How to mix spatial and spectral information when processing hyperspectral images

Gilles Rabatel
Cemagref, France
Presentation summary

⇒ Introduction: Hyperspectral images and spatial information

⇒ Markov Random Fields

⇒ Split-and-merge and Unin-Find strategies

⇒ Anisotropic diffusion

⇒ Conclusion
Hyperspectral imaging at Cemagref

Spectrometry & chemometrics (J-M Roger et al)

Image processing (G. Rabatel et al)

"Proximity" hyperspectral imaging
Hyperspectral imaging in a few words

Compared to standard imagery:
- More pixel information → better classification potential

Compared to spectrometry:
- Spatial mapping of the data

Some applications:
- Image segmentation (urban areas, forests, etc.)
- Object detection (e.g. military applications)
- Mapping of chemical components (pharmacy, biology, geology...)
The basics of hyperspectral processing

Two main approaches

- Supervised learning:
  output/spectra relationship learned from examples
- Unsupervised processing:
  clusterisation, pure spectra extraction…

A common requirement: model parcimony

- Looking for a data subspace that keeps class separability
  PLS (supervised); PCA, projection pursuit (unsupervised)
The present situation

🚇 In almost cases, **pixel-based** processing
  ➢ Pixels are just considered as a large set of samples
  ➢ Their neighbourhood relations are not taken into account:
    • neither in the calibration step
    • nor in the classification step
  ➢ A spatial mixing of the pixels would lead to the same processing and the same results

麻辣 A simple question: can we do better by considering pixels as « picture elements », not only as samples ?
Spatial relationships

Basic notion: neighbourhood

Texture

Homogeneity

Region: set of connected pixels with the same attributes (level or texture)

Shape
Homogeneity constraint

- Every pixel in the image have a high *a priori* probability to have spectral attributes close to its’ neighbour ones

- Can be used to:
  - overcome classification ambiguities
  - remove erroneous pixels (outsiders)
  - Improve pixel clusterisation
Example 1: Markov Random Fields (MRF)
Markov Random Field: definition

Let us consider a 2D image \( I(x,y) \)

\( I: \) attribute vector, attribute scalar, or label

Let us consider \( p( I(i,j) = x_s \mid I(k,l) \{k \neq i, l \neq j\} ) \)

Probability for the pixel \((i,j)\) to have a value \(x_s\), knowing every other values in the image

If \( I(x,y) \) is a Markov Random Field, this probability only depends on \((i,j)\) neighbourhood, and can be written as:

\[
p(x_s) = \frac{1}{Z} \exp(-\sum C V(C))
\]

\( Z: \) normalisation constant

\( V(C): \) potential of the clique \( C \)

\( C: \) clique or neighbourhood configuration

Cliques:

1st order  2\textsuperscript{nd} order  3rd order  4th order

Can be used for texture modelisation
MRF for homogeneity regularisation

The Potts model: limited to cliques of 1st and 2nd orders

 أشهر First order potential:
\[ V(C1) = \log \left[ \Pr(xs / S) \right] \]
Pr(xs / S): probability to have a label xs for a spectrum S

Gaussian hypothesis:
\[ V(C1) = (S - \mu_{xs})^T COV^{-1} (S - \mu_{xs}) \]

 أشهر Second order potential:
\[ V(C2) = V(xs, xt) = -\beta \text{ if } xs = xt \]
\[ = +\beta \text{ if } xs \neq xt \]
\[ \beta \text{ parameter: tuning of the regularisation effect} \]
MRF implementation

⇒ The problem: to find the output image corresponding to the maximal probability $pr(xs)$ for every pixel

- Simulated annealing (optimal, heavy computation)
  - Introduction of a temperature parameter
  - Slow decreasing of $T$ to zero.

$$p(xs, T) = \frac{1}{Z(T)} \exp\left( -\frac{\sum C V(C)}{T} \right)$$

- Algorithm ICM (sub-optimal, faster)
  - Image scanning
  - For each pixel, setting of the $xs$ value with the higher probability $pr(xs)$
  - Re-iteration until a stationary state is reached
Hyperspectral example (Prony et al, 2000)

ICM regularisation

(a) No regulation  (b) $\beta = 3$  (c) $\beta = 5$
Unsupervised segmentation based on split-and-merge

N. Gorretta, Cemagref
(under development)
Basic principles

균: Objective: to partition the hyperspectral image in a set of homogeneous regions

균: Splitting process:
Building of a quad-tree
  ➢ Root = initial image
  ➢ For each tree node: children nodes are created as long as the subimage variance is too high

균: Merging process:
  Adjacent regions are merged if the distance between their average pixel values is below a threshold
Implementation

_articles

⇒ Before the split-and-merge segmentation
  - a PCA is made on the spectral data of the subregion
  - Score images are built using a grey-level renormalisation
  - Subregion variances (split) and Euclidian distances (merge) are computed using the multi-dimensional set of scores

⇒ Local segmentation refinement
  - For each resulting sub-region, the entire process is launched again (PCA computation + split-and-merge process)
  - Refinement is stopped when no more subregions are created
  - This allows a hierarchical image segmentation, where the user can tune the level of segmentation detail required.
Results: synthetic test image

- The test image
  - Four different spectra have been generated using a random process
  - A synthetic image has been built, using a particular concentration pattern
Results: image segmentation
Alternative: Union-Find algorithm (Fiorio, 1996)

¬ Starting with a region associated to each pixel
¬ Image scanning, and comparison of each pixel (region) with the pixel (region) above and on the right.
¬ Merging condition: euclidian distance < S

S = 40
Union-Find result

Application to the PC score vector
(after normalisation 0-255)

S = 100

S = 10

S = 2
Anisotropic diffusion
Basic principle (Perona & Malik, 1990)

- Objective: to smooth an image (noise reduction) while preserving edges

- Method:
  - Implementation of the Gaussian smoothing as an iterative process, using PDE formalism (PDE: Partial Derivative Equation)
  - Local attenuation of the iterative smoothing in high gradient areas
Isotropic diffusion

⇒ Let us consider the temporal equation for the image I(x,y):

\[ \frac{\partial I(x,y)}{\partial t} = \text{div}(\nabla I) = \Delta I \quad \text{« Heat equation »} \]

⇒ The equation solution is a temporal gaussian filtering:

\[ I(x,y,t) = I(x,y,t_0) \ast G(x,y,t) \quad \text{where:} \quad G(x,y,t) = \frac{1}{4\pi t} \exp\left(-\frac{(x^2+y^2)}{4t}\right) \]

Gaussian with a variance \( \sigma^2 = 2t \)

Original image 30 iterations 100 iterations
Anisotropic diffusion

- Introduction of a function $g(\nabla I)$ so that:
  - $g(\nabla I) \approx 0$ if $\nabla I$ high
  - $g(\nabla I) \approx 1$ if $\nabla I$ low
  e.g.: $g(\nabla I) = \exp(-\nabla I^2)$

$$\frac{\partial I(x, y)}{\partial t} = \text{div}(g(\nabla I) \cdot \nabla I)$$

Original image 30 iterations 100 iterations
Extension to vectoriel data

\( \vec{\nabla} I \) has to be redefined

**Di Zenzo analysis** (Di Zenzo, 1986)

\[
\| dI \|^2 = \begin{bmatrix} dx & \frac{g_{11}}{g_{21}} \frac{g_{12}}{g_{22}} \end{bmatrix} \begin{bmatrix} dx \end{bmatrix}
\]

where \( g_{ij} = \left< \frac{\partial I}{\partial x_i}, \frac{\partial I}{\partial x_j} \right> \)

Maximum and minimum vectorial variations (eigen values of \([g_{ij}]\)):

\[
\begin{align*}
\lambda^+ &= g_{11} + g_{22} + \sqrt{(g_{11} - g_{22})^2 + 4g_{12}^2} \\
\lambda^- &= g_{11} + g_{22} - \sqrt{(g_{11} - g_{22})^2 + 4g_{12}^2}
\end{align*}
\]

\( \nabla I \) Then:

- \( |\nabla I| = \sqrt{\lambda^+} \) (Di Zenzo)
- \( \nabla I = \sqrt{(\lambda^+ - \lambda^-)} \) (Sapiro)
Hyperspectral example (Velez-Reyes, 2006)

Anisotropic diffusion

Reduction of the spectra deviation
Conclusion

☞ Taking into account the spatial relationships between hyperspectral imaging pixels can help in:
  - Reduction of classification errors
  - Unsupervised segmentation
  - Spectral noise reduction

☞ An apparent paradox
  - « Classical » hyperspectral processing requires a very accurate instrumental calibration
  - Image processing tools have been developed for years to overcome image signal inaccuracy (8 bits signal level, lighting variations, etc.)

Can spatial information reduce hyperspectral calibration requirements?

  - Unsupervised clusterisation merging smooth spectral variations
  - Chemometric modelisation adapted to these «deviating» clusters (e.g. EPO (J-M Roger, 2003) …
References