Urban ontology for semantic interpretation of multi-source images

Puissant* A., Durand** N., Sheeren*** D., Weber* C., Gançarski** P.

*Image et Ville Laboratory, UMR 7011 CNRS, Louis Pasteur University, rue de l'Argonne
F-67000 Strasbourg (France) {anne.puissant@lorraine.u-strasbg.fr}
**LSIIT-AFD, UMR 7005 CNRS, Louis Pasteur University, Bd Sébastien Brant,
F-67000 Strasbourg (France) {Durand.gancarski@lsiit.u-strasbg.fr}
*** Dynafor INRA-INPT/ENSAT UMR 1201 CNRS, Av. de l'Agrobiopôle,
F-31326 Castanet-Tolosan (France) {David.Sheeren@ensat.fr}

Abstract. The multiplication of High or Very High Resolution (spatial and spectral) remotely sensed images is an opportunity to characterize and identify urban objects. Image analyses methods using object-oriented approaches, based on the use of domain knowledge, are necessary to classify these data. A major issue in these approaches is domain knowledge formalization and exploitation. In this paper, we present a methodology to build an urban ontology adapted to the multi-level interpretation of multi-source images. Domain knowledge is stored independently in the urban ontology which contains a set of pre-defined terms characterizing urban domain concepts. The ontology is then used in a classification method in order to assign segmented regions into semantic objects. A matching process between the regions and the concepts of the ontology is proposed. The method is tested on Very High Resolution images (0.7m) on the urban area of Strasbourg (France).

Keywords: urban object, ontology, Very High resolution remotely sensed images, semantic interpretation

1. Introduction

Urban planners are interested on up-to-date landcover/use information on urban objects at several spatial (1/100,000 to 1/10,000) and temporal scales. Acquiring automatically this information is complex, difficult and time-consuming if traditional data sources (ground survey techniques) are being used. The multiplication of High or Very High Resolution (spatial and spectral) remotely sensed images is an opportunity to characterize and identify these objects in urban and peri-urban areas. Images can be exploited to provide this spatial information, which can also be easily integrated in urban GIS platforms.

Image interpretation is a difficult task which can be defined as the automatic extraction of the image semantic. It consists in obtaining useful spatial and thematic information on the objects by using human knowledge and experience [1]. In this domain, differences between the 'visual' interpretation of the spectral information and the semantic interpretation of the pixels are observed because of differences in the levels of abstraction. The semantic is not always explicitly contained in the image and depends on domain knowledge and on context. This problem is called the *'semantic* *gap'* and is defined as the lack of concordance between low-level information (automatically extracted from the images) and high-level information (analysed by urban experts) [2].

In order to reduce the semantic gap, image analyses methods using object-oriented approaches, based on the use of domain knowledge, are being developed [3,4]. These methods involve the segmentation of the images into homogeneous regions and the characterization of objects with a set of spectral (signature, index), spatial (shape) and topological (adjacency, inclusion) features. These features can be called upon in the classification process. Only few works has focused on the use of domain knowledge for classifying urban objects [5], and a major issue in these approaches is therefore domain knowledge formalization and exploitation.

Building a knowledge base or an urban ontology is a difficult task because, most of the time, the knowledge is implicit and is held by the domain experts. Previous works in the knowledge acquisition field have already proved that it is still difficult to grasp knowledge directly from experts, by means of elicitation technique (interviews, observations). Domain experts are rarely able to provide an explicit description of their knowledge and of their reasoning [6,7].

In this paper, we present a methodology to build an urban ontology adapted to the multi-level interpretation of multi-source images. The work is a part of the French-funded project *FoDoMuSt* which aims at developing a complete process of data mining for interpreting urban remotely sensed images. The methodology associates (1) segmentation of the images and their classification in regions using low-level descriptors (e.g. radiometry, texture, shape, size of the elements), and (2) use of domain knowledge in order to transform the segmented regions into semantic objects. Domain knowledge is stored independently in the urban ontology which contains a set of pre-defined terms characterizing urban domain concepts.

The paper is structured in four sections. First, approaches using ontologies or domain knowledge in image analysis are discussed (Section 2). Second, the methodology to construct the urban ontology is presented (Section 3). Third, the matching process between the regions and the concepts of the ontology is proposed (Section 4). Finally, some experiments on Very High Resolution images (0.7m) on the urban area of Strasbourg (France) are proposed (section 5).

2. Ontologies applied to images: state of the art

An ontology can be defined as a simplified view of the world which is represented for specific purpose [8, 9, 10]. An ontology defines a set of representational terms called *concepts,* their characteristics and their relationships. It is the result of a consensus in an user community to clarify the communication [11]. According to the building process, an ontology can be generic [12] (such as WordNet [13, 14] and Sensus [15, 16, 17]) or domain-dependent.

For our application of landcover/use analysis in urban and peri-urban areas, a domain-dependent ontology has been chosen. The objective is to help urban planners to map and update automatically information at several spatial and temporal scales.

2.1 Context: multi-source remotely sensed images

The conceptualization of landcover/use thematic classes from remotely sensed images depends on the scale of the objects and on the spatial resolution of the aerial or satellite images. Some concepts only exist at a single spatial resolution (for instance, it is difficult to individualize a tree on a 20 m spatial resolution image) while other concepts exist at all spatial resolution (from instance, a lake can be identified from 30 to 1 m spatial resolution images). However, their instantiation change in terms of 'spatial representation'. For instance, if a small street is divided into two branches when observed on a 20 m spatial resolution image, its representation is changed when observed on a 1 m spatial resolution. The conceptualization of thematic classes can be associated to an ontology putting together all the multi-level descriptors used by the urban experts to identify thematic classes.

2.2 Related works

Knowledge based systems have proved to be effective for complex object recognition [18] and for image analysis [19]. For instance, the SIGMA [19] and Schema [20] systems perform image analysis on aerial images by using several descriptors of objects. These systems give access to a high semantic level. Nevertheless, as pointed by [21] such systems are strongly domain-dependent as they integrate prior knowledge on the image. Their drawbacks consist in the fact that the domain knowledge is not clearly separated from the procedure, and that the knowledge bases are difficult to produce.

Therefore, recent works have proposed to use ontologies to describe more clearly the knowledge of the studied domain. In [22] spatial relations between concepts are used to merge regions and to recognize objects. The exclusive use of spatial relations is however not possible in the case of remotely sensed images. This work points out the differences between domain knowledge and procedures. [23] proposed an ontology-based object learning and recognition system for image analysis. An interesting point is the separation of a local matching and a global matching procedure (e.g. the global matching combines the probabilities computed during the local matching). The descriptors used for the matching correspond to "visual concepts" which are learnt during the learning phase. The matching function is then dependent of these visual concepts. The authors state that the global matching should take into account the hierarchy of the ontology. [24] proposed an ontology-based object detection for video analysis, using a segmentation process, while [12] used a neural network method to classify objects in pre-defined classes. Then, the proposed system determines if the image may be classified by a concept from an ontology. [25] proposed a genetic algorithm of ontology-driven semantic image analysis. Some low-level descriptors are extracted from the image and are used to match with the ontology. A set of hypothesis (region, list of possible concepts and their degrees of confidence) are then tested with a genetic algorithm to determine the optimal image interpretation. Only spatial relations (8 directional relations) are used by the system.

Using domain-dependant ontologies for object analysis from muti-source remote-sensed images presents two main challenges: the first is the extraction of the semantic concepts from several spatial and spectral resolution images, and the second is the construction of the ontology. A key issue is to identify appropriate concepts to describe the thematic objects for a multi-scale representation of the territory. As well, the ontology has to be adapted to the multi-scale mapping of urban area (from 1/100,000 to 1/10,000), and should consist in a multi-scale definition of the objects.

3. Construction of the urban ontology

The urban ontology is built in order to assign the segmented regions from remotely sensed images into semantic objects. The knowledge domain has to be modelled as a scene-knowledge detailing how each object appears in each image. The proposed ontology is then adaptative and implies that all urban objects are neither described at all scales, nor classified at all resolutions.

Even if no standard type of ontology exists for each domain of application [26], [27] proposes a 3-steps methodology to construct the ontology. The first step (Section 3.1) is a phase of specification which consists in identifying the concepts used in urban management. The second step (Section 3.2) is a phase of conceptualization which consists in storing this knowledge in a dictionary. The third step (Section 3.3) is a phase of formalization which consists in modelling this knowledge in an ontology and in implementing it in a computer-usable form.

3.1 Phase of specification

The inventory of semantic objects is based on urban object typologies and nomenclatures defined by experts. There is a wide range of urban object nomenclatures for remotely sensed data such as the Corine Land Cover nomenclature defined for Landsat images (30m spatial resolution), the Spot Thema nomenclature defined for Spot images (5m to 20m) or the French national landcover database BDCarto©IGN (defined for aerial photographs and Spot images). A detailed terminological analysis of these nomenclatures showed that there none of this typology is really adapted to map urban areas at 1/10,000. In this context, since Very High spatial Resolution satellite images can be used to extract urban objects at this scale [28], a high level of description has been established according to the users' needs.

The proposed nomenclature distinguishes thirty-five urban objects or thematic classes, merged in five levels. The nomenclature is dedicated to the mapping of the urban characteristics of western cities [29]. Table 1 presents an extract of this common nomenclature based on housing thematic classes. This fifth level is based on dictionary of urban objects defined in urban GIS Platform and on the minimal spatial resolution defined to identify these objects [30].

Tab. 1. Extract of the nomenclature used to map urban area from 1/100,000 to 1/10,000 scale.

Level 1 1/100,000	Level 2 1/100,000 to 1/50,000	Level 3 1/50,000 to 1/25,000	Level 4 1/25,000	Level 5 1/10,000
Artificial/ Mineral surfaces	Housing surfaces	Continuous Urban fabric	High density of continuous Urban fabric	High density ofc continuous Urban fabric
			Low density of continuous Urban fabric	Low density of continuous Urban fabric
		Discontinuous urban fabric	Individual houses	Low density of individual houses
				Medium density of individual houses
				High density of individual houses
			Collective building	Building with less than 4 floors
				Building with more than 4 floors
			Mixed	Mixed (houses and building)
		Specific urban surfaces	Specific urban surfaces	Cemetry
				Surfaces with military buildings
				Surfaces with scholar buildings
				Surfaces with hospital buildings
				Others surfaces

3.2 Phase of conceptualization

A set of landcover/use terms representing the linguistic expression of the urban planner knowledge has been collected. This set of terms is defined as a hierarchy with related definitions stored in a lexicon. Then the conceptualization phase has consisted in organizing and structuring the elements of the lexicon into a taxonomy of concepts.

The taxonomy has then been adapted and enriched because all the urban concepts are not always identifiable on the images and because the spectral responses of the pixels mainly provide information about the landcover properties. Thus, the taxonomy explicitly distinguishes '*image object'(IO)* and *'built object' (BO).*

An '*image object'* (IO) is defined as an '*object directly identifiable on images for a specific spatial resolution'*. Three ranges of spatial resolution are defined [30]: [0.5m-5m[, [5m-15m] and]15m-30m]. Each IO is distinguished into '*simple*' or '*aggregate*' IO. A '*single IO*' (IOs) is '*an object for which only one group of homogeneous pixels (region) is sufficient to identify the concept to which it refers'*. An '*aggregate IO*' (IOa) is defined as '*an object for which several groups of homogeneous pixels (regions) are necessary to identify the concept to which it refers'*. For instance on the range [0.5m-5m], a tree is an IOs because it can be identified with one group of pixel (Figure 1a), and a tree row is an IOa because it is composed of several individual or grouped trees (Figure 1b).

A '*built object*' (BO) requires to use spatial knowledge to be identified. A BO is then always composed of several IO organized within a spatial pattern. For instance, a park is a BO because it is composed of a specific pattern of trees associated to grassland, and water bodies (Figure 1c). Specific relationships between each IO can be set up to identify the BO.

The taxons are classified into three classes in accordance to their spectral response and their thematic interpretation. More precisely a '*thematic code*' (mineral, vegetation, water) is assigned. The class of mineral object distinguishes buildings, ways and equipments. A '*color code*' is also added to describe this class with reference to their 'natural color' (White, Gray/Black, Orange) on the images.

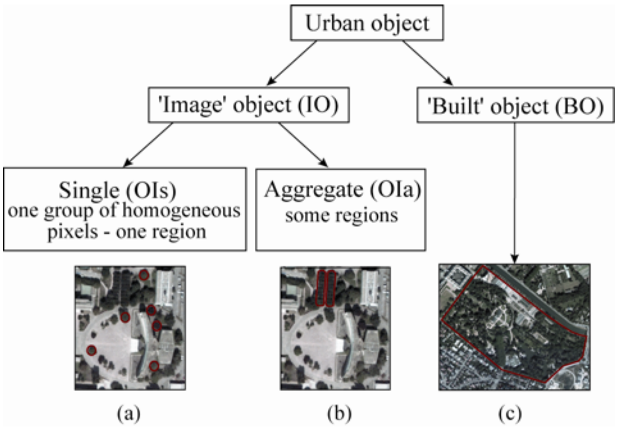

Fig. 1. Conceptualisation of an urban object. Example of: (a) a single 'image object' (IOs) such as a tree, (b) an aggregate 'image object' (IOa) such as a line of trees, and (c) a 'built object' (BO) such as a park.

All the concepts are stored in a dictionary, adapted from [31], which contains three categories of information:

(1) Some characteristics to identify the objects: name, representation in a GIS database (point, polyline or polygone), type of IO (single, aggregate), range of spatial resolution at which the object is identifiable, color code, and 'thematic code';

(2) A qualitative (textual definition) and a quantitative description of the 'real world'-objects in terms of radiometric measures. The textual definition is a result of a consensus between the urban users' needs, and the radiometric measures are obtained from spectra-radiometer measurements [32]. The latest information is only available for some vegetation classes (species) and some urban material (roof, asphalt, etc). A graphic illustration (orthophotograph of the object) is also added;

(3) A description of the IO in terms of low-level descriptors (color, shape, texture, context or spatial relationships). The Table 2 summarizes the proposed descriptors.

Tab. 2: The low-level descriptors identified for each IO.

Descriptors	Comments
Spectral reflectance	Range of observed values in 4 spectral bands: Blue (B) – Green (G) – Red (R) – near-infrared (NIR)
Normalized Difference Vegetation Index (NDVI)	Range of observed values of NDVI
Soil Brightness Index (BI)	Range of observed values of BI
Shape properties	Range of observed values of area, perimeter, elongation, diameter, compactness (Miller index), and solidity

Texture	Range of observed values of the homogeneity index and of the variance derived from the co-occurrence grey-level matrix (Haralick, 1973)
Context (or relationships)	Adjacency, inclusion, composition, neighbourhood

3.3 Phase of formalization

The next step consists in formalizing the concepts and in storing the knowledge. Among the works presented in [33] to model knowledge, *frame-based* approaches seemed to be the most relevant for image classification problems [34]. In these systems [35], knowledge is grouped into "*frames*". A frame is a prototype (e.g. a representative object of a family) composed of a set of slots describing the properties of the prototypes. Thus, domain knowledge is modelled in the ontology which consists in both a set of classes organized in a hierarchy and a set of slots (attributes) associated to concepts.

The ontology has been developed with *Protégé-2000* [26] which is a free open-source platform that provides a suite of tools to construct domain models and knowledge-based applications. *Protégé* is specifically *frame-oriented*. It is based on Java and can be extended by way of a plug-in architecture and a Java-based Application Programming Interface (API) for building knowledge-based tools and applications. *Protégé-2000* facilitates conformance to the Open Knowledge Base Connectivity (OKBC) protocol for accessing knowledge bases stored in knowledge representation systems. For all these points, *Protégé-2000* is widely used to model domain knowledge, for instance in images indexation [36, 37].

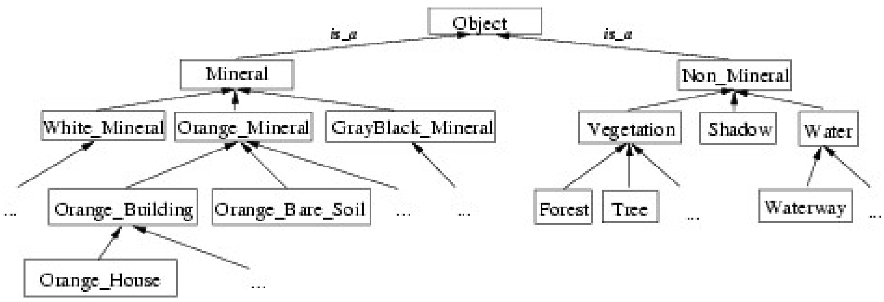

Fig. 2. Extract of the ontology.

The ontology (Figure 2) is composed of 91 concepts and 20 attributes. The depth of the ontological tree is 6. Each concept has a label (e.g. `Orange_House` for individual houses with orange roof tiles) and is defined by attributes corresponding to the low-level descriptors. Each attribute of a concept is weighted according to its importance to recognize the urban object represented by the concept (IO). At this stage of development, the term the spatial relationships are not yet defined in the ontology. The following knowledge formalization is then used:

Definition 1 (concept, sub-concept, depth):

Let Θ *be the set of concepts,* $\leq_{\Theta}$ *is a partial order between concepts.* $\forall (C_i, C_j) \in \Theta^2$, $C_i \leq_{\Theta} C_j$ *means that* C_i *is a sub-concept of* C_j. $\rho(C)$ *is the depth of the concept C in the hierarchy.*

For example, C_i = `Orange_House` is a sub-concept of C_j = `Orange_Building`. $\rho(C_i)$ = 5 (Figure 2).

Definition 2 (classes of attributes):
Let Φ *be the set of attribute classes.* A *is the set of all attributes.* A = *{spectral_signature_Blue, …, area, …, Miller index, …}. For a set of attribute classes* $\alpha \subseteq \Phi$, $A_\alpha \subseteq A$ *is the set of attributes of each class in* α.

For instance, Φ = {`spectral`, `spatial`, `contextual`} (Figure 3).

If α = {`spectral`}, A_α = {`spectral_signature_Blue`, `spectral_signature_Green`, `spectral_signature_Red`, `spectral_signature_NearInfraRed`, `spectral_signature_SBI`, `spectral_signature_NDVI`}.

If $\alpha=\Phi$, $A_\alpha = A$ (all the attributes).

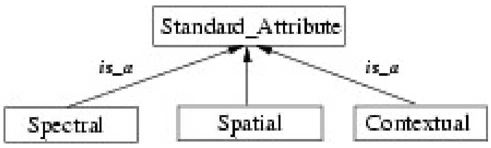

Fig. 3. Hierarchy of the attribute classes.

Definition 3 (specific attributes of a concept):
Let $F_\alpha(C)$ *be the set of attributes of the classes in* α*, specifically associated with the concept* $C \in \Theta$.

For instance, for the concept C = `Orange_House`, if the spectral attributes (`spectral_signature_Blue`, …) and their values are inherited by the `Orange_Building`, they are not present in $F_\alpha(C)$. But an attribute overrided in *C*, is present in $F_\alpha(C)$.

Definition 4 (values and weight of an attribute):
Let $a \in A_\alpha$ *be an attribute of a class in* $\alpha \in \Phi$*. We define* $V_C : A_\alpha \to [\Re, \Re]$ *so that* $V_C(a)$ *is the range of values for 'a' in the concept* $C \in \Theta$*. Let* $\omega(a,C)$ *be the weight associated to the attribute 'a' for the concept C.*

Definition 5 (set of regions):
Let Γ be the set of regions.

Definition 6 (feature value of a region):
Let $a \in A$ be a feature of a region $R \in \Gamma$. We define $V'_R : A_\alpha \rightarrow \Re$ so that $V'_R(a)$ is the value of 'a' for the region R.

This phase of knowledge modelling in the ontology consists of defining Θ, $\leq_\Theta$, $F_\alpha(C)$, $\omega(a,C)$, Φ and $V_C(a)$. This allows reducing the semantic gap between expert knowledge and image level.

4. Use of the ontology in the object-recognition process

In the data mining process, after image segmentation, each region (e.g. group of homogeneous pixels) is characterized by a set of features (e.g. the low-level descriptors described in Section 3.2). The regions and their features are the inputs to the ontology-based object recognition. In this phase, the objective is to match each region with the concepts of the ontology. This is realized by (1) proposing a matching measure, and (2) defining a method to navigate in the ontology.

4.1 Step 1: Matching score

The proposed matching method is a "*feature-oriented*" approach. The matching method verifies the validity of the feature values defined for a region according to the properties and the constraints defined for the concepts. Nevertheless, as a region does not have a semantic structure, measures like MDSM [38] or the measures presented in [39] and [40] can not be directly used. A region can be matched *a priori* with any concepts, but the features of a region allowing the matching could not be identical according to the studied concept. For instance, the concept `Orange_House` is defined by several shape attributes (elongation index, Miller index) and spectral attributes, while the concept `Shadow` is only defined with spectral attributes. Without *a priori* knowledge, this dissymmetry necessitates to compute all the features for each region, even if the majority of these features will not be used in the matching process. In order to take into account all these specificities, a matching measure based on a distance between the extracted features of a region and the observed values of the attributes has been proposed. This measure computes the relevance of a matching and is composed of a local component (representing the inner properties of the concept) and a global component (evaluating the pertinence in the hierarchy of concepts). The local similarity measure compares the features of a region with the specific attributes of a concept of the ontology.

Definition 7 (degree of validity)
Let Valid(a,C,R) be the validity degree of an attribute 'a' between a region R and a concept C.

$$Valid(a,C,R)=\begin{cases}1 \ \ if\ V'_R(a)\in[\min(V_C(a));\max(V_C(a))]\\ \dfrac{V'_R(a)}{\min(V_C(a))} \quad if\ V'_R(a)<\min(V_C(a))\\ \dfrac{\max(V_C(a))}{V'_R(a)} \quad if\ V'_R(a)>\max(V_C(a))\end{cases}$$

Definition 8 (local similarity)

Let be $Sim_\alpha(R,C)$ the local similarity between a region R and a concept C using the attributes of each class in α.

$$Sim_\alpha(R,C)=\frac{\sum_{a\in F_\alpha(C)}\omega(a,C)Valid(a,C,R)}{\sum_{a\in F_\alpha(C)}\omega(a,C)}$$

The global matching score thus evaluates the pertinence of the matching between a region and a concept in the hierarchy of concepts.

Definition 9 (matching score)

Let $Score_\alpha(R,C)$ be the matching score between a region R and a concept C, and $P(C)$ be the path starting from the root of the ontology and ending at the concept C. $P(C)=\{C_j \mid C=C_J \le_\Theta \ldots \le_\Theta C_j \le_\Theta \ldots \le_\Theta C_2 \le_\Theta C_1\}$.

$$Score_\alpha(R,C)=\frac{\sum_{C_j\in P(C)}\rho(C_j)\,Sim_\alpha(R,C_j)}{\sum_{C_j\in P(C)}\rho(C_j)}$$

The matching score is a linear combination of local similarity measures obtained with the concepts of the path starting from the root of the ontology and ending at the studied concept. The local similarities are propagated by inheritance to more specific concepts. In this computation, a specialization coefficient ρ equal to the depth of the concepts is integrated. In this way, the measure favours the specialization of the concepts, considering all additional information giving a new semantic.

4.2 Step 2: Navigation in the ontology

The second step consists to navigate in the ontology to find the best concept(s) according to the score for a region. A level-wise algorithm has been developed to navigate in the ontology using heuristics to reduce the search space. The parameters can be accorded to a specific matching context. The general scheme of the exploration algorithm is defined as follows. If the region matches the current concept, the algorithm will go deeper in the hierarchy defined by the partial order $\leq_\Theta$ in the next level; if the matching fails, the current concept is dropped and its sub-concepts are not explored. The main heuristic corresponds to the selection of the best concepts at each level in order to eliminate some branches for which the starting concept would not be relevant (e.g. with a low matching score value). This strategy supposes that a root concept has more generic properties than its children, and if a few of these properties (or none) are valid, its child will not be relevant.

Two thresholds are also defined: *maxDepth* is the exploration maximal depth (e.g. the degree of detail) and *minScore* is the minimal value of the matching score between a region and a concept to allocate the corresponding label to the region. For instance in Figure 2, if the maximal depth is equal to 3, only the classical categories will be explored (mineral, vegetation, etc.).

Definition 10 (labels identified for a region):

We define $L_\alpha : \Gamma \rightarrow \Theta$ *so that* $L_\alpha(R)$ *is the set of concepts (seen as labels) identified for the region R according to the attributes of* A_α *and the minScore value.*

$$L_\alpha(R) = \{C_i \mid \rho(C_i) \leq \max Depth \ \ and \ \ Score_\alpha(R, C_i) \geq \min Score \ \ and$$
$$\nexists \ C_j (\neq C_i) \ Score_\alpha(R, C_j) > Score_\alpha(R, C_i)\}$$

The procedure of navigation when the heuristic procedure is activated (selection of the best concepts at each level) is presented in Algorithm 1. This can be repeated for each region of the segmented image in order to provide an interpretation of the complete image.

5 Experiments on VHR images

The methodology has been applied to a set of VHR images of the city of Strasbourg (North-East France). In this paper, an experiment on subset of a Quickbird image (©DigitalGlobe2002) representing a typical urban fabric of individual houses is detailed (Figure 4a). The tested image size is 900 x 900 pixels. It has a 0.7m spatial resolution (pan-sharpened process detailed in [41]) and four spectral bands (blue – green – red and near-infrared).

In order to facilitate the phase of assessment, an extract of the ontology is also used. The number of concepts (labels) is thus limited. Only the classes `Vegetation`, `Water`, `Road` and `Orange_House` are being recognized. In the case no label is found for a region, the system uses the label `Unknown`. The evaluation consists in comparing the results of the method with a manually labelled region (by domain expert).

Algorithm 1 Navigation algorithm of the ontology.

Input: a region R, an ontology $(\Theta,\ \Phi,\ \mathcal{V}_C(a),\ \ldots)$, a set of attribute classes (α), $maxDepth$ and $minScore$.
Output: the best label(s) and the matching score value.

$depth = 1;\ \ scoreMax = minScore;$
$\mathcal{L}_\alpha(R) = \emptyset;$
$\mathcal{RC} = \{root\};\ \ scoreDepth = 0;\ \ bestsDepth = \emptyset;$
while $(\mathcal{RC} \neq \emptyset$ and $depth \leq maxDepth)$ **do**
  $scoreDepth = 0;\ \ Best = \emptyset;$
  for all $C \in \mathcal{RC}$ **do**
    $s = Score_\alpha(R, C);$
    if $(s == scoreMax)$ **then**
      $\mathcal{L}_\alpha(R) += \{C\};$
    end if
    if $(s > scoreMax)$ **then**
      $\mathcal{L}_\alpha(R) = \{C\};\ \ scoreMax = s;$
    end if
    if $(s == scoreDepth)$ **then**
      $bestsDepth += \{C\};$
    end if
    if $(s > scoreDepth)$ **then**
      $bestsDepth = \{C\};\ \ scoreDepth = s;$
    end if
  end for
  $\mathcal{RC} = \emptyset;$
  for all $C_j \in bestsDepth$ **do**
    $\mathcal{RC} = \mathcal{RC} \cup \{C_i | C_i \preceq_\Theta C_j\};$
  end for
  $depth++;$
end while
return $\{\mathcal{L}_\alpha(R),\ score\};$

(a) (b) (c)

Fig. 4. a. Extract of the Quickbird image of the city of Strasbourg illustrating a typical individual house urban fabric. b. Segmented image. c. Recognized objects set (the unknown objects are highlighted in white).

The segmentation step is not described here but it is important to note that the global result is extremely dependent of the segmentation. Some researches are being in development to apply a segmentation algorithm adapted to the object recognition [42]. We can however note that the proposed method is not dependent of the segmentation algorithm. Figure 4b presents the segmented image obtained by a supervised segmentation algorithm.

Table 3 presents the average values of Precision, Recall and F-measure (on/in overall, all classes) for several *minScore* values. The best overall results are obtained with *minScore*=0.98 and not with the maximal value of *minScore*. This is due to the decrease of the Recall measure more important than the increase of the precision measure for *minScore*=1. Nevertheless, in our case, a very good precision value is more important than a very good recall value.

Tab. 3. Assessment of the results.

minScore	Precision	Recall	F-Measure
0.85	0.878	**0.861**	0.870
0.9	0.893	0.854	0.873
0.98	0.954	0.823	**0.884**
1	**0.967**	0.771	0.858

Tab. 4. Compared results between *minScore* = 1 and *minScore* = 0.98.

classes	Precision		Recall	
	1	0.98	1	0.98
Orange_House	0.895	0.875	0.435	0.585
Vegetation	0.995	0.994	0.950	0.953
Road	0.980	0.947	0.712	0.762
Water	0.999	0.999	0.988	0.995

Table 4 presents the detailed results for *minScore*=1 and *minScore*=0.98. `Vegetation` and `Water` are very well identified. `Road` has good precision values and the recall values is correct. The precision values for `Orange_House` are relatively good but the recall values are too low. An explanation of this result can be found in the quality of the segmentation (Figure 4b). Indeed, some houses are not correctly segmented and they are merged with other houses. Thus, these houses can not have the features corresponding to the attributes defined in the ontology, especially the elongation indexes. A similar problem is encountered with the road which is over-segmented.

The percentage of recognized objects according to the *minScore* value, and the percentage of the corresponding image (pixels of the recognized objects) according to the *minScore* value are illustrated in Figure 5. They show that a major part of the image is recognized, and thus labelled. For the maximal value (*minScore*=1), 14.8% of the objects are recognized but it corresponds to 62.5% of the image. With *minScore*=0.98, 26.7% of the objects are identified (72.5% of the image). These results are promising even if many small objects can not yet be identified (Figure 4c).

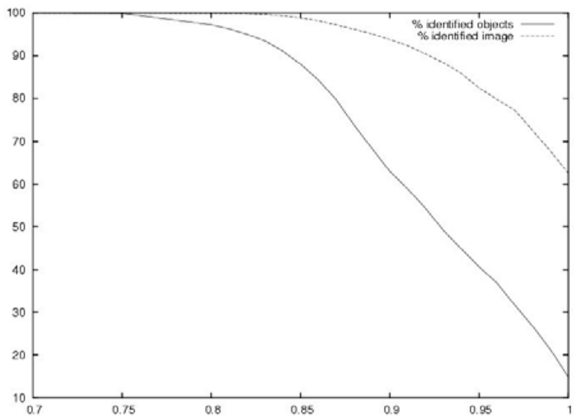

Fig. 5. Percentage of labelled objects and pixels according to the *minScore* value.

6 Conclusion and perspectives

In this paper, the steps to build an urban ontology applied to VHR image analysis have been presented. A new knowledge representation and reasoning method has been detailed. The approach is based on a domain-dependent ontology developed by experts of the domain. A similarity measure and an exploration procedure of the ontology have been presented in order to allocate a semantic to regions of a segmented image. The experiment results have shown effectiveness of the method, despite the fact that the results could be improved using more appropriate segmentation algorithms.

In order to improve and to enrich the urban ontology, machine learning techniques will be used to extract knowledge automatically from the VHR images [7]. Topological relations based on the RCC-8 (Region Connection Calculus) theory will also be integrated in the methodology. Moreover, several experiments on different types of urban images are planned using several segmentation algorithms. Finally, the method will be incorporated in a multi-strategy classification approach in order to guide the process, to label the clusters, and to improve the final classification results.

References

1. Lillesand, T. M., and. Kiefer, R. W: Remote sensing and image interpretation, 3nd Edition, John Wiley and Sons. New York. (1994). 721p.
2. Smeulders, A., Worring, M., Santini, S., Gupta, A., and Jain R.: Content-based Image Retrieval at the End of the Early Years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12): (2000) 1349–1380.
3. Herold, M., Scepan, J., Muller, A., Gunter, S.: Object-oriented Mapping and Analysis of Urban Landuse/Cover using Ikonos Data. In *22nd Earsel Symposium Geoinformation for European-Wide Integration*, Prague, (2002).

4. Benz, U., Hofmann, P., Willhauck, G., Lingenfelder, I., and Heynen M.: Multi-Resolution, Object-oriented Fuzzy Analysis of Remote Sensing Data for GIS-Ready Information. *ISPRS Journal of Photogrammetry & Remote Sensing*, 58, (2004) 239–258
5. Baltasavias, E.P.: Object extraction and revision by image analysis using existing geodatanand knowledge: current status and steps towards operational systems, *ISPRS Journal of Photogrammetry and Remote Sensing*, 58, (2004) 129-151
6. Sester, M., Knowledge acquisition for automatic interpretation of spatial data, International Journal of Geographical Information Science, 14 (2000) 1-24
7. Sheeren, D., Puissant, A., Weber, C., Gancarski, P., Wemmert, C.: Deriving Classification Rules from Multiple Remotely Sensed Data with Data Mining. In *Proc. 1st Workshop of the EARSeL Special Interest Group on Urban Remote Sensing*, Berlin, March (2006), 8p.
8. Bunge, M. A.: Treatise on Basic Philosophy: Ontology: The Furniture of the World", Reidel, Boston, (1977).
9. Gruber, T. R.: Toward Principles for the design of Ontologies used for Knowledge Sharing, in Proc. Of *International Workshop on Formal Ontology*, (1993).
10. Labrou, Y. and Finin, T.: Yahoo! as Ontology: Using Yahoo! Categories to Describe Documents, in Proc. Of *The Eighth International Conference on Information Knowledge Management*, Kansas City, MO. (1999) 180-187
11. Gandon, F.: Distributed Artificial Intelligence And Knowledge Management: Ontologies And Multi-Agent Systems For A Corporate Semantic Web. Doctoral school of sciences and technologies of information and communication, University of Nice Sophia Antipolis(2002).
12. Breen, C., Khan, L., Ponnusamy, A.: Image Classification Using Neural Networks and Ontologies. In *Proc. 13th International Workshop on Database and Expert Systems Applications, co-located with DEXA 2002*, , Aix-en-Provence, France, (2002) 98–102
13. Miller, G.: Wordnet: A Lexical Database for English, in Proc. of *Communications of CACM*, (1995).
14. Miller G.: Nouns in WordNet: a Lexical Inheritance System', *International Journal of Lexicography*, 3(4), (1994) 245-264
15. Hovy, E. and Knight K.: Motivation for Shared Ontologies: An Example from the Pangloss Collaboration, in Proc. of *The Workshop on Knowledge Sharing and Information Interchange, IJCAI-93*, Chambry, France, (1993).
16. Knight, K., Chander, I., Haines, M., Hatzivassiloglou, V., Hovy E. H., Iida, M., Luk, S. K., Whitney, R. A., and Yamada, K.: Filling Knowledge Gaps in a Broad-Coverage MT System," in Proc. of *The Fourteenth IJCAI Conference*, Montreal, Quebec, Canada, (1995)
17. Large Resources Ontologies (SENSUS) and Lexicons, http://www.isi.edu/naturallanguage/projects/ONTOLOGIES.html
18. Liu, S., Thonnat, M., and Berthod, M.: Automatic Classification of Planktonic Foraminifera by a Knowledge-Based System. In *10th Conference on Artificial Intelligence for Applications*, San Antonio, Texas, (1994) 358–364
19. Matsuyama, T. and. Hwang, V.-S.: *SIGMA - A Knowledge-Based Aerial Image Understanding System*. Plenum Press New York USA, (1990)
20. Draper, B., Collins, A., Brolio, J., Hanson, A., and Riseman, E.: The Schema System. *International Journal of Computer Vision*, (1989) 2(3):209–250
21. Crevier, D., Lepage, R.: Knowledge-Based Image Understanding Systems: a Survey. *Computer Vision and Image Understanding*, (1997) 67(2):161–185
22. Zlatoff, N., Tellez, B., and Baskurt, A.: Image Understanding and Scene Models: a Generic Framework Integrating Domain Knowledge and Gestalt Theory. In *Proc. IEEE International Conference on Image Processing (ICIP 2004)*, , Singapore, (2004) 2355–2358

23. Maillot, N., Thonnat, M., and Hudelot, C.: Ontology Based Object Learning and Recognition: Application to Image Retrieval. In *Proc. International Conference on Tools with Artificial Intelligence (ICTAI 2004)*, Boca Raton, Florida (2004). 620–625
24. Dasiopoulou, S., Mezaris, V., Kompatsiaris, I. ,. Papastathis, V.K, and Strintzis M.G.: Knowledge-Assisted Semantic Video Object Detection. *IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Analysis and Understanding for Video Adaptation*, (2005) 15(10):1210–1224
25. Panagi, P., Dasiopoulou, S., Papadopoulos, G.T. Kompatsiaris, I., and Strintzis. M.G.: A Genetic Algorithm Approach to Ontology-Driven Semantic Image Analysis. In *Proc. 3rd IEE International Conference of Visual Information Engineering (VIE 2006)*, Bangalore, India (2006).
26. Noy, N.F., Fergerson, R.W., and Musen, M.A.: The Knowledge Model of Protege-2000: Combining Interoperability and Flexibility. In *Proc. 12th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2000)*, Juan-les-Pins, French Riviera, (2000).
27. Uschold, M. & King, M.: Towards a Methodology for Building Ontologies. In: *Workshop on Basic Ontological Issues in Knowledge Sharing*, Montreal, Canada. **(1995)**
28. Puissant A., Weber C., Vers une démarche orientée-objets pour extraire les objets urbains sur des images THR, *Bulletin de la Société Française de Photogrammétrie et Télédétection*, (2003) 173-174: 123-134.
29. Puissant, A., Weber C.: *The use of image in Geographical Information Market: results of an inquiry on the needs of end-users in urban studies, in* Proceedings of the 7th EC-GI & GIS WORKSHOP, EGII – Managing the Mosaic, June 13-15, 2001, Potsdam, Germany, (2001) 8p.
30. Puissant, A., Hirsch, J. : Télédétection urbaine et résolution spatiale optimale : intérêt pour les utilisateurs et aide pour les classifications, *Revue Internationale de Géomatique* (2004) 14(3-4), 403-415.
31. Pantazis, D., Cornelis, B., Billen, R. & Sheeren, D.: Establishment of a Geographic Data Dictionnary : a case study of UrbIS 2©, the Brussels Regional Government GIS, Computers, Environment & Urban Systems, (2002) 26(1): 3-17
32. Wania et al., 2006
33. Roussey, C.: Une méthode d'indexation sémantique adaptée au corpus multilingues. Thèse Informatique. Insa de Lyon, Lyon. (2001).
34. Marino Drews, O.: Raisonnement classificatoire dans une représentation à objets multi-points de vue. Thèse Informatique. Université J. Fourier, Grenoble. (1993).
35. Minsky, M.: A Framework for Representing Knowledge. *The Psychology of Computer Vision*, Edited by P. Winston. New-York: McGraw-Hill, (1975). 211-281.
36. Schreiber, A, Dubbeldam, B, Wielemaker, J, Ontology-based Photo Annotation. *IEEE Intelligent Systems* (2001) 16(3):66-74.
37. Hyvönen, E, Styrman, A, Saarela, S.: Ontology-Based Image Retrieval. In: *Proceedings of XML Finland 2002 Conference* (pp. 15-27), Helsinski (2002).
38. Rodriguez, M. A. and Egenhofer, M. J.: Determining Semantic Similarity Among Entity Classes from Different Ontologies. *IEEE Transactions on Knowledge and Data Engineering*, (2003) 15(2):442–456
39. Tversky, A.: Features of Similarity. *Psychological Review*, (1977)84:327–352
40. Schwering, A. and Raubal, M.: Measuring Semantic Similarity between Geospatial Conceptual Regions. In *Proc. 1st International Conference on GeoSpatial Semantics (GeoS)*, (2005) 90–106
41. Puissant, A., Ranchin, T., Weber, C., and Serradj, A.: Fusion of Quickbird MS and Pan Data for Urban Studies. In *Proc. European Association of Remote Sensing Laboratories Symposium (EARSeL)*, Gent, Belgium, (2003) 77–83.

42. Derivaux, S., Lefèvre, S., Wemmert, C., and Korczak, J.: Watershed Segmentation of Remotely Sensed Images Based on a Supervised Fuzzy Pixel Classification. In *Proc. IEEE International Geoscience And Remote Sensing Symposium (IGARSS 2006)*, Denver, Colorado, (2006).
43. Derivaux, S., Lefèvre, S., Wemmert, C., and Korczak, J.: Watershed Segmentation of Remotely Sensed Images Based on a Supervised Fuzzy Pixel Classification. In *Proc. IEEE International Geoscience And Remote Sensing Symposium (IGARSS 2006)*, Denver, Colorado, (2006).

Aknowledgments

This works has been performed through the French-funded project FoDoMusT 'Fouille de données multi-stratégies pour extraire et qualifier la végétation urbaine' (2004-2007).