An Extension of an Ontology-Based Land Cover Designation Approach for Fuzzy Rules

Mariana BELGIU¹, Thomas J. LAMPOLTSHAMMER²,³ and Barbara HOFER³

¹Austrian Academy of Sciences, Salzburg/Austria · Mariana.Belgiu2@sbg.ac.at
²Salzburg University of Applied Sciences, Salzburg/Austria
³IFFB Geoinformatics – Z_GIS, University of Salzburg/Austria

This contribution was double-blind reviewed as full paper.

Abstract

Satellite image interpretation requires the assignment of sets of objects (or pixels) that share certain attribute values to object categories. This procedure requires expert intervention and knowledge. An approach has been developed that formalizes expert knowledge in the image interpretation procedure with ontologies. Ontologies provide a definition of object categories and associated attribute values that are known to represent these object categories. A classic ontology has the limitation that the definitions of object categories and their properties need to be crisp, i.e. not overlapping. Practical tests showed that less rigid definitions of class properties make the ontology-based approach more flexible and adaptable to different study areas and satellite images. This paper presents the extension of the ontology-based approach with fuzzy rules and discusses the advantages of this extension.

1 Introduction

Satellite image interpretation requires human intervention to assign labels to image objects. The automatized labelling of image objects is difficult to achieve, due to the semantic gap between image information (spectral information, channel rations, indices etc.) and contextual expert knowledge (high level semantics). An introduction to object-based image analysis and ontologies is given in section 2.

An approach has been prepared that uses a formalization of expert knowledge in an ontology. This ontology is checking the pre-computed properties of the image objects and presenting a recommended label for the objects to the experts. This ontology-based approach is a first step towards an automation of the image interpretation process (section 3).

A challenge for the ontological approach is that the attributes of image objects vary from one application area to the next, since the semantics of the concepts vary depending on the context. Water bodies, for example, do not have the same attribute values in every satellite image. This requires the definition of ranges for attribute values that are related to a certain type of objects.

In classical ontologies definitions of classes and attribute ranges are crisp, i.e., there is no overlap between class definitions and an attribute value falls into a certain range or not even
if the value itself is very close to the threshold defined in the range. In the process of classifying objects/pixels in satellite images, this leads to misclassifications and objects that remain unclassified. To overcome this limitation inherent to classical ontologies, we propose the use of a fuzzy ontology.

In a fuzzy ontology definitions can overlap and membership to is expressed via fuzzy sets. Fuzziness can be used in different stages of the ontology design. It can be applied to class definitions themselves and thereby account for ambiguities in the definitions of classes like “mixed forest”. In our approach, we limit the use of fuzziness to the definition of attribute ranges. That means that we work with probabilities with that a certain attribute value satisfies the property of having an, e.g., “high value”. Thus, it becomes possible to move beyond classical “is or is-not” towards “is-a to some degree” definitions. As a running example we use water bodies and confront the definition of the class and its properties in the classical ontology with the definition in the fuzzy ontology (section 4).

The fuzzy extension of the ontology-based approach makes the labeling of objects in satellite images more flexible (see discussion in section 5). Future work includes the implementation and application of the fuzzy ontology in the web ontology language (OWL) (section 6).

2 Background

This section introduces the concepts and tools involved in the ontology-based approach and its fuzzy extension: object based image analysis (OBIA), ontologies and fuzzy ontologies.

2.1 Object Based Image Analysis (OBIA)

In the last decade, Object Based Image Analysis (OBIA) has been widely used to classify Very High Spatial Resolution (VHSR) imageries (BLASCHKE & STROBL 2001, BLASCHKE 2010). OBIA is a knowledge-driven image interpretation method (HAY et al. 2005) where the expert knowledge is translated into rule sets designed to assign semantics to image objects delineated by means of image segmentation algorithms. Like any other knowledge-based image interpretation method, OBIA is challenged by the absence of an explicit link between low-level image information (spectral information, channel rations, indices etc.) and contextual expert knowledge (high level semantics). This problem is called semantic gap (SMEULDERS et al. 2000) and it impairs the repeatability and transferability of image classification. Ontology-driven image interpretation frameworks are promising approaches to overcome this issue (BITTNER & WINTER 1999, CÂMARA et al. 2001).

2.2 Ontology

Every domain has an agreed understanding of features and phenomena that are of interest to them. Ontologies aim at capturing this understanding in a vocabulary of terms with formally defined meaning and are defined as “…an explicit specification of a shared conceptualization” (GRUBER 1995, 908). Whenever data, information or services from different disciplines should be integrated, exchanged or queried, ontologies may serve as a means for supporting these tasks (GRUBER 1995); ontologies enable semantic interoperability (see JANOWICZ 2009).
Spatial ontologies or geo-ontologies are ontologies providing the description of geographic features. They have to account for the object and field view of physical and social concepts and need to include spatial relations among entities (Fonseca & Câmara 2006). For an overview of work on ontologies in the GIScience field consult Agarwal (2005).

Ontologies with reasoning capabilities that are based on the formal theory of description logics are in the focus of this work. These ontologies allow to reason about the knowledge contained in the ontology. The W3C developed the Web Ontology Language (OWL), which is currently available in its second version (OWL 2) as a recommendation (Hitzler et al. 2009). For an account of the development of OWL we refer to Horrocks, Patel-Schneider et al. (2003).

OWL is a knowledge representation language for providing a specification of a domain of interest. The vocabulary of OWL comprises three main constructs: classes, individuals and properties. Classes are sets of individuals and properties are defining relationships between two individuals or an individual and a data type.

The definition of the constructs of OWL is based on description logics (DL) for the species of the language called OWL DL. Description logics thereby provide the formal theory on which statements in OWL are based and through which the statements can automatically tested by a reasoner. A reasoner is a software program that can infer superclass/subclass relationships from an ontology (subsumption testing), and therefore conduct consistency, equivalence and instantiation testing.

OWL 2 is based on OWL DL and improves the original Web Ontology Language in following points among others (GRAU et al. 2008, HITZLER et al. 2009):

- Increased expressivity for data types and properties: for example, OWL 2 can now express a restriction to a range of data type values.
- Improved annotations: annotation properties allow statements that are not evaluated by the reasoner.
- Clarification of syntax issues: that issue relates to the necessity of an OWL ontology to be provided using different syntaxes.

For preparing an ontology in OWL 2, we use the Protégé 4.1 editor provided by the Stanford Center for Biomedical Informatics Research1. Pellet and Hermit are, amongst others, two reasoners that can be used together with Protégé.

### 2.3 Fuzzy Reasoning in Ontologies

The classical ontology-based approach works fine as long as there are distinct concepts and therefore no ambiguity exists within the domain model. In practical terms, named classes in OWL are generally defined as distinct on each level of the hierarchy and therefore, not supposed to overlap. However, reality is different and ambiguity as well as imprecision issues arises in almost all real-world applications. To overcome these issues fuzzy set theory and fuzzy logic proved to be suitable (Zimmermann 2001). For example, the definition of the property “old” is rather vague and one would end up probably with as much

different definitions as people were asked. One agreed fact is that the definition of “old” has to be a range from a starting point to an upper limit and as the upper limit is approached, the probability to be old increases. A definition of the property “old” based on fuzzy concepts provides flexibility in expressing the ambiguity of the term.

ZADEH (1965) published the initial and ground-breaking ideas as regards fuzzy logic and fuzzy. The fuzzy approach can be seen as a description of any system via fuzzy logics and fuzzy predicates. From a more general perspective, it is a qualitative description of system behaviour. Fuzzy modelling uses fuzzy quantities for defining a system. These quantities can assume shapes such as fuzzy numbers or sets. It is possible to assign linguistic labels in natural language to them, which is, however, not necessary for reasoning only (SUGENO & YASUKAWA, 1993).

An ontology incorporating fuzzy concepts (fuzzy ontology) can be described as a quintuple $FO = <I, FC, FT, NFT, X>$ (FULLÉR 2008). The $I$ represents a set of individuals or objects as they are instances of classes. $FC$ denotes a set of fuzzy concepts. The overall set of entities modelled within the ontology is described as $E = FC \cup I$. The third tuple $FT$ represents the fuzzy taxonomy relations as regards the set $FC$. In contrast to that, $NFT$ stands for non-taxonomy relationships. These are still fuzzy relationships, but across the classical tree structures and not limited to vertical relationships. Finally, $X$ describes a set of axioms for constraining elements within the ontology (e.g. individuals, relationships etc.). These sets are formulated in a logical language.

In order to capture and represent the above-described paradigm we adopted the idea of BOBILLO & STRACCIA (2011) using OWL2 itself to conduct fuzzy reasoning within ontologies. In OWL 2, fuzzy concepts are introduced in the annotation properties and not as elements of the existing crisp ontology. OWL2-based reasoners ignore this property as it is used for the sake of being readable by humans. As a result it becomes possible to perform crisp and fuzzy reasoning with the same ontology with the use of a fuzzy reasoner. An example of a fuzzy annotation can be seen in section 4.

3 Ontology-Based Categorization of Objects from Satellite Imagery

This section presents the ontology-based approach for categorizing objects or pixels in satellite images. It covers an overview on existing literature (section 3.1) and the description of the methodology used (section 3.2).

3.1 Image Interpretation by Ontology

Previous research underlined the importance of using explicit semantics for content based image indexing (CBIR) (DATCU 2003) and (DURBHA 2005). A few research initiatives focused on using the concepts of semantics to enhance image interpretation. FORESTIER et al. (2012) for instance developed a methodology for the automatic matching of objects (segments) extracted from High Spatial Resolution (HSR) imagery against the domain concepts whose semantics are explicitly defined in the knowledge base. This methodological framework has been used to identify single objects (e.g., a house) and aggregated objects.
Fuzzy Rules For Land Cover Designation in an Ontology-Based Approach

(e.g., urban blocks) from QuickBird images. Despite all these efforts to introduce ontology-engineering methods to the remote sensing community, the semantic gap problem is far from being solved. Moreover, the enrichment of ontologies with fuzzy logic formalisms has not been yet discussed in the remote sensing domain. By contrast, in the medical domain, HUDELOT et al. (2008) proposed a fuzzy ontology of spatial relations (both topological and metric relations) for recognizing internal brain structures derived through segmentation from 3D magnetic resonance images. In Geographic Information Systems (GIS), LÜSCHER et al. (2009) underlined the benefits of using ontologies in spatial data enrichment. Furthermore, they compared the urban pattern classification obtained by using a classical ontology with the classifications obtained using supervised Bayesian inference. As expected, Bayesian inference performed better than classical knowledge inference formalisms.

3.2 Methodology Description

This paper extends the ontology-based object recognition described by (authors’ names removed due to blind review) with fuzzy definitions of the target classes. The proposed method enables the classification of a particular object (called Object of Interest – OI) using the definitions and constraints defined in the ontology. It is a loosely coupled object categorization architecture where the input data layer is separated from the reasoning layer and the presentation layer. The architecture’s modularity makes our framework flexible enough to be adapted to different application scenarios.

The methodological framework involves a pre-processing step where the satellite imagery is partitioned into homogeneous regions using one of the available segmentation algorithms (figure 1). After the segmentation, each image object becomes a unit of analysis for which a number of features/attributes can be measured and used during the classification. The image objects (the geometry) together with the computed object features are added to the ontology middleware module, which is responsible for assigning semantics to the image objects. The association between high level concepts (expert knowledge) and low-level image features is performed automatically via the implemented inference engine. The inference engine (reasoner) replaces the classical image classifiers (e.g., Support Vector Machine). It infers new knowledge, based on the class definitions, constraints and axioms defined explicitly in the developed ontology. The class definitions are thereby operating as prototypes (or samples) of objects that shall be identified in the images; the semantics of the prototypes are explicitly defined in the ontology. This methodological framework has been developed in Java. A number of modules have been integrated into the system including the reasoner system (Pellet API), OWL API, WordNet API and KML API. The ontology was expressed in Ontology Web language2 (OWL2).

We applied the proposed object recognition method to classify water features (rivers, lakes and swimming pools) from WorldView2 satellite imagery (DigitalGlobe, Inc., USA) in an exemplary analysis. Object recognition (object categorization) starts with a categorization request which contains the object of interest to be classified. The users send a query request to the system and the system is processing it by firstly checking the availability of the searched concept in the domain ontology. If needed concept is not modeled, the users are asked to use other words related with the foregoing browsed concept. This functionality is supported by WordNet lexical database that has been integrated into our system via the WordNet API. If the object of interest has been modeled in the ontology, the implemented reasoner is checking the match between the queried object whose semantics are described
in the ontology and the image regions are delineated by means of segmentation algorithms. Features used for the concept categorization are computed by feature extraction algorithms. The resulting vectors of features are automatically parsed in the ontology modeling language format. The assignment of domain ontology concepts to the regions extracted by means of segmentation techniques is achieved automatically by the integration of an inference mechanism (reasoning). If the image regions matched the needed concept, the results are sent back to the users in Keyhole Markup Language (KML) format and displayed on Google Earth. If the matching fails, the current classification is dropped.

Fig. 1: Schematic depiction of the ontology-based approach for labelling objects

The assessment of the classification accuracy revealed an overall accuracy of 0.83. The swimming pool class has a producer and user accuracy of only 80%. The reason for this relatively low accuracy is the inappropriate delineation of swimming pools by the used segmentation algorithm. This is a well-known problem of the OBIA method and further research is required in order to improve the segmentation methods.

The key advantages of the proposed methodology are: (1) it is an expert based image analysis framework (ii) the knowledge base is modelled using W3C recommendations; (iii) it achieves the consistent integration of remote sensing data with other GIS information; (iv) the proposed framework can be extended to different sensors.

4 Contrasting the Definition of Water Bodies in OWL and fuzzyOWL

In the following, we shortly describe an example of extending the classic ontology with the fuzzy formalism. As said before, the fuzzy formalism is a way of dealing with the vagueness or non-precise, qualitative information. The vagueness (or imprecision) encompasses two semantic indeterminacies: conceptual ambiguity and threshold vagueness (BENNETT, 2011). In this paper, we used fuzzy logic to define vague qualitative predicates such as “high NDWI”. The threshold vagueness is referred to as sorties’ paradox (BENNETT 2011).
Conceptual ambiguity is also an important issue especially when dealing with indeterminate, vague definitions such as ‘Biotope’, but this discussion is out of the scope of this paper.

The key element of the definition of the class water bodies is: water bodies have a high NDWI value. The Normalized Difference Water Index (NDWI= green-earinfrared1/green+nearinfrared2) has been successfully used to classify and to monitor water features (GAO 1996, XU 2006). Please note that this definition does not express a precise description of water bodies in the satellite imageries. This definition denotes the qualitative expert knowledge regarding the object appearance in the satellite imageries: water features absorbs energy at near-infrared wavelength and reflect energy at green wavelength and therefore have a high Normalized Difference Water Index- NDWI. Furthermore, this definition conveys only a part of the water body properties; additional properties can be used as well: e.g. NIR spectral reflectance.

![Diagram of domain application knowledge and image regions ontology]

**Fig. 2:** Ontologies providing domain application knowledge and knowledge of how this information is linked to image objects.

Reasoning on qualitative knowledge is an important asset of any system that models real world concepts whose underlying meanings are conveyed in non-precise, vague information. In our approach, the non-precise, qualitative information is modeled as ontology classes (e.g. NDWIHigh) which are linked with the assertional knowledge that denotes the concrete objects of the considered domain (concept instances) (figure 2). However, the instantiation of the qualitative concepts occurs by using crisp definitions (if value > value1, then concept = true).

In set notation the water body description looks as follows in Formula 1:

\[
\text{Water Bodies} \equiv \text{hasNDWI.HighNDWI} \tag{Formula 1}
\]

The expression of the rule that a water body has a high NDWI value looks as following in OWL/XML format:
The definition of what “HighNDWI” means, is given in the following statement that says: all NDWI values that are the same or higher than 0.3 are high NDWI values. In OWL/XML syntax this statement looks as follows:

```
<EquivalentClasses>
  <Class IRI="#HighNDWI"/>
  <DataSomeValuesFrom>
    <DataProperty IRI="#Ndwi"/>
    <DatatypeRestriction>
      <Datatype abbreviatedIRI="xsd:double"/>
      <FacetRestriction facet="http://www.w3.org/2001/XMLSchema#minInclusive">
        <Literal datatypeIRI="http://www.w3.org/2001/XMLSchema#double">0.3</Literal>
      </FacetRestriction>
    </DatatypeRestriction>
  </DataSomeValuesFrom>
</EquivalentClasses>
```

Therefore, the image regions are classified as water if the NDWI > 0.3 (NDWI is high) and as non-water if NDWI <=0.3 (NDWI is Low). This definition might not be suitable and robust enough for classifying water features, because the threshold values of the computed features may overlap (HU & WENG 2010). In this case, the threshold selection might lead to an over- or under-estimation of water features (XU 2006). To overcome this problem, we propose the enrichment of ontology with fuzzy logic formalism. Fuzzy classifications developed based on spectral, spatial and texture features extend the object-based image analysis (HU & WENG 2010, HOFMANN et al. 2011).

As mentioned above, in the OWL2 semantics, the expression “HighNDWI” is modeled as a concept whose semantics is constrained by defining a specific datatype value (i.e., 0.3 double). In the fuzzyOWL formalism, the semantics of the “HighNDWI” concepts is constrained by using a fuzzy datatype (e.g. ndwiHigh). For instance, the fuzzy datatype ndwiHigh: [-1, 1]→[0,1] defines the degree of an image region being water body as

\[ \text{ndwiHigh}(object_i) = \text{right}(-1, 1, 0.3, 0.7) \]

where [-1, 1] denotes the NDWI value range, while [0.30, 0.70] defines the range for the fuzzy ndwiHigh datatype. We defined two fuzzy sets for the NDWI feature: low and high.

The following two excerpts show the fuzzy definition of the expression “HighNDWI” in OWL/XML syntax. The first block provides the general definition of the NDWI datatype, which ranges from -1.0 to 1.0 and is double. The second block provides the expression of the membership function of value to the property “HighNDWI” with the increase of membership from 0.3 to 0.7 as shown in figure 3.

```
<EquivalentClasses>
  <Class IRI="#HighNDWI"/>
  <DataSomeValuesFrom>
    <DataProperty IRI="#Ndwi"/>
    <DatatypeRestriction>
      <Datatype abbreviatedIRI="xsd:double"/>
      <FacetRestriction facet="http://www.w3.org/2001/XMLSchema#minInclusive">
        <Literal datatypeIRI="http://www.w3.org/2001/XMLSchema#double">0.3</Literal>
      </FacetRestriction>
    </DatatypeRestriction>
  </DataSomeValuesFrom>
</EquivalentClasses>
```

![Fig. 3: Example of defining two fuzzy sets on NDWI feature](image-url)
The parametrisation of membership involves a deep understanding of the modelled features. For instance, according to figure 3, if the i object has NDWI value = 0.32, the membership to the water body class is 0.1 and the membership to other land cover class is 0.6. If the objectj has NDWI value 0.7, the membership to the water body class is 1 and the membership to other land cover class is 0. The overlap of membership increases or decreases the yielded classification outputs (e.g. the greater the membership overlap, the more vague is the final classification).

5 Discussion

The proposed ontology-based method for image object categorization has been embedded within the OBIA framework. OBIA allows image analysis experts to use their knowledge about the application context and to develop rulesets that convey the objects’ properties (spectral and spatial properties) and the hierarchical and spatial relations that hold them together. These properties have been modelled in the ontology whose underlying semantics have been expressed using OWL2. The feasibility of the proposed methodology has been tested in (BELGIU et al. 2013). Since the spectral behaviour of objects differs depending on the sensing conditions and environmental settings, the class definitions embedded in the ontology are not flexible enough to yield outputs of good accuracy. Therefore, the challenge was to come up with a more flexible data model that performs better in varying environmental settings.

In this paper we proposed the extension of the ontology-based categorization of image with a fuzzy formalism. The methodological framework itself is foreseen to be used in parallel to mainstream image analysis methods, e.g. supervised image classifiers.

Section 4 shows the definition of “water bodies” with the classical ontology and with the fuzzy ontology. The main difference in the definitions lies in the way the expression
“HighNDWI” is defined. In the crisp or classical definition, the value is specified by a range that has a certain lower limit value. A specific NDWI attribute value either falls within the range or not.

In the fuzzy case the expression “HighNDWI” is defined as a membership function for a range of the NDWI values. This definition assigns a probability to an individual and expresses to which degree an individual satisfies the expression “HighNDWI”.

The extension of the ontology for fuzzy expression makes the classification of individuals more flexible and reduces the wrong decisions based on the yes/no logic in the decision making with the crisp ontology. In the case of using fuzzy expressions in the decision logic, the expert can be presented with a list of objects that fulfil the specification to a lower degree and require manual decision making.

6 Conclusions and Future Work

The ontology-based approach is a promising approach towards the automation of the labelling of objects in satellite images. It is a knowledge-driven image analysis approach that models explicitly the knowledge used to extract information from satellite images. The ontology granularity and structure can be extended to new application domains and satellite imageries. By contrast, the available expert based image analysis frameworks are based on rulesets whose semantics are hidden in the form of complex algorithms. Furthermore, they are tailored to a specific application and therefore their transferability is reduced. This work presented the extension of the definition of object categories in the used ontology for fuzzy rules. This extension supports the flexibility of the approach.

Future work includes the implementation of the fuzzy ontology with the fuzzyOWL plugin for the Protégé OWL editor. The ontology that is extended for fuzzy rules will be linked to a fuzzy reasoner. To our knowledge, there are two reasoners available that can infer new knowledge from the semantics defined in the ontology: fuzzyDL\(^2\) and the DeLorean reasoner\(^3\) whose functionality is, however, restricted to the translation of a fuzzy ontology into classical ontology language.

FuzzyOWL is an important formalism to deal with the uncertainty inherent in the information extracted from satellite imageries. Our approach focused on the fuzzy definition of class properties, whereas the class definitions themselves are still crisp. Indeed, the uncertainty problem applies to domain vocabulary as well: e.g., land cover/land use, biotope or recreational areas concepts. These concepts are intrinsically vague and imprecise, their semantics varying according to the analysis scale and the geographic and socio-cultural context where the objects are embedded. The inclusion of fuzzy rules in class definition is yet another step for making the ontology-based approach more flexible and easily transferrable.


\(^3\) [http://webdiis.unizar.es/~fbobillo/delorean](http://webdiis.unizar.es/~fbobillo/delorean)
Acknowledgment

This research was partly funded by the Austrian Science Fund (FWF) through the Doctoral College GIScience (DK W 1237-N23).

References


