Hierarchical segmentation of multiresolution remote sensing images

Application to urban environments

Camille Kurtz[1], Nicolas Passat[1], Anne Puissant[2], Pierre Gançarski[1]

[1]LSIIT – Image Sciences, Computer Sciences and Remote Sensing Laboratory
[2]LIVE – Image, City and Environment Laboratory
University of Strasbourg

ISMM 2011 (6-8 July 2011)
10th International Symposium on Mathematical Morphology

Outline

Outline

Image and knowledge

Context

Extraction of knowledge about the Earth surface from remote sensing data

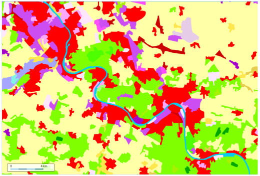
Purpose: automatic information retrieval from satellite images

To provide a set of (semi)-automatical tools enabling to extract relevant information (*e.g.*, land uses/covers, urban structures) from satellite images

Domains

- Data mining
- Image analysis

Raw image

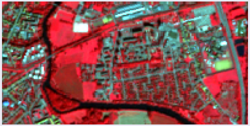


Land cover map

Multiresolution images/complex objects of interest

Data

Numerous kinds of images (multisource, multiresolution, etc.)



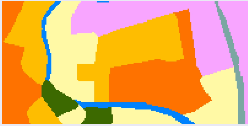
MSR (10m)

HSR (2,4m)

VHSR (60cm)

Objects of interest

Different complex objects of interest have to be extracted



Districts

Blocks

Buildings

Issues

Thematical issues

1. The (manual) extraction of objects of interest is a complex process
2. A level is not always linked to a particular spatial resolution

Computational issues

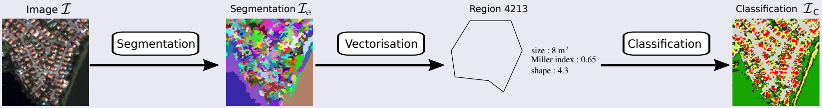
1. (V)HSR images: huge volume of data (several GB!)
2. The objects of interest (districts, blocks, simple objects) are complex and heterogeneous

Urban blocks on a VHSR image

Classical approaches and their limits

1. **Pixel-based approaches** *vs.* **Object-based approaches**
2. **Supervised approaches** *vs.* **Unsupervised approaches**

Object-based approaches



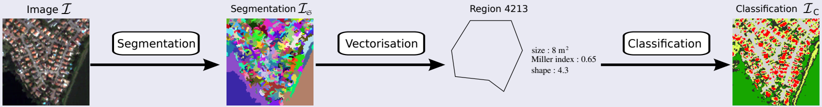
Unsupervised approaches

- To discover the data structures in order to extract relevant information
- Does not require complex *a priori* knowledge about the considered data (labelled examples, number of classes...)

Classical approaches and their limits

1. **Pixel-based approaches** *vs.* **Object-based approaches**
2. **Supervised approaches** *vs.* **Unsupervised approaches**

Object-based approaches



Unsupervised approaches

- To discover the data structures in order to extract relevant information
- Does not require complex *a priori* knowledge about the considered data (labelled examples, number of classes...)

Purpose

Main purpose

To extend the **object-based** approaches to extract complex/multilevel urban elements from satellite images

- To use the complementarity of the information available in all the resolutions (MSR, HSR, VHSR): **multiresolution analysis**
- To use an **unsupervised** approach

Advantages

- To consider all the available data
- To not required complex *a priori* information
- To propose to the user homogeneous sets of objects of interest which can be labelled by using his background knowledge

Outline

Extraction of complex patterns

Grouping approaches: to extract complex objects by grouping several basic ones: to look for patterns within the **RAG** [Barnsley and Barr, 1997]

Hierarchical segmentations: to provide a series of partitions of an image with an increasing (or decreasing) level of details

- **Top-down:** Graph partitioning [Shi and Malik, 2000]
- **Bottom-up:** Region merging [Baatz and Schape, 2000], Connected operators [Serra and Salembier, 1993]

Multiresolution approaches: to extract a specific level of information/semantic using a specific spatial resolution

- To process synthetic **degraded images** (Fourrier transforms, wavelets) [Aksoy and Akcay, 2005] or directly **multi-sensors ones**
- To process a coarser resolution than the original: the largest and complex structures of interest may appear more **homogeneous** [Goffe et al., 2010]

Extraction of complex patterns

Grouping approaches: to extract complex objects by grouping several basic ones: to look for patterns within the **RAG** [Barnsley and Barr, 1997]

Hierarchical segmentations: to provide a series of partitions of an image with an increasing (or decreasing) level of details

- **Top-down:** Graph partitioning [Shi and Malik, 2000]
- **Bottom-up:** Region merging [Baatz and Schape, 2000], Connected operators [Serra and Salembier, 1993]

Multiresolution approaches: to extract a specific level of information/semantic using a specific spatial resolution

- To process synthetic **degraded images** (Fourrier transforms, wavelets) [Aksoy and Akcay, 2005] or directly **multi-sensors ones**
- To process a coarser resolution than the original: the largest and complex structures of interest may appear more **homogeneous** [Goffe et al., 2010]

Extraction of complex patterns

Grouping approaches: to extract complex objects by grouping several basic ones: to look for patterns within the **RAG** [Barnsley and Barr, 1997]

Hierarchical segmentations: to provide a series of partitions of an image with an increasing (or decreasing) level of details

- **Top-down:** Graph partitioning [Shi and Malik, 2000]
- **Bottom-up:** Region merging [Baatz and Schape, 2000], Connected operators [Serra and Salembier, 1993]

Multiresolution approaches: to extract a specific level of information/semantic using a specific spatial resolution

- To process synthetic **degraded images** (Fourrier transforms, wavelets) [Aksoy and Akcay, 2005] or directly **multi-sensors ones**
- To process a coarser resolution than the original: the largest and complex structures of interest may appear more **homogeneous** [Goffe et al., 2010]

Purpose

Objectives of this work

1. To extend the **multiscale/hierarchical** (connected operators) approaches to deal with multiresolution remote sensing images
   - To extract hierarchies of segments
   - To use all the available data
2. To use **a top-down approach through the resolution** (to analyse the content of an image at a coarser resolution and then progressively increase this resolution)
   - Similar to the photo-interpretation process
   - To avoid the problems due to the analysis of VHSR images
3. To apply this approach to extract **complex objects** at different semantic levels (Ex: districts, blocks, simple objects)

Purpose

Objectives of this work

1. To extend the **multiscale/hierarchical** (connected operators) approaches to deal with multiresolution remote sensing images
   - To extract hierarchies of segments
   - To use all the available data
2. To use **a top-down approach through the resolution** (to analyse the content of an image at a coarser resolution and then progressively increase this resolution)
   - Similar to the photo-interpretation process
   - To avoid the problems due to the analysis of VHSR images
3. To apply this approach to extract **complex objects** at different semantic levels (Ex: districts, blocks, simple objects)

Purpose

Objectives of this work

1. To extend the **multiscale/hierarchical** (connected operators) approaches to deal with multiresolution remote sensing images
   - To extract hierarchies of segments
   - To use all the available data
2. To use **a top-down approach through the resolution** (to analyse the content of an image at a coarser resolution and then progressively increase this resolution)
   - Similar to the photo-interpretation process
   - To avoid the problems due to the analysis of VHSR images
3. To apply this approach to extract **complex objects** at different semantic levels (Ex: districts, blocks, simple objects)

Outline

Methodology: a top-down extraction approach

Input / Output

- **Input**: n images of the same scene with different spatial resolutions
- **Output**: n levels of segmentation

Principle

The extraction methodology performs ***n* successive steps** (one step per resolution) from the lowest resolution to the highest one, enabling different scales of interpretation

Each step is composed of:

1. **a monoresolution hierarchical segmentation approach**
2. **a multiresolution clustering approach**

At each resolution/step r, **the output** (a set of regions gathered into c clusters) **is embedded** into the resolution $r+1$ **to be treated as input** of the next step

Methodology: a top-down extraction approach

Input / Output

- **Input**: n images of the same scene with different spatial resolutions
- **Output**: n levels of segmentation

Principle

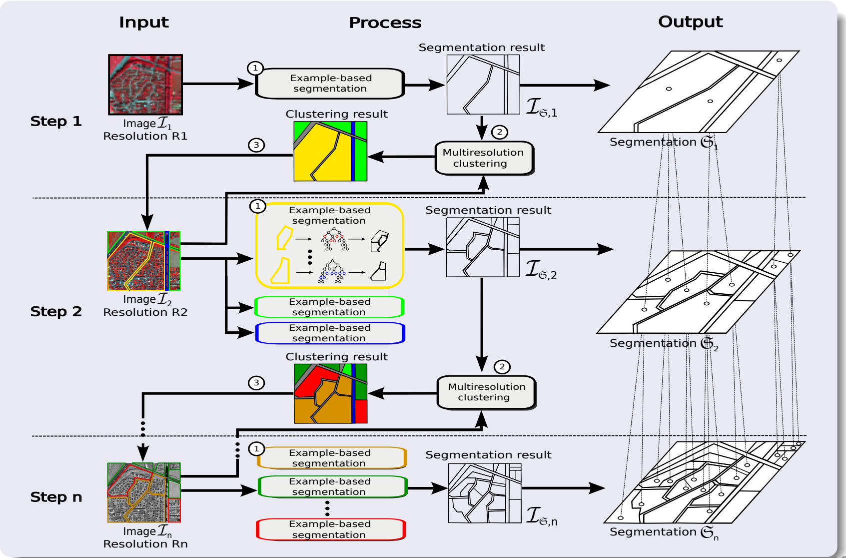
The extraction methodology performs ***n* successive steps** (one step per resolution) from the lowest resolution to the highest one, enabling different scales of interpretation

Each step is composed of:

1. **a monoresolution hierarchical segmentation approach**
2. **a multiresolution clustering approach**

At each resolution/step r, **the output** (a set of regions gathered into c clusters) **is embedded** into the resolution $r + 1$ **to be treated as input** of the next step

Methodology: a top-down extraction approach



A monoresolution hierarchical segmentation approach

Ideas

- **To adapt/divide the segmentation process** (and/or the segmentation parameters) **to local areas** of homogeneous classes of radiometric intensity
- To provide **an interactive segmentation tool** using the advantages of the BPTs

Principle

1. BPT segmentations are defined interactively by the user on different parts of the images (with different sematical contents)
2. These segmentation results are then learnt and automatically reproduced on the whole remaining data

A monoresolution hierarchical segmentation approach

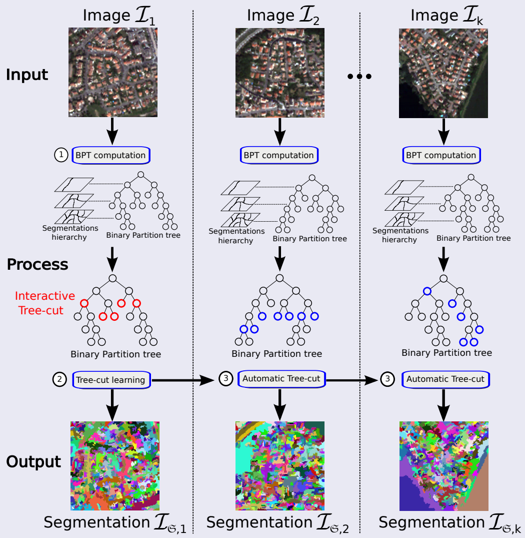
Input / Output

- **Input**: $k \geq 2$ parts of a same image representing k different (but specific) areas
- **Output**: k segmentations with a similar level of scale

Methodology

1. For one of the k images, the user first interactively performs a segmentation, by providing a cut in its BPT. This cut is assumed to correctly characterise the user-defined segmentation
2. This cut is then considered as an example to reproduce in the BPTs of the $k-1$ other images

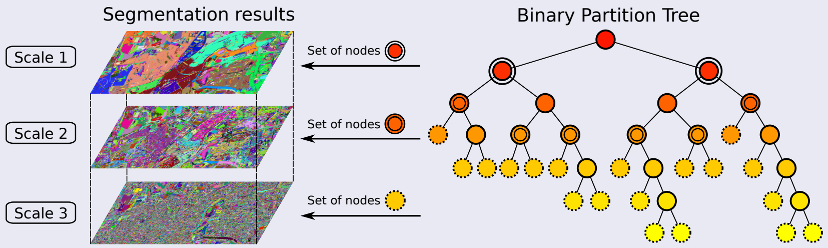
A monoresolution hierarchical segmentation approach



Binary Partition Tree

Definition

- The tree leaves correspond to the initial pixel level partition
- The remaining tree nodes represent the regions formed by the merging of two children regions
- The tree construction is performed by keeping track of merging steps of an iterative region merging algorithm



Building a BPT

The creation of BPT implies two important notions:

- **Region model** $M(R_i)$
  The region model specifies how a region is represented / modelled
- **Merging criterion** $O(R_i, R_j)$
  The similarity between neighboring regions determines the merging order

Building a BPT

The creation of BPT implies two important notions:

- **Region model $M(R_i)$**
  The region model specifies how a region is represented / modelled
- **Merging criterion $O(R_i, R_j)$**
  The similarity between neighboring regions determines the merging order

The region model $M(R_i)$

A region $R_i \in \mathcal{N}$ ($R_i \subseteq E$) is modelled by the couple

$$M_r(R_i) = \langle (v_b^-(R_i), v_b^+(R_i)) \rangle_{b=1}^{s}$$
$$M_g(R_i) = (e(R_i), a(R_i))$$

1. $v_b^\star$ refers to the extremal values in the b^{th} spectral band of $\mathcal{I}$ (*i.e.*, in $\mathcal{I}_b$)
2. e and a represent respectively the elongation and the area of the region

Building a BPT

The creation of BPT implies two important notions:

- **Region model $M(R_i)$**
  The region model specifies how a region is represented / modelled
- **Merging criterion $O(R_i, R_j)$**
  The similarity between neighboring regions determines the merging order

The merging criterion $O(R_i, R_j)$

At each step, the algorithm determines the pair of most similar connected regions minimizing the increase of the ranges of the intensity values and having low elongation/area properties

$$O_r(R_i, R_j) = \frac{1}{s} \sum_{b=1}^{s} \frac{\max\{v_b^+(R_i), v_b^+(R_j)\} - \min\{v_b^-(R_i), v_b^-(R_j)\}}{v_b^+(E) - v_b^-(E)}$$

$$O_g(R_i, R_j) = \frac{1}{2}(e(R_i \cup R_j) + a(R_i \cup R_j))$$

Building a BPT

The creation of BPT implies two important notions:

- **Region model $M(R_i)$**
  The region model specifies how a region is represented / modelled
- **Merging criterion $O(R_i, R_j)$**
  The similarity between neighboring regions determines the merging order

The merging criterion $O(R_i, R_j)$

The similarity measure between 2 regions R_i and R_j is computed as

$$O(R_i, R_j) = \alpha . O_r(R_i, R_j) + (1 - \alpha) . O_g(R_i, R_j)$$

Building a BPT

The creation of BPT implies two important notions:

- **Region model $M(R_i)$**
  The region model specifies how a region is represented / modelled
- **Merging criterion $O(R_i, R_j)$**
  The similarity between neighboring regions determines the merging order

The merging criterion $O(R_i, R_j)$

The similarity measure between 2 regions R_i and R_j is computed as

$$O(R_i, R_j) = \alpha . O_r(R_i, R_j) + (1 - \alpha) . O_g(R_i, R_j)$$

How to determine $\alpha \in [0, 1]$?

- In practice, the closer the nodes are to the root, the less relevant O_r is
- Consequently, the weight α can be defined as a function depending directly on the value of O_r (and decreasing when O_r increases)

Building a BPT

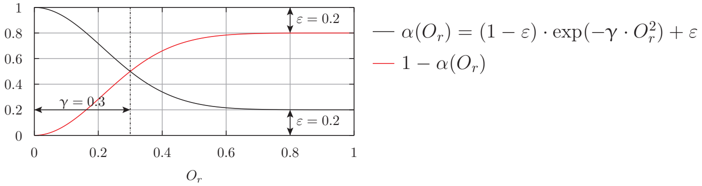
The creation of BPT implies two important notions:

- **Region model $M(R_i)$**
  The region model specifies how a region is represented / modelled
- **Merging criterion $O(R_i, R_j)$**
  The similarity between neighboring regions determines the merging order

The merging criterion $O(R_i, R_j)$

The similarity measure between 2 regions R_i and R_j is computed as

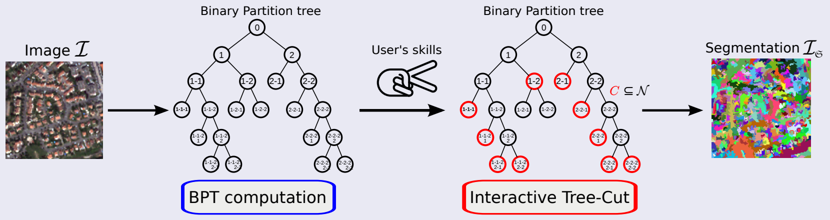
$$O(R_i, R_j) = \alpha . O_r(R_i, R_j) + (1 - \alpha) . O_g(R_i, R_j)$$



Defining a cut interactively

Interactive cut

- It enables to obtain a segmentation adapted to the user requirement
- It is possible to interactively browse the tree in order **to extract a cut** "example" C_j



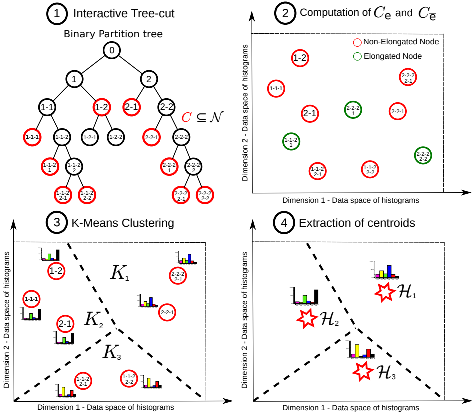
Learning of the cut

Principle

To enable the reproduction of the cut example: ➪ It is necessary **to learn this cut** by extracting the most relevant features

Learning algorithm

1. Find u coherent/homogeneous groups into the set of nodes of the cut example
2. **Extract u centroids modelled by their histograms** $\{\mathcal{H}_i\}_{i=1}^{u}$ which summarize/characterize the cut example



Reproducing the cut

Principle

To reproduce the cut example on the remaining data

Climbing algorithm

- For each tree to cut, the algorithm looks for a cut $\widehat{C}_j$, minimizing a scattering measure computed between the set of histograms of the u centroids $\{\mathcal{H}_i\}_{i=1}^{u}$ (which summarized the cut example) and the set of nodes of the current cut C_j
- The scattering measure $\zeta(C_j)$ associated to the cut C_j is defined as

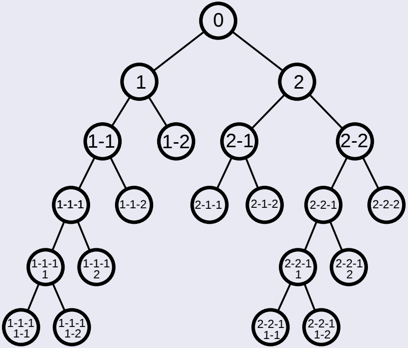
$$\zeta(C_j) = \sum_{i=1}^{u} \frac{|\bigcup_{X \in C_j^i} X|}{|\bigcup_{X \in C_j} X|} \cdot d(\overline{\mathcal{H}_{i,j}}, \mathcal{H}_i)$$

Reproducing the cut

Climbing algorithm: initialisation

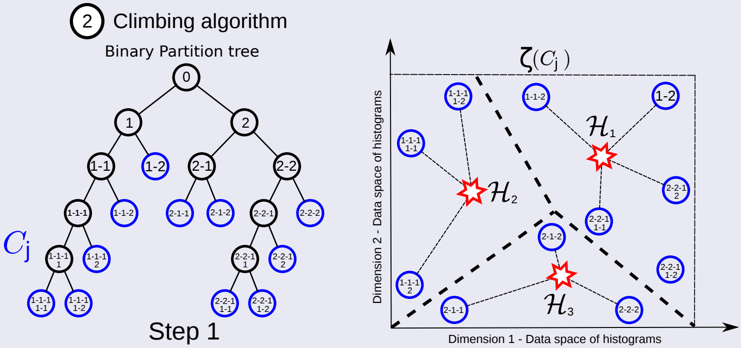
① Partition-tree to cut associated to $\mathcal{I}_j$

Binary Partition tree



Reproducing the cut

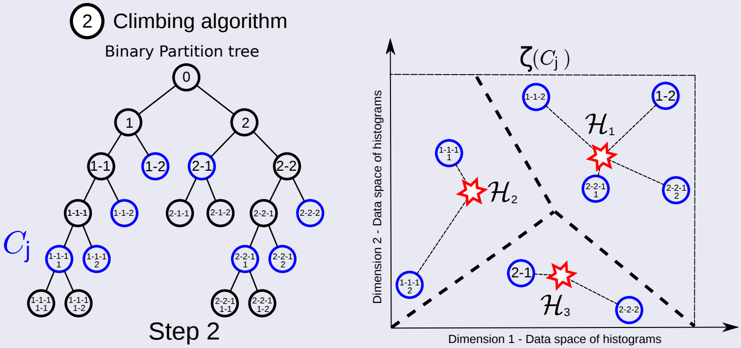
Climbing algorithm



- Does the cut C_j minimize the ζ function ?

Reproducing the cut

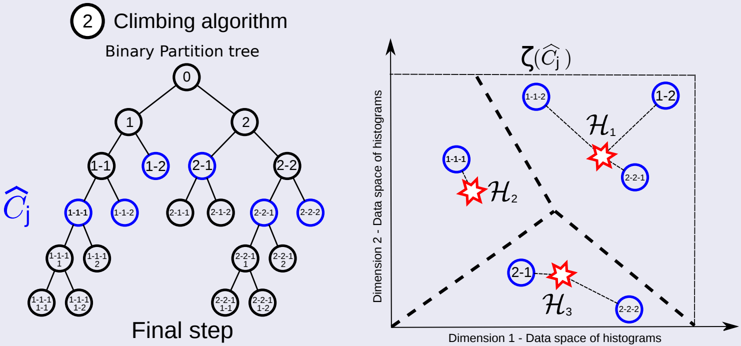
Climbing algorithm



- Does the cut C_j minimize the ζ function ?

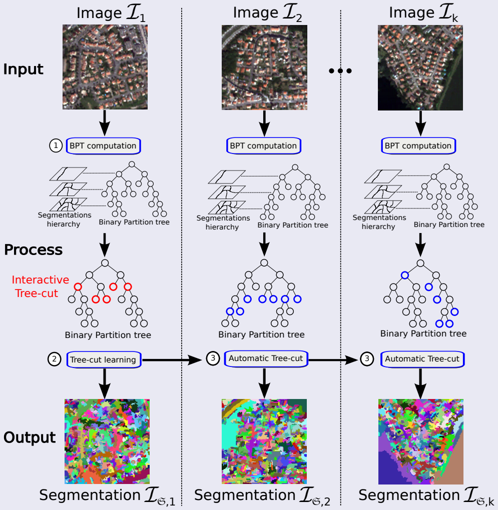
Reproducing the cut

Climbing algorithm



- End: $\widehat{C}_j$ minimizes the ζ function

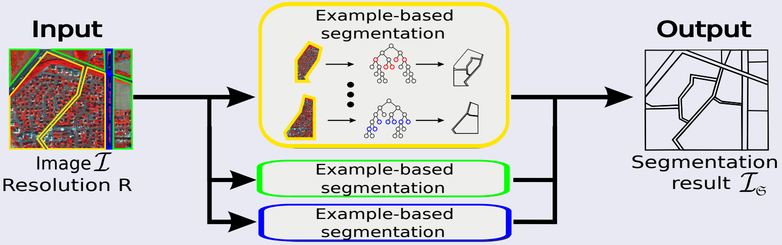
Reproducing the cut



Application

To segment a whole image

1. Find s sets of similar semantic area through the image
2. Apply the methodology for each set of similar regions
3. Gather the s sub-partitions obtained



A multiresolution clustering approach

Ideas

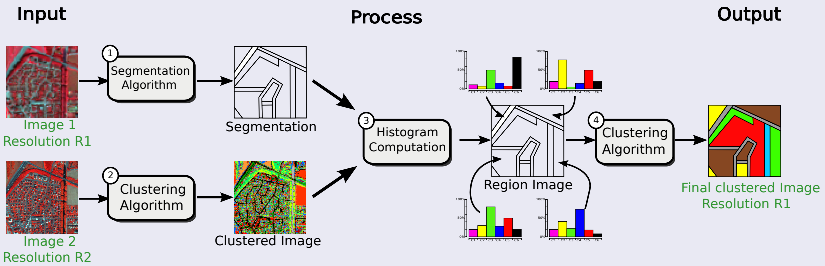
- To **fuse the information** provided by the analysis of the regions of the image at the lower resolution with the clustering result of the image at the higher resolution
- To consider the spatial context of the objects of interest and their semantic relations through the different resolutions available

Principle

- To use a clustering algorithm to gather **the segments extracted at a resolution r into c homogeneous sets (clusters)**
- In order to do that, these segments are caracterized using their composition into the resolution $r + 1$

A multiresolution clustering approach

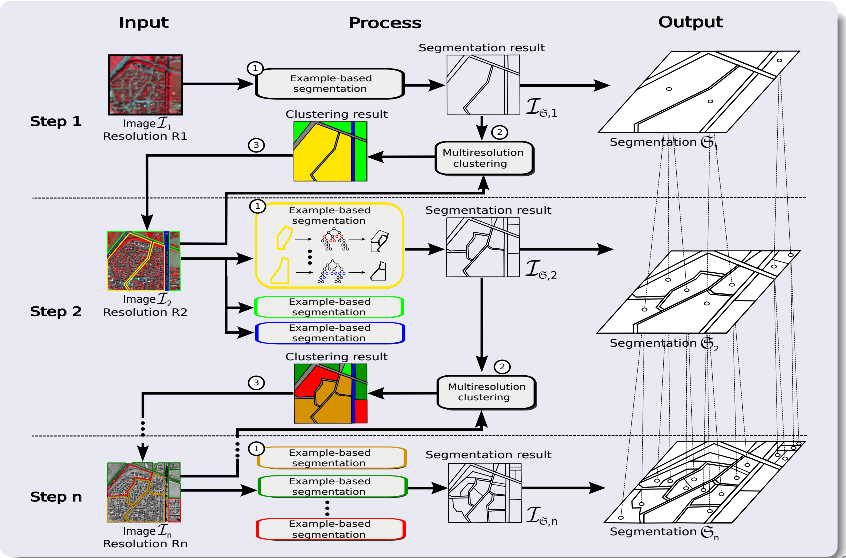
Workflow



Methodology

Multiresolution region-based clustering for urban analysis [Kurtz et al., 2010]

Global methodology



Outline

Methodology

Tools

The proposed top-down approach has been used to extract 3 levels of complex objects from urban scenes:

1. urban districts
2. urban blocks
3. urban objects

Level Scale	Urban areas 1:100 000–1:25 000	Urban blocks 1:10 000	Urban objects 1:5 000
Objects of interest	⋆ High-density fabric ⋆ Low-density fabric ⋆ Industrial areas ⋆ Forest zones ⋆ Agricultural zones ⋆ Water surfaces ⋆ Bare soil	⋆ Continuous urban blocks ⋆ Discontinuous urban blocks - Individual urban blocks - Collective urban blocks ⋆ Industrial urban blocks ⋆ Urban vegetation ⋆ Forest ⋆ Agricultural zones ⋆ Water surfaces ⋆ Roads	⋆ Building/roofs - Red tile roofs - Grey residential roofs - Light commercial roofs ⋆ Vegetation - Green vegetation - Non-photosynthetic veget. ⋆ Transportation areas - Streets - Parking lots ⋆ Water surfaces - Rivers - Natural water bodies ⋆ Bare soil ⋆ Shadows

Data

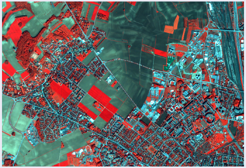
Three datasets

Each dataset is composed of:

- one MSR image, 1 pixel = 9.6 m × 9.6 m, 4 spectral bands
- one HSR image, 1 pixel = 2.4 m × 2.4 m, 4 spectral bands
- one VHSR image, 1 pixel = 60 cm × 60 cm, panchromatic

Hautepierre dataset

MSR



HSR

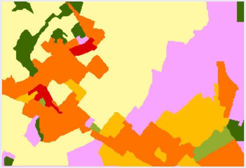
VHSR

Validation protocol

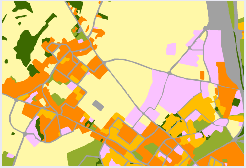
Results evaluation

- At each step, the extracted result (a classification map) has been compared to a certified ground thruth map
- Computation of the **Kappa** and **F-Measure** indexes

Ground-thruth maps of the **Hautepierre** dataset



Districts (5 cl.)

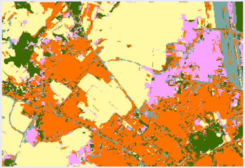


Blocks (7 cl.)

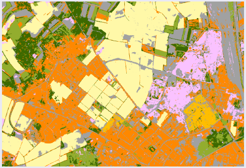
Buildings (1 cl.)

Results

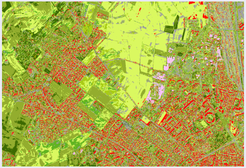
Results for the **Hautepierre** dataset



MSR result



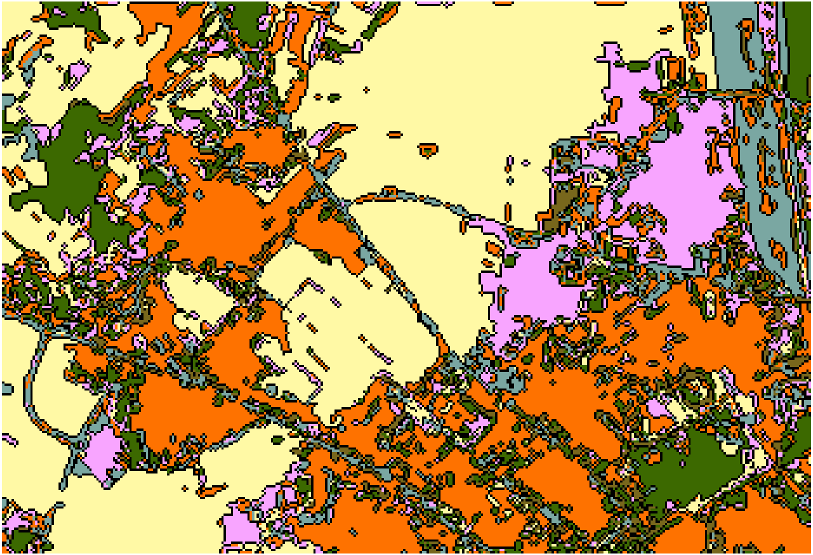
HSR result

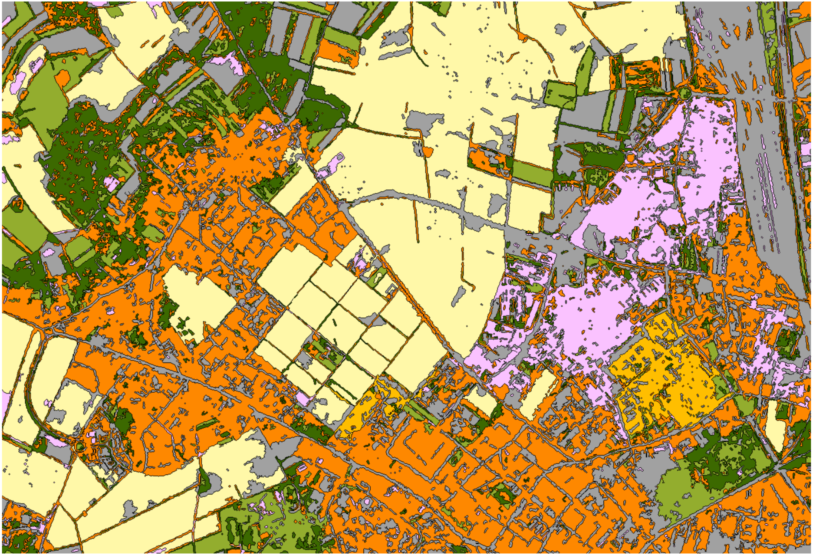


VHSR result

Percentage of pixels correctly classified

- Kappa: (MSR - 72%), (HSR - 78%), (VHSR - 76%)
- F-Measure: (MSR - 56%), (HSR - 69%), (VHSR - 72%)





Outline

Contributions

Methodological contributions

- Extension of an approach based on **connected operators** to deal with **multiresolution data**
- **Interactive segmentation** approach based on BPTs, interactively defined by the user on a part of an image, and then automatically reproduced on the remaind of the data
- **Top-down methodology**: unsupervised classification and then segmentation of the obtained clusters

Applicative contributions

- Developement of a top-down methodology to extract complex and **multilevel objects** from multiresolution satellite images
- Application to the **extraction of urban objects**

Theoritical and methodological perspectives

Theoretical perspectives

- To introduce the knowledge of the expert in the segmentation and in the clustering processes
- To enable the correction of the borders of the objects extracted at the coarser resolutions (by using an ascendant climbing approach)

Methodological perspectives

- To try another hierarchical segmentation models
- Integration of thematical knowledges
- Application to other domains (Landslides monitoring)

Thank you for your attention

Aksoy, S. and Akcay, H. G. (2005).
Multi-resolution segmentation and shape analysis for remote sensing image classification.
In International Conference on Recent Advances in Space Technologies, pages 599–604.

Baatz, M. and Schape, A. (2000).
Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation.
In Verlag, W., editor, Angewandte Geographische Informations-Verarbeitung XII, volume 58 of Karlsruhe, pages 12–23.

Barnsley, M. J. and Barr, S. L. (1997).
Distinguishing urban land-use categories in fine spatial resolution land-cover data using a graph-based, structural pattern recognition system.
Computers, Environment and Urban Systems, 21(3):209–225.

Goffe, R., Damiand, G., and Brun, L. (2010).
A causal extraction scheme in top-down pyramids for large images segmentation.
In International Workshop On Structural and Syntactic Pattern Recognition, volume 6218 of Lecture Notes in Computer Science, pages 264–274. Springer.

Serra, J. C. and Salembier, P. (1993).
Connected operators and pyramids.
In Dougherty, E. R., Gader, P. D., and Serra, J. C., editors, Image Algebra and Morphological Image Processing IV, volume 2030, pages 65–76. SPIE.

Shi, J. and Malik, J. (2000).
Normalized cuts and image segmentation.
IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8):888–905.