AN ONTOLOGY-BASED MODELING OF AN OCEAN SATELLITE IMAGE RETRIEVAL SYSTEM

Jesús M. Almendros-Jiménez, José A. Piedra-Fernández, Manuel Cantón-Garbín

University of Almeria
Department Lenguajes y Computación
Ctra. Sacramento s/n CP 04120, Almería, Spain

ABSTRACT

The need to access information in large volumes of image data, e.g. big images, large image archives, distributed image repositories, etc. has stimulated the research in the field of *content-based image retrieval* during last decade [DJLW08, LZLM07, Han08]. Many new concepts have been developed and prototyped. However the dramatic increase in volume, details, diversity and complexity and the user demand for simultaneous access to multi-domain data urgently requires new approaches for *image information mining*, *multi-domain information management*, and *knowledge management* and *sharing*.

In this work we present the elements of the *architecture of a satellite image retrieval system*. In order to have a *flexible and dynamic* image retrieval system we have designed an architecture which integrates *Web technologies* with modern *pattern recognition systems* applied to satellite images [PGC07]. Our proposal aims to develop a satellite image retrieval system in which the kind of image, sensor, satellite, together with the kind of *classifiers* can be *customized* by the system user. More concretely, image, sensor and satellite features can be *imported/exported* and *updated*, and the algorithms can be *imported/exported* and *refined/improved* from the system. In particular, the image retrieval system allow the *data and knowledge exchange*. In order to equip the system with importing/exporting facilities for data and knowledge exchange, well-known Web technologies are used: the *XML (Extensible Markup Language)* [W3C07] for data exchange, and the *Web Ontology Language (OWL)* [W3C04a] for knowledge exchange.

The database community has been interested in the last years in the development of mechanism for *data exchange* between applications. Data exchange is crucial for *data integration* between applications, and it is crucial in the case of *Web tools*. The development of image retrieval systems allowing data exchange seems to be crucial for the integration of such image resources.

However, data exchange and integration is not an easy task. Images can be described by means of a set of *features* like name, date, etc. However, images can be also described by means of the so-called *"semantic content"*. The semantic content represents properly features of the images based on the *image content*. The image content can be expressed by means some kind of *keyword*, however, image content might represent *complex knowledge*. This is the case of *ocean satellite images*, in which the semantic content represents certain kind of *ocean structure* like an *upwelling* or an *eddy*. Such semantic concepts are complex concepts which are not easy to define, but they can be detected by applying some techniques of pattern recognition. In this context, images can be described by means of a set of features, and some of the visual elements of the image can be classified according to semantic concepts.

Now, the question is how to exchange image features and semantic content between image retrieval systems? Image features which do not correspond to semantic content can be stored in a database by means of *records*, and they can be exchanged by

THIS WORK HAS BEEN PARTIALLY SUPPORTED BY THE SPANISH MICINN UNDER GRANT TIN2008-06622-C03-03

importing/exporting the records from (and to) *XML format*. An XML document is plain text in which the image features are stored together with the name of the feature.

In the case of semantic content, the image content has been detected by applying a certain *classification criterium*. Such classification criterium is not directly applied to the image, rather than the image is *processed*, and, in the case of ocean satellite images, from the image a set of *regions* are obtained from *segmentation* (i.e. *region based segmentation*). Afterwards, the regions are analysed according to the classification criterium, and the semantic content is obtained. However, the classification criteria are not standard, that is, many classification criteria can be considered depending on the kind of image, satellite, sensor, etc, and, in addition, the knowledge of the user about the kind of image to be analysed plays a crucial role in the classification. Therefore, the semantic content depends on many factors.

In this context, data exchange and integration has to take into account that semantic content of images, and the *exchange of knowledge*. The knowledge can be described by means of several formalism like *rule-based expert systems*, *neuronal networks*, *decision trees*, among others. The exchange of such knowledge should consider some kind of data format for knowledge representation.

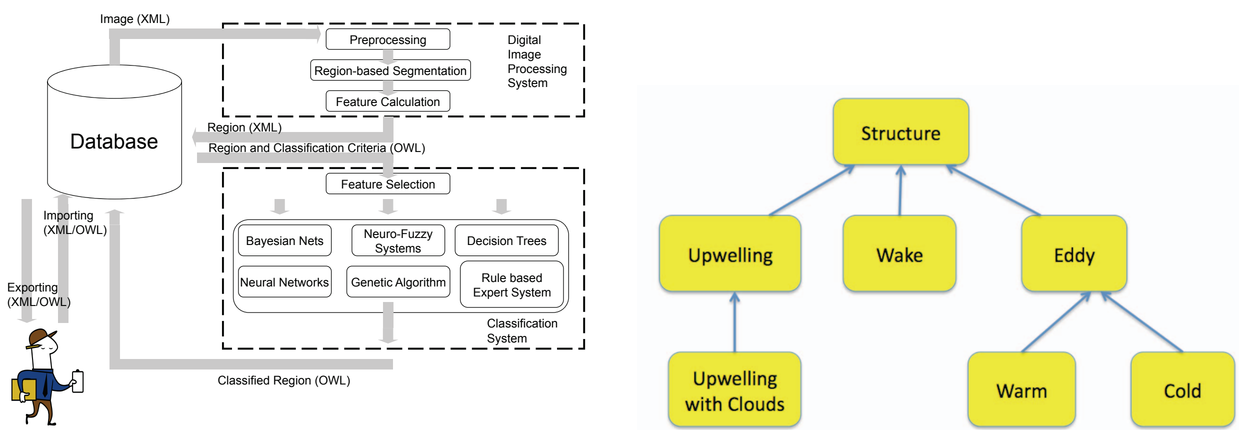

Fig. 1. Architecture and Ontology Modelling

The database community has been also interested in the representation of knowledge for exchanging. In this line, the so-called *Semantic Web* [BLHL01] approach aims to provide mechanisms for representation and exchange of knowledge through the net. In this framework, two formalisms, named *RDF (Resource Description Format)* [W3C04b] and *OWL (Web Ontology Language)* [W3C04a] have been proposed. This later is an extension of the former, and they provide languages for describing semantic information about a domain of interest. The semantic information is considered as an *ontology* about the domain. An ontology expresses the elements of the domain and the intended meaning of the elements.

In the case of ocean satellite images, the ontology would include the concepts of upwelling and eddy, among others, together with the meaning of such concepts. Such meaning is given by the classification criterium used for recognizing such ocean structures. OWL allows, similarly to XML, to represent in plain text, the elements of an ontology and its meaning. For such a reason, OWL is a suitable format for exchanging semantic content of images between image retrieval systems.

Finally, the use of data exchange formats in a image retrieval system allows the user to customize his (her) image features, classification criteria and semantic content by importing/exporting mechanisms. Therefore, such importing/exporting mechanisms allow to have a flexible and dynamic system in which new kinds of images, new kinds of sensors, new kinds of classification criteria can be added and tested.

In summary, the proposed architecture (see Figure 1) of the satellite image retrieval system can be summarize as follows:

(a) A *database* which contains records of the satellite images in which the image file name is specified together with image

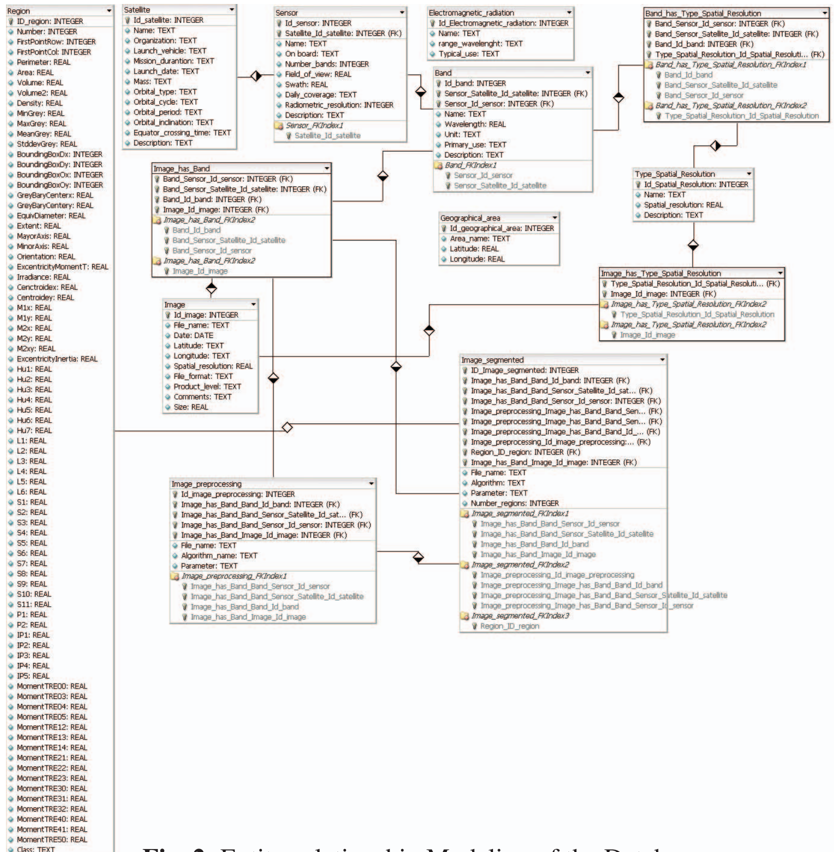

Fig. 2. Entity-relationship Modeling of the Database

format, date, sensor, etc. In addition, the database contains records about the features of the satellites: sensors, orbit, launch date, etc. For each sensor, the database contains the bands, the field of view, etc. And finally, for each band, the wavelength, the spatial resolution, etc. In Figure 2 we can see the *entity-relationship* modeling of the database.

(b) A *digital image processing system* which retrieves a given image from the database and obtain a segmentation given by a set of regions. Regions are also stored in the database together with its computed features and the name of the image they correspond (see Figure 2). In order to achieve the segmentation process, the database provides information to the digital image processing system about the satellite, sensor and band of the processed image.

(c) A *classification system* which retrieves the regions of a image from the database, and applies a classification criterium to the regions. For each region, it obtains a classification result which updates the database of regions. Regions are stored in the database. The classification criterium is described by means of an ontology. Ontologies are also stored in the database.

(d) An *XML-import/export facility* which takes as input an XML document containing the features of new images to be added to the database, and stores them in the database. The XML-import/export facility can also add to the database new sensors, satellites, etc, which have to be in XML format. The XML-import/export facility is able to generate in XML format the features of a image, satellite and sensor stored in the database, and in addition, it can also reports the result of the classification of a given image in XML format.

(e) An OWL-import/export facility which takes as input an OWL document containing the knowledge of new classifiers to be added to the database to be stored in the database. The OWL-import/export facility is able to generate in OWL format the classifiers used in a given classification process.

For instance, an image can be represented by means of XML as follows, and the ontology of Figure 1 can be represented by means as OWL as follows:

```
<image Id_image=”00001”>
<File_name> ”900805.hdf” </File_name>
<Date> 90/8/05 </Date>
<Bands> <Band> 2 </Band>
<Band> 4 </Band>
<Band> 5 </Band>
</Bands>
<Latitude> 30N-20N </Latitude>
...
</image>
```

```
<owl:Class rdf:about=”#Structure”/>
<owl:Class rdf:about=”#Eddy”>
<rdfs:subClassOf rdf:resource=”#Structure”/>
</owl:Class>
<owl:Class rdf:about=”#Wake”>
<rdfs:subClassOf rdf:resource=”#Structure”/>
</owl:Class>
<owl:Class rdf:about=”#Upwelling_with_Clouds”>
<rdfs:subClassOf rdf:resource=”#Upwelling”/>
</owl:Class>
```

Finally, a classification criterium of *upwelling*, given by the values of Hu’s invariant, mean grey, grey barycenter, centroid, etc of the given region can be described by means of OWL. For instance, in OWL we can express as follow that an upwelling has Hu’s invariant between 19.0 and 98.0:

```
<owl:Class rdf:about=”#Upwelling”>
<owl:equivalentClass>
<owl:Class>
<owl:intersectionOf rdf:parseType=”Collection”>
<rdf:Description rdf:about=”#Region”/>
<owl:Restriction>
<owl:onProperty rdf:resource=”#InvHu1”/>
<owl:someValuesFrom>
<rdf:Description>
<xsd:minInclusive rdf:datatype=”&xsd;double”>19.0</xsd:minInclusive>
</rdf:Description>
...
<owl:onProperty rdf:resource=”#InvHu1”/>
<owl:someValuesFrom>
<rdf:Description>
<xsd:maxInclusive rdf:datatype=”&xsd;double”>98.0</xsd:maxInclusive>
</rdf:Description>
...
```

1. REFERENCES

[BLHL01] T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web – A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, May:36 pages, 2001.

[PGC07] J.A. Piedra, M. Cantón, F. Guindos. Application of Fuzzy Lattice Neurocomputing (FLN) in Ocean Satellite Images for Pattern Recognition. *Studies in Computational Intelligence (SCI), Springer-Verlag*, 67, pages 215–232, 2007.

[DJLW08] R. Datta, D. Joshi, J. Li, and JZ Wang. Image Retrieval: Ideas, Influences, and Trends of the New Age. *ACM Computing Surveys*, 40(2), 2008.

[Han08] A. Hanbury. A survey of methods for image annotation. *Journal of Visual Languages and Computing*, 19(5):617–627, 2008.

[LZLM07] Y. Liu, D. Zhang, G. Lu, and W.Y. Ma. A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 40(1):262–282, 2007.

[W3C04a] W3C. OWL Ontology Web Language. Technical report, www.w3.org, 2004.

[W3C04b] W3C. Resource Description Framework (RDF). Technical report, www.w3.org, 2004.

[W3C07] W3C. Extensible Markup Language (XML). Technical report, www.w3c.org, 2007.