Improving spatial information extraction for local and regional decision makers using VHR remotely sensed data (SPIDER)

STEREO Project SR/00/02

Frank Canters, VUB
Robert De Wulf, RUG
Jean-Paul Donnay, ULg
Rudi Goossens, RUG
Eléonore Wolff, ULB

Arsenal, Namur, 9 May 2003
Research teams

- Centre for Cartography and GIS (Brussels, VUB)
  - Frank Canters
  - Tim Van de Voorde
  - William De Genst

- Laboratory of Forest Management and Spatial Information Techniques (Ghent, UG)
  - Robert De Wulf
  - Koen Mertens

- SURFACES (Liège, ULg)
  - Jean-Paul Donnay
  - Marc Binard

- Department of Geography (Ghent, UG)
  - Rudi Goossens
  - Dennis Devriendt

- IGEAT (Brussels, ULB)
  - Eléonore Wolff
  - Nathalie Stephenne
  - Marie Sintzoff
Major objectives

- Investigate how EO-technology can support local and regional decision-making, particularly in Belgium, with emphasis on urban and suburban areas
- Focus on technical and user-oriented issues
- Major objectives:
  - Define optimal methods for improved spatial information extraction from high- and very-high-resolution data
  - Identify useful EO-applications at the level of local and regional decision-making
  - Define and develop value-added products that will support these applications
Overall structure of the project

Year 1

- Geometric aspects of VHR data (TG)

Year 2

- Data management and technology watch (DM)
- Classification of VHR data (TC)
- Multi-resolution strategies (TMC, TMS)

Year 3

- Identification of user needs (UN)

Year 4

- Product and application development (UP)

Project co-ordination and diffusion of results (CD)
Study areas, test zones and confidence sites

Study area: Ghent

Location of the test zones and confidence sites on the Quickbird image (23rd August, 2002)

Only one third of the area is free of clouds and shadow
Ghent area: test zones and confidence sites

Test zone 5 (155 ha)

CS 3 (2.4 ha) Residential buildings

CS2 (1.5 ha) High density built-up area

CS1 (3 ha) Old centre
Geometrical aspects of VHR-data processing

- Objective:
  - Evaluate the geometric accuracy of satellite-derived DSMs and ortho-corrected image data in comparison with similar products obtained by large-scale aerial photography

- Topics:
  - Develop reference DSMs from large-scale aerial photography of urban and sub-urban areas
  - Perform ortho-rectification of VHR satellite data based on reference DSMs
  - Define an optimal procedure for the derivation of DSMs and ortho-photoplans from VHR satellite data
  - Study the effect of the oblique viewing angle of VHR data on image displacements caused by building height and relief
Development of reference DSMs

- **Collection of GCPs**
  - Differential GPS in real-time mode
  - 6 points/stereo-pair
  - Total number of points for Ghent study area: 51

- **Reference DSMs for:**
  - **Test zones**
    - Output resolution: 1m
    - Source: aerial photographs 1/12000, resolution 14cm
  - **Confidence sites**
    - Output resolution: 20cm
    - Source:
      - aerial photographs 1/4000, resolution 8cm
      - aerial photographs 1/12000, resolution 14cm
Development of reference DSMs

- Process of DSM generation

  - Step 1: definition of breaklines
    - Rooftop level DSM
  - Step 2: editing of height contours around buildings
    - Ground level DSM
DSM editing step 1
DSM editing step 2
Development of reference DSMs

- Problem in DSM generation:
  - Editing takes about 90% of the time that is needed for DSM creation. The amount of editing depends on:
    - Resolution of output DSM and ortho-photo
    - Height and shape of objects
    - Position of objects in the image (centre or border)
Ortho-rectification of VHR data

- VHR image + Rational Polynomial Coeff. + DSM = ortho-image

The **slant effect is corrected**: the top of the tower is centred on the longitudinal axis of the church.

The **slant effect is NOT corrected**: the top of the tower is moved with respect to the longitudinal axis of the church.
Classification of VHR data

- **Objective:**
  - Extract detailed LULC-related information from VHR-data that is useful for local and regional management and planning purposes

- **Topics:**
  - Evaluate different approaches for VHR urban land-cover classification, using a common reference data set:
    - Probabilistic and non-probabilistic methods
    - Pixel-based and region-based methods
    - Spectral, textural and contextual information
  - Develop strategies to infer land use from land-cover classification results, using rule-based techniques
## LULC classification scheme

<table>
<thead>
<tr>
<th>Land use 1 (LU1)</th>
<th>Land use 2 (LU2)</th>
<th>Land cover (LC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Buildings</td>
<td>1.1 Isolated house</td>
<td>Grey surface</td>
</tr>
<tr>
<td></td>
<td>1.2 Block of houses</td>
<td>Orange/red surface</td>
</tr>
<tr>
<td></td>
<td>1.3 Low building</td>
<td>Green copper</td>
</tr>
<tr>
<td></td>
<td>1.4 High building</td>
<td>Glass or plastic</td>
</tr>
<tr>
<td></td>
<td>1.5 Other</td>
<td>Bare soil</td>
</tr>
<tr>
<td>2. Road and rail network</td>
<td>2.1 Road</td>
<td>Water</td>
</tr>
<tr>
<td></td>
<td>2.2 Parking</td>
<td>Grass</td>
</tr>
<tr>
<td></td>
<td>2.3 Railway</td>
<td>Crops</td>
</tr>
<tr>
<td></td>
<td>2.4 Square</td>
<td>Shrub and trees</td>
</tr>
<tr>
<td>3. Hydrology</td>
<td>3.1 Water body</td>
<td>Mixed</td>
</tr>
<tr>
<td></td>
<td>3.2 Watercourse</td>
<td></td>
</tr>
<tr>
<td>4. Vegetation</td>
<td>4.1 Urban green area</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.2 Agriculture</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.3 Forest</td>
<td></td>
</tr>
<tr>
<td>5. Miscellaneous</td>
<td>5.1 Sport or recreative area</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.2 Graveyard</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.3 Construction site</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.4 Other</td>
<td></td>
</tr>
</tbody>
</table>
Collection of training and validation data

Choice of « training polygons » on aerial photographs

Selection of training pixels on Quickbird image
Visual interpretation of confidence sites

Aerial photograph

Quickbird image

Land-use 2

Land-cover
Land-cover classification approach

- Comparison of results obtained with:
  - Different classifiers
    - Maximum-likelihood classification (ML)
    - Neural network classification (NN)
    - Region-based classification (E-cognition)
  - Different classification variables
    - Spectral variables (R, G, B, IR, PAN, NDVI)
    - Spectral + textural variables
      - Measures derived from Haralick co-occurrence matrices, calculated for different window sizes (NN)
      - Segment-based texture measures (E-cognition)
  - Different training approaches
    - 4 typical pixels per polygon
    - 4 typical and 2 atypical pixels per polygon
    - 6 blocks (3x3 pixels) per polygon
Pixel-based classification

True-color composite

Neural network classification
Best scenario: Kappa = 0.83
Region-based classification

Scale parameter = 4.7
Kappa = 0.80

Scale parameter = 15
Kappa = 0.74
Classification results

- **Different classifiers**
  - Differences in overall classification performance for maximum-likelihood, neural network and per-region classification are very small (best Kappa’s around 0.80)

- **Classification variables**
  - Adding the PAN-band to the four spectral bands substantially increases classification performance
  - Adding window-based texture measures in per-pixel classification slightly increases the performance of the classifier (from 0.79 to 0.83 for the best approach)

- **Different training approaches**
  - Including atypical pixels in the training phase improves the overall accuracy of the classification with a few percent for some classification scenarios
  - The use of 3x3 training blocks does not improve the accuracy of per-pixel classification
From land cover to land use

- Two-step approach:
  - Classification of land cover, followed by:
    - Post-classification filtering (per-pixel approach)
    - Aggregation of image segments belonging to the same land-cover class (segmentation-based approach)
  - Inference of land use from land cover:
    - Rule-based classification or grouping of land-cover regions
    - Using:
      - Region-based metrics (area, shape,...)
      - Properties of neighbouring regions
      - Ancillary data, e.g. DSMs, vector maps
Post-classification based on DSM

NN-classification

Intersection with DSM

Postclassification (rule-based)

Identification of buildings
Assessment of GI needs in Belgium

- Objective: assess the GI needs of Belgian local and regional authorities in order to define useful products or applications of HR/VHR data, in an urban or suburban context
- Survey of a carefully selected group of users
  - Approach:
    - Detailed written survey (82 closed questions), followed by in-depth interview to gain more insight into:
      - Use and treatment of geographical data
      - Products/applications based on these data
      - Specific land-use/land-cover information needs
      - Use of satellite data
  - Targets:
    - 20 to 30 key representatives of various local and regional authorities in Flanders, Brussels and the Walloon region
Multi-resolution approaches

- Objective: combine VHR-data with HR-data for cost-effective production of detailed information on land-use/land-cover for extended areas

- Focus on two distinct, yet closely related issues:
  - Sub-pixel classification:
    - Estimation of sub-pixel class proportions for HR-pixels (ETM+), using VHR-data as a source for calibration
  - Sub-pixel mapping:
    - Use of sub-pixel class proportions to predict the spatial distribution of classes at smaller pixel sizes
Methodology

High resolution images

Sub-pixel classification (soft)

Fraction images (high resolution)

Sub-pixel mapping

Hard classification (high to very high resolution)
Sub-pixel mapping

<table>
<thead>
<tr>
<th>Land cover 1</th>
<th>Land cover 2</th>
<th>Land cover 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>32% 100% 40%</td>
<td>0% 0% 0%</td>
<td>68% 0% 60%</td>
</tr>
<tr>
<td>12% 40% 16%</td>
<td>16% 60% 12%</td>
<td>62% 0% 72%</td>
</tr>
<tr>
<td>0% 0% 0%</td>
<td>40% 100% 32%</td>
<td>60% 0% 68%</td>
</tr>
</tbody>
</table>

Soft classification (high resolution)

Hard classification (high to very high resolution)
Sub-pixel mapping: approaches

The different techniques:

- **Simplex**: solving a set of linear equations
- Use of **Genetic Algorithms** to optimize configuration
- **Neural Networks**: learning spatial configuration
Example: Sub-pixel mapping of degraded VHR-classification on Gent

- Scale = 2
  - Hard classification: Kappa = 0.900
  - Simplex inv. sq. dist. : Kappa = 0.967

- Scale = 4
  - Hard classification: Kappa = 0.813
  - Simplex inv. sq. dist. : Kappa = 0.887

- Scale = 10
  - Hard classification: Kappa = 0.649
  - Simplex inv. sq. dist. : Kappa = 0.707
Hard classification, scale 4, resolution 2.44 m
Sub-pixel mapping, scale 4, resolution 0.61 m
Final goal

VHR-classification (QuickBird)

Intermediate resolution classification

Evaluation

Degradation

Degraded classifications used as reference

Sub-pixel mapping

HR sub-pixel classification (Landsat)

- Note: Introduction of extra error due to:
  - Errors in the sub-pixel classification
  - Co-registration errors