Definition (Morphological profile)

The Morphological Profile of size n is a $(2n+1)$-dimensional vector such as:

$$\mathsf{MP}(\mathbf{x}) = \Big[\mathsf{CP}_n(\mathbf{x}), f(\mathbf{x}), \mathsf{OP}_n(\mathbf{x})\Big].$$



For a given pixel $\mathbf{x}$, information include in the $\mathsf{MP}(\mathbf{x})$ are:

- Contrast: Is the structure to which the pixel belongs to darker or lighter than his surrounding neighbors?
- Size: Is the structure to which the pixel belongs to small or big compared to B_i?

Derivative of the MP

Definition (Derivative of the Morphological profile)

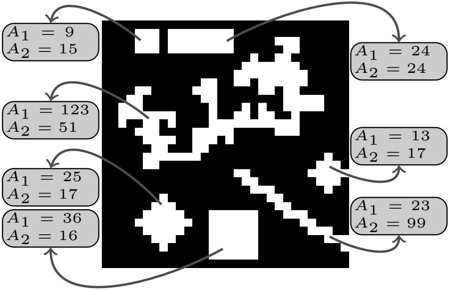
The Derivative of the Morphological Profile of size n is a $(2n)$-dimensional vector such as:

$$\mathrm{DMP}(\mathbf{x}) = \Big[|\phi_n(\mathbf{x}) - \phi_{n-1}(\mathbf{x})|, \ldots, |\gamma_{n-1}(\mathbf{x}) - \gamma_n(\mathbf{x})|\Big].$$

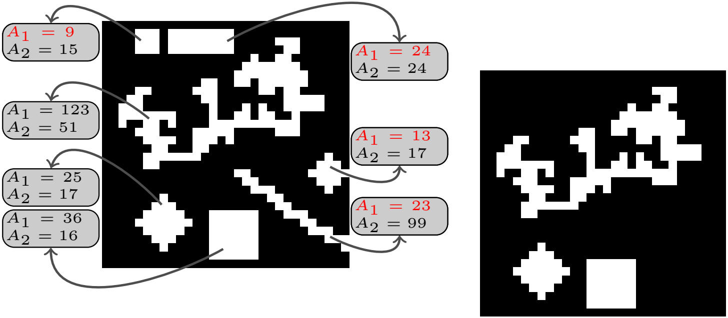
Attribute filters 1/3

- Objects are filtered according to a given attribute (no SE anymore):
  - Area
  - Perimeter
  - Standard deviation
  - Moment of inertia
  - ...
- Combination of attributes are possible !
- Filtering procedure: Attribute filters are based on the following operations:
  - Compute *attribute* for each connected component in the image
  - Keep the components that satisfy the *criterion*.

Attribute filters 2/3

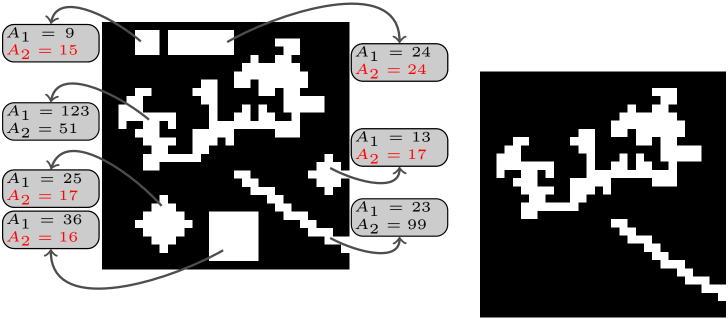


Attribute filters 2/3



$$T_1 = A_1 \geq 25?$$

Attribute filters 2/3

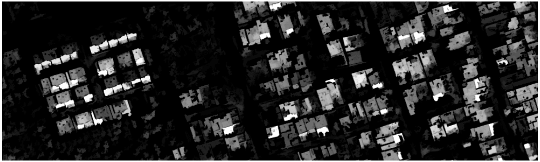


$$T_1 = A_2 > 30?$$

Attribute filters 3/3

Panchromatic image (f)

Filter by rec. $f - \gamma_R^B(f)$ (B: disk radius 7 pixels)

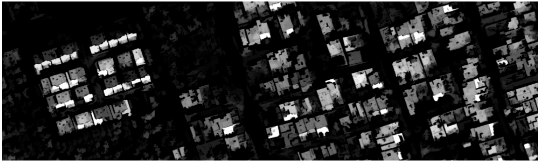


Attribute filter $\gamma^T(f)$ with $T = (R > 0.3) \wedge (I < 0.5) \wedge (50 < A < 5000)$

Attribute filters 3/3

Panchromatic image (f)

Filter by rec. $f - \gamma_R^B(f)$ (B: disk radius 11 pixels)

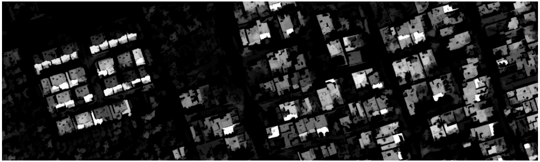


Attribute filter $\gamma^T(f)$ with $T = (R > 0.3) \wedge (I < 0.5) \wedge (50 < A < 5000)$

Attribute filters 3/3

Panchromatic image (f)

Filter by rec. $f - \gamma_R^B(f)$ (B: disk radius 15 pixels)

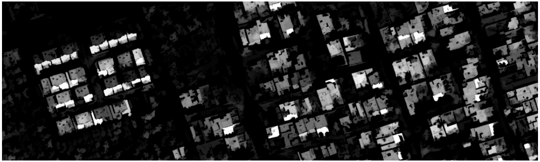


Attribute filter $\gamma^T(f)$ with $T = (R > 0.3) \wedge (I < 0.5) \wedge (50 < A < 5000)$

Attribute filters 3/3

Panchromatic image (f)

Filter by rec. $f - \gamma_R^B(f)$ (B: disk radius 19 pixels)



Attribute filter $\gamma^T(f)$ with $T = (R > 0.3) \wedge (I < 0.5) \wedge (50 < A < 5000)$

1. Introduction
2. Spatial information from MM
   - Basic tools
   - Remote sensing tools
   - Advanced tools
   - Extension to multivalued images
3. Classification
   - Support Vectors Machines
   - Classification of spectral-spatial features
4. Experiments
   - Data
   - Results-ROSIS-03
5. Conclusions

Ordering relation

- MM is based on $\inf$ and $\sup$ operators
- No unambiguous $\inf/\sup$ for pixel/vector:

$$\begin{bmatrix}1\\5\\2\end{bmatrix} \overset{?}{\lesseqgtr} \begin{bmatrix}0\\6\\1\end{bmatrix}$$

- Marginal ordering $\Rightarrow$ by band filtering
- Reduced ordering $\Rightarrow h : \mathbb{R}^d \rightarrow \mathbb{R}$

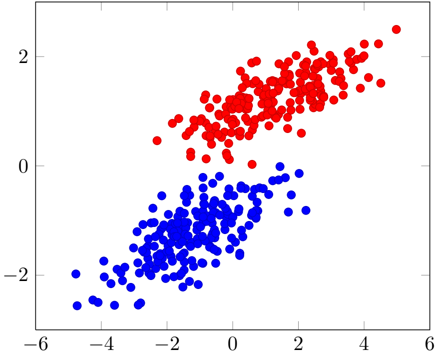
$$\mathbf{x} \mapsto h(x)$$

Principal Component Analysis

- $h(\mathbf{x}) = \langle \mathbf{x}, \mathbf{v}_m \rangle$ such as $\mathbf{v}_m$ is the m^{th} principal component of the spectral data
- $\mathbf{v}_m$ is computed as the eigenvector corresponding to the m^{th} largest eigenvalue of the covariance matrix of the spectral data

$$\lambda_m \mathbf{v}_m = \mathbf{\Sigma} \mathbf{v}_m$$

- The number of retained PCs correspond to a certain percentange of the cummulative variance

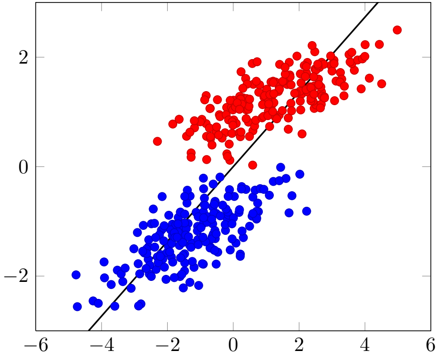


Principal Component Analysis

- $h(\mathbf{x}) = \langle \mathbf{x}, \mathbf{v}_m \rangle$ such as $\mathbf{v}_m$ is the m^{th} principal component of the spectral data
- $\mathbf{v}_m$ is computed as the eigenvector corresponding to the m^{th} largest eigenvalue of the covariance matrix of the spectral data

$$\lambda_m \mathbf{v}_m = \mathbf{\Sigma} \mathbf{v}_m$$

- The number of retained PCs correspond to a certain percentange of the cummulative variance

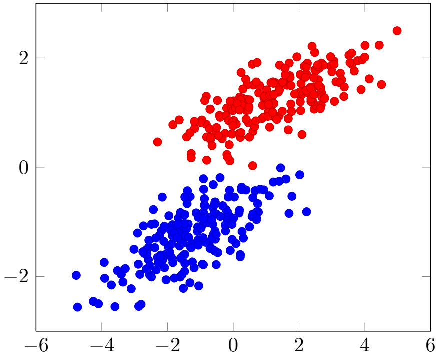


Linear Discriminant Analysis

- $h(\mathbf{x}) = \langle \mathbf{x}, \mathbf{v}_l \rangle$ such as $\mathbf{v}_l$ maximizes the separability of the classes
- $\mathbf{v}_l$ is computed by solving the generalized eigenvalue problem

$$\lambda_l \mathbf{S}_w \mathbf{v}_l = \mathbf{S}_b \mathbf{v}_l$$

where $\mathbf{S}_w = \sum \mathbf{\Sigma}_c$ and $\mathbf{S}_b = \sum (\boldsymbol{\mu}_c - \boldsymbol{\mu})(\boldsymbol{\mu}_c - \boldsymbol{\mu})^T$

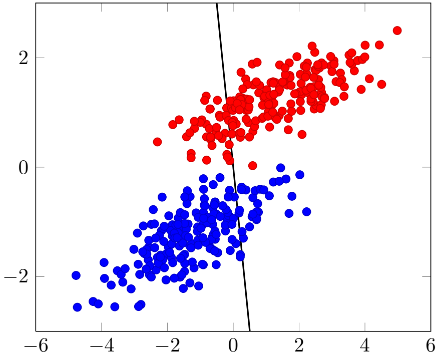


Linear Discriminant Analysis

- $h(\mathbf{x}) = \langle \mathbf{x}, \mathbf{v}_l \rangle$ such as $\mathbf{v}_l$ maximizes the separability of the classes
- $\mathbf{v}_l$ is computed by solving the generalized eigenvalue problem

$$\lambda_l \mathbf{S}_w \mathbf{v}_l = \mathbf{S}_b \mathbf{v}_l$$

where $\mathbf{S}_w = \sum \mathbf{\Sigma}_c$ and $\mathbf{S}_b = \sum (\boldsymbol{\mu}_c - \boldsymbol{\mu})(\boldsymbol{\mu}_c - \boldsymbol{\mu})^T$

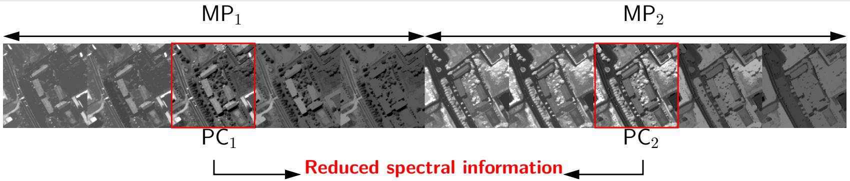


Extended Morphological Profile

Definition (Extended Morphological Profile - EMP)

The EMP of size $n \times p$ is a $(2n+1)p$-dimensional vector made of the MP build with the p first principal components:

$$\text{EMP}(\mathbf{x}) = \Big[\text{MP}_1(\mathbf{x}), \dots, \text{MP}_p(\mathbf{x})\Big].$$

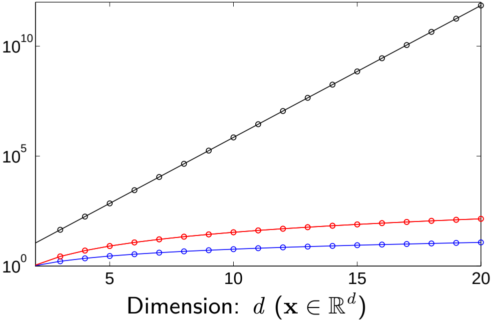


- Fusion of morphological and spectral features

High dimensional data

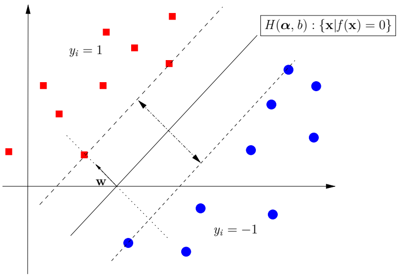
- Number of samples for *accurate* classification:

Linear	$\gg$	d
Quadratic	$\gg$	d^2
Non-parametric	$\gg$	$\exp(d)$



- Hyperspectral data: $d > 100$

Principle

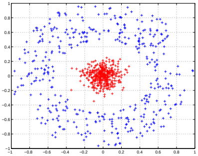


- Training set: $\mathcal{S} = \left\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_\ell, y_\ell)\right\} \in \mathbb{R}^n \times \{-1; 1\}$
- Learn a function $f(\mathbf{z}) = \text{sgn}\left(\sum_{i=0}^{\ell} \alpha_i k(\mathbf{z}, \mathbf{x}_i) + b\right)$
- Solve:

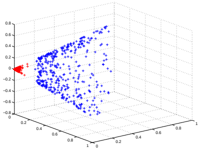
$$\begin{aligned} \max_{\alpha} \; g(\alpha) \quad &= \quad \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{\ell} \alpha_i \alpha_j y_i y_j k(\mathbf{x}_i, \mathbf{x}_j) \\ \text{subject to} \quad & \quad 0 \leq \alpha_i \leq C \text{ and } \textstyle\sum_{i=1}^{\ell} \alpha_i y_i = 0 \end{aligned}$$

Kernel trick

- Kernel function k: semi-definite positive $k(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle_{\mathcal{H}}$

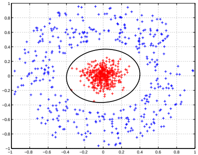


$\xrightarrow{\Phi}$



$\longrightarrow \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle_{\mathcal{H}}$

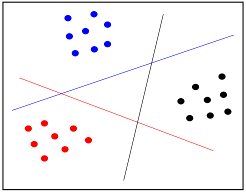
$k(\mathbf{x}_i, \mathbf{x}_j) \longrightarrow$



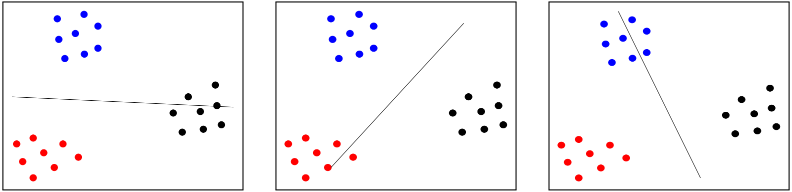
- Some kernels:
  - ⋆ Linear kernel: $k(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{x}_i, \mathbf{x}_j \rangle_{\mathbb{R}^n}$
  - ⋆ Polynomial kernel: $k(\mathbf{x}_i, \mathbf{x}_j) = \Big(\langle \mathbf{x}_i, \mathbf{x}_j \rangle_{\mathbb{R}^n} + q \Big)^p$
  - ⋆ Gaussian kernel: $k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(- \frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2_{\mathbb{R}^n}}{2\sigma^2} \right)$

Multiclass problem

- One versus All: m binary classifiers

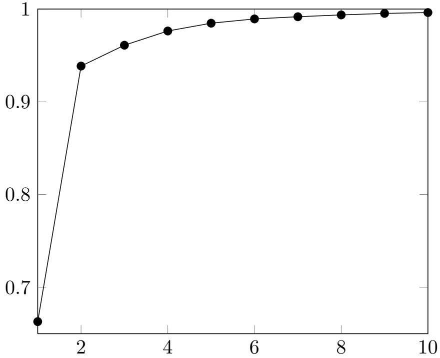


- One versus One: $m(m\text{-}1)/2$ classifiers



EMP

- Inputs: $\text{EMP}(\mathbf{x})$, $\dim\big(\text{EMP}(\mathbf{x})\big) = (2n+1)p$
- Parameters:
  - ⋆ Numbers of PCs $\rightarrow p$
  - ⋆ Size of MP $\rightarrow n$
  - ⋆ Hyperparameters



Cumulative variance

ROSIS-03

University Area, Pavia - Italy

- Airbone
- [H W]=[610 340]
- 103 channels
- 1.3 m/pixel
- 43 923 referenced samples
- 9 classes : Asphalts, Meadow, Gravel, Tree, Metal Sheet, Bare Soil, Bitumen, Brick and Shadow.

Quantitative analysis

Let's generate some results!

Discussion

- Mathematical Morphology:
  - Morphological operators extract relevant spatial features for classification
  - Connectivity is a key property
  - Extended profiles should be constructed with advance FE algorithms
  - The use of attributes profile allows the extraction of relevant spatial features
- SVM:
  - Ability to include additional spatial information
  - Robust to the dimensionality
- Classification:
  - Analyst intervention is minimized
  - Classification accuracy
  - Works with other geographical area

Perspectives

- Supervised or un-supervised FE?
- Best attribute?
- Time processing?

Some references

M. Pesaresi and J. A. Benediktsson.
A new approach for the morphological segmentation of high-resolution satellite imagery.
IEEE Trans. Geosci. Remote Sens., 39(2):309–320, February 2001.

P. Soille and M. Pesaresi.
Advances in mathematical morphology applied to geoscience and remote sensing.
IEEE Trans. Geosci. Remote Sens., 40(9):2042–2055, September 2002.

J.A. Benediktsson, J.A. Palmason and J. R. Sveinsson.
CLassification of hyperspectral data from urban areas based on extended morphological profiles.
IEEE Trans. Geosci. Remote Sens., 43(3):480-491, Mars 2005.

M. Fauvel, J. Chanussot, and J.A. Benediktsson.
Kernel Principal Component Analysis for the Classification of Hyperspectral Remote Sensing Data over Urban Areas.
EURASIP Journal on Advances in Signal Processing, pages 1–14, 2009.

M. Dalla Mura, J.A. Benediktsson, B. Waske and L. Bruzzone.
Extended profiles with morphological attribute filters for the analysis of hyperspectral data.
International Journal of Remote Sensing, 31(22)5975–5991, 2010.

Image analysis of hyperspectral data using mathematical morphology

M. Dalla Mura[1] and M. Fauvel[2]

[1] GIPSA-Lab, Grenoble Institute of Technology - France
[2] Universitï¿½ de Toulouse, INP-ENSAT, UMR 1201 DYNAFOR, France

Tutorial WHISPERS 2014

24-27 June 2014, Lausanne, Switzerland