Knowledge-based region labeling for remote sensing image interpretation

Germain Forestier, Anne Puissant, Cédric Wemmert, Pierre Gançarski

Knowledge-based region labeling for remote sensing image interpretation

G. Forestier[1], A. Puissant[2], C. Wemmert[1], P. Gançarski[1]

University of Strasbourg (France)
[1] *LSIIT - Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection UMR 7005 CNRS - UdS*
[2] *LIVE - Laboratoire Image, Ville, Environnement ERL 7230 CNRS - UdS*

Abstract

The increasing availability of High Spatial Resolution (HSR) satellite images is an opportunity to characterize and identify urban objects. Thus, the augmentation of the precision led to a need of new image analysis methods using region-based (or object-based) approaches. In this field, an important challenge is the use of domain knowledge for automatic urban objects identification, and a major issue is the formalization and exploitation of this knowledge. In this paper, we present the building steps of a knowledge-base of urban objects allowing to perform the interpretation of HSR images in order to help urban planners to automatically map the territory. The knowledge-base is used to assign segmented regions (*i.e.* extracted from the images) into semantic objects (*i.e.* concepts of the knowledge-base). A matching process between the regions and the concepts of the knowledge-base is proposed, allowing to bridge the semantic gap between the images content and the interpretation. The method is validated on Quickbird images of the urban areas of Strasbourg and Marseille (France). The results highlight the capacity of the method to automatically identify urban objects using the domain knowledge.

Keywords: Urban object, knowledge base, high resolution, remote sensing images, semantic interpretation, region labeling

1. Introduction

Urban planners are interested in up-to-date land cover and land use information on urban objects at several spatial (1:100,000 to 1:5,000) and temporal scales. Acquiring automatically this information is complex, difficult and time-consuming if traditional data sources (*e.g.* ground survey techniques) are used. The increasing availability of remotely sensed images with Medium Spatial Resolution (MSR) from 30 to $10m$ or High Spatial Resolution (HSR) from 5 to $1m$ is an opportunity to characterize and identify these objects into urban and peri-urban areas [44]. 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 and can be defined as the extraction of the image semantic. It consists in obtaining useful spatial and thematic information on the objects by using human knowledge and experience [24, 27]. In this domain, differences are observed between the *visual* interpretation of the spectral information and the semantic interpretation of the pixels, mainly due to different levels of abstraction. The semantic is not always explicitly contained in the image and depends on domain knowledge and on the context. This problem is known as the *semantic gap* [37] and is defined as the lack of concordance between low-level information (*i.e.* automatically extracted from the images) and high-level information (*i.e.* analyzed by urban experts). In order to reduce the semantic gap, image analysis methods using *region-based* (or *object-based*) approaches with domain knowledge are developed [23, 5]. These methods involve the segmentation of the images into homogeneous regions and the characterization of the regions with a set of spectral (*e.g.* spectral signature, spectral index), spatial (*e.g.* shape index) and topological (*e.g.* adjacency, inclusion) features. Region-based classification is known to achieve better results than pixel-based classification [11] for processing HSR images. However, only few initiatives have focused on the use of domain knowledge for classifying urban objects [4], and a major issue in these approaches is therefore domain knowledge formalization and exploitation. Building a knowledge-base is a difficult task because the knowledge is most of the time implicit and held by the domain experts.

The aim of this paper is to highlight the benefits of using a knowledge-base (KB) for automatic regions labeling in order to store expert knowledge and to use it to automate image interpretation. The contribution of this paper is twofold. First, we present the building steps of a knowledge-base adapted to the interpretation of HSR images. A key issue is to identify appropriate concepts in terms of external structure (*i.e.* a hierarchy) and, in terms of internal definition (*i.e.* the attributes and their domain values) to describe the thematic objects for mapping the territory. In particular, we describe an attributes-filling
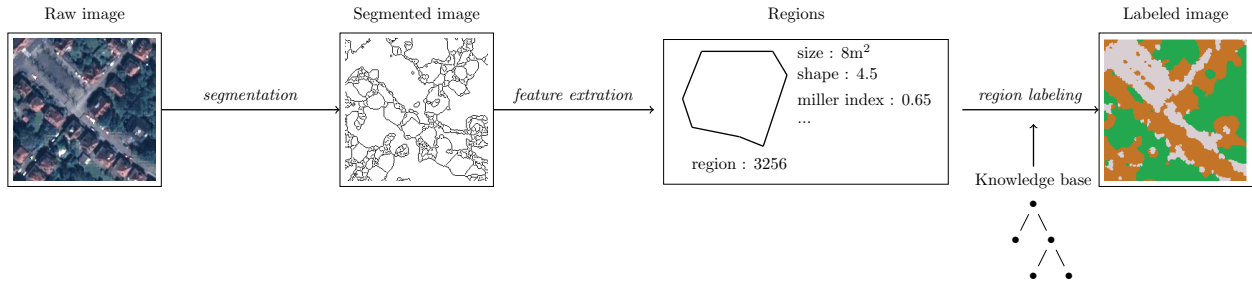
Figure 1: The region labeling workflow: from a raw image to a labeled image.

mechanism used to feed the knowledge-base. The second contribution lies in the validation of a matching method which uses the knowledge-base for automatic image interpretation. The purpose of this method is to label regions extracted from remotely sensed images. The method starts by associating a set of low-level characteristics to each region built using a segmentation algorithm. Then, the knowledge-base is used to assign a semantic to the regions. The figure 1 illustrates the different steps of the approach. We also present experimental results to highlight the relevance of our method on multiple HSR images.

The paper is organised in six sections. First, approaches using domain knowledge in image analysis are discussed (Section 2). Second, the steps to build the knowledge-base adapted to image interpretation is presented (Section 3). Third, the knowledge based region labeling process is detailed (Section 4). Then, some experiments on Quickbird (Digital Globe$^{©}$) images with a spatial resolution of $0.61m$, on the urban areas of Strasbourg and Marseille (France), are proposed (Section 5). Finally, we conclude and present some perspectives (Section 6).

2. Knowledge-based systems for image analysis

Knowledge-based systems (KBS) are becoming more and more important in various domains despite the fact that they are still complex to produce [20]. Indeed, acquiring and representing the knowledge of a domain is often a tedious process and the multiple steps involved in the creation of the knowledge-base can be very different according to the studied domain. This heterogeneity led to an abundance of propositions and the expert is often lost when the time comes to choose a solution. However, the advantages of representing and storing domain knowledge are undeniable. Indeed, it is then possible to produce intelligent systems based on the use of the acquired knowledge and to better explain and understand the domain under consideration.

Knowledge-based systems have proved to be effective for complex object recognition and for image analysis. For instance, the Sigma [26] and Schema [15] systems performed image analysis on aerial images by using several descriptors of the objects. These systems give access to a high semantic level but are strongly domain-dependent as

they integrate prior knowledge on the image [12]. Their main drawback is that the knowledge is not clearly separated from the procedure. Alternatively, [10] proposed a model of knowledge in the framework of landscape planning, with a particular emphasis on cultural landscape, to resolve conceptual misunderstandings and semantic ambiguities, and to provide a precise and accurate description of the current state of the knowledge. In the domain of image segmentation and object labeling, their exist some previous work trying to benefit from a representation of expert knowledge on the objects to extract and label. For example, [2] proposed a new framework for automatic image annotation, guided by expert knowledge represented by an ontological knowledge base. A region growing segmentation algorithm is driven by new similarity measures and merging criteria defined at a semantic level. In [9], the authors presented a fuzzy shape annotation approach for automatic image labeling. The method is based on a fuzzy clustering algorithm, partially supervised by information on the shape of the object and textual labels related to semantic categories. In the remote sensing field, the Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung[1] made a lot of effort since many years for incorporating a priori knowledge into the image interpretation process [39, 7, 8]. Their GeoAIDA system uses a semantic net to model a priori knowledge on the objects in the studied scene. A multi-level semantic segmentation is proposed, built by the collaboration of multiple segmentation algorithms controlled by external operators evaluating the interpretation hypothesis made by the different methods.

A classical way to build a knowledge-base is to use an ontology. An ontology can be defined as a simplified view of the world, which is represented for specific purpose [21]. It defines a set of representational terms called concepts, their characteristics and their relationships. It is the result of a consensus in a user community to clarify the communication. An ontology can have a different representation according to its level of expressivity. It can simply be composed of a taxonomy but can also carry complex axioms about the domain concepts. Depending on the building process, an ontology can be generic or domain-dependent. Therefore, recent works have proposed to use

---

[1]TNT, University of Hannover, Germany

ontologies to describe more clearly the knowledge of the studied domain. In [45], 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. In a same way, [25] 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 (*i.e.* the global matching combines the probabilities computed during the local matching). The descriptors used for the matching correspond to *visual concepts* which are acquired 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. Although, this kind of system needs a time consuming learning step, and also requires the expert to produce examples for each of the concept he is looking for.

Many other works on image analysis tried to benefit from building an ontology. In [13], an ontology-based object detection using a segmentation process for video analysis is proposed. [6] used a neural network method to classify objects in pre-defined classes. Both systems determine if the image may be classified by a concept from an ontology. In [30], the authors 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 (*i.e.* a 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. In [1], the authors present a framework for simultaneous image segmentation and object labeling using an ontology in the domain of multimedia analysis.

In the field of remote sensing several propositions involving the construction of an ontology exist. For example, [17] presented a reflexion about the construction and the use of ontologies at different levels of Geographic Information System (GIS). They proposed an ontology-driven GIS that acts as a system integrator. In this system, an ontology is a component, such as the database, cooperating to fulfill the system's objectives. In another initiative, [40] proposed a framework for ontology-based geographic data set integration, an ontology being a collection of shared concepts. Components of this formal approach are an ontology for topographic mapping (*i.e.* a domain ontology), an ontology for every geographic data sets involved (*i.e.* the application ontologies), and abstraction rules (*i.e.* capture criteria). It is common in GIS to use multiple ontologies to represent different levels of knowledge. The main advantage is to efficiently separate the different kind of knowledge but it leads to complex systems which are difficult to understand as a whole.

Although these work using ontologies are interesting, they rarely tackle the problem of actually identify the concepts present in the created ontology. Indeed, they often describe in details meta-data about the representation, the hierarchy of concepts but often omit an important part: does the modeled knowledge can be used in remote sensing image interpretation ? Our goal in this paper is to propose an actionable representation of the knowledge for image interpretation. In the following sections we present the different steps of the construction and the use of our knowledge-base.

3. Construction of the knowledge-base

The use of domain-dependent knowledge-base (KB) for object analysis from HSR images presents two main challenges: the first is the extraction of the semantic (or thematic) concepts adapted to HSR images and the second is the actual construction of the KB. There are no standard type of KB available for all the domains of application [29, 43]. In agreement with [41], we used a 3-steps methodology to construct our KB. We started by identifying the concepts needed for mapping the urban territory from HSR images. In Section 3.1, we detail the gap between these concepts and their identification in HSR images. Then, we describe in Section 3.2 the urban objects identifiable in such images. Finally, we present in Section 3.3 an implementation of the KB in a computer-usable form.

3.1. Step1: Identification of the concepts

A lot of land cover/land use terms exist, which represents the linguistic expression of the urban scene knowledge. Nevertheless, several terms correspond to urban objects which are not always identifiable on the images depending of their spatial resolution. In fact, there is a wide range of 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 BD-CARTO IGN$^{©}$ (defined for aerial photographs and SPOT images). All these nomenclatures built from MSR images are adapted to map urban areas from 1:100,000 to 1:50,000 (Table 1). A fourth level is commonly added by users to map urban area with a scale of 1:25,000 allowing for instance to specify the density of an urban fabric [3] (Table 2, left column). Nowadays, it is possible to extract urban objects (*e.g.* house, garden and road) from HSR images. This allows to map individual objects with their material (*e.g.* houses with orange tile roof) corresponding to a scale near of 1:5,000 (Table 2, right column).

In the domain of urban planning and management, some users also need to map the territory at the scale of the urban blocks (*i.e.* which can be defined as a minimal cycle closed by communication way) corresponding to a scale near of 1:10,000. In this case, there is no existing available land cover/land use product. The MSR images

Table 1: Extract of the Corine Land Cover Nomenclature used to map urban area

1:100,000 Level 1	**1:100,000** Level 2	**1:50,000** Level 3
1. Artificial surfaces		
	1.1. Urban fabric	1.1.1. Continuous urban fabric
		1.1.2. Discontinuous urban fabric
	1.2. Industrial, commercial and transport units	1.2.1. Industrial or commercial unit
		1.2.2. Road and rail networks
		1.2.3. Port areas
		1.2.4. Airports
	1.3. Mine, dump and construction sites	1.3.1. Mineral extraction sites
		. . .
	1.4. Artificial, non agricultural vegetated areas	1.4.1. Green urban areas
		1.4.2. Port and leisure facilities
2. Agricultural areas		
3. Forest and semi-natural areas		
4. Wetlands		
5. Waterbodies		
	5.1. Inland waters	5.1.1. Water courses
		5.1.2. Water bodies
	5.2. Marine waters	. . .

have a too coarse spatial resolution and HSR images have a too fine spatial resolution to map urban blocks. Thus, it is necessary to add an intermediate level (Table 2, middle column).

3.2. Step 2: Formalization of the concepts

These new urban concepts based on HSR images have to be translated into objects directly identifiable on images at this specific spatial resolution. These objects are called single object if one group of homogeneous pixels (referred here as region) is sufficient to identify one of them. For example, at a metric resolution, an object can correspond to a house, level 6 (Fig. 2(a)). Alternatively, it is called an aggregate object if several groups of homogeneous regions are necessary to identify it. For example, at a metric resolution, an aggregate object can correspond to an individual urban block, level 5, composed of houses, gardens, streets, etc. (Fig. 2(b)).

In this work, we described these objects using a dictionary, adapted from [31], which contains three categories of information:

1. Some characteristics to identify the objects: name, representation in a GIS database (*e.g.* point, polyline or polygon), type of object (*e.g.* single, aggregate), range of spatial resolution at which the object is identifiable ;
2. A qualitative description of the object (*e.g.* textual definition) ;
3. A list of relevant indicators used in the photo-interpretation domain to characterize these objects, classed by their relative importance respectively: color, shape, texture [28], context or spatial relationships (Table 3, left column)

These qualitative descriptors have to be translated into quantitative low descriptors according to the analyzed image. For example, the low-level descriptors associated to the color information depend on the radiometric reflectance

(a) Single objects, each corresponding to one group of pixels (houses).

(b) Aggregate object composed of some groups of homogeneous pixels (houses, gardens, road).

Figure 2: Single and aggregate objects from Quickbird image.

of the objects and on some indexes calculated from this reflectance. The table 3 presents the low-descriptors we used in this work.

Each object type is associated to an urban concept and each qualitative descriptor is associated with a low-level quantitative attribute. We focused our work on identify single object as it is the first step before considering trying to identify aggregate objects. We present in the following the implementation of the knowledge-base.

3.3. Step 3: Implementation of the knowledge-base

Experiments carried out at the 6th level shown that it is very difficult on the one hand, to define exactly the right range of accepted values for each attribute for each concept, and on the other hand, to extract objects from the HSR images. In fact, some concepts can be difficult to discriminate. For example, it is difficult to discriminate between house with orange tile roof and orange bare soil from tennis court or between water and shadow. The ability to discriminate forest and tree depends for example of the quality of the image segmentation. To address these problems, we built a KB which allows to generalize the urban concepts that are difficult to discriminate. To

Table 2: Extract of the taxonomy added to map urban area on MSR and HSR images

1:25,000 Level 4: Area level	1:10,000 Level 5: Block level	1:5,000 Level 6: Urban object level

• High-density urban fabric
• Low-density urban 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
	• Road
		• Building/roofs:
		  orange tile roof, ...
		  light gray residential roof,
		• Vegetation: green vegetation,
		  non-photosynthetic veg., ...
		• Transportation:
		  street, parking lots, ...
		• Water surfaces:
		  river, natural water bodies, ...
		• Bare soil
		• Shadow

Table 3: The descriptor classes and the low-level descriptors identified to characterize the regions

Class of descriptor	Numbers of low level desc.	Low-level descriptors
Color	4	Reflectance: Range of observed values in 4 spectral bands: - Blue (B) - Green (G) - Red (R) - near-infrared (NIR)
	2	Spectral index : Range of observed values of - Normalized Difference, Vegetation Index (NDVI) - Soil Brightness Index (BI)
Shape	5	Range of observed values of area, perimeter, elongation,
Texture	2	Range of observed values of the homogeneity index and the variance derived from the co-occurrence grey-levelmatrix [22]
Context	4	Relationships: Adjacency, inclusion, composition, neighborhood

build the hierarchy of concepts, we used the order of importance of the descriptors. The KB we created (Fig. 3) corresponds to the 6th level and is composed of 91 concepts. 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. To precisely describe how the hierarchy of concepts is built and used, let us introduce some notations and definitions [16].

Definition 1 (concept, sub-concept, depth). *Let* Θ *be the set of concepts,* $\preceq_\Theta$ *is a partial order between concepts.* $\forall(C_i, C_j) \in \Theta^2$, $C_i \preceq_\Theta C_j$ *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`. $\rho(C_i) = 5$ (see Fig. 3).

Definition 2 (specific attributes of a concept). *Let* $\mathcal{F}_\alpha(C)$ *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 $\mathcal{F}_\alpha(C)$. But an attribute overridden in C is present in $\mathcal{F}_\alpha(C)$.

Definition 3 (values and weight of an attribute). *Let* $a \in \mathcal{A}_\alpha$ *be an attribute of a class in* $\alpha \in \Phi$. *We define* $\mathcal{V}_C : \mathcal{A}_\alpha \to [\mathbb{R}; \mathbb{R}]$ *so that* $\mathcal{V}_C(a)$ *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 4 (set of regions). *Let* Γ *be the set of regions.*

Definition 5 (feature value of a region). *Let* $a \in \mathcal{A}$ *be a feature of a (segmented) region* $R \in \Gamma$. *We define* $\mathcal{V}'_R : \mathcal{A}_\alpha \to \mathbb{R}$ *so that* $\mathcal{V}'_R(a)$ *is the value of 'a' for the region* R.

The conception phase of the KB consisted in defining Θ, $\preceq_\Theta$, $\mathcal{F}_\alpha(C)$, $\omega(a, C)$, Φ and $\mathcal{V}_C(a)$. For all the concepts C, all the attribute values $\mathcal{V}_C(a)$ have to be provided by the expert or using learning algorithms. This allows to reduce the semantic gap between expert knowledge and image content. An example through the concept `Orange_House`, is described in table 4. Note that in general, it is rather difficult to draw knowledge from domain experts. The experts are rarely able to directly supply an explicit description of the knowledge they use for objects identification. In addition, acquiring knowledge this way is usually time consuming. This is a well-known problem within the artificial intelligence community. Thus, in order to ease the creation of the KB, we used machine learning techniques to automatically extract knowledge from the raw images. For example, to learn interpretable rules and build a reusable knowledge base, we used symbolic tools [35, 36]. This step was very important for the discussion with the experts, and helped to create the geographical KB content. The proposed KB was developed using Protégé [29], a free open-source software that provides tools to construct domain models and knowledge-based applications.

4. Knowledge-based region labeling

The proposed method which associates each region of an image to a concept of the KB (*i.e.* to assign a semantic label to each region) is composed of two main steps: the construction of the regions (Section 4.1) and the matching
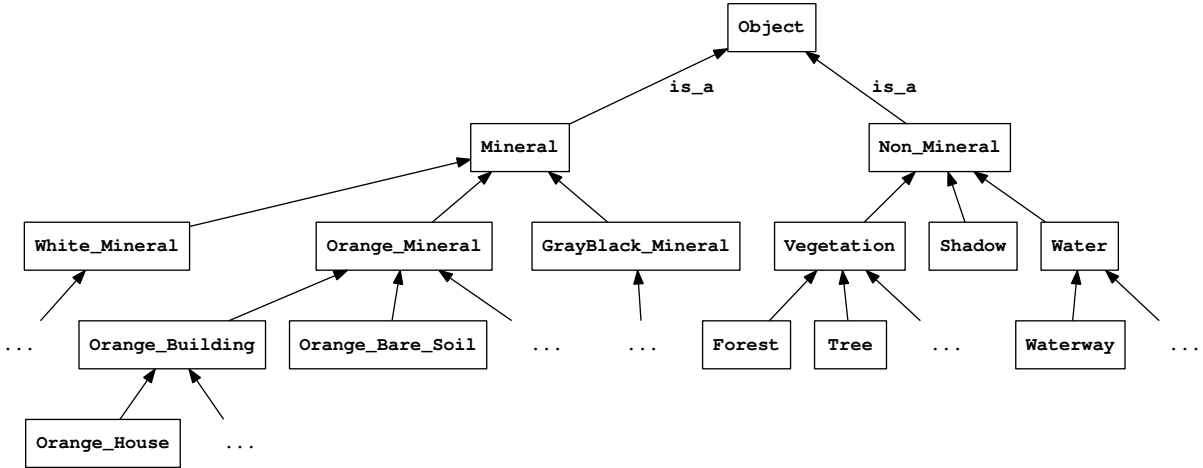
Figure 3: Excerpt of the hierarchy of concepts.

Table 4: Concept `Orange_House`.

Descriptor	Desc-associated attribute	Weight	Values	
			min	max
Color	Blue	1	21.7	62.3
	Green	1	19.4	80.1
	Red	1	29.7	135.1
	NearInfaRed	1	34.8	139
	NDVI	1	50.2	108
	SBI	0.5	14.6	60.1
Shape	diameter(m)	0.8	13	61
	area (m^2)	1	10	600
	perimeter (m)	1	28	116
	elongation (m)	0.6	1	3.1
	Miller index	0.5	0.5	0.8
	Solidity index	1	0.85	1

of the regions with the KB to assign a semantic to each region (Section 4.2).

4.1. Regions building using a segmentation algorithm

A segmentation algorithm is applied on the image in order to obtain a set of regions. A region is a set of connected and spectrally homogeneous pixels. The regions are then characterized by assigning a set of low-level descriptors to each of them. A numerical value is calculated for each attribute. It is important to note that any segmentation method can be used. However, this step is a critical point of the global identification method. Indeed, the quality of the produced segmentation is very important and is strongly linked to the quality of the identification process. This point is discussed in further details in Section 5.1.

4.2. Regions labeling using the knowledge base

The regions and their features are the inputs of the KB-based object recognition. The aim of this step is to find the concepts of the KB that best match the regions. To carry out this comparison, we defined a matching measure and a traversing method of the hierarchy of concepts.

Matching score. The proposed matching mechanism is a feature-oriented approach. It consists in check-ing the validity of feature values of the region, according to the properties and the constraints defined in the concepts. However, as a region does not have a semantic structure, we cannot directly use measures like MDSM [33], or other matching measures [34]. A region can be matched with any concepts and the features of a region allowing the matching are not identical according to the studied concept. For example, the concept `Orange_House` is defined by several indexes (*e.g.* elongation, shape, etc.) and spectral attributes, while the concept `Shadow` is only defined with spectral attributes. Without *a priori* knowledge, this asymmetry involves to compute all the features for each region, even if the majority of them will not be used by 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 descriptors was proposed. The measure computes the relevance of a matching and is composed of a local component and a global component (*i.e.* evaluating the pertinence in the hierarchy of concepts).

The *matching score* $Score(R, C_i)$ between a region R and a concept C_i is based on the definition of a local similarity measure, that evaluates the similarity between a region and a specific concept of the hierarchy. Each attribute of the concept is compared to the corresponding attribute calculated on the region.

Definition 6 (degree of validity). *Let $Valid(a, C, R)$ be the validity degree of an attribute 'a' between a region R and a concept C.*

$Valid(a, C, R)$ is equal to:

$$\begin{cases} 1 & if \ \mathcal{V}'_R(a) \in [min(\mathcal{V}_C(a)); max(\mathcal{V}_C(a))] \\ \\ \frac{\mathcal{V}'_R(a)}{min(\mathcal{V}_C(a))} & if \ \mathcal{V}'_R(a) < min(\mathcal{V}_C(a)) \\ \\ \frac{max(\mathcal{V}_C(a))}{\mathcal{V}'_R(a)} & if \ \mathcal{V}'_R(a) > max(\mathcal{V}_C(a)) \end{cases}$$

Algorithm 1 Traversing algorithm of the KB.

Input: a region R, a KB (Θ, Φ, $\mathcal{V}_C(a)$, ...), 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 $\{\mathcal{L}_\alpha(R), score\}$;

Definition 7 (local similarity). *Let $Sim_\alpha(R, C)$ be the local similarity between a region R and a concept C using the attributes of each class in α.*

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

Definition 8 (matching score). *Let $Score_\alpha(R, C)$ be the matching score between a region R and a concept C, and $\mathcal{P}(C)$ be the path starting from the root of the hierarchy and ending at the concept C. $\mathcal{P}(C) = \{C_j \mid C \preceq_\Theta ... \preceq_\Theta C_2 \preceq_\Theta C_1\}$.*

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

Traversing the hierarchy of concepts. To match a region with the KB, it is necessary to navigate in the hierarchy to find the best concept(s) for a region. A level-wise algorithm (Alg.1) was developed to navigate in the hierarchy of concepts using heuristics to reduce the search space: if the region matches the current concept, the algorithm will go deeper in the hierarchy; if the matching fails, the current concept is dropped and its sub-concepts will not

be explored. The $maxDepth$ value defines the exploration maximal depth (*i.e.* the degree of detail). The $minScore$ threshold is the minimal value of the matching score between a region and a concept to allocate the corresponding label to the region.

Definition 9 (labels identified for a region). *We define $\mathcal{L}_\alpha : \Gamma \to \Theta$ so that $\mathcal{L}_\alpha(R)$ is the set of concepts (seen as labels) identified for the region R according to the attributes of $\mathcal{A}_\alpha$ and the $minScore$ value.*

$$\mathcal{L}_\alpha(R) = \{C_i \mid \rho(C_i) \leq maxDepth \text{ and } Score_\alpha(R, C_i) \geq minScore \text{ and } \nexists C_j (\neq C_i) \ Score_\alpha(R, C_j) > Score_\alpha(R, C_i)\}$$

5. Experiments on remote sensing images

In order to illustrate how the knowledge-base can be used for automatic image interpretation, we carried out two series of experiments on three urban districts of Strasbourg (North-East of France) and on a district of Marseille (South of France) using Quickbird images (Digital Globe©). The Quickbird sensors produce two kind of images: panchromatic images with low spectral resolution but a high spatial resolution, and multispectral images with a good spectral resolution but a low spatial resolution. Consequently, each panchromatic image (at $0.61m$ spatial resolution) were merged with the multispectral image (at $2.44m$ spatial resolution), using the UWT-M2 method [32] to obtain an image at $0.61m$ spatial resolution with four spectral bands (Fig. 4).

All the districts are mainly composed of road (or parking), vegetation, water and small houses with gray or orange roofing tiles. Consequently, we focused our analysis to recognize the regions belonging to the concepts Vegetation, Water, Road, Orange_House and Gray_House of the KB. These concepts are the most relevant concepts to identify in urban areas according to geographer experts. For both cities, a set of manually labeled regions (*i.e.* ground truth) given by the expert were available. Some of the samples along with external information (*i.e.* topographic databases, expert knowledge, etc.) were used to fill the knowledge-base as described in the previous sections.

The Strasbourg knowledge-base was directly used for the labeling of the regions of the three districts of Strasbourg. For the district of Marseille, we performed the region labeling step by using three different knowledge-bases. Firstly, with the knowledge-base already used for the experiment on the Strasbourg districts. Secondly, with a knowledge-base created from information about the Marseille district. And finally, with a knowledge-base where the knowledge from Strasbourg and Marseille were merged. This experiment aimed at highlighting the ability of the KB to evolve and to leverage from new knowledge sources.

In the experiments, the segmentations of the images were computed using a supervised segmentation algorithm

(a) Strasbourg I

(b) Strasbourg II

(c) Strasbourg III

(d) Marseille

Figure 4: The districts used for the experiment extracted from Quickbird images.

[14]. This segmentation algorithm uses a fuzzy pixel classification using a k-Nearest Neighbor classifier and the watershed transform [38] applied on the fuzzy classification result.

5.1. Experiment on Strasbourg districts

The first step of the identification process is to segment the image in order to produce the set of regions to identify. An extract of the segmentation obtained from the District I and corresponding to the red square on figure 4(a) is presented on figure 5(b). For a better understanding of the scene, an aerial photography corresponding to the considered area is displayed on figure 5(c).

Once this segmentation produced, we used the KB to identify the regions. In order to evaluate if the results obtained using our method were in agreement with the ground truth given by the expert (Fig. 6(a)), we computed the *precision*, *recall* and the *F-measure* [42], accord-

(a) Raw Image

(b) Segmented Image.
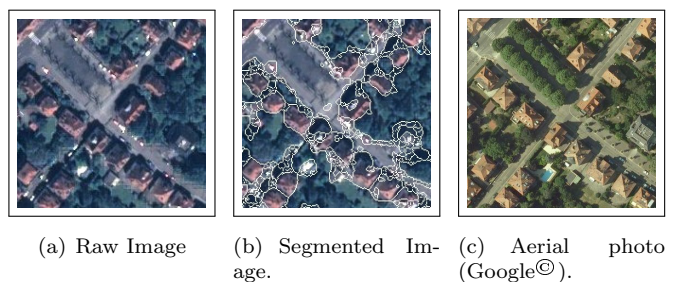
(c) Aerial photo (Google©).

Figure 5: Extract of the segmentation of the District I.

ing to different values of the *minScore* parameter (Section 4). For the F-measure, a value of 1 means that the result is in agreement with the ground truth. The table 5 presents the average values of the three indexes for the three dis-

8

Table 5: Assessment of the results of the identification according to the different Strasbourg districts.

District (Fig. 4)	minScore	Precision	Recall	F-Measure
Strasbourg I	0.75	0.859	**0.859**	**0.859**
	0.80	0.859	0.858	**0.859**
	0.85	0.861	0.857	**0.859**
	0.90	0.864	0.854	**0.859**
	0.95	0.876	0.837	0.856
	1.00	**0.881**	0.660	0.755
Strasbourg II	0.75	0.824	**0.824**	0.824
	0.80	0.826	**0.824**	0.825
	0.85	0.829	0.821	0.825
	0.90	0.836	0.816	**0.826**
	0.95	0.858	0.777	0.816
	1.00	**0.999**	0.533	0.695
Strasbourg IIII	0.75	0.862	**0.861**	**0.861**
	0.80	0.864	0.858	**0.861**
	0.85	0.864	0.855	0.860
	0.90	0.915	0.576	0.707
	0.95	0.956	0.164	0.281
	1.00	**1.000**	0.067	0.126

Table 6: Results according to different minScore values for District I.

index \\ class	Precision					
	1.00	0.95	0.90	0.85	0.80	0.75
Building	0.708	0.690	0.699	0.695	0.695	0.696
Vegetation	0.993	0.991	0.985	0.980	0.977	0.976
Road	0.850	0.843	0.832	0.826	0.824	0.823
Water	0.972	0.978	0.942	0.941	0.941	0.941

index \\ class	Recall					
	1.00	0.95	0.90	0.85	0.80	0.75
Building	0.595	0.620	0.675	0.690	0.694	0.695
Vegetation	0.969	0.973	0.976	0.976	0.976	0.976
Road	0.801	0.815	0.823	0.823	0.823	0.823
Water	0.276	0.940	0.941	0.941	0.941	0.941

index \\ class	F-Measure					
	1.00	0.95	0.90	0.85	0.80	0.75
Building	0.647	0.653	0.687	0.692	0.694	0.695
Vegetation	0.981	0.982	0.980	0.978	0.976	0.976
Road	0.825	0.829	0.827	0.824	0.823	0.823
Water	0.430	0.959	0.941	0.941	0.941	0.941

tricts of Strasbourg with in bold, the maximal value of each index. As the method is deterministic, with the same parameters set, two runs provide exactly the same results. From these results, one can see that the precision increases with the $minScore$ while the recall decreases. This result is consistent as the method tends to be more restrictive when the $minScore$ increases. This result means that the method identified a fewer number of regions but with a higher confidence. One can also observe from the results that the best value for $minScore$ is in $[0.75, 0.85]$ regardless to the image.

5.1.1. Detailed results for District I.

The table 6 presents the detailed results for the District I: for each concept, for different values of $minScore$, it presents the values of the indexes according to the ground truth. From this table, one can see that the concepts Vegetation and Water are very well identified, except when $minScore= 1$. In that case, the recall rate for the Water class is 0.276. It can easily be explained: when the expert defined the ground truth, he made one and only one area representing the river visible on the right of the image. In addition, reflections of the sun and turbulence in the water show the surface clearer than defined in the knowledge-base. The concept Road has good precision values and the recall values are acceptable. The precision values for the Building concept are relatively good but the recall values are very low.

The percentage of recognized objects and the percentage of the corresponding area in the image (*i.e.* the number of pixels from all the recognized objects) according to the $minScore$ values, are illustrated in figure 7. The curves show that a major part of the image is recognized, and thus labeled. With $minScore= 1$, 18.9% of the objects are recognized corresponding to 53.7% of the image area. With $minScore= 0.98$, 37.9% of the objects are identified and 66.2% of the image area. These results are promising: the majority of unlabeled objects correspond to small objects built from not properly segmented regions. The figure 6(b) shows the result of the identification of the District I with a $minScore$ of 0.98.

It is also important to note that some houses are not correctly segmented: the corresponding regions are sometimes composed of some pixels from shadow and vegetation. Thus, these houses could present features which do not correspond to the values defined in the knowledge-base, especially for the elongation indexes. Furthermore, very close buildings are sometimes grouped into only one single region and consequently, these regions cannot match with any concept of the knowledge-base. The opposite problem is encountered with the roads which are often over-segmented. In the following, we studied the influence of the segmentation step on the quality of the identification results.

5.1.2. Influence of the segmentation on identification results

As introduced previously, the identification results depends on the quality of the segmentation. Thus, in order to study and evaluate the influence of the segmentation, we carried out experiments on the extract of District I presented on figure 8(a) using five different segmentation approaches :

1. The watershed algorithm [38] (Fig. 8(b)) ;
2. A supervised segmentation algorithm [14] (Fig. 8(c)) ;
3. A supervised segmentation algorithm with manual corrections an expert made by splitting or merging regions (Fig 8(d)) ;

(a) Available ground truth.
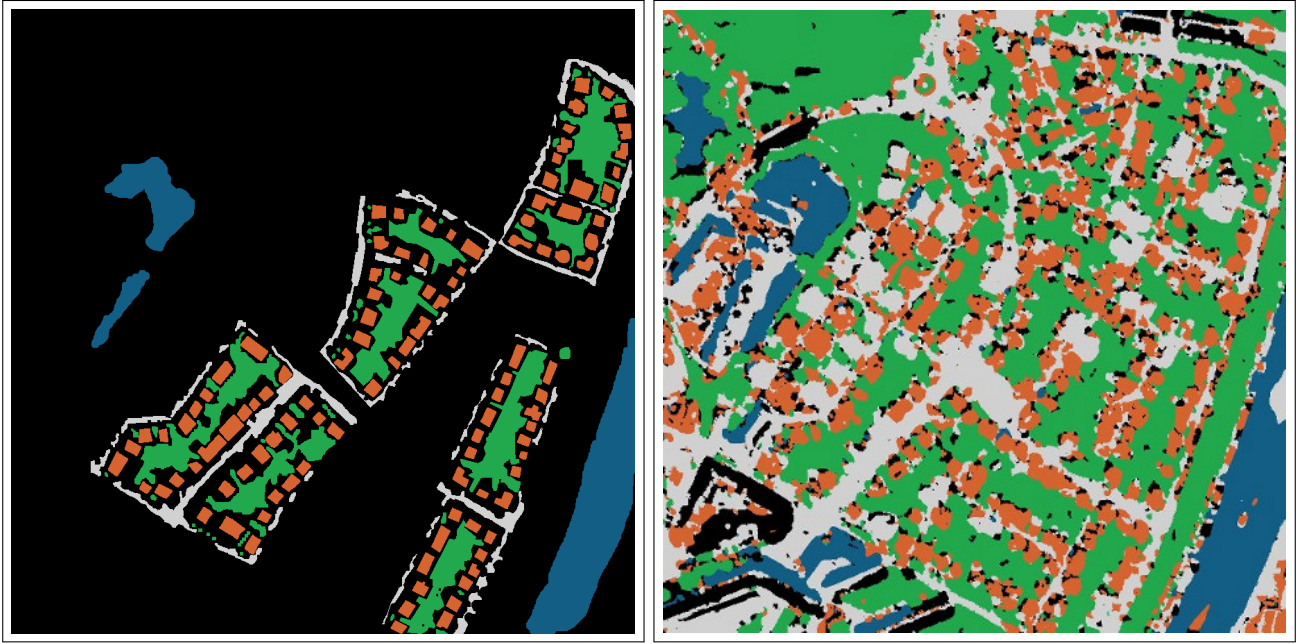
(b) Result of the identification.

| Road | Water | Unknown |
| Building | Vegetation | |

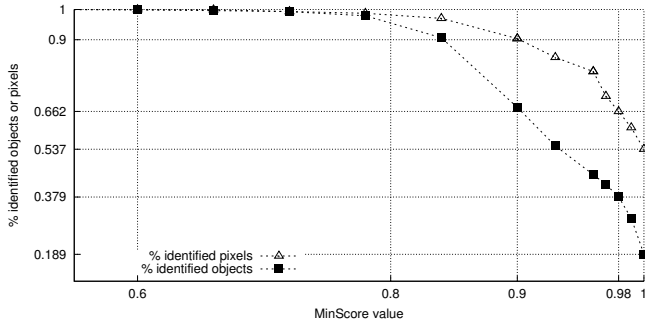Figure 6: Available ground truth and result identification of District I.

Figure 7: Percentage of labeled objects and pixels according to the *minScore* value.

4. The eCognition 5.0 software [2] (Fig 8(e)) (parameters: scale=50, color=0.7, shape=0.3);
5. The ENVI EX 4.8 software [3] (Fig 8(f)) (parameters: scale=30, merge=90).

The table 7 presents the results of the identification using the five segmentations according to different values of the *minScore* parameter. The best F-Measure value is obtained using the supervised segmentation algorithm with expert modifications, followed by the segmentation from ENVI EX, the supervised segmentation, the watershed, and finally the segmentation from eCognition. Note that the segmentations obtained using commercial softwares (eCognition and ENVI EX) offered good identification accuracies but low recalls when the *minScore* was high. This result means that few regions are identified but they are identified with a high degree of confidence. These results confirm that the better the segmentation, the better the results. Furthermore, our method turned out to be highly generic and the results were not as dependent of the segmentation as expected. Indeed, even with the over-segmented result proposed by the eCognition software, our identification method performed well. Finally, for all of the studied segmentations, the best F-measure value is obtained with a *minScore* value of 0.75 or 0.80 which means that the method is able to leverage from this parameter to soften the matching with the knowledgebase leading to a better identification. A trade-off between quality of the identification and the amount of recognized regions is thus easily obtained.

---

[2]http://www.ecognition.com/
[3]http://www.ittvis.com/

(a) Raw Image   (b) Watershed segmentation

(c) Supervised segmentation   (d) Supervised segmentation with manual corrections

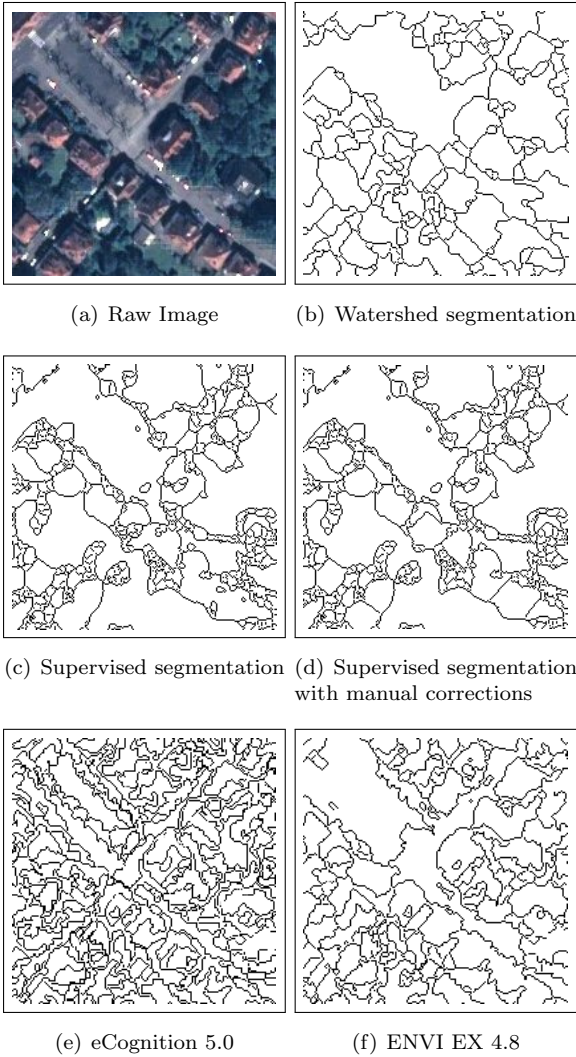(e) eCognition 5.0   (f) ENVI EX 4.8

Figure 8: Five segmentation extracts of the District I.

5.2. Experiment on Marseille district

In this section, experiments carried out on the district of Marseille are presented. The aim of these experiments is to highlight the ability of our approach to reuse the knowledge acquired from previous experience. Consequently, we used the knowledge acquired from the Strasbourg images to identify regions in the image of Marseille.

The ground truth provided by the expert for the Marseille district suffered of two problems. First, the number of examples was very low and, second, there was no example of the water class. To evaluate the ability of our approach to deal with these data, we first segmented the Marseille district using the samples from Strasbourg. Then, the regions were labeled firstly, with the KB already used for the experiment on the Strasbourg districts. Secondly, with a KB created from information about the Marseille district. And finally, with a KB where the knowledge from Strasbourg and Marseille were merged. The figure 9 presents the results obtained in that last experiment.

Table 7: Results from different segmentations.

Segmentation (Fig. 8)	$minScore$	Precision	Recall	F-Measure
Watershed	0.75	0.815	**0.815**	**0.815**
	0.80	0.815	0.814	**0.815**
	0.85	0.815	0.813	0.814
	0.90	0.834	0.758	0.794
	0.95	0.838	0.734	0.783
	1.00	**0.853**	0.538	0.660
Supervised segmentation	0.75	0.842	**0.842**	0.842
	0.80	0.842	0.841	0.841
	0.85	0.843	0.840	**0.842**
	0.90	0.847	0.836	0.841
	0.95	0.857	0.819	0.837
	1.00	**0.860**	0.642	0.735
Supervised segmentation with some user modifications	0.75	0.859	**0.859**	0.859
	0.80	0.859	0.858	**0.859**
	0.85	0.861	0.857	**0.859**
	0.90	0.864	0.854	**0.859**
	0.95	0.876	0.837	0.856
	1.00	**0.881**	0.660	0.755
eCognition 5.0	0.75	0.806	**0.815**	0.810
	0.80	0.805	0.813	0.809
	0.85	0.805	0.812	0.808
	0.90	0.813	0.681	0.741
	0.95	0.851	0.625	0.721
	1.00	**0.930**	0.319	0.475
ENVI EX 4.8	0.75	0.848	**0.861**	0.854
	0.80	0.848	0.860	**0.855**
	0.85	0.848	0.858	0.853
	0.90	0.855	0.818	0.837
	0.95	0.871	0.776	0.821
	1.00	**0.919**	0.546	0.685

The quality of the results were evaluated using the samples from Marseille. Note that, as there were no example of water in these examples, this class was not evaluated. The results presented in table 8 show that even if there is no knowledge available on the studied image, our approach can be used and rely on the knowledge acquired in the past. Indeed, by using only the knowledge extracted from Strasbourg images, we were able to identify regions in the Marseille image. Furthermore, the results show that when we enriched the KB with the Marseille knowledge, the detection performed even better. This result is consistent as the KB including also the Marseille knowledge had more information about regions extracted from Marseille district.

6. Conclusion

In this paper, the steps to build an urban knowledge-base applied to HSR image analysis were presented and a new knowledge representation was introduced. The approach is based on a domain-dependent knowledge-base developed by experts of the domain. A similarity measure and an exploration procedure of the knowledge-base were used in order to affect a semantic to the regions of a segmented image. The experimental results highlighted the effectiveness of the method, and the obtained results were compared using different segmentation approaches, including commercial softwares. The results also showed that even if there was no knowledge available on a studied area, our approach could be used and rely on the knowledge acquired in the past.

(a) Available ground truth.
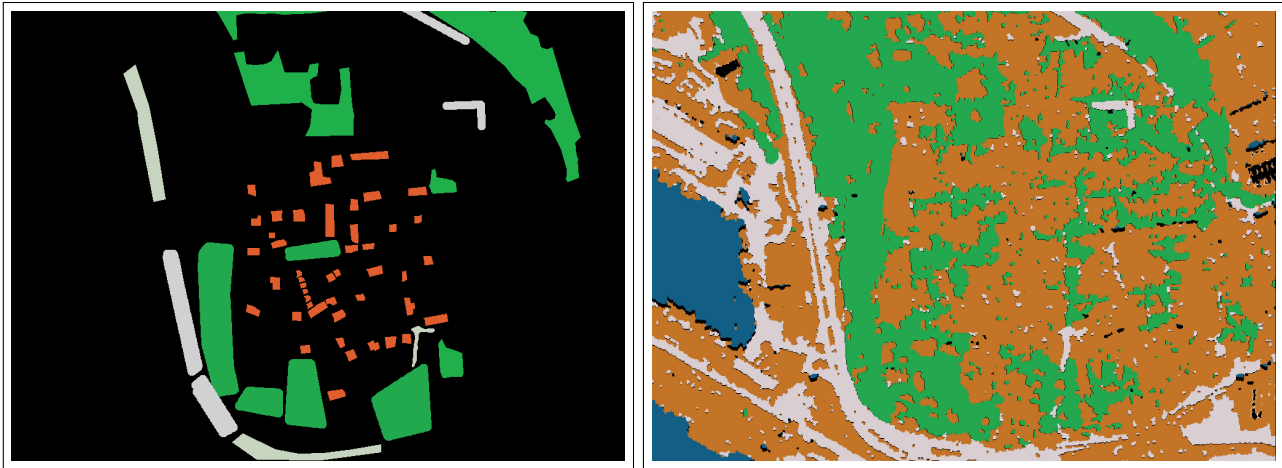
(b) Result of the identification.

| Road | Water | Unknown |
| Building | Vegetation | |

Figure 9: Results for the Marseille district using both knowledge-bases.

Table 8: Precision (Prec.), recall (Recall) et F-Measure (F-M) from results on Marseille.

KB used: index class	Marseille			Strasbourg			Both cities		
	Prec.	Recall	F-M	Prec.	Recall	F-M	Prec.	Recall	F-M
Building	0.353	0.892	0.506	0.589	0.751	0.660	0.600	0.963	0.739
Vegetation	0.976	0.725	0.832	0.995	0.873	0.930	0.988	0.905	0.945
Road	0.991	0.927	0.958	0.732	0.971	0.835	0.996	0.891	0.940
Means	0.774	0.848	0.809	0.772	0.865	0.816	0.861	0.919	0.889

In the future, we will plan several experiments on different types of urban images using other segmentation algorithms. We also wish to integrate the method into a framework of collaborative clustering. Indeed, in [19], collaborative clustering and the knowledge extracted from a knowledge-base were used together. Furthermore, we also plan on using directly the knowledge during the segmentation step [18] in order to build regions easily identifiable. In order to improve and to enrich the content of the knowledge-base, machine learning techniques continues to be developed in order to automatically extract information from the HSR images. In particular, we are focusing on topological relations based on the RCC-8 (Region Connection Calculus) theory.

References

[1] Athanasiadias, T., Mylonas, P., & Avrithis, Y. (2007). Semantic image segmentation and object labeling. *IEEE Trans. on Circuits and Systems for Video Technology*, *17*, 298–312.

[2] Athanasiadis, T., Mylonas, P., Avrithis, Y., & Kollias, S. (2007). Semantic image segmentation and object labeling. *Circuits and Systems for Video Technology, IEEE Transactions on*, *17*, 298–312.

[3] Autran, J. (2007). Extension de la nomenclature corine land cover pour la description de l'occupation du sol urbain à grande échelle. In *Journée francophone sur les ontologies*.

[4] Baltasavias, E. (2004). 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*, 129–151.

[5] Benz, U., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for gis-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, *58*, 239–258.

[6] Breen, C., Khan, L., & Ponnusamy, A. (2002). Image classification using neural networks and ontologies. In *Proc. 13th International Workshop on Database and Expert Systems Applications, co-located with DEXA 2002* (pp. 98–102). Aix-en-Provence, France.

[7] Bückner, J., Pahl, M., Stahlhut, O., & Liedtke, C. (2001). Geoaida a knowledge based automatic image data analyser for remote sensing data. In *ICSC Congress on Computational Intelligence Methods and Applications*. Citeseer.

[8] Bückner, J., Pahl, M., Stahlhut, O., & Liedtke, C.-E. (2002). A knowledge-based system for context dependent evaluation of remote sensing data. In L. J. V. Gool (Ed.), *DAGM-Symposium* (pp. 58–65). Springer volume 2449 of *Lecture Notes in Com-*

puter Science.

[9] Castellano, G., Fanelli, A. M., & Torsello, M. A. (2011). Fuzzy image labeling by partially supervised shape clustering. In A. Knig, A. Dengel, K. Hinkelmann, K. Kise, R. J. Howlett, & L. C. Jain (Eds.), *KES (2)* (pp. 84–93). Springer volume 6882 of *Lecture Notes in Computer Science*.

[10] Cataldo, A., & Rinaldi, A. M. (2010). An ontological approach to represent knowledge in territorial planning science. *Computers, Environment and Urban Systems*, *34*, 117 – 132.

[11] Cleve, C., Kelly, F. R., M. abd Kearns, & Moritz, M. (2008). Classification of the wildland-urban interface: A comparison of pixel- and object-based classifications using high resolution aerial photography. *Computers, Environment and Urban Systems*, .

[12] Crevier, D., & Lepage, R. (1997). Knowledge-based image understanding systems: a survey. *Computer Vision and Image Understanding*, *67*, 161–185.

[13] Dasiopoulou, S., Mezaris, V., Kompatsiaris, I., Papastathis, V. K., & Strintzis, M. G. (2005). Knowledge-assisted semantic video object detection. *IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Analysis and Understanding for Video Adaptation*, *15*, 1210–1224.

[14] Derivaux, S., Forestier, G., Wemmert, C., & Lefévre, S. (2010). Supervised image segmentation using watershed transform, fuzzy classification and evolutionary computation. *Pattern Recognition Letters*, *31*, 2364 – 2374.

[15] Draper, B., Collins, A., Brolio, J., Hanson, A., & Riseman, E. (1989). The schema system. *International Journal of Computer Vision*, *2*, 209–250.

[16] Durand, N., Derivaux, S., Forestier, G., Wemmert, C., Gançarski, P., Boussaid, O., & Puissant, A. (2007). Ontology-based object recognition for remote sensing image interpretation. In *IEEE International Conference on Tools with Artificial Intelligence* (pp. 472–479). Patras, Greece: IEEE Computer Society volume 1.

[17] Fonseca, F., Egenhofer, M., Agouris, P., & Camara, G. (2002). Using ontologies for integrated geographic information systems. *Transactions in GIS*, .

[18] Forestier, G., Derivaux, S., Wemmert, C., & Gançarski, P. (2008). An evolutionary approach for ontology driven image interpretation. In *Tenth European Workshop on Evolutionary Computation in Image Analysis and Signal Processing* (pp. 295–304). Napoli, Italy: Springer volume 4974 of *Lecture Notes in Computer Sciences*.

[19] Forestier, G., Wemmert, C., & Gançarski, P. (2008). On combining unsupervised classification and ontology knowledge. In *IEEE Geoscience and Remote Sensing Symposium*. Boston, Massachusetts.

[20] Gomez, F., & Segami, C. (2007). Semantic interpretation and knowledge extraction. *Knowledge-Based Systems*, *20*, 51 – 60.

[21] Gruber, T. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human Computer Studies*, *43*, 907–928.

[22] Haralick, R. (1979). Statistical and structural approaches to texture. *Proceeding of the IEEE*, *67*, 45–69.

[23] Herold, M., Scepan, J., Muller, A., & Gunter, S. (2002). Object-oriented mapping and analysis of urban landuse/cover using ikonos data. In *Proc. 22nd Earsel Symposium Geoinformation for European-Wide Integration* (pp. 531–538). Prague.

[24] Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2003). *Remote Sensing and Image Interpretation*. Wiley.

[25] Maillot, N., & Thonnat, M. (2008). Ontology based complex object recognition. *Image and Vision Computing*, *26*, 102–113.

[26] Matsuyama, T., & Hwang, V.-S. (1990). *SIGMA - A Knowledge-Based Aerial Image Understanding System*.

[27] Moller-Jensen, L. (1997). Classification of urban land cover based on expert systems, object models and texture. *Computers, Environment and Urban Systems*, *21*, 291 – 302.

[28] Momm, H., Easson, G., & Kuszmaul, J. (2009). Evaluation of the use of spectral and textural information by an evolutionary algorithm for multi-spectral imagery classification. *Computers, Environment and Urban Systems*, *33*, 463 – 471.

[29] Noy, N. F., Fergerson, R. W., & Musen, M. A. (2000). The knowledge model of protege-2000: Combining interoperability and flexibility. In *Proc. 12th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2000)* (pp. 17–32). Juan-les-Pins, France.

[30] Panagi, P., Dasiopoulou, S., Papadopoulos, G. T., Kompatsiaris, I., & Strintzis, M. G. (2006). A genetic algorithm approach to ontology-driven semantic image analysis. In *Proc. 3rd IEEE International Conference of Visual Information Engineering (VIE 2006)* (pp. 132–137). Bangalore, India.

[31] Pantazis, D., Cornelis, B., Billen, R., & Sheeren, D. (2002). Establishment of a geographic data dictionnary : a case study of urbis 2, the brussels regional government gis. *Computers, Environment and Urban Systems*, *26*, 3–17.

[32] Puissant, A., Ranchin, T., Weber, C., & Serradj, A. (2003). Fusion of quickbird ms and pan data for urban studies. In *Proc. European Association of Remote Sensing Laboratories Symposium (EARSeL)* (pp. 77–83). Gent, Belgium.

[33] Rodriguez, M. A., & Egenhofer, M. J. (2003). Determining semantic similarity among entity classes from different ontologies. *IEEE Transactions on Knowledge and Data Engineering*, *15*, 442–456.

[34] Schwering, A., & Raubal, M. (2005). Measuring semantic similarity between geospatial conceptual regions. In *Proc. 1st International Conference on GeoSpatial Semantics (GeoS)* (pp. 90–106). Mexico City, Mexico volume 3799 of *Lecture Notes in Computer Science*.

[35] Sheeren, D., Puissant, A., Weber, C., Gançarski, P., & Wemmert, C. (2006). 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*. CDROM 9p, Berlin.

[36] Sheeren, D., Quirin, A., Puissant, A., Gançarski, P., & Weber, C. (2006). Discovering rules with genetic algorithms to classify urban remotely sensed data. In *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS'2006)* (pp. 3919 – 3922).

[37] Smeulders, A., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*, 1349–1380.

[38] Soille, P. (2003). *Morphological Image Analysis*. 2nd ed. Springer-Verlag.

[39] Tiinjes, R., Glowe, S., Biicknel, J., & Lledtke, C. (1999). Knowledge-based interpretation of remote sensing lmages using semantic nets. *Photogrammetric Engineering & Remote Sensing*, *65*, 811–821.

[40] Uitermark, H., van Oosterom, P., Mars, N. J. I., & Molenaar, M. (1999). Ontology-based geographic data set integration. In *Spatio-Temporal Database Management* (pp. 60–78). Springer volume 1678 of *Lecture Notes in Computer Science*.

[41] Uschold, M., & King, M. (1995). Towards a methodology for building ontologies. In *Workshop on Basic Ontological Issues in Knowledge Sharing*. Montreal, Canada.

[42] van Rijsbergen, C. (1979). *Information Retrieval*. London, Butterworths.

[43] Waterson, A., & Preece, A. (1999). Verifying ontological commitment in knowledge-based systems. *Knowledge-Based Systems*, *12*, 45–54.

[44] Wu, H., Li, Y., Li, Q., & Chen, X. (2009). Research on fractal model of urban land use considering the appropriate spatial resolution for remote sensing imagery. In *MIPPR 2009: Remote Sensing and GIS Data Processing and Other Applications* (p. 749816). International Society for Optics and Photonics volume 7498.

[45] Zlatoff, N., Tellez, B., & Baskurt, A. (2004). Image understanding and scene models: a generic framework integrating domain knowledge and gestalt theory. In *Proc. IEEE International Conference on Image Processing (ICIP 2004)* (pp. 2355–2358). Singapore.