

# Ontology-based Object Recognition for Remote Sensing Image Interpretation

Nicolas Durand, Sébastien Derivaux, Germain Forestier, Cédric Wemmert, Pierre Gançarski  
LSIIT (CNRS UMR 7005), Université Louis Pasteur  
Bd Sébastien Brant,  
67400 Illkirch, France  
{durand,derivaux,forestier,wemmert,gancarski}@lsiit.u-strasbg.fr

Omar Boussaïd  
ERIC (EA 3083), Université Lumière  
5 avenue Pierre Mendès-France  
69676 Bron, France  
omar.boussaid@univ-lyon2.fr

Anne Puissant  
IDEES-GEOSYSCOM (CNRS UMR 6228)  
Université de Caen, Esplanade de la paix  
14032 Caen, France  
anne.puissant@unicaen.fr

## Abstract

*The multiplication of Very High Resolution (spatial or spectral) remote sensing images appears to be an opportunity to identify objects in urban and periurban areas. The classification methods applied in the object-oriented image analysis approach could be based on the use of domain knowledge. A major issue in these approaches is domain knowledge formalization and exploitation. In this paper, we propose a recognition method based on an ontology which has been developed by experts of the domain. In order to give objects a semantic meaning, we have developed a matching process between an object and the concepts of the ontology. Experiments are made on a Quickbird image. The quality of the results shows the effectiveness of the proposed method.*

## 1. Introduction

Image interpretation is a difficult problem which can be defined as the automatic extraction of semantic data from an image. However, these semantic data are not always explicitly in the image and depend on domain knowledge and the context. In the document indexation and analysis field, the use of ontologies is usual. Nevertheless, in the image interpretation domain, a problem arises because the perceptual response (pixels or voxels) is not at the same level than the semantic interpretation. This problem is called *semantic gap* and is defined as the lack of concordance between low-level information which can be automatically extracted from images and the interpretation by a user in a given situa-

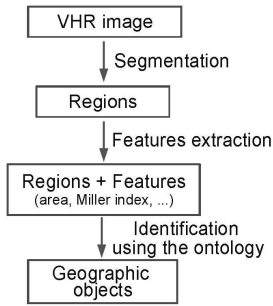
tion [22]. Thus, a lot of approaches use ontologies to introduce a semantic point of view and some knowledge in order to reduce the semantic gap. An ontology [8] is a specification of an abstract, simplified view of the world represented for some purpose. Ontology defines a set of concepts, their characteristics and their relations to each other. These definitions allow to describe and to use reasoning on the studied domain.

Since the appearance of VHR-images (Very High Resolution), the current tendency is the development of object-oriented methods for image analysis [2] [9]. These approaches involve segmenting images into homogeneous regions which are then characterized by a set of features related to spectral signatures, geographical features such as shape and length, and topological properties such as adjacency, inclusion, etc. These features are used in the recognition and the classification process. While there are some studies comparing object-oriented and pixel-based classification techniques [17] [25], few works focus on the development and the use of the domain knowledge for identifying and classifying urban areas. The main difficulties for this task are to define, for instance, ontologies taking into account the expert knowledge, and how to use them.

In this paper, we propose an object recognition method based on an ontology built by experts (see Figure 1). This work is a part of the FoDoMuSt project<sup>1</sup>. The input of the method is a segmented image. The regions are characterized by features related to the spectral, spatial and contextual properties. These features are used in the recognition procedure. In order to allocate to each region a semantic meaning, we have developed a matching process between

---

<sup>1</sup><http://lsiit.u-strasbg.fr/afd/sites/fodomust/>



**Figure 1. Treatment steps.**

an object and the concepts of the ontology. An original matching measure and a traversing procedure of the ontology are presented.

The paper is organized as follows. Section 2 gives an overview of existing approaches. The object features are presented in Section 3. The ontology and the proposed object-recognition method are detailed respectively in Section 4 and 5. Section 6 presents the image segmentation algorithm used in this work and some experiments on a VHR-image of a district of Strasbourg (France). We finally conclude in Section 7.

## 2. Related work

Knowledge based vision systems have proved to be effective for complex object recognition [11] and for scene understanding [13]. For instance SIGMA [13] or Schema [7] perform image understanding tasks on aerial images, based on several descriptions of objects which are bound to appear. These systems give access to a high semantic level. Nevertheless, as pointed by [4] such systems are strongly domain-dependent as they integrate prior knowledge about the scene in the algorithms of image understanding. In fact, the domain knowledge is not clearly separated from the procedure. Moreover, they rely on knowledge bases which are difficult to produce.

Recent works use ontologies to describe clearly knowledge of the studied domain. In [26], spatial relations between concepts (inclusion, adjacency, neighborhood, left, above, ...) are used to merge regions and to recognize objects in thessalian graves images. The exclusive use of spatial relations seems to be not relevant to analyse remote sensing images. Moreover, the definition of these relations is very difficult. Nevertheless, this work separates well domain knowledge from procedures. We claim that it is a very important point.

Maillot [12] proposes an ontology-based object learning and recognition system for image retrieval. An interesting point is the separation of the local matching to the global matching. The global score combines the probabilities com-

puted during the local matching. The descriptors use for the matching correspond to “visual concepts” which are learnt during the learning phase. The matching function is dependent of these visual concepts. We think that the global matching should take into account the hierarchy of the ontology.

Dasiopoulou [5] proposes an ontology-based object detection in videos. This work only corresponds to a segmentation process. Breen [3] uses a neural network to classify objects in pre-defined output categories. Then, the proposed system determines if the image may be associated to a concept from an ontology. In [15], Panagi proposes a genetic algorithm to ontology-driven semantic image analysis. Some low-level descriptors are extracted from the image and are used in a matching process between regions of the image and concepts of an ontology. A set of hypothesis (region; list of possible concepts with their degrees of confidence) are then passed in a genetic algorithm to determine the optimal image interpretation. But only spatial relations (8 directional relations) are used by the system.

## 3. Features extraction

In our approach, after an image segmentation, each region is characterized by means of a set of features. These features were selected by the experts in order to allow object recognition. The spectral reflectance of pixels composing the objects is the first criterion we used. Several features were retained: the mean spectral value of the objects in the four bands of the image (red, green, blue and near infrared), the mean value of the NDVI index (Normalized Difference Vegetation Index), and the mean value of the SBI index (Soil Brightness Index) [10]. However, spectral signatures are not sufficient to determine, for example, the functional character of the buildings in urban environment. The corresponding classes have the same spectral values and therefore overlap in the feature space. Spatial and contextual information are more relevant to separate residential buildings, collective buildings, industrial and commercial buildings, and continuous built-up areas. Thus, we also compute this kind of features to recognize the building objects. Several shape properties were selected: area, elongation, diameter (length of the major axis), compactness (Miller index), solidity (ratio of the area to the convex hull area), etc. The spatial features computation is based on Geoxygene<sup>2</sup> developed by the French National Geographic Institute (IGN).

<sup>2</sup><http://oxygene-project.sourceforge.net>

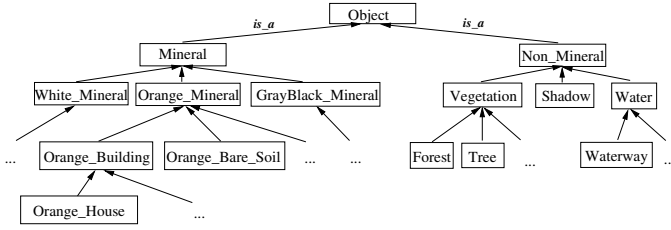


Figure 2. Excerpt of the ontology.

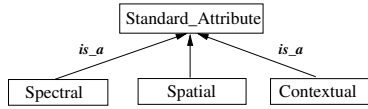


Figure 3. Hierarchy of the attribute classes.

#### 4. Geographical ontology and knowledge formalization

An ontology [8] models a domain in a formal way. It defines a set of concepts (buildings, water, etc), their characteristics and their relations to each other.

In our case, each concept has a label (e.g. Orange\_House) and is defined by some attributes (corresponding to low-level descriptors) associated to an interval of accepted values. Each attribute of a concept is weighted according to its importance to recognize the geographic object represented by the concept. Let us note that we use “features” instead of “attributes” when we talk about regions.

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 in this way takes usually a long time. This is a well-known problem within the artificial intelligence community. So, in order to facilitate the creation of the ontology, we used machine learning techniques to extract knowledge automatically from the raster data (raw images). To learn interpretable rules and build a reusable knowledge base, we used symbolic supervised machine learning tools [20] [21]. This step is very important for the discussion with the experts, and have made easier the work of the experts to improve and to finalize the geographical ontology.

The ontology (see Figure 2) is composed of 91 concepts, 20 attributes (in total) and 66 leaf concepts. The depth of the ontological tree is 6. The development of the ontology have been realized with Protégé 2000 [14]. Let us note that the contextual attributes are not defined yet in the current ontology.

In the rest of the paper, we use the following knowledge formalization.

Table 1. Concept “Orange\_House”.

Class	Attribute name	Values
spectral	spectral_signature_Blue	[21.7-62.3]
	spectral_signature_Green	[19.4-80.1]
	spectral_signature_Red	[29.7-135.1]
	spectral_signature_NearInfraRed	[34.8-139]
	spectral_signature_SBI	[14.6-60.1]
	spectral_signature_NDVI	[50.2-108]
spatial	diameter(m)	[13-61]
	area (m <sup>2</sup> )	[10-600]
	perimeter (m)	[28-116]
	elongation (m)	[1-3.1]
	Miller index	[0.5-0.8]
	Solidity index	[0.85-1]

**Definition 1 (concept, sub-concept, depth)** Let  $\Theta$  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,  $C_i = \text{Orange\_House}$  is a sub-concept of  $C_j = \text{Orange\_Building}$ .  $\rho(C_i) = 5$  (see Figure 2).

**Definition 2 (classes of attributes)** Let  $\Phi$  be the set of attribute classes.  $\mathcal{A}$  is the set of all attributes.  $\mathcal{A} = \{\text{spectral\_signature\_Blue}, \dots, \text{area}, \dots, \text{Miller index}, \dots\}$ . For a set of attribute classes  $\alpha \subseteq \Phi$ ,  $\mathcal{A}_{\alpha} \subseteq \mathcal{A}$  is the set of attributes of each class in  $\alpha$ .

For instance,  $\Phi = \{\text{spectral}, \text{spatial}, \text{contextual}\}$  (see Figure 3).

If  $\alpha = \{\text{spectral}\}$ ,  $\mathcal{A}_{\alpha} = \{\text{spectral\_signature\_Blue}, \text{spectral\_signature\_Green}, \text{spectral\_signature\_Red}, \text{spectral\_signature\_NearInfraRed}, \text{spectral\_signature\_SBI}, \text{spectral\_signature\_NDVI}\}$ .  
If  $\alpha = \Phi$ ,  $\mathcal{A}_{\alpha} = \mathcal{A}$  (all the attributes).

**Definition 3 (specific attributes of a concept)** Let  $\mathcal{F}_{\alpha}(C)$  be the set of attributes of the classes in  $\alpha$ , specifically associated with the concept  $C \in \Theta$ .

For instance, for the concept  $C = \text{Orange\_House}$ , if the spectral attributes ( $\text{spectral\_signature\_Blue}, \dots$ ) and their values are inherited by the  $\text{Orange\_Building}$ , they are not present in  $\mathcal{F}_{\alpha}(C)$ . But an attribute overridden in  $C$  is present in  $\mathcal{F}_{\alpha}(C)$ .

**Definition 4 (values and weight of an attribute)** Let  $a \in \mathcal{A}_{\alpha}$  be an attribute of a class in  $\alpha \in \Phi$ . We define  $\mathcal{V}_C : \mathcal{A}_{\alpha} \rightarrow [\mathbb{R}; \mathbb{R}]$  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 5 (set of regions)** Let  $\Gamma$  be the set of regions.

**Definition 6 (feature value of a region)** Let  $a \in \mathcal{A}$  be a feature of a (segmented) region  $R \in \Gamma$ . We define  $\mathcal{V}'_R : \mathcal{A}_\alpha \rightarrow \mathbb{R}$  so that  $\mathcal{V}'_R(a)$  is the value of 'a' for the region  $R$ .

Conception phase of the ontology consists of defining  $\Theta$ ,  $\preceq_\Theta$ ,  $\mathcal{F}_\alpha(C)$ ,  $\omega(a, C)$ ,  $\Phi$  and  $\mathcal{V}_C(a)$ . This allows to reduce the semantic gap between expert knowledge and image level.

## 5. Ontology-based object recognition

The regions and their features are the input the ontology-based object recognition. Our method consists of matching each region with the concepts of the ontology. We have defined a matching measure and a traversing method of the ontology.

### 5.1 Matching score

The proposed matching method is a “feature-oriented” approach. It corresponds to verify the validity of feature values of the region according to the properties and the constraints defined in the concepts. Nevertheless, a region does not have a semantic structure and thus we cannot directly use the measures like MDSM [18] or the measures presented in [24] and [19]. A region can be matched a priori with any concepts. 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 a lot of indexes (elongation, Miller, ...) and some spectral attributes, while the concept `Shadow` is only defined with spectral attributes. Without a priori knowledge, this dissymmetry involves to compute all the features for each region, even if the majority of these features will be unused for the matching process with a concept. In order to take into account all these specificities, we propose a matching measure based on a distance between the extracted features of a region and the values of the attributes of the concepts of the ontology. Being inspired by the literature, our measure computes the relevance of a matching and is composed of a local component (dealing with the inner properties of the concept) and a global component (evaluating the pertinence in the hierarchy of concepts).

#### 5.1.1 Local similarity

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  $C$ .

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

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

**Definition 8 (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  $\alpha$ .

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

#### 5.1.2 Global score

The matching score (global score) 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  $\mathcal{P}(C)$  be the path starting from the root of the ontology and ending at the concept  $C$ .  $\mathcal{P}(C) = \{C_j \mid C \preceq_\Theta \dots \preceq_\Theta C_2 \preceq_\Theta C_1\}$ .

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

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, we integrate a specialization coefficient  $\rho$  equals to the depth of the concepts. In this way, the measure favours the specialization of the concepts, considering all additional information give a new semantic meaning.

From another point of view, the measure evaluates the pertinence of a region to be an instance of a concept.

Let us take an example (concept in Table 1), for a region  $R = \{\text{spectral\_signature\_Blue} = 40.2 ; \text{spectral\_signature\_Green} = 32.3 ; \text{spectral\_signature\_Red} = 50.1 ; \text{spectral\_signature\_NearInfraRed} = 66 ; \text{spectral\_signature\_SBI} = 58.6 ; \text{spectral\_signature\_NDVI} = 79.9 ; \text{area} = 120 ; \dots ; \text{elongation} = 3.5 ; \dots\}$  and  $\alpha = \{\text{spectral}\}$ , the matching score is  $Score_\alpha(R, \text{Orange\_House}) = 1$ . So `Orange_House` is a possible label for the region  $R$ .

## 5.2 Traversing the ontology

The matching score between a region and a concept has been defined. It remains to traverse the ontology to find the best concept(s) according to the score for a region. We have developed an level-wise algorithm to traverse the ontology using heuristics to reduce the search space. The parameters can be accorded to a specific matching context.

The general sketch of the exploration is: if the region matches the current concept, the algorithm will go deeper in the hierarchy defined by the partial order  $\preceq_{\Theta}$  in the next level. If matching fails, the current concept is dropped and its sub-concepts will not be explored.

The main heuristic corresponds to the selection of the best concepts at each level in order to prune some branches which the starting concept would not be relevant (with a poor matching score value). This strategy is based on the property that an internal concept is more general than its children.

We have also defined two thresholds. Let  $maxDepth$  be the exploration maximal depth (i.e. the degree of detail) and  $minScore$  be the minimal value of the matching score between a region and a concept to allocate the corresponding label to the region. For example, 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  $\mathcal{L}_{\alpha} : \Gamma \rightarrow \Theta$  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.

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

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

## 6. Experiments

The proposed method have been evaluated using a Quickbird image of a urban district of Strasbourg, France. The sensors of the Quickbird satellite return one panchromatic channel (resolution of 0.70 meter) and three spectral channels (resolution 2.8 meters). The panchromatic channel has been merged with the other channels using a UWT-M2 method studied in [16]. Finally, we have obtained an image with a spatial resolution of 0.70 meter and four channels (blue, green, red and near infra-red). The size of the image

---

### Algorithm 1 Traversing algorithm of the ontology.

---

**Input:** a region  $R$ , an ontology  $(\Theta, \Phi, \mathcal{V}_C(a), \dots)$ , a set of attribute classes  $(\alpha)$ ,  $maxDepth$  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
         $\mathcal{RC} = \mathcal{RC} \cup \{C_i \mid C_i \preceq_{\Theta} C_j\}$ ;
    end for
    depth ++;
end while
return  $\{\mathcal{L}_{\alpha}(R), score\}$ ;

```

---



Figure 4. A district of Strasbourg (France).



Figure 5. Excerpt of the segmented image.

is of 900x900 pixels. Figure 4<sup>3</sup> corresponds to the channels red, green and blue of the image.

In order to make easy the evaluations, we have used an excerpt of the ontology. The number of concepts (labels) has been limited. In this way, the work of the expert has been less laborious to labeled manually the reference image. We have only tried to recognize *Vegetation*, *Water*, *Road* and *Orange\_House* (personal houses with orange roofing tiles). Let us note that if no label is found for a region, the system uses the label *Unknown*. The evaluation has consisted in comparing the results of the method with the manually labeled regions.

### 6.1. Image segmentation

For these experiments, the segmentation has been built using a supervised segmentation algorithm [6]. This algorithm consists of two steps. In the first step, pixel samples are used to compute a fuzzy classification (k-Nearest Neighbor) [1]. We have used three classes: road, building and vegetation. The second step corresponds to apply a segmentation algorithm, namely the watershed transform [23], on the fuzzy classification result. An example of the segmentation result is shown in Figure 5.

Currently, no perfect segmentation methods exist. Let us remark that the used segmentation suffers from two problems. First, a lot of small regions are generated. The second problem is that very close buildings are grouped into one region. These regions have some geometrical indexes “not in line” with characteristics in the ontology for the building concept. Let us note that the proposed method is not as-

<sup>3</sup>All the images of the paper are available in color at: <http://dpt-info.u-strasbg.fr/~durand/obor/>.

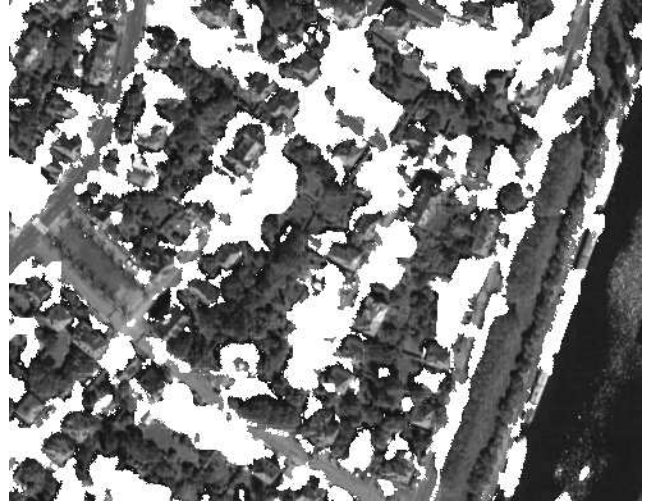


Figure 6. Excerpt of the recognized objects set (unknown objects in bright).

Table 2. Results.

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

sociated to a segmentation algorithm. Other segmentation algorithms could be used.

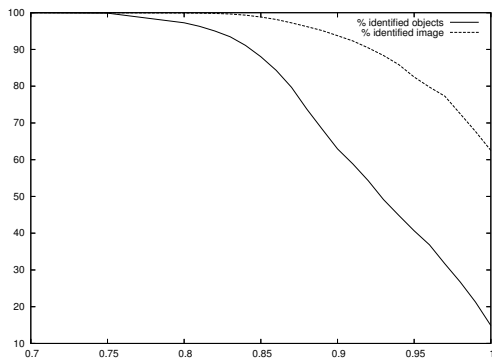
### 6.2. Object recognition results

Table 2 presents the average values of precision, recall and F-measure (in overall), according to the *minScore* value (see Section 5). The best F-measure value is obtained with *minScore*=0.98. This is due to the decrease of the recall value which is more important than the increase of the precision value for *minScore*=1. Nevertheless, in our case, we consider that a very good precision value is more important than a good recall value.

Table 3 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 are correct. The precision values for *Orange\_House* are relatively good but the recall values are too low. This can be explained. Let us recall that the results strongly depend on the quality of the segmentation. As said in Section 6.1 some houses are not correctly segmented and they are merged with other houses. Thus, these houses could present features which do not correspond to the values defined in the ontology, especially for the elongation

**Table 3. 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



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

indexes. We face a similar problem with the road which is over-segmented.

Contrary to what we could expect, a large part of the image has been recognized. Figure 7 presents 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. For the maximal value (*minScore* = 1), 14.8% of the objects are recognized but they correspond to 62.5% of the image. With *minScore* = 0.98, 26.7% of the objects are identified (72.5% of the image). These are very sufficient values. Let us note that many small objects have not been identified.

## 7. Conclusion

In this paper, we have proposed a new knowledge representation and reasoning method applied on remote sensing image interpretation. The approach is based on an ontology developed by experts helped by machine learning tools. A similarity measure and an exploration procedure of the ontology have been presented in order to associate a semantic meaning to regions of a segmented image. The experimental results have shown the effectiveness of the proposed method. Moreover, we believe that the results could be improved using a better segmentation results.

In future work, we will perform other experiments with different images and using several segmentation algorithms. We will incorporate the method in a multi-strategy classification approach in order to guide the process, to label the clusters, and to improve the final classification results.

## Acknowledgment

This work was supported by the FoDoMuSt project (multi-strategies data mining to extract and identify urban elements from remote sensing image database) of the ACI “Masses de Données” (2004-2007).

The authors would like to thank David Sheeren, Rémy Brisson and Mehdi Kaytoue-Uberall, contributors to the project.

## References

- [1] D. W. Aha, D. F. Kibler, and M. K. Albert. Instance-based Learning Algorithms. *Machine Learning*, 6:37–66, 1991.
- [2] U. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen. Multi-Resolution, Object-oriented Fuzzy Analysis of Remote Sensing Data for GIS-Ready Information. *ISPRS Journal of Photogrammetry & Remote Sensing*, 58:239–258, 2004.
- [3] C. Breen, L. Khan, and A. Ponnusamy. Image Classification Using Neural Networks and Ontologies. In *Proc. 13th International Workshop on Database and Expert Systems Applications, co-located with DEXA 2002*, pages 98–102, Aix-en-Provence, France, September 2002.
- [4] D. Crevier and R. Lepage. Knowledge-Based Image Understanding Systems: a Survey. *Computer Vision and Image Understanding*, 67(2):161–185, 1997.
- [5] S. Dasiopoulou, V. Mezaris, I. Kompatsiaris, V. K. Papatathis, and M. G. Strintzis. 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(10):1210–1224, October 2005.
- [6] S. Derivaux, S. Lefèvre, C. Wemmert, and J. Korczak. 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.
- [7] B. Draper, A. Collins, J. Brolio, A. Hanson, and E. Riseman. The Schema System. *International Journal of Computer Vision*, 2(3):209–250, 1989.
- [8] T. Gruber. Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human Computer Studies*, 43(5/6):907–928, 1995.
- [9] M. Herold, J. Scepan, A. Muller, and S. Gunter. Object-oriented Mapping and Analysis of Urban Landuse/Cover using Ikonos Data. In *22nd Earsel Symposium Geoinformation for European-Wide Integration*, Prague, June 2002.
- [10] T. M. Lillesand, R. W. Kiefer, and J. W. Chipman. *Remote Sensing and Image Interpretation*. Wiley, 2003.

- [11] S. Liu, M. Thonnat, and M. Berthod. Automatic Classification of Planktonic Foraminifera by a Knowledge-Based System. In *10th Conference on Artificial Intelligence for Applications*, pages 358–364, San Antonio, Texas, March 1994.
- [12] N. Maillot, M. Thonnat, and C. Hudelot. Ontology Based Object Learning and Recognition: Application to Image Retrieval. In *Proc. International Conference on Tools with Artificial Intelligence (ICTAI 2004)*, pages 620–625, Boca Raton, Florida, November 2004.
- [13] T. Matsuyama and V.-S. Hwang. *SIGMA - A Knowledge-Based Aerial Image Understanding System*. Plenum Press New York USA, 1990.
- [14] N. F. Noy, R. W. Ferguson, and M. A. Musen. 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, October 2000.
- [15] P. Panagi, S. Dasiopoulou, G. T. Papadopoulos, I. Kompatsiaris, and M. G. Strintzis. A Genetic Algorithm Approach to Ontology-Driven Semantic Image Analysis. In *Proc. 3rd IEE International Conference of Visual Information Engineering (VIE 2006)*, Bangalore, India, September 2006.
- [16] A. Puissant, T. Ranchin, C. Weber, and A. Serradj. Fusion of Quickbird MS and Pan Data for Urban Studies. In *Proc. European Association of Remote Sensing Laboratories Symposium (EARSeL)*, pages 77–83, Gent, Belgium, June 2003.
- [17] L. Rego and B. Koch. Automatic Classification of Land Cover with High Resolution Data of the Rio de Janeiro City Brazil: Comparison Between Pixel and Object Classification. In *Proc. 4th International Symposium on Remote Sensing of Urban Areas*, pages 1682–1777, Regensburg, Germany, June 2003.
- [18] M. A. Rodriguez and M. J. Egenhofer. Determining Semantic Similarity Among Entity Classes from Different Ontologies. *IEEE Transactions on Knowledge and Data Engineering*, 15(2):442–456, 2003.
- [19] A. Schwering and M. Raubal. Measuring Semantic Similarity between Geospatial Conceptual Regions. In *Proc. 1st International Conference on GeoSpatial Semantics (GeoS)*, pages 90–106, 2005.
- [20] D. Sheeren, A. Puissant, C. Weber, P. Gançarski, and C. Wemmert. 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.
- [21] D. Sheeren, A. Quirin, A. Puissant, P. Gançarski, and C. Weber. Discovering Rules with Genetic Algorithms to Classify Urban Remotely Sensed Data. In *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS'2006)*, Denver, Colorado, 2006.
- [22] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based Image Retrieval at the End of the Early Years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.
- [23] P. Soille. *Morphological Image Analysis*. 2nd ed. Springer-Verlag, 2003.
- [24] A. Tversky. Features of Similarity. *Psychological Review*, 84:327–352, 1977.
- [25] T. Whiteside and W. Ahmad. A Comparison of Object-oriented and Pixel-based Classification Methods For Mapping Land Cover In Northern Australia. In *Proc. Spatial Sciences Institute Biennial Conference (SSC 2005)*, pages 1225–1231, Melbourne, Australia, September 2005.
- [26] N. Zlatoff, B. Tellez, and A. Baskurt. Image Understanding and Scene Models: a Generic Framework Integrating Domain Knowledge and Gestalt Theory. In *Proc. IEEE International Conference on Image Processing (ICIP 2004)*, pages 2355–2358, Singapore, October 2004.