From Slums detection to slum definition ...

Urban remote sensing: recent technological and methodological developments

Christiane Weber
CNRS – Strasbourg – France
Christiane.weber@lorraine.u-strasbg.fr

Laboratoire Image et Ville UMR 7011
Presentation

Context

Satellite imagery
  Detection & Extraction
  Identification & Analyse

Slums detection towards slums definition
  Direct link?
  Relevant spatial characteristics?
Venezuela

Kurla, Mahārāshtra (India)

Argentina

Dharavi slums, Mumbai, India 2007.
Photograph: David Levene.
Various situations

Informal settlements on slopes
Chaotic structures without networks
Elevation and regular forms
Diversity of materials
Size of elements and surrounding buildings

→ Variety of situations, difficulty to generalize
Context: Urban specificities

Urban management: Strategic, Tactical, Operationnal

Decision makers
- Planners
- Technicians

Urban management = - need of various types of information
- at various scales

<table>
<thead>
<tr>
<th>Niveau</th>
<th>Global level</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Niveau I</td>
<td>Global level</td>
<td>&gt; 1/25 000e</td>
</tr>
<tr>
<td>Niveau II</td>
<td>Urban Planning</td>
<td>du 1/5 000e au 1/10 000e</td>
</tr>
<tr>
<td>Niveau III</td>
<td>Urban mapping</td>
<td>du 1/1000e au 1/2000e</td>
</tr>
<tr>
<td>Niveau IV</td>
<td>Technical applications</td>
<td>du 1/200e au 1/500e</td>
</tr>
</tbody>
</table>
Spatial, spectral (mineral/vegetation)  
Texture, structure (homogeneity/heterogeneity)  
Material (tiles, concrete, vegetation etc.)  
Geometry (size, shape, orientation)  
Radiometry  

Parameters  

Scales  

Urban area  

Urban structure = Sum of geographic elements  

Urban Objects = (street, building, garden …)  

Pixel  
1 element  
Or  
A combination of several Elements  

➜ Information extraction at various levels (urban area, urban structure, urban object) with HSR images?
Context: Detection

What can we do? Detection & Extraction

1) Very high spatial resolution? (or multi-resolution solution)
2) RS usual image processing concepts or morphological concepts or knowledge based concepts?

Identification & Analyse

3) GIS integration?
4) Methodological efficiency?
Satellite imagery
Increasing availability of satellite images

The use of several sources of data is often required to identify urban elements:
- Landsat (30m)
- SPOT (20m)
- ASTER (15m)
- Quickbird (2.8m)

Urban purposes ➔ The use of several sources of data is often required to identify urban elements
New paradigm?

• From pixel to object
  changing classical rules
  defining optimal spatial resolution

• From object to region
  adding more (different) knowledge:
  ancillary data, logical or spatial rules …
Detection: new paradigms

Modification of semantic signification of spectral values

Inappropriate usual classification methods

The use of spectral information is no more sufficient

HR ➔ VHR
pixel ➔ objet
<table>
<thead>
<tr>
<th>Spatial Resolution</th>
<th>Applications</th>
</tr>
</thead>
</table>
| VHR : 0.5 -1.5 mètre | Identification, cartography of objects (cars, trees, urban materials…)  
Classification of vegetal species and strata  
Detection of small grassland areas |
| HRS : 1.5 - 5 mètres | Distinction of buildings  
Identification, cartography of objects (constructions)  
Classification of strata and shrubby areas  
Detection small areas, plants diseases, small agricultural areas |
| HRS : 5 - 10 mètres | Location/cartography of buildings, roads, agricultural lands, streets  
Classification of vegetation strata  
Distinction of vegetation species, plants disease  
Classification of land parcels |
| HRS : 10 -20 mètres | Location et geometry of large infrastructures (airports, city centres, suburbs, commercial malls, industrial areas)  
Global classification |
Detection: characteristics

Image heterogeneity increase

SPOT XS 1998 (20 m)

Quickbird MS 2002 (2.8 m)


© CNES (1998)

Detection: characteristics

- Same object type composed with various material,
  versus same material for various objects

Ex: Buildings with different roofs material
Detection & Extraction approaches

1. Integrate spatial or structural information
   - Optimal spatial resolution
   - Local contextual information, relationships
   - Extraction and integration of knowledge

2. Improve the results of traditional extraction methods
   - Combining spectral data with measures of texture or mathematical morphology.
   - Developing new algorithms
   - Combining ancillary data and expert knowledge

3. Focus on the development of knowledge extraction
   - Formalization of the knowledge
   - Learning tools
   - Urban ontology, data mining approaches
1. Definition of an Optimum Spatial Resolution?

**Minimal** resolution / **Functional** resolution

Protocole: identification of the objects

Pixel = 0.8 m
1. Definition of an Optimum Spatial Resolution?

Protocole:
- identification of the objects
- variance analysis at different spatial resolutions
1. Definition of a **Optimum Spatial Résolution**?

**Minimal resolution / Functional resolution**

**Results:**

- **areas type « grassland »**
  \[ < = 0.8 - 5 \text{ m} \quad 0.8 - 1\text{m} \]

- **objets type « building » (rectangular)**
  \[ = 2 - 3 \text{ m} \quad 0.8 - 3\text{m} \]

- **objets type « road » (linear)**
  \[ = 0.8 - 1 \text{ m} \quad 1 - 2\text{m} \]

- **objets type « residential building » (square)**
  \[ = 0.8 - 5 \text{ m} \quad 0.8 - 3\text{m} \]
Integrate spatial or structural information

- Urban core 20-30 m
- Industrial area 50-60 m
- Residential area 30-40 m
- New residential area 30-40 m

Da Nang Vietnam
(D. Bin Tran Thi, 2007)

(Images: Google Earth)
Integrate spatial or structural information

Fractal dimension «D» to identify build areas

D Dimension
0 = heterogeneity
2 = homogeneity
Integrate spatial or structural information

Fractal behavior of Da Nang 1990 et 2001

- Two processes: densification and urban sprawl

- Physical obstacles
<table>
<thead>
<tr>
<th>LULC categories Level 4</th>
<th>Urban categories</th>
<th>D - Surface correlation measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1990</td>
</tr>
<tr>
<td>Urban core</td>
<td>Very dense</td>
<td>1,73 – 1,89</td>
</tr>
<tr>
<td>Less densely urbanized</td>
<td>dense</td>
<td>0,96 – 1,73</td>
</tr>
<tr>
<td>Residential areas</td>
<td>less dense</td>
<td>0 - 1,63</td>
</tr>
<tr>
<td>Surbubs</td>
<td></td>
<td>&lt; 1,0</td>
</tr>
</tbody>
</table>

Fractal behavior of Da Nang 1990 et 2001
1. Integrate spatial or structural information
Optimal spatial resolution
Local contextual information, relationships
Extraction and integration of knowledge

2. Improve the results of traditional extraction methods
Combining spectral data with measures of texture or mathematical morphology
Developing new algorithms probabilities, soft classifiers
Combining ancillary data and expert knowledge

Image processing improvements
Combining spectral data with measures of texture or mathematical morphology or ancillary data

Development of new algorithms using

(i) a priori probabilities or a posteriori processing known as “soft classifier”, based on Bayesian probabilities or fuzzy set theorie or believes theorie

(ii) combining ancillary data and expert knowledge (AI): machine learning, ontology, data mining
Combining ancillary data and expert knowledge

1) To integrate spatial or structural information
   Local contextual information
   (Geometry and spatial relationships)

Urban Objects

Characteristics

Segmentation "regions"
Combining ancillary data and expert knowledge

3 relevant criteria:
(a) Spectral values
(b) Dimensions and shape
(c) Spatial relationships
« If a region has spectral characteristics associated to a roof with rectangular shape and is linked with a second region having the same characteristics, then regions 1 and 2 are aggregated into one, labelled as « building ».
Combining ancillary data and expert knowledge.

Urban Objects

Characteristics

1

Relevant criteria

Rules definition

Integration and classification

Object analysis
Combining ancillary data and expert knowledge.

Urban Objects

1. Rules for object definition

Characteristics

2. Functionnal or biophysic characteristics (1/10 000°)

7 aggregation criteria
- Buildings
- Open space
- Linearity
- Parallelism

Rules for object recognition
Combining ancillary data and expert knowledge

1. Urban Objects
   Characteristics

2. Functionnal or biophysic characteristics (1/10 000e)

Recognition rules definition
= Multicriteria attributs

Ilot n°3

- critère n°1: 70% de surfaces artificialisées
- critère n°2: bâtiments de type (dense - tuile)
- critère n°3: pas d'orientations définies
- critère n°4: distance entre bâtiments nulle
- critère n°5: alignement non-identifiable
- critère n°6: 14% espace vide - forme carré IH = 0.63
- critère n°7: 6% arbre + 4% pelouse + 20% route

⇒ Tissu urbain moyennement dense
Combining ancillary data and expert knowledge:

1. Satellite Image → Regions
2. Regions → Characteristics
   - Extraction of knowledge: “More than the numerical values.”

Functionnal or biophysical characteristics (1/10 000°)
Extraction and integration of Knowledge
1. Integrate spatial or structural information
   - Optimal spatial resolution
   - Local contextual information, relationships
   - Extraction and integration of knowledge

2. Improve the results of traditional extraction methods
   - Combining spectral data with measures of texture or mathematical morphology.
   - Developing new algorithms
   - Combining ancillary data and expert knowledge

3. Focus on the development of knowledge extraction
   - Formalization of the knowledge
   - Urban ontology, data mining approaches ...
   - Learning tools

(Fodomust project: ANR resp. P Gançarski)
What kind of knowledge is useful to identify the urban objects in the images?
- **Domain:** End-users needs / typology
- **Expert attributes:** color, shape, size, texture, spatial relationships ...

**How to acquire** this knowledge?

→ Learning methods
Focus on the development of knowledge extraction
« concept definition

Classifiers:
Quantitative approaches or artificial neural network, collaborative or genetic classifiers, AI approaches

Questions?
Number of classes
Results assessment
Choice of « attributs »

« Concept » definition:
set of attributs relevant for classification
Focus on the development of knowledge extraction « concept definition

Geographic Object

Image Object

Simple
- one element

Composite
- some elements

Complex Object
- spatial organisation of image objects, simple or composite

Tree

Group of trees

Park
Focus on the development of knowledge extraction - Spatial Ontology

How to describe clearly the objects of the domain in an understandable language?

Identify objects to be extracted => 'concepts inventory'
Definition of these concepts => attributes

Design of a 'spatial ontology'
Dictionary of geographic objects
Translation into an understandable language (Protege2000)
Focus on the development of knowledge extraction - Spatial Ontology

Objects Dictionary

**Fiche 1 : Pavilion**

**A. Identification de l'objet**

<table>
<thead>
<tr>
<th>Type</th>
<th>Nom de l'objet</th>
<th>Type de objet élémentaire</th>
<th>Type de objet image</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poligone</td>
<td>Pavilion</td>
<td>Bâtiment</td>
<td>Objet image simple</td>
<td>THR1</td>
</tr>
</tbody>
</table>

**B. Description de l'objet dans le monde réel**

**B.1 Définition textuelle**

L'objet « pavillon » ou « maison individuelle » appartient à la catégorie d'objets élémentaires « bâtiment ». Il désigne une construction durable destinée à abriter l'activité humaine liée à l'habitat.

La portée de cette définition est restreinte par les séries suivantes. En général, un pavillon :
- est situé dans un lot physique (domaine privé);
- a une emprise au sol d'au moins 12 m².

Ces séries visent à exclure notamment les ateliers, les cordonniers, les maisons d'outre, cabanes de jardin etc. de la définition de l'objet « pavillon ».

Le pavillon ou maison individuelle est le plus souvent organisé en intercom aux côtés d’une autre (cf. objet construct). 

**B.2 Illustration graphique : THR**

<table>
<thead>
<tr>
<th>TIR1 = 0.85 mS (2.6 m)</th>
</tr>
</thead>
</table>

![Image of pavilion types](image)

Empreinte sur satellite sur une image satellite

**C. Description de l'objet dans l'image**

**C.1 Nature de l'objet**

Objet physique - objet image simple identifiable à THR1

**C.2. Définition textuelle**

L'objet pavillon ou « maison individuelle » est représenté graphiquement par un polygone dont la surface correspond à l'emprise au sol du bâtiment.

**C.3. Principales relations**

<table>
<thead>
<tr>
<th>Adjacence</th>
<th>Objet de type « végétation »</th>
<th>Objet de type « autre route »</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alignement</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Distance entre l'objet</td>
<td>Table de pertinence à une distance fixe</td>
<td>Mouvement reposée</td>
</tr>
<tr>
<td>Densité</td>
<td>10°000 polyvalent à 100°000</td>
<td>10°000 polyvalent à 100°000</td>
</tr>
</tbody>
</table>

**C.4. Attributs**

<table>
<thead>
<tr>
<th>Signification spectrale</th>
<th>Bleu 1° [20.0–250] avec histogramme de 0 à 200</th>
<th>Bleu 1° [10.0–200] avec histogramme de 0 à 200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Longueur ou diamètre</td>
<td>12 à 204</td>
<td>12 à 204</td>
</tr>
<tr>
<td>Largeur (m)</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Profondeur</td>
<td>38 à 282</td>
<td>38 à 282</td>
</tr>
<tr>
<td>Surface (m²)</td>
<td>55 à 2447</td>
<td>55 à 2447</td>
</tr>
<tr>
<td>Infos de Méri</td>
<td>0.05 ± 0.72</td>
<td>0.05 ± 0.72</td>
</tr>
<tr>
<td>Surface Poly-Convex (Sc)</td>
<td>93 ± 205</td>
<td>93 ± 205</td>
</tr>
<tr>
<td>Surface (Sc)</td>
<td>0.95 ± 1</td>
<td>0.95 ± 1</td>
</tr>
</tbody>
</table>

| Infos de Morin           | 0.6 à 0.93                                   | 0.6 à 0.93                                      |
| Téneur (séparation)     | Téneur presser                                | Téneur presser                                  |

---

*Note: This document appears to contain a mix of French and English text, possibly discussing the identification and description of pavilions in a spatial context, with diagrams and tables.*
Focus on the development of knowledge extraction - Spatial Ontology

Spatial Ontology

=> Written in an « open-source » software

« open-source » software (Protege 2000)
Focus on the development of knowledge extraction - Learning to

Supervised Machine Learning

- Learning *from examples* given by the expert:
  description and classes

\[ \{x_i, y_i\} \]
\[ y = f(x) \]

Data

Selection and classification of a sample by an expert

Learning Algorithm

Classification function
Focus on the development of knowledge extraction - Learning to

Supervised Machine Learning

- Learning *from examples given by the expert*: description and classes built the rules from these examples to explain the classification from the description

- **Apply the rules** (C4.5 algorithm [Quinlan 93] symbolic algorithm providing a decision tree > shortest optimal description for classification)

![Diagram of supervised learning process](image)
Focus on the development of knowledge extraction - Learning to

Steps of the **rules acquisition** process

1. **Segmentation**
   - **Image (bands)**

2. **Definition of examples**
   - **Regions**

3. **Learning procedure**
   - **Training examples**

4. **Classification**
   - **Classified regions**

Focus on the development of knowledge extraction
1 Segmentation

A region growing approach (eCognition software)

Use of ancillary data to constrain the segmentation procedure and obtain homogeneous regions

Vector data from Topographic database: BDTopo (IGN) – metric precision
Experiments (1)

1. Segmentation

- Image
- Vector data
- Segmentation without ancillary data
- Segmentation with ancillary data
2. Definition of Training examples

Retained classes:

Level 1

- Description of examples with spectral features

Level 2

Level 3

- Description of examples with spatial and contextual features
Definition of Training examples

- Spectral features:
  - 4 bands (R,G,B,NIR): mean by region
  - 2 index (NDVI, SBI): mean by region

- Spatial and contextual features:
  - perimeter
  - area
  - diameter
  - compacity (Miller’s index)
  - solidity (convexity)
  - % of vegetation around buildings (20m)
2 Definition of Training examples zones

*Continuous*

*Segmented image*

*Residential building*

*Social high-rise building*
Learning procedure

- Several steps:

Level 1
- Water
- Vegetation
- Bare Soil
- Mineral
- Shadow

Level 2
- Grass
- Tree
- Building
- Road

Level 3
- Residential
- High rise buildings
- Continuous
- Activity
Learning procedure

- Learned Rules:

**SPECTRAL RULES:** Entire range of values [0..255]

Class Hierarchy - Level 1:

Rule 1: IF NDVI < 38.23 and IBS > 14.67 THEN Class = Water
ELSE Class = Non Water

Rule 2: IF NDVI > 169.14 THEN Class = Vegetation
ELSE Class = Non Vegetation

Rule 3: IF GREEN < 15.65 THEN Class = Shadow
ELSE Class = Non Shadow

Rule 4: IF NIR > 59.25 and BLUE < 57.86 THEN Class = Bare Soil
IF RED > 101.24 THEN Class = Bare Soil
ELSE Class = Mineral

Class Hierarchy - Level 2:

Rule 5: IF 60.2 < BLUE < 130.8 THEN Class = Road
ELSE Class = Building

Rule 6: IF GREEN > 30.4 THEN Class = Grass
ELSE Class = Tree
Learning procedure

Learned Rules:

**SPATIAL RULES:** Entire range of values [0...255]

Class Hierarchy - Level 3:

Rule 7: IF AREA > 5203 m² and IM > 0.3 THEN Class = Building of A.
      IF AREA < 436.8  THEN Class = Residential B.
      IF AREA < 1254.9 THEN Class = Collective B.
      IF PV > 11.9 and AREA < 1803.2 THEN Class = Collective B.
      IF IS < 0.43 THEN Class = Collective Building
      ELSE Class = Continuous Built-up Area
Classification

Experiments (}
Experiments

How to integrate? Exemple QB MS (eCognition)

Global Accuracy: 79.7%
Kappa: 75.5%

Global Accuracy: 80.2%
Kappa: 79%

For Buildings:
Results Analyses:

- Enhancement of the classification results (> 5%)
- Rely on the segmentation quality
- Influence of identification order of the objects
- Influence of the integration approach of the knowledge rules
Slums detection towards slums definition
Detection versus SLUM definition

- **Detection**
  - Scale
    - Environment
    - Structure
    - Object

- **Definition**
  - Criteria
    - Poverty
    - Water accessibility
    - ...
  - Measurements
    - Surveys
    - ...

Direct link?

Relevant spatial characteristics
Detection versus *SLUM* definition

**Detection & Extraction**

**Urban element**
- Urban element:
  - “Pencil house” (Vietnam)
  - “Bloc” (Mexico)
  - ...

**Environment characteristics**
- vegetation?, water, slope?
- Network?
- Risk?
- ...

**Analyse & Identification**

**Urban production mode**
- “Selfmade” house
- Community action
- Urgency
- ...

**Social characteristics**
- Poverty
- Water Access
- Unemployment
- ...

Detection versus *SLUM definition*

Detection & Extraction

**Scale:** HR or/and VHR → availability, cost, date

**Methodology:**
- **structure** → heterogeneous, dense, without network specific geometry or morphology dimensions
- **spectral** → Material : adapted spectral library?
- **spatial** → ratio between resolution/objectives/means

**Generalisation:**
- **Ontology** → to be defined (ground survey and comparison); *dictionary* adapted to the country or continent
- **concept definition** → experiments with ancillary data (spatial relationships, contextual features…)
- Environmental characteristics (slopes, watershed, derelict areas…)
- **rules** → to be defined
Detection versus SLUM rules definition

Relevant criteria:
Spectral & geometry
Spectral values (material)
Structure characteristics (morphological or fractal dimension)
Dimensions and shape (urban fabric)

Spatial relationships:
Open space
Linearity or orientation
Distance between the buildings
Parallelism

Environment Rules:
« If the element is located on a slope (n%) or in a watershed or floodable area or over the water pressure capacities than the location might be potential for a slum location”

Object Rules:
« If the element is characterised by these Spectral footprint types (x1, xn) and if the density is > 75% and the element belongs to this Urban fabric type than it can be characterized As potential slum habitat »
From slums detection to slum definition

Multidisciplinary issues: task groups
Step by step procedure
   Need ground truth investigations
   Need to know the urban model production
   Need to test the design of a specific dictionary
Thank you ...
Focus on the development of knowledge extraction - Learning to

Supervised Machine Learning

- Learning *from examples*

Use of the C4.5 algorithm [Quinlan 93]:
symbolic algorithm providing a decision tree > shortest optimal
description for classification using the concept of information
Entropy
Dynamics of the urban slum problem

Population
- Rural
- Urban slums
- Urban other

Solutions
- A Regional development
- B Slum upgrading
- C Urban development

Source: UN-HABITAT
Integrate spatial or structural information

Spectral Fusion

Complementarity of high spatial and spectral information