Knowledge representation for satellite imagery

Représentation des connaissances en imagerie satellitaire

Marine Campedel

Khaled Aouini
Hakim Hadioui
Haiming Liu
Prisca Plesel

2009D026

Décembre 2009

Département Traitement du Signal et des Images
Groupe TII : Traitement et Interprétation des Images

”Knowledge Representation for satellite imagery”

Représentation des connaissances en imagerie satellitaire

Action R&T R-S08/OT-0004/033
Bon de commande : 4500026290 / DCT094 du 02-10-2008
Sur marché 05/2409/00

Marine CAMPEDEL

involved students : Khaled AOUINI
Hakim HADIOUI, Haiming LIU, Prisca PLESEL

December 2009

Département Traitement du Signal et des Images
Competence Center CNES - DLR - Telecom ParisTech
Institut Telecom - Telecom ParisTech

Table des matières

Chapitre 1

Résumé

Ce rapport de recherche est le résultat d'un contrat intitulé *Représentation des connaissances pour l'imagerie satellitaire*, dans le cadre du Centre de Compétence (COC) CNES/DLR/Télécom ParisTech. Il part du constat que le COC a étudié et développé, depuis sa création, de nombreux outils d'analyse des (séries d') images satellitaires (outils d'extraction et sélection d'attributs, outils de classification) dans le but de faire émerger une information sémantique. Ces outils génèrent ainsi de nouvelles données, qui enrichissent la description initiale (sous forme de métadonnées) des images. Cependant, ces données sont-elles porteuses d'une information utile pour les usagers des images satellitaires, à savoir les photo-interprètes ? et permettent-elles de faire émerger de nouvelles connaissances ?

Ainsi nous sommes intéressés d'une part aux interprètes et d'autre part à la notion de connaissance, telle qu'elle émerge actuellement dans un domaine qui se fait appeler "ingénierie des connaissances". En ce qui concerne les interprètes, nous avons rencontré deux types d'interprètes très différents, dont nous avons tenté d'analyser le savoir-faire, à l'aide de la Gestion Mentale. Cette théorie, issue des travaux d'A. de la Garanderie, tend à décrire les cinq gestes mentaux que nous sommes tous amenés à faire plus ou moins naturellement (attention, mémorisation, réflexion, compréhension, imagination créatrice) afin d'en déduire une méthode d'apprentissage de ces gestes et de les optimiser. Plus précisément, nous avons utilisé cette théorie pour tenter une description de ce que peut être la tâche d'interprétation des images satellitaires, à savoir compréhension de l'image puis restitution sous forme d'une carte annotée. A la suite de cette analyse, nous avons proposé un système (grossier) d'assistance aux interprètes dans leur tâche d'annotation.

Dans un second temps, nous nous sommes intéressés à la notion de connaissance, telle qu'elle émerge dans le web sémantique. La connaissance devient une entité partageable, interopérable, compréhensible à la fois des machines et des hommes, qui structure les informations issues des données brutes (documents du web, images satellitaires). Soutenus de plus par le projet ANR DAFOE, nous avons construit deux

ontologies en vue du développement d'un outil d'aide à l'annotation des images satellitaires. La première ontologie, dite "de l'image" s'inspire des travaux de la littérature effectués sur des images photographiques classiques et permet de décrire le processus d'extraction d'information sur les images satellitaires, ainsi que les résultats, notamment en termes d'objets dans l'image (régions, points, ...) et de leurs caractéristiques (couleur, textures, formes). La seconde, dite "de la scène", permet de décrire les objets de la scène et donc de nommer effectivement la nature des objets visibles dans les images. Comme pour toute ontologie, nous avons donc défini et décrit un ensemble de concepts, des relations sémantiques les reliant (relations de hiérarchie, mais également relations spatiales et de composition) ; le peuplement de ces ontologies s'effectue au travers de l'application d'annotation (nommée SISA) en cours de développement avec l'entreprise Mondeca (sur la scène de Marseille).

Cette étude a donc pour ambition de réduire le fossé existant entre les interprètes et les traiteurs d'image, en tentant de mieux connaître l'expertise des premiers afin que les seconds puissent les assister. Les perspectives sont nombreuses :

- méthodologiques : en s'appuyant sur des aspects cognitifs issus de Gestion Mentale,
- applicatives : SISA n'est qu'un démonstrateur qu'il serait possible de développer plus avant afin de véritablement répondre à des enjeux de traitement des données géographiques dans leur globalité, ie raster et vecteurs,
- scientifiques : notre choix de modélisation par des ontologies est criticable ; il serait intéressant d'étudier des représentations alternatives (graphes/hypergraphes sémantiques par exemple).

Chapitre 2

Introduction

2.1 Context of this project

The Competence Center on Information Extraction and Image Understanding (CoC) is involved in problematics related to the content analysis of very high resolution satellite images. These images are particularly difficult because of their size and semantical richness. Nowadays lots of interpretation works directly involve human interpreters and are thus time consuming. In this action we are interested in easing the access to the image content for interpreters (or even non expert user) in order to allow explicit knowledge extraction from these images.

It appears that the CoC is able to automatically extract many information from the images at a low level (thanks to feature extractors) or at a upper semantical level (using learning approaches). Moreover in very recent works, we are interested in explicating (temporal, spatial or purely numerical) relations between observed informations. However the richness, the interest of these information is difficult to evaluate without a human intervention, using for example a relevance feedback approach. Hence, this evaluation needs a representation step of all extracted information in interaction with a human evaluator (designed also as *user* in the following).

Among all information, many are easy to represent : for example a road network can be represented as a layer on the original image ; specific colors can be used to represent different landscapes (Cf Corine Land Cover definitions), ... However, considering the high diversity of the scene objects, a representation based on additional layers could lead to an heavy and not so informative representation of the extracted information. Maps are also considered since they use symbolic representations of objects that are topologically related. In this action we would like to represent a higher diversity of relations issued from the automatic processing or from external source of knowledge.

Moreover all extracted information are not relevant for a given user at a given time. There is a necessity to adapt the representation to this user and to consider the image interpretation as a temporal task. It is quite natural to consider interpretation as a subjective task but how could we ensure that the final model of the image content is valid for other observers ? How ensuring some kind of consensus about the image content ?

2.2 Project organization

The project has been organized in 4 different tasks, with, at the end, different contributions which will be detailed in the conclusion of this report.

Task 0 : Knowledge engineering Since now 5 years the CoC produced tools to characterize satellite images content (PhD theses of B. Luo, L. Gueguen, A. Batthacharya, I. Kyrgyzov, H. Chaabouni) and then to exploit it in learning tasks (M. Lienou, J.B. Bordes, M. Costache) However we did not succeed in demonstrating that these tools are able to assist an end user and to produce knowledge. One part of this study is hence dedicated to the notion of knowledge, as defined by engineers and to how simple data can become knowledge. We do not claim that we answered the question but this project looks at what is done in the knowledge engineering community in order to propose a new point of view in the satellite image domain.

Task 1 : How do interpreters work ? Since photo-interpreters are satellite images' main users, we are interested in studying their behavior in order to identify the phases of their work that need to be assisted. This study can be extended to non-expert user.

The project should have involved many interpreters but in fact none have been seriously contacted. Primary results about the kind of work interpreters do, where they publish and a description of the tools they use is simply given. The main result is that interpreters use lots of heterogeneous data (not only coming from the images) and various processing tools to produce an image interpretation (map). This is why our representation tool should be able to integrate a large variety of information related to the image or to the imaged scene. Moreover the interface is limited to a traditional computer screen.

Task 2 : what are the information to be represented ? In this project we are not interested in developing new information extractors but in finding a good way to exploit already extracted information. *Already extracted* means that we will focus on what the actual CoC tools are able to do (or not). This task should have been based on *Soϕia* project which has not begun. But this was not really a problem and our modelling approach lead to an XML format used to synthesize the whole extracted/interpreted information related to one image. The main contribution of this project is the definition of 2 ontologies, one related to the information extracted from the image and the other one to the scene content. This contribution will be used in the project DAFOE that plans to develop an interactive satellite image annotation tool based on these two ontologies.

Moreover it was planned at the beginning of the project to deal with memory aspects. In fact we will prove that our representation is not rigid and can evolve in time according to interpreters actions and keeps tracks of past actions. But the final implementation will be demonstrated in DAFOE project in summer 2010. We did not deal at all with ”knowledge levels” as presented in the first proposal. We did not use the approach proposed by Mihai Costache in his PhD thesis as the memory modelling process.

Task 3 : proposal of a representation tool The final tool can be decomposed into two steps. The *modelling step* is based on the two above-mentioned ontologies. Then the *representation step* uses two kinds of approaches : i) the first one is related to DAFOE project and propose to use ITM solution to query the image content using for example SPARQL queries, ii) the second one is based on hypergraphs representation. Different scenarii are proposed to evaluate the interest for these approaches.

Contrarily to what was firsly announced, the tool cannot be based on Soϕia viewer because it does not exist. Other approaches are proposed but only non-expert users are involved as previously mentioned.

2.3 Participants

The whole study is decomposed into well-separated Task, as presented above. Different internships have been proposed to Master 2 Level students with different degrees of success. 5 students have been involved :

- Hakim Hadioui (under Marine Campedel supervision) was in charge of studying interpreters and proposing an interactive application to interact with them. He brought some elements about interpreters work and tools. No final report has been written. But this task has been completed after this studentship by the meeting of Julien Andrieu in PRODIG lab and the analysis of an old meeting with SERTIT during project EXITER 08. The proposal of a complete interactive application has failed because these meetings analyzes arrived too late but perspectives are presented in conclusion of this work to be able to design such an application.
- Khaled Aouini (under Marine Campedel supervision) was in charge of the image processing ontology. He studied state-of-the-art elements and proposed an ontology. Moreover he demonstrates that such an ontology can be used to provide a description of any image content, in conjunction with a scene ontology. This description is encapsulated in an XML file-format.
- Prisca Plesel (under Marine Campedel supervision) was in charge of the scene ontology. Contrarily to the 2 first students, she is not engineer but linguist. Her main task was to study Corine Land Cover Taxinomy and derive an ontology from this study. We will put in evidence the interest and difficulties of such an approach.
- Haiming Liu (under Marine Campedel supervision) was in charge of a different subject related to feature selection but he helped the other students to get access to the diversity of image processing that can be used to extract information. His own report is not associated to this work because it deals with a different subject.

As already mentioned as students supervisors, Marine Campedel is involved in this project as engineer, researcher and coordinator. Students reports are given in appendices.

2.4 Remarks on this report

This report is written partially in French and in English. It has been done on purpose. French is used when the expressivity of the author is too limited by the use of English or when it corresponds closely to a student work. A summary in English is given for each section written in French.

This report is organized in 5 chapters with different importance. The first chapter is introductory and should help the reader to understand the context of this study and its organization. The second chapter deals with the notion of knowledge and the necessity to collaborate with interpreters. An original enlightening is given using a philosophical approach based on the mental gesture theory (*Théorie de la Gestion Mentale*) initiated and developed by A. de la Garanderie. The third chapter is dedicated to the study of one specific representation tool called "ontology" that we used to provide reference knowledge about : i) how information is extracted from images, ii) what kind of objects can be observed in the image and iii) how all these elements are related. The last chapter presents the different contributions of this work and its perspectives.

Chapitre 3

Notion of knowledge

3.1 Notion of knowledge

Studying Knowledge (with big K) is not simple nor innovative. Philosophers were interested in this notion well before engineers. For example Plato defined Knowledge as an intersection between truths and beliefs. There is a theory called "Knowledge Theory" who studies nature, origins, contents, means and limits of Knowledge and particularly human knowledge.

Many works are dedicated to knowledge analysis, which means the determination of necessary and sufficient conditions to get knowledge. It consists more specifically to establish relations between notions of truth, belief and knowledge and to define procedures to make the difference between *true belief* and knowledge. We are not expert in that domain but it is quite interesting to look at competitive theories about the origin of knowledge : is knowledge coming from experiment (Percept) or from reasoning (Concept) ? There are also many other questions that occur in Philosophy domain and that we also have as researchers or engineers like : what can be known ? Does a method/procedure exist to access to knowledge ?

In French, our domain of interest would be translated as "connaissances" with a plural and not as "Connaissance" like philosophers. Many domains are interested in this notion : philosophy, epistemology (that studies sciences), psychology, cognitive sciences, sociology, ... In our actual framework we restrain ourselves to the study of knowledge(s) related to satellite images and more specifically to the imaged scene and to the image processing results (manually or automatically produced).

3.2 Data, Information and Knowledge

Main idea : there is a growing number of available data. How do we exploit these data to get knowledge ?

3.2.1 Data

Considering the applicative success of numerical data, it is obvious that the quantity of stored data is growing very rapidly. People have access to efficient sensor (cameras, films) to capture numerical images and are able to spread these images all over the world thanks to widely developed communication means. On another scale, space agencies send satellites to get more and more precise images... The global results is that researchers, as well as common people, have access to enormous amount of data that we cannot manage. Even at the "personal photographies" level, lots of people are lost in front of their thousands of images : we all need specific tools to access the desired photographies. In the context of satellite imagery, the problem is of higher importance because of the image size, their quantity (problem of storage capacity evolving quickly) and their "sensible" content. Moreover these images are weakly exploited by common people, certainly because of a lack of tools : we observe that Google Earth and similar services receive a great success because they are intuitive and simple.

3.2.2 Information

At the beginning of the 20th century, scientists (biologists, mathematicians, ...) and engineers converged around the notion of Information. Ronald Fisher proposed in 1936 the first criteria to compute the quantity of information relative to a data feature. His task was to classify flower species and based on a statistics criterion, he developed a new approach able to characterize the discriminative power of any feature. He observed that the less an observation was probable, the more it contained information.

In 1945, on the engineers side, Shannon (and others) proposed to quantify the information contained in a message using the notion of Entropy. This notion was already known in the fluids mechanics domain. Information is now computable and is related to the number of bytes used to encode a given message. All terms in the message are not equivalent and the notion of redundancy becomes central : to reduce the storage cost, it is necessary to get rid of the redundancy but to ensure a good transmission and be robust to noise, some redundancy is necessary. This illustrates the idea of a compromise between coding distorsion and rate, that is still up-to-date in nowadays compression methods.

However, what do information become ? The information exploited in a compression process is generally not visible : on the contrary, image compression is based on a perceptively similar encoding/decoding process. This is not sufficient since final user of the encoded documents (images) is looking for a higher level of information whereas Shannon Information is referring to a very low level one. Many recent studies, including Lionel Gueguen PhD work, try and exploit encoded information to make semantics emerge. This is designated as joint indexing and compression. The relationships between low level representations and semantical contents is not obvious and lots of works have been dedicated to fullfill the ”semantic gap”, mainly based on statistical learning approaches.

3.2.3 From information to knowledge : first intuitions

Structured information leads to knowledge Using classification process, we are able to structure low-level information to sum up this information and make it usable at a new interpretation level.

Unsupervised classification In the context of unsupervised classification, the goal is to group similar data and to separate at most the obtained groups. This is the basis of partition algorithms like KMeans. The obtained groups (clusters) can be related by hierarchical relations (”is-a”), or other types of relations (spatial or not). The global result of this process is a structure based on groups and relations between them. Considering image processing, it is classical to use information such as colors, textures, shapes, to produce a segmentation into homogeneous regions according to these information ; one of the expected result is to extract the scene objects. At least, as already mentioned in the previous section, interpreters use it to obtain a summary of the image land cover.

Unsupervised results are often disappointing since the final user have already advanced expectations that generally do not correspond to a given mathematical model. Any classifier is based on a specific model or fitness to be optimized and this model is barely never corresponding to relevant semantics. However, when following Kyrgyzov PhD thesis [10], we got the intuition that the confrontation of several algorithms can make semantical classes emerge.

Supervised classification and a priori knowledge Generally target classes are known in the context of image interpretation. For example, when dealing with major catastrophes like an Earthquake, the main objective is to help safety teams and provide information about buildings and ways to these buildings (roads). These target classes can be considered as some kind of knowledge about the observed scene, related to some specific expectations about what should be observed. These knowledge can be textually defined using semantical descriptions (taxonomies like Corine Land Cover) or using visual examples, or even statistical models. Then the annotation task is simply the process used to put class labels on the image.

This annotation can then be used to produce a map in the context of cartography. In this case, another knowledge reference is used using symbolic representation. This representation exploits our visual analysis capacity and obey semitics. The main rule is that the used signs/symbols should be consistent in time and from one map to another. Moreover this map is addressed to a potentially high diversity of people who are supposed to be able to decode it easily and get knowledge from it.

Knowledge is evolving and has too be shared Annotation can also be used lated and this supposes that the original elements that produced the result have to be stored (for example, extracted features and classifiers). The retrieval process can be difficult since the context could be different from the initial one. For example, when dealing with manual annotation it could be usefull to keep information about who produced the labels, when, for what purpose, ... During his PhD thesis, Costache [3] put in evidence the interest of getting into account user interest and of having some kind of selective memory. This ask the questions of how to select the stored components, what should be kept in time, with what degree of knowledge. No definite answer is proposed by the literature. Moreover another user could differently formulate his query to retrieve the same target image : for example, is someone is looking for forest, the system should be able to propose him coniferous forest. But this is not possible if the link between both concepts is not explicitly defined. Hence it could be interesting to semantically define the knowledge associated to one applicative domain in order to have a reference knowledge shared by the users.

More generally, we conclude that a common (reference) knowledge representation would be usefull. It has to be consistent over time and being able to evolve.

3.3 Knowledge engineering

A new professional domain Knowledge engineering (also called knowledge management) gather methodologies and technologies to access, identify, analyze, organize, memorize and share knowledge between members of the same organization [1]. This notion of organization (mostly related to firms) reveals that the main users of knowledge management are interested by the management of internal knowledge (generated by Marketing or R&D departments for example) as well as external ones (for example in the context of economic intelligence). The goal is mainly to localize and make visible the firm knowledge and being able to store them, access them and update them, as well as to communicate them, share them, being able to better use them, promote them. This approach is confirmed by the development of specific formations like Master "Systèmes d'Information et de Connaissance de Paris I" (http ://siciae.univ-paris1.fr/) which gives access to this recent domain called knowledge engineering.

Considering that COC is such an organization, it is natural to look into this new domain, which is related to older ones, like artificial intelligence.

Artificial intelligence Knowledge engineering is seen as a subdomain of artificial intelligence (AI). From http ://bat710.univ-lyon1.fr/ñguin/IA/IC.pdf, we read that the goal of this new domain is to increase the intelligence of the pair (man-computer) and should then provide methodologies, tools, softwares, work organization strategies, ...

Following AI, knowledge engineering relies on cognitive sciences. An example of such science is given by Gesture Mental Theory that we used in chapter 4.2 to define a new system able to assist interpreters.

Experiments and Reasoning In the knowledge engineering domain we observe the two main tendencies about where do the knowledge come from i) observation (experiments) or ii) reasoning. Everybody seems to agree on Gruber definition the fact that knowledge databases should be stored, reused and then interoperable, transmitted, shared and allowing reasoning (using for example logic rules). The normalization consortium W3C actually recommends to use OWL ontologies for that purpose. In [2], Bachimont explains the use of ontologies in their formal way and insist on the fact that 3 main steps are necessary to define them : i) the linguistic definition (conceptual task), ii) the formal definition (using for example owl language) and iii) the functional definition.

Research in knowledge engineering Researchers in this domain are not well-known by signal processing researchers, but there are bridges between these two domains. As

[1]This has been translated from French definition on Wikipedia

already mentioned researchers in knowledge engineering are mainly computer scientists specialized in formal logic and artificial intelligence. Numerous conferences are dedicated to this domain, such as KM-based conferences, with KM for Knowledge Management, or KD-based conferences with KD for Knowledge Discovery. Examples of such (international) conferences are : KEOD, OLP (ECAI), ACM, SAMT, CVPR, NIPS, ECML, PKDD, InfoVis, EGC, IFAC Symposium on Automated Systems Based on Human Skill and Knowledge, ... and journals are : "revue d'Intelligence Artificielle", "Revue française de linguistique appliquée", IEEE Trans. Software Eng. Computer & graphics, IEEE trans. On Visualization and Computer Graphics, Journal of the American Society for Information Science and Technology (JASIST), Revue des Nouvelles Technologies de l'Information (RNTI) and many more.

The researchers in this domain are producing methodologies and technologies to access, analyze and structure information to produce knowledge ; they are also interested in platforms modelling but also in theoretical modelling aspects mainly related to linguistics.

3.4 Knowledge representation

Ontologies are not the only knowledge representation tools that can be used. Different attempts have been performed in the past. Conceptual graphs, introduced by Sowa in 1984, were based on description logic able to deal with natural language structures. The graph representation facilitates visibility and can be equivalently represented using descriptive logic. It is based on concepts and relations definitions. Using semantic network, concepts and relations are related to context, language, emotion and perception.

3.4.1 Conceptual scheme and databases

The definition of concepts and relations is related to what is now called "conceptual modelling". This is the basis of database management system and the reference for all people implementing or using the produced database. PLATO is based on a specific and original model which then exploited using PostGreSQL management solution and Python multimedia applications.

When developing such a model, normalized tools can be used like UML (Unified Modeling Language) diagrams.

The problem related to (relational) database management system is that no flexibility, no evolution in the structure, nor reasoning are possible. They are only use for static representations and data storage.

3.4.2 Visualization

People are used to access information using external representations like images. For example, geographic maps are usefull to get an overview of a territory and graphics are generally proposed to sum up numerical values. We are able to quasi instantly and without effort assimilate information that are graphically represented [2] ; this assimilation allows fast error detection, it facilitates hypotheses formulation, and it could assist internal working memory.

Visualization tools adapted to satellite images and extracted information are proposed in Hichem Sahbi 2010 project. These are still part of research areas. Many approaches have been proposed to project multidimensional data onto a 2D or 3D space (for example Principal Components Analysis, Independent Component Analysis or MultiDimensional Scaling, ...). And new work on graph and hypergraph for satellite knowledge representation are also proposed by Soufiane Rital (and Pr. Alain Bretto, Caen University).

3.5 Conclusion

In this chapter we put in evidence the difference between knowledge, data and information. We demonstrated that this notion corresponds to an active scientific field called "knowledge engineering" whose main research actors are from the Artificial Intelligence domain.

We already identified specific features of knowledge : it should be sharable, evolutive. The management of knowledge is not obvious and necessitates a structure allowing reasoning in interaction with a final user. In our case the target task is satellite image interpretation and the final users are i) experts and the goal of the knowledge-based system is to assist him or ii) ordinary people and the goal is to propose interpretation elements able to guide them navigating in the image.

In the following chapter we study interpreters and propose a system able to assist them.

[2] Cf Colin Ware's latest book is Visual Thinking for Design http ://ccom.unh.edu/vislab/CWBio.html.

Chapitre 4

Interpreters

4.1 How to interact with interpreters ?

As final consumers of the major parts of teledetection images, interpreters are interesting us. Who are they ? where do they come from ? How do they work ? Do they need our help as signal processing experts ? Answers are given in this section based on two meetings with SERTIT (december 2007) and PRODIG/Pôle Image laboratory (october 2009), as well as public information mainly obtained on the web. Our main concern in this study is to answer the first question : how could we interact with interpreters ? How could our image/machine learning/reasoning processing tools be usefull to the satellite images end-users ?

Definition

photointerpretation (or photographic interpretation)
(Encarta) : science of identifying photographed objects : the science of identifying objects in photographs, especially in order to determine their potential military or topographic importance
(Merriam Webster dictionnary) : the science of identifying and describing objects in photographs.
(Universalis.fr) : Le terme « photo-interprétation », qui présente l'avantage d'être identique en français, en anglais, et de s'être imposé en espagnol (à l'orthographe près), désigne l'interprétation des photographies aériennes et des images spatiales.

The photo-interpreters are the scientists associated to the photointerpretation domain. In this study, we will only consider applications in the teledetection domain but an opening thought is proposed at the end of this section (cf section 4.1.3). These interpreters are major actors in the geographical information domain, as consumer and producer of such an information. Their role is to create and formalize the link between what is measured by the spatial sensors and the measured objects' characteristics.

4.1.1 Who are the interpreters ?

While considering the high diversity of photointerpretation applications, we observe also the high diversity of interpreters origin : geography, geomarketing, hydrography, urban and land settlement, agriculture, risk management, ... Most of interpreters we met, or heard about, were geophysicists, geologists, geo-technicians, earth scientists, some of them were mathematicians, architect, ... none of them came from the image processing domain.

What kind of teaching ?

Someone who calls himself interpreter is often a geoscientist specialized in interpretation. This specialty is obtained in engineering schools or at Master level in universities. In Table 4.1, you can find some examples of such school and Masters in

France. lots of teaching exchanges exist involving Telecom ParisTech, Telecom Bretagne, ENSTA, AgroParisTech, UPMC, Ecole Normale Supérieure (ENS-Ulm and Cachan), Ecole Navale, Ecole Nationale des Sciences géographiques (ENSG) and many universities.

It is observed that an interpreter should be able to work with different kind of tools : mainly teledetection tools (including image acquisition and processing tools) and GIS/cartographical (i.e. geomatics) tools in order to i) manipulate the image data, ii) confront it to other localized data and iii) produce an interpreted map. These teaching are often involving many schools with an interdisciplinary approach corresponding well to the interpreters specificity. There teachings are completed by a specific practical project leading to a Master Thesis.

Level	Name	Involved schools	Brief description
Licence,Master	SIG	École Supérieure de Cartographie Géographique (ESCG)	cartographic softwares
Master	Cartagéo	Paris I	
Master	Télédétection et Géomatique (TGAE)	Paris VII	
DESU	Systèmes d'information géographiques (SIG)	Paris VIII	website [1]
M2pro	Géomatique,Géomarketing et multimédia	Paris VIII	
Master	STEP (parcours TTS)	Paris VIII	Teledetection et Technique spatiale
Master	OACT	Paris VI	
Master	Interaction Climat-Environnement et Télédétection (ICE)	Versailles St Quentin	
Master	Télédétection et Imagerie Numérique (TIN)	Paul Sabatier (Toulouse III, ISTN)	website [2]
Master	Systemes d'Information - Spécialité Géomatique	INSA (Strasbourg)	

TAB. 4.1 – List of some masters proposed to future interpreters - Only in France.

Here are some examples of such projects proposed under the term "photointerpretation" :
- SPOT images supervised classification and spectral indexes computation.
- Analysis of relations texture-structure in aerial photographies on forest landscapes.
- Color analysis : relate spectral responses and natural as well as artificial (vegetation, building, roads, ...) objects.

Under the term ”Geographical Information System”, we find examples like :

– Spatial reasoning on a collection of segmented imagesR.
– Analysis of agricultural lanscapes - Combination of different data sources (teledetection, aerial photographies, geographical databases) and search for particular objects.

It is obvious that such tools are closely related as it is illustrated in the cited projects. In teledetection and geomatics, there are also specific challenges like for example ”how to construct the sensors in order to get the most precise information ?” (teledetection) and ”how to ensure geographical data to be compatible and well-exploited ?” (geomatics).

As a conclusion we observe that interpreters acquire their specialty at Master level and become researchers or engineers, able to manipulate images and understand how they were produced, and able to produce an analysis of the image content considering lots of other related information thanks to spatial tools.

Where do interpreters work ?

In France many national organisms employ (directly or not) interpreters : CNES, CEA, CNRS, DGA, IFREMER, INRA, INRIA, ONERA, IGN (more specifically COGIT and MATIS labs). Indirect employment is done using ”Bureau d'études” (design offices) mandated for specific applications, or agencies (for example Agence Française de Sécurité Sanitaire de l'Environnement et du Travail, Fédération Française de Randonnée Pédestre) and local authorities (for example Communauté d'Agglomération Plaine Commune). Direct employement as researchers or engineers is also possible in research laboratories as well as companies : ESRI((http ://www.esri.com/), Blom (http ://www.fm-kartta.fi/aerofilms/en), Infoterra (http ://www.infoterra-global.com/), Fugro (http ://www.fugro.com/), EADS and SPOT Images are examples of big companies employing a high diversity of people including interpreters. In France the AFIGEO association (http ://www.afigeo.asso.fr) collects information related to the management of geographic information. Several (national and international) companies are AFIGEO partners and listed on the website http ://www.geo-entreprises.fr.

AFIGEO recently proposed a seminar [3] on innovations and development of geographic information all over the world. In the presentation of Aude Areste Lamendour (IGN FI) and Hervé Halbout (Halbout Consultants), it is shown that this development is increasing inspite of the international crisis content and that this development has to be managed by the states to ensure diffusion and exploitation of the spatial data. M. Craglia (European Commission Joint Research Centre) insists on international initia-

[3] AFIGEO seminar, October 9th 2009 ”Innovations et développement de l'information géographique”, http ://www.afigeo.asso.fr/page_287.html

tives like the proposal of a spatial data infrastructure [4], Digital Earth [5], GEOSS [6] to capitalize on similar aims (sustainable development, environmental and societal applications) and services (data sharing, interoperability, normalization).

4.1.2 Interpreters methodology

Two Examples

We met two different kinds of interpreters who explained their motivation, context and methodology. Two examples of task are described in Table 4.2. The methodology is quite similar in both cases even if applications are not the same, with very different constraints. We identified 4 main tasks (steps) in the methodology ; the two first are related to one major task : get all data relative to the observed scene.

- Step1 : study images
- Step2 : collect (relevant) data related to the observed scene
- Step3 : confront data using spatial knowledge
- Step4 : produce an analysis of the observed scene (map)

Laboratory	Julien Andrieu (PRODIG)	SERTIT
Task	Study evolution of landscapes in Africa	Rapid mapping Boumerdès EarthQuake
Images	Landsat image series around 30/pixel	SPOT and QuickBird Before and after the EarthQuake around 1m/pixel
Used Data (not exhaustive)	Terrain campaign Vegetal physionomy Phenology Landscape typology	web public data photos before catastrophy (old) maps (toponimy) Damage typology
Tools	TNT mips (ACP - KMeans)	Definiens registration (OTB) change detection (OTB) ArcGIS

TAB. 4.2 – Two examples of interpretation tasks used to promote the diversity of applications related to teledetection.

[4] For example INSPIRE, http ://inspire.jrc.ec.europa.eu/ with the last status online since the beginning of October 2009 precising the rules used to store metadata and datasets and technical guidance for the development of network services to ensure sharing and interoperability.

[5] "Digital Earth is a visionary concept, popularized by former US Vice President Al Gore, for the virtual and 3-D representation of the Earth that is spatially referenced and interconnected with digital knowledge archives from around the planet with vast amounts of scientific, natural, and cultural information to describe and understand the Earth, its systems, and human activities." http ://www.isde5.org

[6] http ://earthobservations.org/

Step1 : Image Study The first task of an interpreter getting a new image (or set of images) is to look at it. Since many data are multi(hyper)-spectral, this task is not obvious and many interpreters have a preferred ”view” (using false colors) of the data giving him a first idea of the image content. This step is performed using teledetection (i.e. satellite image processing) tools. At this level the objective is to get familiar with the image content using processing able to extract ”point-of-views” on the data. The processing results are usually considered as poor but they are involved in this first discovery task. Moreover it seems that they are used to get a rough partition of the scene into different major landscapes.

Step2 : scene data collection All data that could be usefull to interpret the scene are gathered. According to the final application, specific knowledge can be introduced (like phenology principles when studying vegetation evolution or damage typology when dealing with damage assessment). Interpreters try and get an idea about the scene landscape and history using many exogen data (public web photographies, maps, social and historical data, ...). Their expertise is required to filter relevant information from all available data. At this step we see the global interest to get an up-to-date geographical information/knowledge available and interoperable at any time.

Step3 : spatial alignment (confrontation) After getting an overall idea of the scene, interpreters find correlations between their different datasets. For example, in the case of the Earth Quake they confront population settlement to (manually or automatically) detected buildings in order to relate the image information to an emergency measure. In the case of landscapes evolution, correlations can be performed between the observed area in time (using NDVI profil in our example) and war sites explaining deforestation phenomenons. The ”confrontation” is used to validate the gathered information and produce new knowledge using simple correlation measures, learning or reasoning approaches. At this level GIS are required as well as user expertise to analyze the produces results.

Step4 : produce maps The last step is related to the way the expertise is promoted. The interpreter activity is generally centered on a wide application, going far beyond his personal scientific interest. It is clear that the interpreter is used as a bridge between image producer and decision makers. Their works tend to produce analyzes that are understandable by these decision makers. Maps are nowadays the traditional way to present interprete’s results. Specific cartographic tools can be used to assist map creation with problems of symbols choice, color layers, quantity of information that can be represented, ... It is also important to ensure that this synthetic and specific view of the scene can be reused in other applications, then to ensure its durability in time and availability through services (geoportals for examples).

Multimodal Information

Various kind of information are considered by experts. Of course they can deal with a high diversity of images (optical, SAR), with different spectral contents (panchromatic, multi(hyper)-spectral), different resolutions, pointing on very different types of sites, at various dates, ... Products obtained from image providers (like SPOT image) are combining images and metadata information related to the image production (sensor and related parameters, location, date, applied processing).

This information is not sufficient for an interpreter ; he needs to create his own mental map of the scene, be able to describe this scene in terms of landscapes and landcover classification, in terms of identified objects and their relative positions. Considering his own expertise and application goals, he will his attention to specific knowledge that can be more or less easy to acquire. This process is completely expert-depending as well as resource-availability depending. For example toponymy is usefull to produce localized maps in case of major catastrophe but this type of information is not available nor up-to-date at any time all over the world. This is one great challenge of our world : let geographic information be available to anybody at any time anywhere in the world. Google participates to such challenge while developing Google Map and Google Earth services and as already mentioned international organizations like GMES are also involved. GPS (Ground Point Systems) embedded in individual cars are also consequences of such challenge and international policies.

All these information are intrinsically multimedia. Some of them are images : satellite images but also simple geolocalized photographies. Other are contained in (numerical) atlas, or in GIS (Geographic Information Systems). Some information are specific to an application domain and are only accessible to an expert through high level publications. We can imagine also that in case of major catastrophes, videos, recordings can be obtained from the population on the damaged site (using mobile phones for example) or recorded before the catastrophe... Dealing with such a high diversity of information supports is quite challenging and no tool is available at this time to store and automatically analyze all these information : the interpreter expertise is necessary and unavoidable.

Usual Tools

As already mentioned when considering interpreters teaching, two main tools are referred by interpreters : teledetection (including image processing) tools and GIS/cartographic tools. Table 4.3 refers to mainly used tools. Teledetection tools are used to get an overview of the image content, while GIS tools are used to spatialize and confront information sources. Cartographic tools, often related to GIS (like ArcGIS tool) are used to produce the final map for an end-user. Contrarily to what was planned at the beginning of this project, the tools were not individually tested, only public information on websites have been considered up-to-now.

It is important to note that interpreters use image processing tools without being

Image processing Teledetection	OTB ENVI-IDL TNT mips Definiens ER-Mapper	http ://www.orfeo-toolbox.org http ://www.microimages.com
Geomatics	MapInfo ArcGIS Grass TNT mips Savane	 http ://www.microimages.com

TAB. 4.3 – Tools usually used by interpreters at different analysis steps. This list is not exhaustive ; it is based on tools mentionned by interpreters we met.

expert users and they usually consider automatically produced results as unsatisfactory. For example, Julien Andrieu applies KMeans for two main reasons : i) KMeans is very easy to use and tune, ii) KMeans is well-known in his scientific domain but he knows that the obtained classification is poor. However he will base his work on it to avoid using "black-box" processes and being criticized on his scientific methodology. At the end, using a specific processing chain (based on a cascade of KMeans and PCA tools), he will be able to produce concise indicators (NDVI curves for each landscape class) reflecting his scene analysis with a highly reproducible experiment. Using these simple tools is a way for him to get familiar with the data and to know how to extract significant information from the image : he can then prove that this information was included in the original image.

Simplicity and communicability

The use of tools is motivated by their simplicity, availability and peers assessment.

For example, KMeans is the only classification algorithm presented in Paris VII Master level. This is understandable because the students are usually not coming from scientific classes but mainly from geo-or social sciences and KMeans is considered as a well-understood reference algorithm. Hence new generations of interpreters will prefer this kind of classification tools and promote it to the next generation.

Considering peers assessment is related to the same idea. When using new tools, an interpreter is submitted to the judgement of his colleagues who do not use and are then suspicious about it. Was it correctly used ? How can I ensure to reproduce it by myself ? ... Moreover an interpreter usually needs to be confident that he does not loose his time while employing something new. This question of time could be critical in case of rapid mapping application for example.

The processing chain used by an interpreter should then have two major characteristics : i) it has to be simple and understandable by his community, ii) it should support the expert's analysis. The goal of this chain is generally to produce a concise representation of the image content ; it helps the user not to say too many things about the observed scene but to restrict himself to what is represented in the image.

We already see that proposing new tools to interpreters is a hard task. It should be done with a very interactive and didactic approach in order to give the impression of simplicity and easiness of use.

Publications and promotion

In Table 4.4 we refer some of conference and journal used by interpreters. The main idea is they are application-oriented publications, as well as methodological articles.

Name	In French	In English
bulletin SFPT	x	
VertigO	x	
CyberGeo	x	
revue de Teledetection et photogrammetrie	x	
congrès de stéréologie	x	
transactions in GIS "Remote sensing and urban analysis		x
Remote Sensing of Environment (Elsevier)		x
Photochemical & Photobiological Sciences		x
International Journal of Remote Sensing		x
Applied Optics		x
IGARSS		x

TAB. 4.4 – Examples of journals and conferences used by interpreters to promote their works. This list has been established using public bibliography of interpreters. This is not exhaustive.The main idea is that interpreters publish their work in applicative domains as well as interpretation (photo or teledetection) domain.

It appears globally that interpreters are highly specialized in one applicative domain and use teledetection as a tool to produce image interpretation related to what they know about the reality of the geographic scene. Considering this expertise, specific knowledge is required. For example, hydrographs and urbanists using photo-interpretation or teledetection will need specific knowledge databases (specific vocabulary, taxonomies, literature, ...) to produce a valuable result ; and these databases will certainly be different. Such knowledge bases are difficult to translate in other languages and this could have an effect on the promotion of their work : we were said that some geoscientists consider their work cannot be translated without loss of information because of the terms specificity, hence they prefer not to publish at the international level, which intrinsically restricts the promotion process as well as the reusability of their results.

In this conditions, interpreters could appear as mysterious people, badly known but extremely usefull in the interpretation of spatial data. The geoscientists who do not publish at the international level consider more fruitfull to interact with inter-disciplinary researchers/engineers than confronting themselves to other international interpreters.

4.1.3 Suggestion of interactions

We recall that the object of our study on interpreters is to be able to identify possible ways of interaction. It has already been mentioned that the tools used by interpreters have to be easy to use, simple to understand and highly interactive so that interpreters have the impression to master the tool. Considering our own activities inside the Competence Center, it appears that we could propose such tools at different steps of the interpreter analysis.

Step 1 : Image mining tools

During Step 1, the interpreters get used to the image data. He uses simple visualization tools as well as fast classification algorithms (like KMeans) to get a first idea on the image content. This step is close to the product "zonage grossier" (like rough landscape classification) we proposed during project EXITER08. The alternative we can propose to PCA+KMeans approaches are based on feature extraction, feature selection, model selection, supervised and unsupervised learning approaches, interactive or not processes.

Considering the high diversity of our tools, interpreters could get lost and it is then necessary to be able to propose a priori processing chains when considering a given set of images. This a priori selection is now manually performed by the more advances experts (like in SERTIT) based on previous experiments. One of EXITER 2009 project is dedicated to the evaluation of such chains considering several input image types.

Step 2-3 : Spatio-Temporal knowledge representation and management

The current project is related to this step : knowledge gathering and representation of the scene. Interpreters are facing lots of information related to one scene and they currently have no tool to create a global/complete representation of the available information. GIS are actual tools to represent the spatial information and knowledge but they cannot ingest anykind of data, and only give access to ingested data through visible layers.

What should be the main characteristics of such a representation ?

- Complete : it should contain all the necessary information selected by the interpreter ;
- Easy to manage : services should be proposed to access the stored information and manage it ;
- Alive : the representation can be updated and used as a memory (past experiments are stored) ;

– Adaptive : the representation and related services should adapt to the user needs and data.

In this project we will study the ability of two main tools in order to provide representation and managing tools at knowledge level. We will deal with ontologies and hypergraphs representations.

Interaction with interpreters vs non-expert users

The two types of interactions we propose are not specific to a user type. In fact photo-interpreters, teledetectors or non-expert users could deal with such a vague description. The main concern will be to give the right tools to the right persons.

For non-expert users simple chains (i.e. without parameters) with high interaction will be recommended.

4.2 Mental gestures and knowledge acquisition

4.2.1 Mental gestures Theory in few words

Summary : The theory of Mental Gestures (called Gestion Mentale and introduced around 40 years ago by Antoine de La Garanderie [6]) concerns how/why/when people are able to formalize their own mental representations, to think, memorize, imagine, understand. These actions are called mental gestures.

La théorie de la Gestion Mentale (GM) est donc apparue dans les années 70. Antoine de la Garanderie [6] l'a développée et continue de la porter ainsi qu'en témoigne le colloque d'octobre 2009 [7], au cours duquel il a apporté de nouveaux éclairages. L'objet de cette section de notre rapport n'est pas de faire l'éloge de cette théorie mais de montrer en quoi cette approche philosophique et pédagogique, nous permet de mieux comprendre comment travaillent les interprètes et comment interagir avec eux.

La GM a comme objectif principal la valorisation de la capacité de tout être humain à se développer au cours de sa vie, en particulier en apprenant à mieux exploiter ses capacités mentales. C'est un don à la portée de tous. La GM ne fait pas qu'énoncer des grands principes et trouvent sa force dans l'application méthodologique de ces principes : des projets locaux (portés par quelques enseignants ou thérapeutes isolés), nationaux (essentiellement en Belgique où une association internationale de Gestion mentale s'est créée, elle s'appelle IIGM) et internationaux (cf par exemple les projets européens Co-nai-sens [8] et Signes Et Sens [9]) se développent depuis donc environ 40 ans et apportent des solutions pratiques généralement liés à des problèmes d'apprentissage.

La GM est l'étude des gestes mentaux que sont : la mémorisation, l'imagination, la compréhension, la réflexion et l'attention. Afin de répondre à des objectifs issus du "monde extérieur" chaque être fait appel à une stratégie mentale (donc "intérieure") mettant en oeuvre, de façon unique et originale, ces gestes mentaux. Bien que chaque stratégie soit unique, cette opération repose sur 4 paramètres principaux qui sont : le vécu, les codes (nombres, mots, ...), les liens logiques (raisonnement par exemple), les liens imaginaires (associations inédites). Diversifier et enrichir les éléments personnels dont nous disposons au niveau de chacun de ces paramètres permet d'être plus performant dans l'utilisation des gestes mentaux. La GM donne des méthodologies dans ce sens.

4.2.2 How can we get access to knowledge ?

Summary : The GM theory provides practical methodologies to access knowledge.

Sens et référence Dans [4], Antoine de la Garanderie détaille comment la Gestion Mentale permet de comprendre les chemins de la connaissance. Il met en évidence que

[7] http ://www.iigm.org
[8] http ://www.conaisens.org/
[9] http ://www.signesetsens.eu/

la clé vient du "sens" à donner aux actes de connaissance. Ce qui est particulièrement intéressant est que cette quête de sens est individuelle et personnelle et a priori non consensuelle.

Cette vision du sens rejoint tout a fait celle A. Bentolila pour qui, le signifié désigne une conceptualisation, une idéalité sémantique déclenchée par le choix d'un mot plutôt qu'un autre ; il est unique et partagé par tous sans être explicité, c'est une référence commune qui se construit au fur et à mesure de notre apprentissage du langage. Le signifié n'est pas une image du réelle, ni une représentation qui implique une action personnelle. Le sens, quant à lui, n'est pas partagé par tous mais est coloré par la vie de chacun et il ne peut être exprimé richement qu'à l'aide du langage.

Donner du sens est donc personnel et nécessaire à l'appropriation personnelle des objets et êtres qui constituent le monde.

Importance du langage et confrontation à l'autre Selon Alain Bentolila, lors de sa conférence au colloque de gestion mentale [10], "le langage est un conventionnel commun qui va vers l'intime" ou encore "le langage est une pure abstraction sur laquelle chacun construit quelque chose de spécifique, d'individuel", "le langage sert à dire ce que nous pensons du monde". Plus encore, le linguiste affirme que le langage est construit pour que nous puissions exprimer nos visions individuelles du monde. Il constate que là où règne le consensus, on constate un appauvrissement du langage. Le langage est donc l'outil de communication fait pour le partage de représentations singulières et pour tirer de cette confrontation une intelligence collective, la connaissance.

"L'objet de connaissance est le fruit de l'acte d'attention", [4], p17 Antoire de la Garanderie s'est attaché à donner forme à ces gestes mentaux, ie à la caractériser précisément et à en détailler les mécanismes de telle façon qu'il est possible à un enseignant d'apprendre à une élève à être attentif, à mémoriser, à comprendre... et non simplement de lui dire "concentre-toi" ou "si tu le veux vraiment, tu peux" car ces deux dernières injonctions ne sont pas explicites pour l'élève et peuvent le bloquer dans son acquisition des connaissances.

Il n'est pas possible de détailler tous les gestes mentaux mais il est intéressant de préciser l'acte d'attention qui "aurait pour tâche de mettre en évocation les choses qui sont perçues" ([4], p17). Mettre en évocation signifie (un peu rapidement) *rendre présent dans sa tête*, faire exister une chose ou un être du monde réel dans son monde intérieur, mental. Plus l'attention est importante et plus l'évocation est riche ; cette évocation est nécessaire à tous les gestes mentaux. La GM permet de travailler les évocations et de rendre conscient un processus que nous faisons généralement sans y penser vraiment ; cette prise de conscience est la première étape vers la connaissance. La mémorisation est un processus qui installe ces évocations dans le futur, dans le but

[10]Titre de l'intervention de Alain Bentolila au colloque international de gestion Mentale : "le verbe pour dépasser l'oeil".

d'une réutilisation ultérieure. La compréhension, quant à elle, résulte de la comparaison entre les objets mémorisés et la chose perçue à un instant, dans le but de faire surgir des intuitions de sens, ie des hypothèses (identité, similitudes, différences, rapports de cause à effet, ...) sur ce nouvel objet. Le geste de réflexion permet de se servir de ces évocations maintenues par mémorisation ou en présence par la compréhension ; alors que l'imagination permet de faire apparaître des aspects nouveaux.

Il est intéressant qu'en GM l'acte d'attention "procure le sentiment de co-naissance, de naissance à la présence de ce quelque chose [...] Le sentiment que l'on éprouve est un sentiment d'agrandissement, d'un enrichissement d'un plus-être de soi".([4],p77) Il n'y a pas a rechercher une motivation particulière, l'acte lui-même d'attention procure un plaisir à être présent aux choses perçues. Ce ressenti est important en Gestion Mentale et aide à effectuer les gestes mentaux.

4.2.3 Comparison with interpreters methodology

Summary : Interpreters are highly expertised and their cognition process is unique.They construct their interpretation using their own capacity to evocate, memorize, understand, think, imagine. The richness of this interpretation is related to the richness of their own mental evocations and the quality of the final map depends on their language and to their faculty to communicate with other people.

Du commun vers l'intime L'objet examiné par l'interprète est une image satellitaire dont il doit produire une carte, reflet de son interprétation. L'objectif et l'objet perçu sont donc bien définis. Dans la section 4.1.2 de ce rapport, nous avons examiné le comportement des interprètes et les différentes étapes auxquelles ils procèdent. Nous pouvons y retrouver tous nos actes mentaux.

Dans un premier temps, l'interprète se confronte à l'image ; c'est l'étape d'attention. Il la découvre, la manipule à l'aide de divers outils afin de la rendre présente dans sa tête et de la mémoriser (la garder en réserve pour le futur). Il fait des allers-retours entre son image mentale et la réalité afin d'être le plus fidèle possible à l'objet et plus encore, à la scène imagée.

Il la confronte ensuite à d'autres éléments dont il a connaissance. Il essaye donc de comprendre à l'aide de retours sur ses propres connaissances ou d'autres éléments qu'il explore en parallèle. Comprendre l'image n'est pas simplement lui coller des étiquettes, mais plutôt émettre des hypothèses d'interprétation, que sa réflexion peut mettre ensuite à l'épreuve par du raisonnement logique, de l'inférence. Tous les gestes mentaux interviennent successivement et itérativement selon un itinéraire mental qui est propre à l'interprète.

Une fois satisfait (sentiment de compréhension) par sa vision mentale de l'objet, l'interprète doit encore produire une carte, ce qui nécessite une action d'imagination créatrice afin de rendre perceptible le résultat de son travail mental.

L'ensemble de ce processus est personnel, il repose évidemment sur le socle de connaissances déjà acquises par l'expérience et est motivé par une recherche personnelle de sens envers cet objet. Il est considéré comme expert à cause de la richesse et de la fiabilité de la carte produite finalement et/ou de la notice explicative qui va avec.

Nous voyons clairement que les outils de traitement de l'image peuvent servir à produire de nouveaux points de vues sur l'image, que l'interprète pourra ou non considérer. Ces points de vues ne sont pas directement mentaux et n'accélèrent pas l'itinéraire mental de l'interprète qui doit intégrer ces nouveaux objets : ils peuvent même le ralentir s'ils sont trop différents de ce qu'il connaît déjà. En outre, ces points de vue n'ont d'intérêt que s'ils permettent à l'interprète de générer de nouvelles hypothèses, ie s'ils présentent des caractères différents. Ces observations issues de la gestion mentale expliquent pourquoi les outils innovants ont du mal à être utilisés par les interprètes : c'est parce qu'ils nécessitent une phase de mémorisation et de compréhension de la part de l'interprète, phase qui peut être coûteuse en temps. En outre ils expliquent également pourquoi les interprètes aiment utiliser des outils simples même peu efficaces : ces outils sont bien intégrés et présentent des points de vue qui leur posent question (car ils produisent un résultat différents des attentes a priori de l'interprète) et les aident à progresser dans la compréhension de ce qu'ils voient dans l'image.

La spécificité du langage L'étape finale, qui consiste à produire une carte est également problématique. En effet nous avons mentionné que seul le langage, en tant que référence partagée, permettait d'accéder à l'individualisation et à la confrontation des vues du monde. Dans le cadre du laboratoire PRODIG, il nous a clairement été dit que certains interprètes français revendiquaient un lexique, un vocabulaire précis, développé au cours des années et qui n'étaient pas transposable facilement en anglais, ce qui justifiaient une faible volonté de publier dans la littérature internationale. Ce langage particulier donne une cohésion au groupe d'interprètes qui le partagent et qui permet de les qualifier d'experts, cependant si l'on suit le raisonnement de monsieur Bentolila, il va s'appauvrir s'il n'est pas confronté à des réalités suffisamment différentes.

En outre, le passage à une carte nécessite est un acte créatif d'inventeur, selon la Gestion Mentale ie qu'il donne naissance à un objet réel qui n'existait pas avant et qui résulte d'un processus mental complexe et personnel dont l'interprète a une représentation mentale. Ainsi que nous l'avons vu, à l'aide d'un langage enrichi, l'expert est en mesure d'expliquer sa pensée, mais que se passe-t-il lorsque le résultat attendu est une carte ? Une carte est généralement constituée d'une image de référence sur laquelle sont superposées des informations symboliques. Ces symboles sont également les objets de nombreuses années de recherche puisque le sens véhiculé doit être clair (sans ambigüité), précis (pas de nouvelle interprétation à faire de la part des utilisateurs de la carte) ; la carte est un support à la prise de décision ; c'est une représentation permettant à des experts et des non-experts en interprétation d'images de communiquer. Les symboles doivent donc être connus par tous ou accessibles à tous. Nous n'avons hélas pas approfondi la question, mais il semble intéressant de poser la question si la façon actuelle de faire des cartes est l'unique façon de procéder. La sémiologie tente certainement de répondre à cette question, mais hélas nous n'avons pas eu le temps de nous y intéresser plus.

Nous conclurons ce paragraphe sur une ouverture innovante issue des propos du professeur Bentolila : "c'est le verbe qui porte l'intelligence humaine et non le nom [car] l'explication demande le déploiement qui ne peut pas être donnée par le nom". Ceci tend à nous faire prendre conscience que mettre simplement des noms communs sur une image (par exemple des noms de paysages "land cover") est stérile et ne porte pas suffisamment d'information. Par contre, une recherche des verbes nécessaires à la description d'une image nous aiderait sans doute plus et c'est ce que nous avons tenté d'approcher en décrivant les processus d'extraction de l'information contenue dans les images. Mais il est certain que nous ne sommes pas allés assez loin et que cela constitue une perspective intéressante de notre travail de conceptualisation de la tâche d'interprétation.

Consensus et connaissance Enfin, il nous semble important d'insister sur le fait que la différence permet d'accéder à la connaissance. Ceci rejoint l'idée de Fisher qu'une observation est plus informative lorsqu'elle est rarement observée. Ainsi divers savants nous proposent la même assertion : c'est dans l'autre, le différent que nous apprenons, que nous évoluons ; le consensus (la "connivence", dirait le professeur Bentolila) entraîne plutôt un appauvrissement. Ceci est un encouragement pour nous à approfondir le travail initié par la thèse de Ivan Kyrgyzov [11, 10] et à étudier la richesse obtenue en comparant différents résultats de classification par exemple. Le premier résultat était l'observation de ce qui était consensuel, commun aux différents résultats de classification mais nous avons toujours revendiqué, dans ce travail, l'intérêt de comparer chaque classification à ce "commun" afin d'en extraire la différence. Ceci avait produit une nouvelle information, appelée "mesure de stabilité" associée à chaque cluster et à chaque pixel de l'image étudiée.

4.2.4 Could a computer provide image interpretation ?

Summary : the main idea in this section is to try and answer the question : is a system able to learn how to interpret images ? Using the Mental Gestures (GM) Theory, we propose and construct an ideal learner and we observe that the main questions are : "how can the system interact with objects and people ?", "how to give a meaning to the system actions ?", "how could he be evolving ?".

Relation entre le monde et le système Nous avons vu qu'une étape essentielle de l'accès à la connaissance est le fait d'être mis en présence du monde (les êtres et les choses) à l'aide nos sens, de nos perceptions, donc de notre corps et d'être capable d'interagir avec ce monde. Dans le cadre d'un système, d'un ordinateur, il est classique de considérer que les périphériques de l'ordinateur sont ses capteurs perceptifs : la "webcam" représente la vue, le microphone est son ouïe, la souris et le clavier seraient liés à ses capacités tactiles. Il est remarquable de constater qu'actuellement ces périphériques permettent au système d'ingérer, de percevoir mais comment l'ordinateur *répond*-il à ces perceptions ? la webcam peut suivre la personne qui parle, le microphone peut adapter la plage de fréquences perçues à la personne qui parle, et l'écran peut être utilisé pour afficher un résultat de traitement interne, les haut-parleurs peuvent produire des sons.

Mis en présence du monde (geste d'attention), nous nous faisons ensuite une représentation personnelle de ce monde (appelée évocation en GM). Pour le système, il nous faut donc travailler à une représentation interne des choses et des êtres avec lesquels il est amené à interagir. Nous l'appellerons base de connaissance. Nous avons déjà dit que cette représentation reposait sur 4 paramètres : P1 = le vécu (expériences personnelles ou non, du réel), P2 = les codes (mots, nombres, ...), P3 = les relations logiques et P4 = les relations imaginaires. Le système est plutôt aguerri et efficace en paramètre P2 et P4 du fait de sa construction interne : au final, tout n'est que code et agencement de codes selon des règles précises. Cela suffit-il ? nous ne le savons a priori pas mais la GM dit que plus les différents paramètres sont sollicités, meilleure sera le résultat de l'apprentissage. Le vécu pourrait simplement être représenté par les résultats de la *perception* du système ie des images, textes et sons i.e. les données brutes qu'il peut ingérer et que des processus de traitement (indexation, classification) pourraient organiser : cette partie du système pourrait donc ressembler à un PLATO ou un KIM étendu au multimédia. Du point de vue des êtres, ie des personnes interagissant avec ce système (utilisateurs finals, techniciens liés à sa maintenance, les ingénieurs liés à sa conception, ...) il devrait également conserver une trace : cette trace peut être réduite (login/mot de passe) ou plus complexe si l'on pense à des systèmes qui s'adaptent à leur utilisateur, ainsi que l'avait par exemple pensé Mihai Costache [3] ou encore en suivant implicitement les actions effectuées par l'utilisateur sur le système (traces).

Cela signifie qu'il faut donc conserver l'historique des perceptions des choses et des êtres en vue d'une réutilisation future (geste de mémorisation). La capacité a faire des liens inédits ne semble pour le moment par utile à notre système (ce qui n'est pas le cas de tout apprenant).

Quelle peut être la quête de sens pour un système ? Nous avons également précédemment mis en évidence une quête de sens pour l'interprète, nécessaire à la tâche de compréhension notamment. Cette quête (intérieure) est présente quel que soit l'objectif (extérieur à soi) à atteindre (ici apprendre à interpréter une image), elle est au coeur de tout désir d'apprendre. Notre système devrait donc se doter d'un tel désir. Cela ne paraît pas du tout trivial puisque ce sens doit ensuite motiver (mettre en mouvement) les autres gestes mentaux et l'ensemble des actions du système. Nous proposons, provisoirement, que la motivation du système soit "être utile à un utilisateur" [11].

Nous faisons ce choix "être utile à un utilisateur" car pour l'utilisateur il correspond au désir de récupérer une information pertinente de ce système. Cette notion

[11] Notons que ceci est légèrement différent de "utile à tout utilisateur", qui sous-entend que les utilisateurs sont tous équivalents, ce qui est incompatible avec notre idée d'unicité des êtres.

de pertinence est très importante actuellement dans la littérature et fait la fortune de Google : c'est bien le choix d'un ordre (stratégie originale de ranking) sur les documents retournés lors d'une requête qui a permis a Google d'être le premier moteur de recherche [12]. Des chiffres comparatifs concernant la fréquentation des sites Google vs d'autres sont présentés par comscore http ://www.comscore.com) de recherche mondial et de le rester.

Cependant "être utile à un utilisateur" requiert intrinsèquement le besoin de définir, à l'aide de la représentation interne du système, la notion d'utilité et la notion d'utilisateur, qui sont liées ne serait-ce que par la racine des mots employés. Nous avons déjà remarqué que l'interprète utilise des outils de traitement d'image afin qu'ils lui offrent un point de vue différent sur l'image à interpréter, et l'aident à explorer son contenu. Nous avons également relevé qu'un outil est abandonné s'il requiert trop de temps de calcul ou si ses résultats sont "trop erronés" (ce qui rejoint la pertinence pour un utilisateur). Notre système doit, pour être utile, i) soit être capable d'anticiper le résultat d'un interprète ie le remplacer sur certaines tâches ce qui permet d'accélérer et de fiabiliser les résultats, ii) soit présenter des vues innovantes (porteuses d'une information nouvelle selon lui) des données et les confronter à ce qu'il connaît déjà (ie une information structurée et validée) de la scène imagée pour en mesurer l'intérêt. Dans le premier cas, cela suppose que le système est devenu interprète, ce qui nécessite, selon nous, qu'il soit passé par une phase où il apprend à être un interprète et donc à dialoguer avec des utilisateurs, ce qui implique le second point.

Il est nécessaire pour cela de disposer d'une mémoire riche, construite sur les différents paramètres précités dont des outils de comparaison (mesures de similarités, machines d'apprentissage, classification, outils de raisonnement) qui permettront au système de produire des hypothèses d'explication qui seront ensuite vérifiées et confrontées à la question "ce résultat est-il utile à l'utilisateur ?" (gestes de réflexion et de compréhension).

Si l'utilisateur est déjà très performant seul, l'apport du système pourrait être lié à sa capacité d'étonner l'utilisateur et donc à suggérer des explications inédites ne résultant pas de la logique, mais d'un acte d'imagination. L'idée ici serait donc que le système puisse produire des chaînes de traitement imprévues, en utilisant par exemple de la programmation génétique, mais cela reste à étudier.

La nécessité d'évoluer Nous avons dors et déjà mentionné que la richesse des gestes mentaux s'appuie sur la diversité (et la fiabilité) des éléments qui composent la représentation mentale et tous les liens logiques ou imaginaires qui peuvent être effectués dessus. Pour un système automatique, cela implique qu'il puisse ingérer des données et les stocker après validation. Ce processus ne peut pas être instantané, il implique une temporalité et une interaction pour la validation. Nous appelons ce processus "évolution".

Cette évolution concerne à la fois des éléments d'information et des outils de

[12] Plus d'1 requête sur 2 sur internet se fait sur Google, à l'échelle mondiale et plus de 80% des requêtes françaises sont adressées à Google.

traitement. Par exemple, nous pouvons imaginer facilement l'ingestion de nouvelles images liées à une même scène, ainsi que les metadonnées qui leur correspondent, et éventuellement des résultats de segmentation obtenus avec plusieurs algorithmes et des résultats de filtres de Gabor. Il est intéressant dans ce cas de donner de l'information au système sur ce que sont les algorithmes de segmentation et de filtrage qui ont été appliqués et éventuellement sur les paramètres employés. Où s'arrêter dans la précision des informations n'est pas clair mais certaines informations disponibles au temps t peuvent disparaître et se révéler utile en t+1 alors qu'elles ne sont plus disponibles. C'est le challenge de la mémoire de faire le tri et de se projeter dans l'avenir i.e. dans la réutilisabilité des informations stockées à plus ou moins long terme.

L'ingestion de données fiables est facilité si la base de connaissances du système peut s'interfacer avec d'autres bases de connaissances, donc en assurant un aspect consensuel, partageable, normalisé à sa structure. C'est cet argument central qui nous a amené à étudier les ontologies.

L'évolution passe également par une validation des données ingérées, en interaction avec l'utilisateur. Ceci rejoint les approches par boucle de pertinence mises en oeuvre par Mihai Costache [3], mais à des niveaux plus variés que sur des résultats de classification : il faut en effet valider des résultats de traitement tels que des classificateurs mais peut-être également les classificateurs eux-mêmes ainsi que tous les outils d'extraction d'information. En fait ceci rejoint la problématique d'EXITER 2009 et nous ne nous y intéresserons donc pas dans ce rapport.

4.3 Conclusion

In this chapter, we studied interpreters and propose to highlight their interpretation task using Mental Gesture Theory. Our primary goal was to observe how interpreters work and then to understand their behavior. The secondary goal was to be able to propose a system able to assist them in their task. The interpretation task is not limited to a labelling task but necessitates to understand why the image scene is the one observed.

The main conclusions are :

- Interpreters use image processing tools and exogenous data in order to form their own mental representation of the scene ; we could assist them in providing innovative and relevant point of views on the data.
- Interpreters are able to interpret an image as soon as they consider that they understand it, which first corresponds to a personal meaning research ; a system could stimulate this mental activity by suggesting explanations hypotheses that are validated by the interpreter or by helping the interpreter to test personal hypotheses.
- Interpreters have to translate this personal meaning on a map ; this task is also assisted by specific tools. We do not propose anything at this level.

The result of this study is the idea of a system able to be usefull to any interpreter. This system has not been developed in the context of this project but it could be

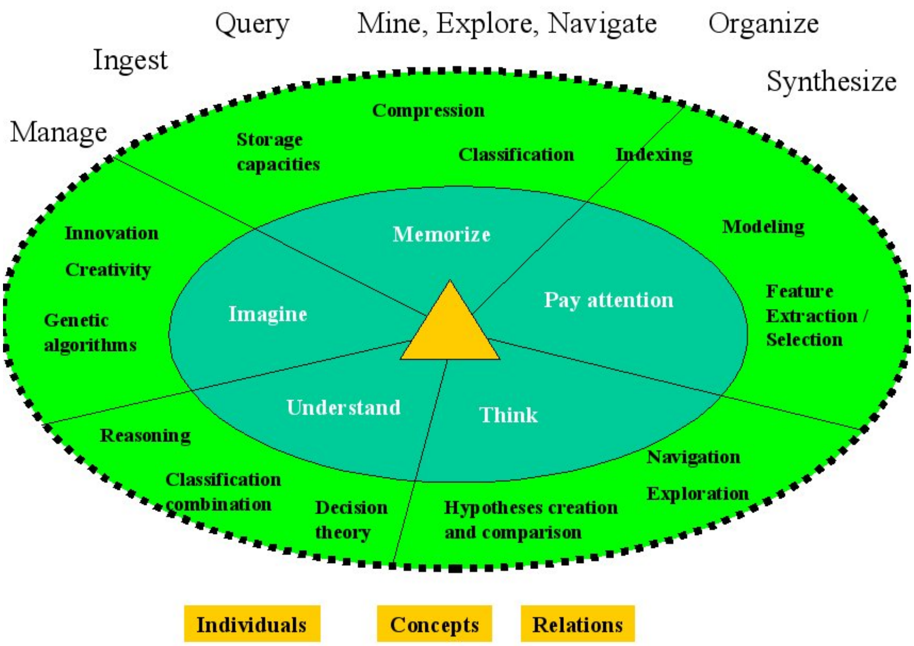


FIG. 4.1 – Mental Gestures Theory-based system. At the heart of the system (figures as a triangle), there is an evolving knowledge representation. Computer abilities and internal tools are figured inside the circle, while functionalities offered to users are figured outside.

a perspective of this project. This system needs a component that we call ”knowledge representation” tool which is the object of the next chapter ; it corresponds to the internal representation of the system knowledge, able to store data and tools about objects of the satellite image world.

It is worthy to note that this system is generic and can be used by non-expert users. But the final objective will be different since the notion of ”image interpretation” does not mean anything to a non-expert. In this case, the final task is merely related to a user-adaptive navigation task allowing the knowledge database exploration.

Chapitre 5

Ontologies as knowledge representation tool

5.1 Introduction

In the preceding chapters of this report, we insisted on the fact that the growing number of satellite images leads to the development of (semi-)automatic procedures to manage their content and assist potential users (including expert ones, called interpreters). The challenge is then to be able to store, to transmit, to query, to index, to retrieve these images while anticipating future needs. Hence we proposed a framework in Section 4.2.4 to design some kind of *interpreter assistant* able to adapt to the final user. In the heart of this system we need a knowledge representation tool.

5.1.1 Semantic Web

In fact many people face similar problem in the world (and particularly in the World Wide Web community) and the organization of data and related information and/or knowledge has become a central problem, leaded by the ”semantic web” idea.

This term has been introduced by Tim Berners-Lee in 1994. It is supposed to be an extension of the actual WWW allowing navigation and use of numerical resources thanks to automatic agents or process. Computer should be able not only to store but to understand information and to bring relevant answers to the user. The semantic aspect comes from the fact that the web resources should be interpretable by people and computers as well.

5.1.2 Ontologies as semantic web tools

The main central tool used to model knowledge is *ontology*. As we will show it in the following, using ontologies is not a new idea, even in the geographical/spatial community and we will propose, in conjunction with DAFOE project [1], to develop 2 ontologies in order to catch knowledge extracted from satellite images (satellite image processing ontology) and knowledge associated to the imaged scene (satellite scene ontology). These 2 ontologies are the tools we want to develop.

Developing ontologies is not obvious, as we will demonstrate it in the following. We got the help of Mondeca and INSERM experts (in the context of DAFOE project) and studied the associated literature. Moreover, considering that ontologies are now well recognized, we used normalized language (OWL, recommended by W3C consortium) and classical editor (Protege). The collaboration with DAFOE will ensure that a complete satellite image annotation tool will be developed for July 2010, based on these two ontologies.

5.1.3 Notion of ontology

Ontologies are only one way to represent knowledge. There are many other (like semantic graphs, expert systems, ...). All these knowledge representation tools have

[1] http ://www.dafoe4app.fr

in common their basic structure ; they all have three main components and are called knowledge bases :

- Concepts : these concepts can also be called classes, or terminology ; it constitutes the long-term memory ;
- Relations : also called properties that relate concepts. These relations are semantical, spatial, ... and can have different classical properties (symmetric, transitive, inverse-of) that can be exploited with a logical reasoner.
- Individuals : called facts or examples ; they are used to populate the classes ; they constitute the short-term memory of a knowledge base.

This reinforces the fact that knowledge bases, and particularly ontologies, organize information, not only by structuring the information space using classes as classifiers do, but also by specifying relations between the identified classes.

According to Gruber [7], ontologies are defined as «explicit formal specifications of the terms in the domain and relations among them». In [8], Hudelot gives her ontology definition (originally written in French [2]) : "An ontology is a formal and explicite specification of a shared conceptualization of a knowledge domain".

- **conceptualization** : this is clearly a central point in the knowledge extraction process. As already mentioned, it should not be reduced to a too simple classification process.
- **explicit** : all concepts and relations are explicitly defined using semantical descriptions and related attributes.
- **formal** : this term is related to formal logic and to the fact that a computer is able to "understand", access this knowledge. Hence formal description is performed using specific expressive languages (RDF, OWL, ...).
- **shared** : this gives us an idea on how to identify knowledge by confronting point of views and making consensus emerge. We already mentioned this idea in chapter 3.
- **domain** : the conceptualization task is reduced to a domain of interest or a given task, problem. The idea of having an ontology of the whole world is now discarded and knowledge scientists prefer to define domain specific shared, reusable ontologies and to connect them using *top* ontologies to relate them.

5.1.4 Conceptual scheme and relational databases

The definition of concepts and relations is related to what is now called "conceptual modelling". This is the basis of relational database management system and the reference for all people implementing or using the produced database. PLATO is based on a specific and original model which then exploited using PostGreSQL management solution and Python multimedia applications.

[2]Une ontologie est une spécification formelle et explicite d'une conceptualisation partagée d'un domaine de connaissance.

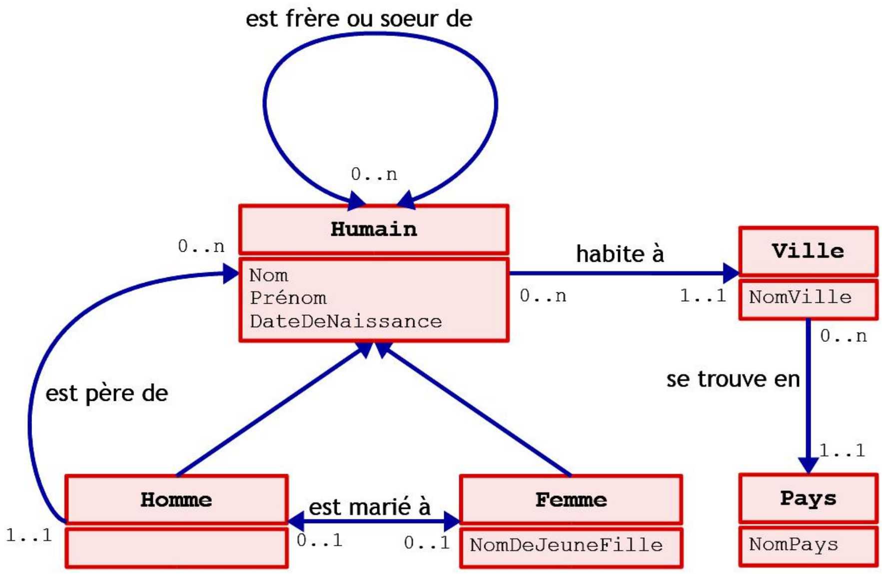


FIG. 5.1 – Example of ontology. Note the diversity of relationships : hierarchical, composition and semantical relations are illustrated.

When developing such a model, normalized tools can be used like UML (Unified Modeling Language) diagrams.

The problem related to (relational) database management system is that no flexibility, no evolution in the structure, nor reasoning are possible. They are only use for static representations and data storage. On the contrary, ontologies can evolve using specific reasoning rules. Moreover even if they can also be used to describe things, ontologies can now be used in conjunction with logical reasoners (Pellet, RacerPro, ...) and allow extraction of new (never-stored) information.

5.1.5 Summary

This chapter is dedicated to ontologies as knowledge representation tools. Two ontologies have been developed and can be used in a further annotation tool that will be developed until July 2010. The next section is dedicated to state-of-the-art ontologies exploited in the image processing domain, including spatial imagery. Then we will present a summary about the ontology construction process and the obtained results. Details can be obtained in the studentship's reports included as appendices.

5.2 Ontologies for image processing

In this section we draw a brief overview about how ontologies are already used in the image processing domain.

5.2.1 Different sources of knowledge

Ontologies are tools to represent knowledge. In our case, we are interested in knowledge related to images, and more specifically satellite images. Different types of knowledge can be identified and related to specific information.

- Domain knowledge : for example spatial or medical domain. The idea is that this specific knowledge induce the use of specific processing tools. Taxonomies can be used to express domain knowledge : in our case, we will exploit Corine Land Cover taxonomy as a basic conceptual structure to construct our scene ontology.
- Sensor knowledge : specific information related to image acquisition can be usefull to pilot processing. In our case, this knowledge is concentrated on spectral (number of spectral bands) and spatial resolutions and induce specific extractors and interpretation tools.
- Meta-knowledge : author, date, scene name, ...
- Image's content knowledge : this is the most challenging knowledge to be extracted. It is related to an image interpretation task, which is detailed in the next section.

5.2.2 Image content analysis - Interpretation

Being able to explicit the knowledge contained in images is still challenging ; but several studies already put in evidence the interest of a mixed knowledge-based and processing-based approach.

The first problem is known as "semantic gap". It is the gap between the perception we have of an image and its content's formal representation (in feature space for example). The image content can be modeled using different abstraction levels (cf Figure 5.2) :

- Image level : this level is the perceptual level, firstly accessible to our eyes and to automatic processing. Pixels can be grouped in segments and features can be extracted. Here is our classical feature space including color, texture, shape (geometry) characterizations.
- Visual level : this level is a semantical description of the image level. It usually implies data classification to put names on group of observations. Spatial relations are used to describe the segments relationships.
- Semantic level : at this level, objects can be identified and related to each other using semantic relations.

In the literature we can find several applications based on both visual and semantic ontologies (called domain ontologies) and exploiting automatic processing as well as

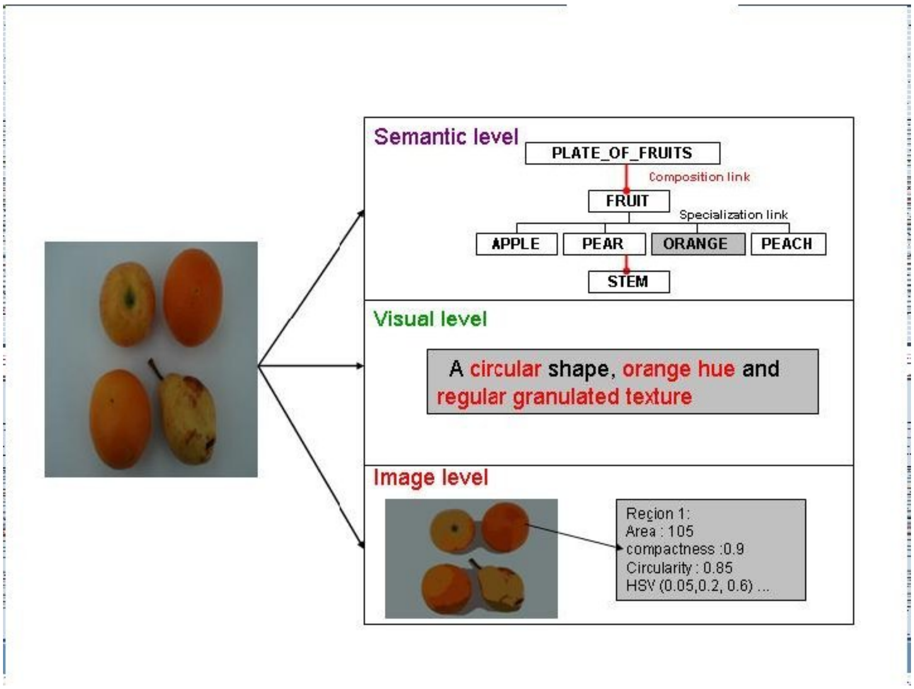


FIG. 5.2 – Illustration of image interpretation levels (extracted from [15]).

reasoning tools. These ontologies can be used as descriptive tools or in conjunction with reasoning ones.

5.2.3 Applications

Visual ontologies

In Maillot thesis [14] and Mezaris "object ontology" [16], we find the same idea of a visual ontology used to describe images content. Figure 5.3 illustrates how low-level feature vectors are related to the visual ontology. This ontology put in evidence object attributes like intensity (color), position, shape, etc... as well as relative position. Values are grouped into classes named with "low", "high" and "medium" terms to enable a given granularity in the semantic description. Unfortunately formal ontologies cannot be obtained, we did not get RDF nor OWL files corresponding to these ontologies. But these two ontologies inspired our own visual ontology designed for satellite imagery.

Annotation tools

As soon as 2001, ontologies were used to annotate images. One interesting example is given by Schreiber [21]. No visual ontology is used but many information on the image's capture, the scene and the object of interest : in this case, the goal is to annotate specific images and being able to query these images using the vocabulary used

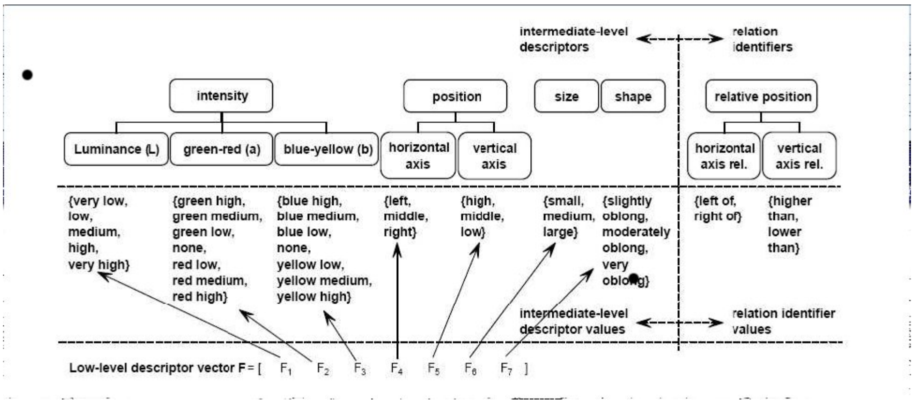


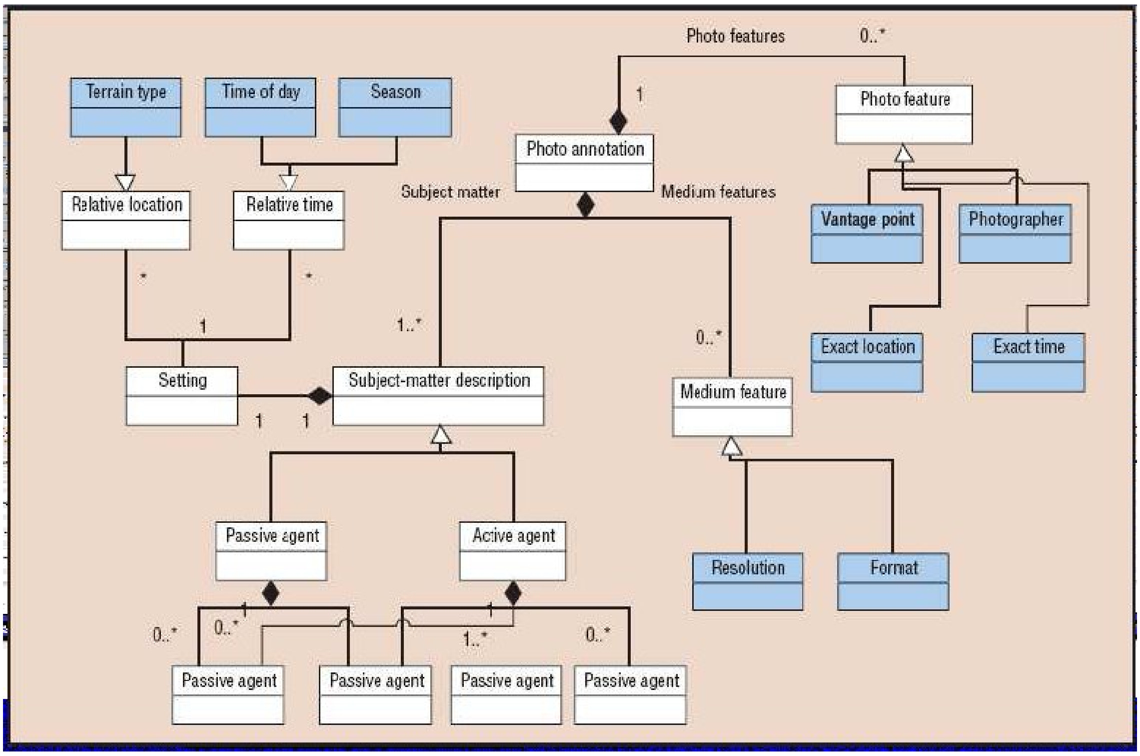
FIG. 5.3 – Sample of Mezaris "object ontology" [16]

in the ontology definitions. There is no reasoning tool but the idea of a comprehensive description at a semantic level.

Annotation is a key application in the context of ontologies applied to image processing. During KSpace NOE project, AceMedia tool has been revised. AceMedia [18] is a project motivated by mining into images and videos databases at a semantic level. It is based on a visual ontology based on MPEG-7 descriptors. Using manual annotation of examples, the system is then able to relate low-level features to semantic concepts defined in the domain ontology. For example, in the sport domain, tennis ball and racket are defined. Manual interaction is used to relate the ball to dominant color and shape defined as MPEG-7 features. This relation (rule) is then used to annotate any image containing a tennis ball and query can be performed to retrieve images of interest.

Grafip : project at Telecom ParisTech

At Telecom ParisTech, in TSI laboratory, projects exploiting ontologies have already been performed under I. Bloch supervision. The central idea is to performed image description while exploiting both knowledge bases and automatic processing. The goal is to provide a relevant description of the image content (using graph representation) based on segmentation tools constrained by knowledge aspects. In the medical domain, many knowledge bases are available like FMA (Foundational Model of Anatomy ontology) and taxinomies like Neuronames that define more than 15 000 terms related to human and macaque brains. Grafip project (cf Figure 5.7) demonstrates the interest of exploiting a priori positions of brain components and relate them to what can be observed through RMI medical images. This necessitates the definition of spatial relations between image regions, what has been done during the collaboration with Hudelot [9]. This spatial ontology is now used in many different projects, including ours.



5.2.4 Ontologies for spatial imagery

Plesel report is dedicated to the creation of a "scene ontology" ie to the semantic definition of scene objects and relations.

Projects

The Towntology project [3] was part of the COST ACtion C21 ; it ended in 2006. The goal in this Action was to produce a taxonomy of ontologies in the Urban Civil Engineering (UCE) field. To achieve this goal an ontology tool suite call Towntology Tool Suite and a set of ontologies have been developed. The first objective was to develop an urban ontology and the second one to deal with the difficulty of defining formal ontologies. That is why a cooperative tool had been designed to develop ontologies insisting on the fact that ontologies are first preconsensual while evolving to become postconsensual ; experts are required to define concepts and annotate images. When looking in details at these tools, we observe that they do not develop strictly formal ontologies but XML files in order to describe urban scenes. This ontology is not clearly adapted to our satellite image interpretation problem since, as illustrated in figure 5.8, the referred objects are not visible from space. But this is a good example to illustrate the capacity of knowledge representation to describe images and facilitate

[3] http ://liris.cnrs.fr/ townto/

FIG. 5.4 – Schreiber annotation tool.

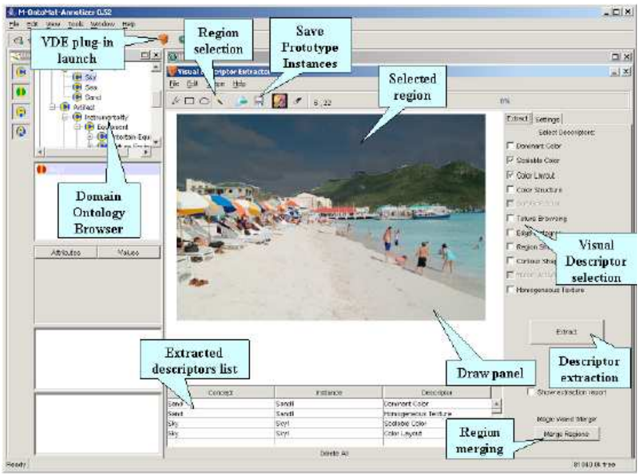


FIG. 5.5 – Ace Media annotation tool.

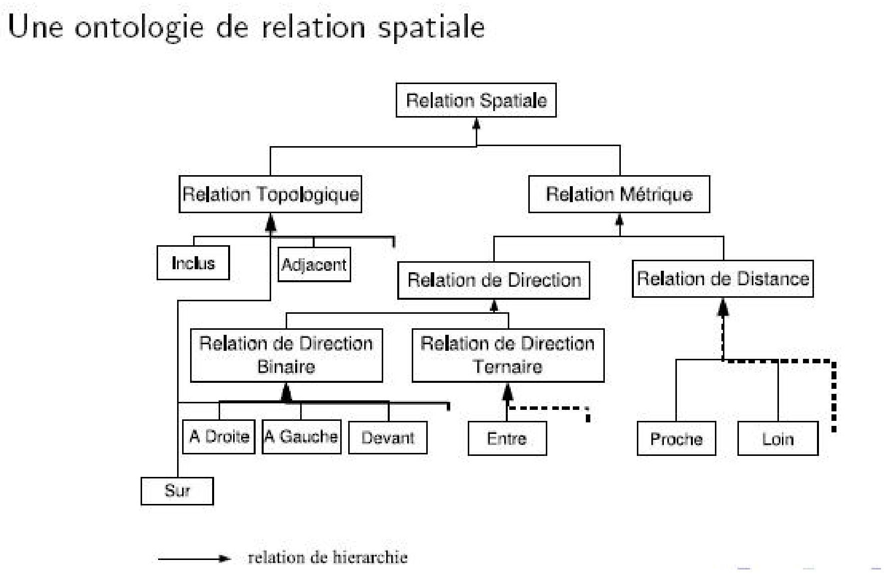


FIG. 5.6 – Relations ontology.

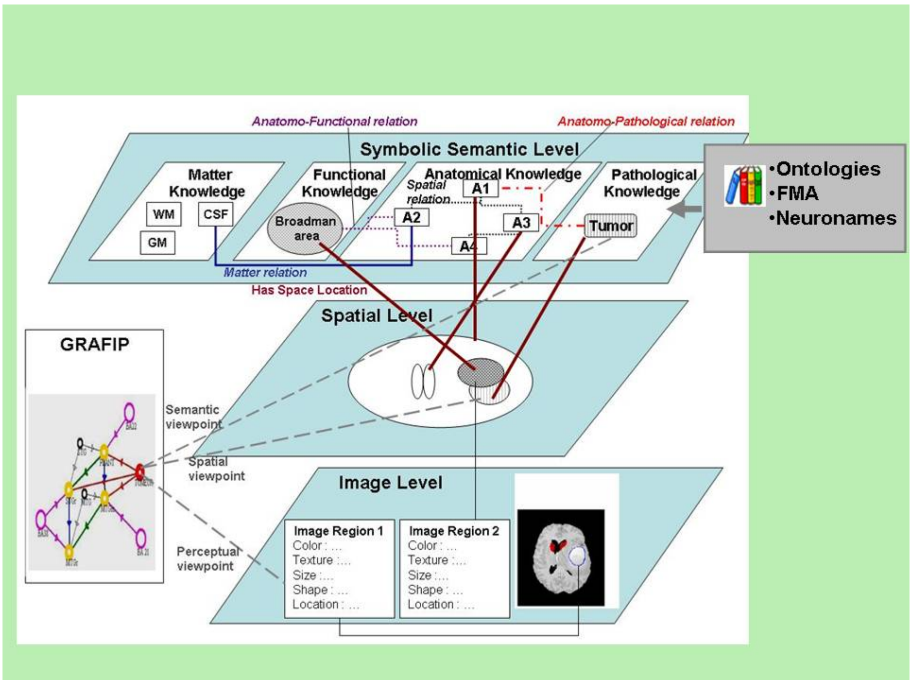


FIG. 5.7 – Grafip project : joint segmentation process and diagnostic tool using knowledge bases.

their browsing.

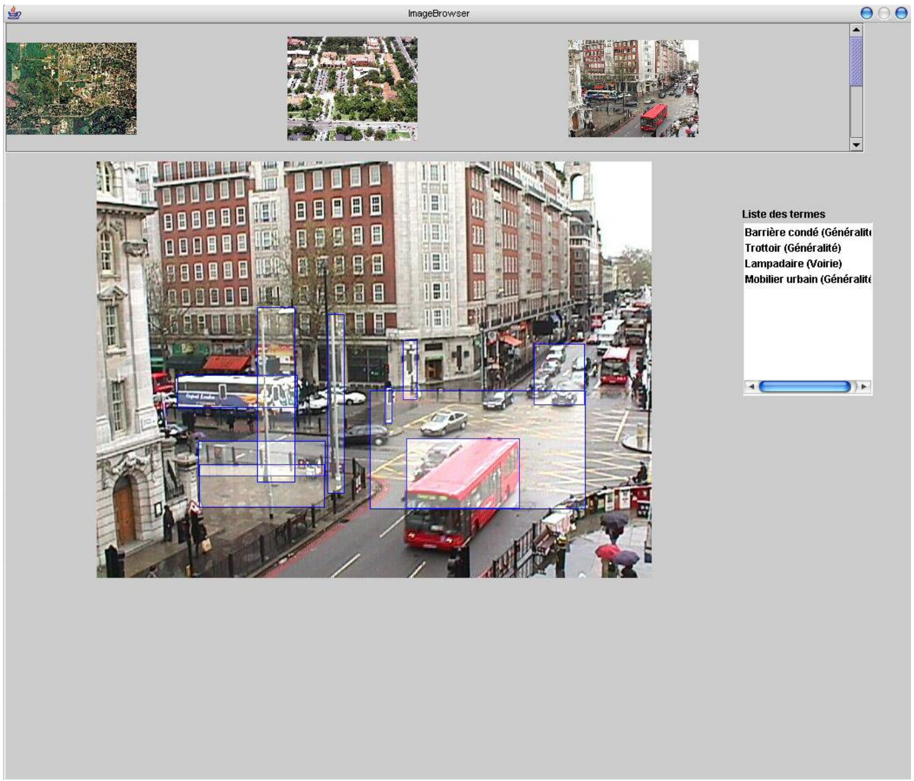


FIG. 5.8 – Example of use for Towntology project.

FoDoMust [5] [4] is another example of ontology development for urban scenes. It was supported by an ANR project between 2004 and 2007. The author idea is to exploit a domain ontology to enhance the VHR satellite image (QuickBird is given as a reference) classification. The domain ontology has been developed by experts and the idea is to give image objects a meaning. A matching process between an object and the concepts of the ontology is proposed. An illustration of this ontology is given in Figure 5.9. We clearly see the two components of the ontology : visually-based (reference to feature extraction process and colors as semantic terms, for example *spectral vs spatial* in the right table) and domain-based concepts are referred in the ontology. In this ontology each domain 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. These values are obtained by machine learning algorithms and are part of concept explicit definitions. When a new image is classified, segmentation and feature computation are performed ; then a matching score is computed between any region and concept pair. The final decision is taken while traversing the ontology from rough to coarse level. We observe again that reasoning is not used and that no complex (nor spatial nor semantical) relation are used.

[4]http ://fodomust.u-strasbg.fr/ ; http ://paristic.loria.fr/content/masse_de_donnees/posters/FoDoMuST.pdf

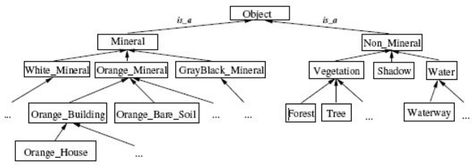


Figure 2. Excerpt of the ontology.

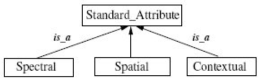


Figure 3. Hierarchy of the attribute classes.

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_NearInfaRed	[34.8-139]
	spectral_signature_SBI	[14.6-60.1]
	spectral_signature_NDVI	[50.2-108]
spatial	diameter(m)	[13-61]
	area (m^2)	[10-600]
	perimeter (m)	[28-116]
	elongation (m)	[1-3.1]
	Miller index	[0.5-0.8]
	Solidity index	[0.85-1]

FIG. 5.9 – FoDoMust project. This figure has been copied from [5]. We observe the two main components of the ontology on both sides : visual (right table) and domain concepts (left tables) are used to annotate VHR satellite images.

Corine Land Cover (CLC) *ontology*

Note that when dealing ontologies, many people, like authors of [13] refer any knowledge base as *ontology* : for example Corine Land Cover and Wordnet [5] are often characterized as ontologies. But they are not formal ontologies in the sense we want to consider it ; they are from one side a detailed taxonomy, and on the other side a lexical database.

However the case of CLC is interesting since detailed maps are provided with textual and explicit definitions of landcover classes. Such definitions are available on-line [17] Both coarse and fine definitions are given as illustrated on Figures 5.10 and 5.11. There are multiple interests in this taxonomy :

- Concepts definitions result from a methodological expertise : photointerpreters all around Europe constructed them through verbal, textual exchanges ie point of views confrontation ;
- The concepts are chosen as visually identifiable by photointerpreters : the CLC classes correspond to usefull geographical objects that can be identified at a given scale through images and geographical databases ;
- From the detailed definitions, many terms appear that are not defined as concepts (for example *structure, transport network, road in another concept, etc.* as well as many relations between terms and appearance in images : we recall that this taxonomy has been designed to define a consensual hierarchy of terms to produce landcover maps for the whole Europe. Hence definitions contain information about how to visually identify regions (according to their colors, size, spatial context, etc.).

[5] http ://wordnet.princeton.edu/

Table 2.2. CORINE land cover nomenclature

Level 1	Level 2	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 units 1.2.2. Road and rail networks and associated land 1.2.3. Port areas 1.2.4. Airports
	1.3. Mine, dump and construction sites	1.3.1. Mineral extraction sites 1.3.2. Dump sites 1.3.3. Construction sites
	1.4. Artificial non-agricultural vegetated areas	1.4.1. Green urban areas 1.4.2. Sport and leisure facilities

FIG. 5.10 – Corine Land Cover (CLC) project, example of rough definition at each level of the taxonomy. This figure has been copied from [17], part 1.

A CLC viewer is available online [6]. We observe the interest to navigate on the map jointly with satellite image representations and to be able to select classes of interest in order to visualize specific objects. This also put in evidence the rough scale of CLC objects : on VHR images like QuickBird, Ikonos or event SPOT5 images, we would like to be able to annotate buildings, runabouts, etc. that are not available on these maps. Moreover the navigation facilities are simple and do not allow complex queries using spatial configuration of classes.

CLC is then a high value resource and it will be for us the basis for a formal ontology definition.

5.2.5 Conclusion

Since the beginning of 2000 years, ontologies for image description, browsing, annotation have been developing. We observe that many concept organizations are qualified as ontologies without being formal ontologies. The consequence is that their use is limited to content description. Moreover the defined relations are often limited to hierarchical relations instead of exploiting the expressivity offered by ontology languages.

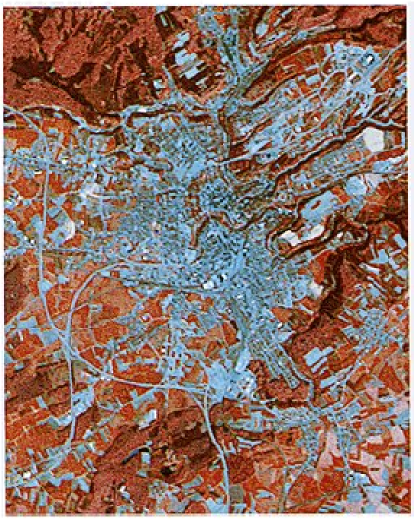
For image annotation, we identified a standard way to proceed when it is desired to exploit knowledge from the image content : define two ontologies, one at the visual level and the other at the scene level. We did not find perfectly suited ontology for satellite image annotation but attempts have been done that can inspire our own work.

The next section is dedicated to the construction of our ontologies.

[6] http ://dataservice.eea.europa.eu/clc/eeaclc.asp

1.1.1. Continuous urban fabric

Most of the land is covered by structures and the transport network. Buildings, roads and artificially surfaced areas cover more than 80% of the total surface. Non-linear areas of vegetation and bare soil are exceptional.



1.1.1. Luxembourg/Area: Luxembourg City
Landsat TM 4.3.2. 1:100 000, August 1989

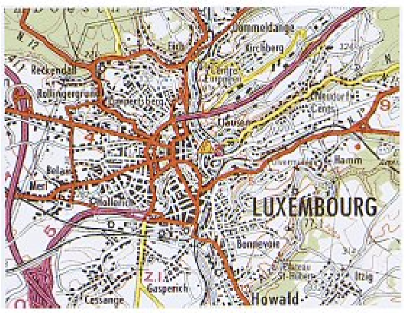
Interpretation

Continuous urban fabric appears blue or a darkish bluegrey on satellite images. Centres of urban districts can easily be identified on satellite images by reference to topographic maps.

In some cases, distinguishing between continuous urban fabric and discontinuous urban fabric can be difficult. The boundary can be set principally by determining the presence and quantity of vegetation.

If an urban district is crossed by a river or road less than 100 m wide, such features are ignored. The area is classified a single unit.

In the case of linear urban construction, even where the constructions situated on either side of the road and the road itself are only 75 m wide, and provided that the total surface area exceeds 25 ha, the area is as continuous urban fabric (or discontinuous urban fabric if the areas are not adjacent).



Topographic and tourist map (scale 1:100 000)

FIG. 5.11 – Corine Land Cover (CLC) project, example of a detailed definition corresponding to the urban fabric concept. This figure has been copied from [17], part 2.

5.3 Ontologies definition

Our goal in this project was to construct two ontologies : one at the image processing level (visual ontology) and the other one at the scene level (scene ontology), based on CLC description. In this section we discuss the employed methodology and the encountered difficulties.

5.3.1 Methodology

Since Knowledge is made explicit through point of views confrontations ; it is necessary to get a multidisciplinary team to work on this project. 4 Master students have been involved with research, engineer and linguistic profiles and a strong collaboration with Mondeca engineers (ontologies specialists) under Jean Charlet supervision (DAFOE leader) has been organized. This team has regular meetings and this dynamic process will go on until the end of DAFOE project and the development of an human-assistant satellite image annotation system. One of the (engineer) students (K. Aouini) has been hired to develop the system until July 2010.

Methodologically, knowledge explicitation takes time ; it is an iterative process and lots of meetings were necessary to produce the two actual ontologies. Moreover Telecom ParisTech members were not accustomed to Protege editor nor OWL formalism. Mondeca partners have helped a lot to recall the expressivity constraints in OWL language and ITM (Mondeca) solution to exploit it. The two produced ontologies are now quite stable and have been integrated into ITM tool. The actual problems are related to rules construction in order to enrich the ontologies ; this will be integrated in DAFOE project.

Summary of actions to construct an ontology, proposed by Plesel [20]

- Terms definition
- Point of view definition
- Associate different definitions to a term according to the point of view
- Illustrate definitions using multimedia supports (images) to clarify and help memorization
- Visualize the concepts network
- Store authors, sources, date for each definition
- Define relations between concepts
- Propose research functionalities for one term or concept
- Propose research, mining and filtering facilities in a concepts network

More details can be obtained from Master theses [20, 1].

5.3.2 Resources

Visual ontology construction When defining ontologies, it is necessary to first look at available resources. We already mentioned the existence of visual ontologies but we did not find them in OWL nor RDF language. We decided to use associated literature and define our own ontology based on PLATO database scheme. PLATO is able to manage both multimedia documents and processing tools ; hence it was quite logical to exploit the already defined classes and extend them to image objects.

The visual ontology (cf Figure 5.12) contains *ImageObject*, *ImageProcessing* classes, as well as *ClassifierModel*, *Sensor*, *Label* and ContainedType classes that are used to make links between the two first classes or to enable a finer description. Image Objects are pixels, regions, and organization as list or graph of such objects. They are used to describe the image content. Moreover spatial relations ontology (cf [9]) is used to describe image objects topological organization. Processing are actually not exhaustive but can easily be extended : we defined *Extractor*, *Classifier*, *Segmentation*, ... classes. These classes are populated by operational implementations (for example libsvm tool is an individual in the *Classifier* class). We wanted to include processing in the ontology to keep in memory how the numerical data associated to images were produced and then to better understand the link between identified objects and automatic processing. This means that this ontology is not only designed for an annotation task ; it can also be used to mine the feature space associated to images.

Relations have been defined to be able to relate the concepts. This is illustrated in Figure 5.14 ; the whole feature extraction process can be described (from the sliding window to the extracted features types and data).

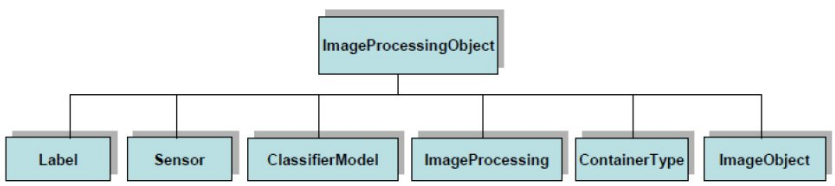


FIG. 5.12 – Visual ontology : illustration of main concepts

The final ontology is available on demand (Marine.Campedel@telecom-paristech.fr) as an OWL file. It is now quite stable but we are aware that it is susceptible to slightly change, particularly at the property level mainly because of an enrichment of the defined relations.

Scene ontology construction Concerning the scene ontology, several geographical databases can be considered. Plesel considered many resources as mentioned in table 5.3.2.

We took Corine Land Cover taxonomy as knowledge basis. As linguist scientist, Plesel studied CLC definitions in order to extract not only the concept hierarchy but also usefull relationhips and new basic concepts (like *building, field, water, etc.*) that are used in CLC definitions. In DAFOE project, a platform is under development to enable the use of Automatic Language Processing tools to help the identification of main concepts and relations. It was not available at the time of our own project and the

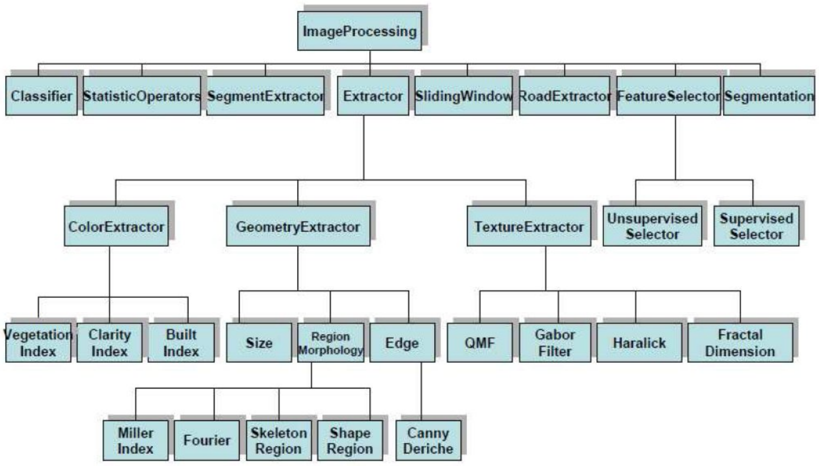


FIG. 5.13 – Visual ontology : illustration of Extractor hierarchy

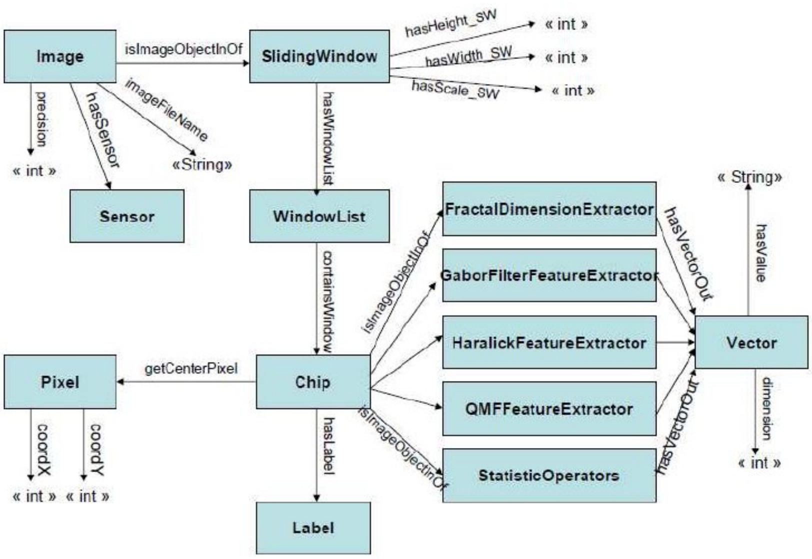


FIG. 5.14 – Visual ontology : illustration of defined relations, when considering the features extracted on a sliding window.

CLC	taxinomy and landcover maps
Cadastral registry	
BD CARTO	vectorial database (IGN) - landcover maps
BD TOPO	topographic reference, 3D database
BD ALTI	MNT data
BD ORTHO	orthophotographic data
BD NYME	toponymes database (IGN)
Geonames	3 level hierarchy, 6M concepts

linguistic extraction process has been manual done. The main structure of this ontology is composed by the classical CLC hierarchy. Semantic and spatial relations between CLC concepts have been added.

Figure 5.15 gives a sample of concepts. We observe that clouds are treated separately since they are not landcover classes but they should be mentioned since they can occlude these objects of interest.

Figure 5.16 gives an example of concept definition. The idea was to preserve at most CLC definitions (given as comments) and to explicit relations between concepts. For example here *ContinuousUrbanFabric isStructuredBy building* .

The final ontology is available on demand (Marine.Campedel@telecom-paristech.fr) as an OWL file. It is now quite stable but we are aware that it is susceptible to slightly change, particularly at the property level mainly because of an enrichment of the defined relations and of the definition of more complex rules.

5.3.3 Difficulties

Methodological and technical problems

Generalities As already mentioned we faced methodological as well as technical problems during the conception of ontologies. In fact this is not even sure that the actual ontologies will not evolve : one solution to this problem is to consider that ontologies should be as simple as possible while allowing the creation of many logical rules that will enrich the final ontologies. Hence rules editor and reasoner become clearly essential to ensure the evolution and exploitation of our ontologies. DAFOE project, and more particularly Mondeca, will treat this problem during the next months.

Technical problems are related to both processing tools access and reusability, as well as ontology management tools that we had to learn. COC tools are currently listed and commented in EXITER project to ensure their reusability. Concerning ontology tools, we had to discover it, as well as ITM Mondeca tool : using two different tools implies interaction between them and we discovered that ontology exploitation tools have different expressivity levels. Even if they can ingest OWL-format files, they cannot always deal with the whole expressivity offered by this language.

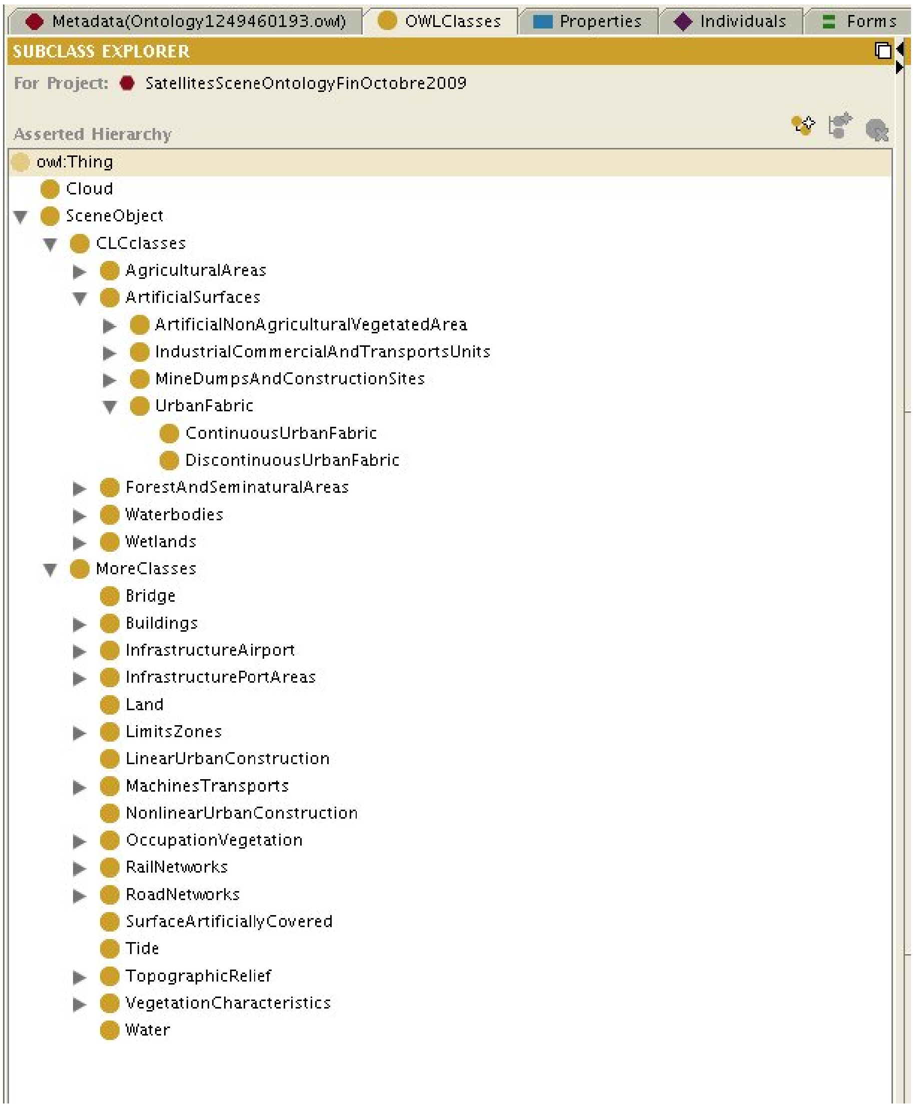


FIG. 5.15 – Scene ontology sample from Protege editor

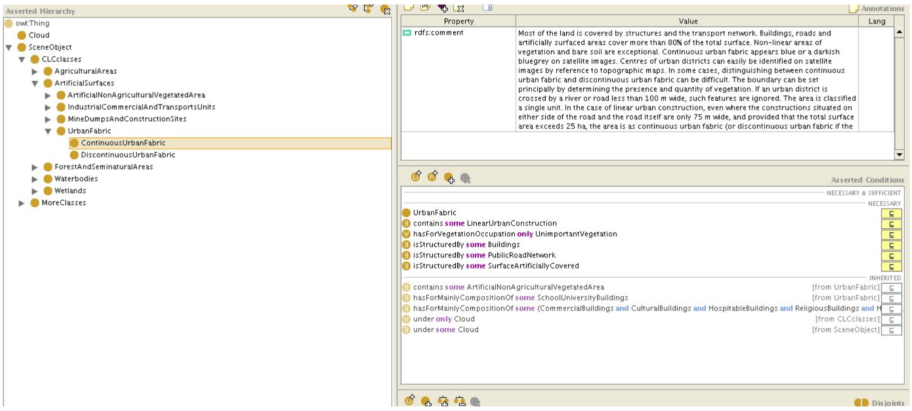


FIG. 5.16 – Scene ontology sample from Protege editor - Details on urban fabric concept.

Concepts vs individuals The set of individuals in a class is called class extension. Each individual is also a class instance. There is a major difficulty in defining the granularity in the concept definition. For example : should we define church as an instance of *Building* or as a specific concept ? should we define NDVI extractor as an instance of *ColorFeatureExtractor* or as a specific concept ? This is clearly a problem depending on the exploitation that will done. The choices results from discussion about the interest of defining and separating instances of such concepts ; this has to be discussed at each concept level. In our case, church is seen as a specific concept because we could be interested by differentiating between simple houses and churches. We also chose to be as specific as possible when considering feature extractors (cf Figure 5.12).

In our scene ontology, instances are not given. It has been defined with Mondeca that instances will be available in separate files. This is described in section 5.4.2.

Labels and Scene Objects In fact to annotate is the process making the link between an *ImageObject* instance with a *SceneObject* instance. For example, we want to link Region#id (instance of *Region* class in image marseille.tif with MarseilleAirport, instance of *Airports*. The relation *hasLabel* can be used to associate both classes. This is particularly interesting when dealing with several images of the same scene : in this case MarseilleAirport is the same instance but the regions are different because the referred images are different. This allows to navigate and query between different images of the same scene without explicit image registration.

The notion of label is used to make the link between Image Objects and Scene objects but also with classifiers. In the context of unsupervised clustering, we want to be able to group ImageObjects using a classification process and to put a tag on the result.

Linguistic expertise

In appendix, a detailed comment in French is provided to better illustrate the difficulty faced at a terminologic level (cf Section [19]). The comments cannot be directly translated in English because homonyms, synonyms are not the same. Automatic Language processing tools could have been used ; this is part of DAFOE platform to provide the users with automatic tools helping the development of ontologies.

Explicit definitions The first problem comes at the concept and relation definition level. Concepts should be defined as explicitly as possible and definitions should be able to differentiate between different objects. Terminological ambiguities should be cleared thanks to the concepts organization (context). For example, the term *quay* will not be confused in the contexts of docks or railway station because it will be introduced as two different concepts (with different terms) with different properties (relations).

The relation definition is clearly more difficult because we are not used to define this kind of property. We use Hudelot spatial relations to be able to describe relative positions of image objects. But there are many other semantic relations that can be established. OWL make a distinction between two property's types :

- Object property : it can be shared by instances of concepts ;
- Data property : it allows to link individuals (instances) to data (simple values that can be textual, numerical, ...).

It can be noticed that the defined object properties are essentially binary. Specific characteristics can be precised like symetry, inverse of (another property), transitivity. We had specific problems on relations like *belongsTo, Contains, isStructuredBy, etc.*. *BelongsTo* is different from a subsumption relation, it is more related to a component in a structure : for example « greenUrbanAreas » BelongsTo « UrbanFabric » comes from the CLC explicite definition *Green Urban Areas = Areas with vegetation within the urban fabric [...]*. Here is another example : *pastures BelongsTo AnnualCropsAssociatedWithPermanentCrops and complexCultivationPattern*. In this case the two classes are sisters, but other sisters do not share this relation.

Expressivity We recall that definition of concepts are not only textual but also formal, in OWL language. For most of classes, this formalization is automatically done by Protege editor. But in some cases, we need to construct more complex classes for example *AnnualCropAssociatedWithPermanentCrops* which is clearly a mixed concept. Marie Lienou [12] proposed a completely machine learning-based approach to deal with such concepts, we propose here to construct logic rules when it is possible.

In *AnnualCropAssociatedWithPermanentCrops* concept [7] there is a constraint of the type "two among three". The idea is that to be an instance of at least 2 classes among three A, B, C, it is necessary and sufficient to be instance of (A and B) or (B and C) or (A and C) (cf Figure 5.17).

[7]CLC definition = Non-permanent crops (arable land or pasture) associated with permanent crops on the same parcel.

This can be easily be translated into OWL format(Figure 5.18). But Mondeca ITM tool is not able to deal with such definition at the ontology definition level. This will be integrated as a specific rule.

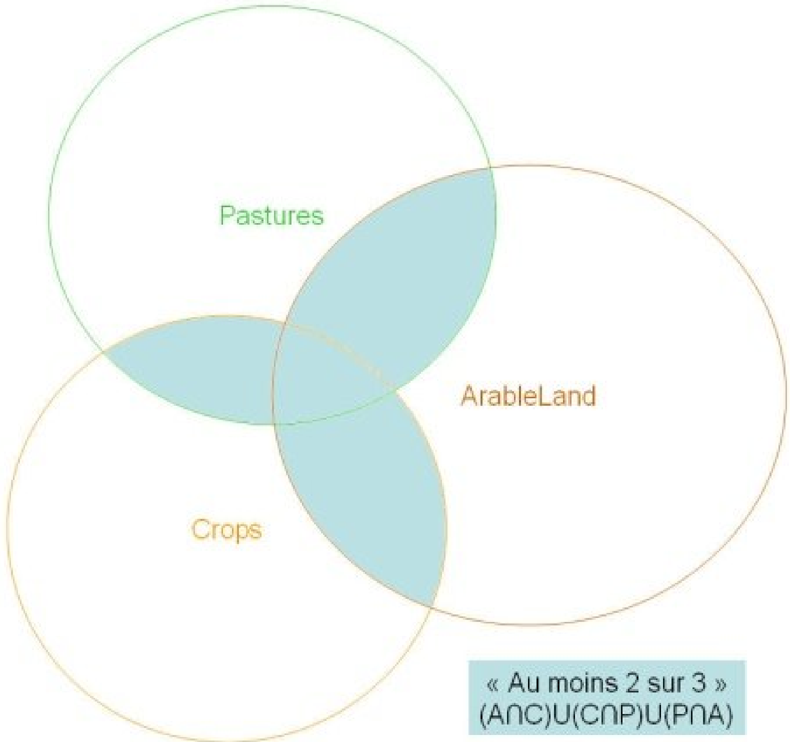


FIG. 5.17 – Modeling problems.

5.3.4 Quality criteria

At the end of the creation stage, it should be possible to qualify our ontology. They are already defined criteria (cf Gruber article [7] : extracted sentences are cited in italic) :

- Clarity : *An ontology should effectively communicate the intended meaning of defined terms*. Natural language can be used to help understanding and memorization of the defined concepts. We documented concepts and particularly CLC-based concepts. The idea is not to be limited by definitions in the logic sense, but to explicit as much as possible the defined concepts in our application domain.
- Coherence : *An ontology should be coherent : that is, it should sanction inferences that are consistent with the definitions. (...) If a sentence that can be inferred from the axioms contradicts a definition or example given informally, then the ontology is incoherent*. This can checked by a reasoner, for example Pellet in Protege editor.

```
owl:Class rdf:about="#AnnualCropsAssociatedWithPermanentCrops">
        <rdfs:subClassOf rdf:resource="#HeterogeneousAgriculturalAreas"/>
        <rdfs:subClassOf>
            <owl:Restriction>
                <owl:onProperty rdf:resource="#belongsTo"/>
                <owl:someValuesFrom
rdf:resource="#ComplexCultivationPatterns"/>
            </owl:Restriction>
        </rdfs:subClassOf>
        <rdfs:subClassOf>
            <owl:Restriction>
                <owl:onProperty rdf:resource="#contains"/>
                <owl:someValuesFrom>
                    <owl:Class>
                        <owl:unionOf rdf:parseType="Collection">
                            <owl:Class>
                                <owl:intersectionOf rdf:parseType="Collection">
                                    <rdf:Description rdf:about="#ArableLand"/>
                                    <rdf:Description rdf:about="#Pastures"/>
                                </owl:intersectionOf>
                            </owl:Class>
                            <owl:Class>
                                <owl:intersectionOf rdf:parseType="Collection">
                                    <rdf:Description rdf:about="#ArableLand"/>
                                    <rdf:Description rdf:about="#PermanentCrops"/>
                                </owl:intersectionOf>
                            </owl:Class>
                            <owl:Class>
                                <owl:intersectionOf rdf:parseType="Collection">
                                    <rdf:Description rdf:about="#Pastures"/>
                                    <rdf:Description rdf:about="#PermanentCrops"/>
                                </owl:intersectionOf>
                            </owl:Class>
                        </owl:unionOf>
                    </owl:Class>
                </owl:someValuesFrom>
            </owl:Restriction>
        </rdfs:subClassOf>
    </owl:Class>
```


FIG. 5.18 – Modeling problems.

– Extendibility : *An ontology should be designed to anticipate the uses of the shared vocabulary. It should offer a conceptual foundation for a range of anticipated tasks (...)*. We will demonstrate in the following how our ontologies can be linked and how they can evolve using rules definitions.
– Minimal encoding bias : *The conceptualization should be specified at the knowledge level without depending on a particular symbol-level encoding. An encoding bias results when representation choices are made purely for the convenience of notation or implementation*.
– Minimal ontological commitment : *An ontology should require the minimal ontological commitment sufficient to support the intended knowledge sharing activities. An ontology should make as few claims as possible about the world being modeled, allowing the parties committed to the ontology freedom to specialize and instantiate the ontology as needed.*. In practice it is difficult to define the “minimal” commitment. This is related to the problem we faced concerning the notions of individuals and concepts (cf Section 5.3.3).

The most important point is that ontologies have to be helpfull in human to human communication or human to machine interaction. The explicitation of concepts and relations lead to knowledge management, it gives a consensual and precise description of systems that can be exploited by both humans and machines. They provide a solid base for more complex descriptions. The next section is dedicated to examples of use of our ontologies.

5.4 Example of Use

5.4.1 introduction

In this section, we illustrate how the two ontologies will be used through an annotation application. This application is under development with Mondeca collaboration in DAFOE project, it will be delivered in July 2010.

There are many applications that could be derived from both ontologies :

– At visual level : the visual ontology can be used to describe the image content as analyzed by automatic processing. Hence we propose an XML format in section 5.4.2.
– At interpretation level : based on CLC taxonomy, the scene ontology can be used to navigate into CLC maps associated to any image while using semantic queries ;
– Ontologies combination : the joint use of both ontologies aims at interacting with both the visual and interpretation levels. Hence the queries can be expressed as semantic queries (keywords, rules, ...) or as actions in the image (region selection, clicks, zoom in/out, drawing, ...) since both levels are related.

5.4.2 Descriptive tools

Derived XML format

We can associate an XML format to any satellite image using the ontologies as descriptive tools. This process is detailed in Aouini report [1]. More precisely, several XML files can be attached to an original image, describing the applied processing and its result. We have done that for two reasons :

– it was necessary to populate ITM (Mondeca tool) while using a simple XML-based format. The simplest way was to derive the XML format from the ontologies.
– CNES was interested in SoΦia project by developing a format able to store all date associated to an image. In fact SoΦia was supposed to be finished at the beginning of this project but is was not the case and we decided to develop our own basic format.

Aouini developed a plugin in Java language : it takes arff (weka format) files as entries and produce the xml corresponding file. This plugin will be further developed in the next months to take libsvm format as entries and be able to represent any types of information.

At this time, for illustration purposes, the plugin has been used on only one portion of image (Madrid scene, SPOT5, 3000x3000 pixels, scene 10, arbitrarily chosen). Sliding window and basic texture extractors have been tested. An example of XML samples associated to one pixel is given in Figure 5.19. Each entity is identified by its type (which refers to an ontology concept), a unique id (which identifies the instance inside the database) and URI (which uniquely identifies the instance in the world), and list of properties (which depend on the concept).

```
− <entity>
    <id>PX1_0_1616</id>
    <type>http://www.telecom-paritech.fr/dafoe/sat#Pixel</type>
    <URI>http://www.telecom-paritech.fr/dafoe/sat#pixel1616</URI>
  − <properties>
    − <property>
        <type>http://www.telecom-paritech.fr/dafoe/sat#coordX</type>
        <datatype>int</datatype>
        <value>16</value>
      </property>
    − <property>
        <type>http://www.telecom-paritech.fr/dafoe/sat#coordY</type>
        <datatype>int</datatype>
        <value>16</value>
      </property>
    − <property>
        <type>http://www.w3.org/2000/01/rdf-schema#label</type>
        <datatype>label</datatype>
        <value>pixel1616</value>
      </property>
    </properties>
  </entity>
```

FIG. 5.19 – Example of XML description used for one Pixel.

The idea in not only to have classical feature extractors and classifiers represented in the XML files but also spatial relations between image objects. Carolina Vanegas (PhD) provided relations extraction between regions : as soon as classifications or segmentations are provided, these extractors can be applied and relations like adjacency and relative positions (with given distance and orientation) can be identified and stored in the XML format. As soon as available, parallelism and any extractible spatial relation will be inserted in the Java plugin.

Perspective : Sofia, XediX collaboration

The idea of having a descriptive format of the image content is not new. For videos, MPEG-7 has been created and for satellite imagery Soϕia project has been proposed. This format should encapsulate information at the image, visual and semantic (scene) levels. Our preliminary XML format can do that.

The perspective is related to the exploitation of such a format in the context of satellite images collections. Efficient XML databases are now available and demonstrated their capacity to handle millions of data. XediX is one storage solution to manage and query such data. Contacts have been taken with the firm developing XediX to demonstrate the ability of XediX platform to handle collection of satellite images.

However, this will be usefull to exploit XML stored data, ie to navigate into the already existing data. To go a step further, reasoning tools can also be used and give access to new knowledge thanks to logical rules application.

5.4.3 Reasoning tool

Example of use

Annotation checking and desambiguisation The idea is to use the reasoner to check an annotation result. For example, in Figure 5.4.3, we observe a vegetation region inside an urban region. According to CLC definitions, this region should be qualified as , but concept *ArableLandAgriculturalLand* has been used. A reasoner should detect automatically this process if the relation “inside” has been identified and if all the knowledge about concept has been explicitly defined as a vegetation area inside a urban one.

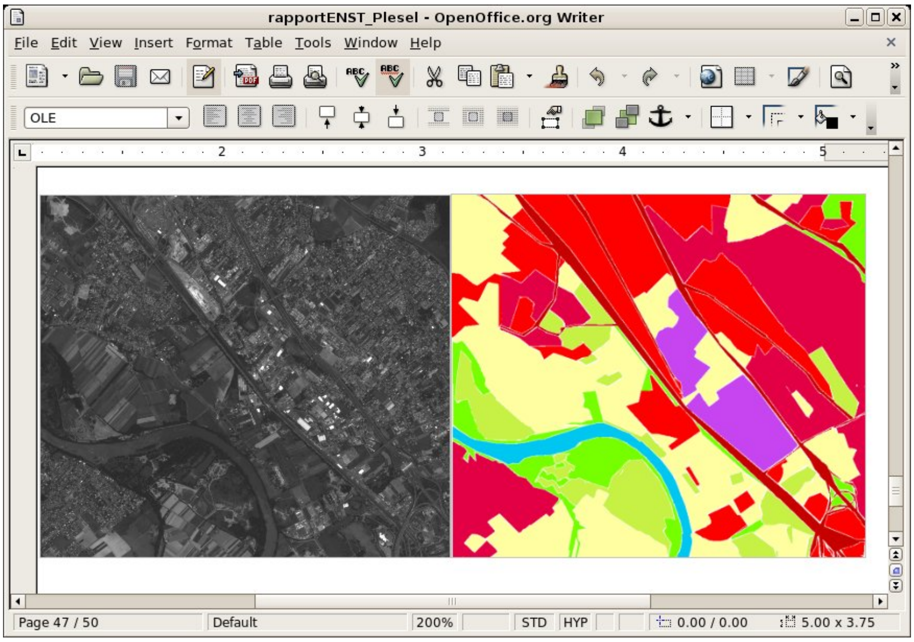


FIG. 5.20 – Toulouse SPOT scene and manual annotation.

Annotation specialization Automatic learning processing have successfully been used in the COC to identify landscapes and networks in SPOT5 panchromatic images, as well as QuickBird spectral images. Smaller objects are more difficult to identify and interaction process could be used to enhance the annotation details. For example, it could be usefull to first apply a rough forest detector before using user manual annotation or available ground truth to distinguish between trees species. Another example comes from the road identification : road network can automatically be extracted but nothing is usually said about road types and how to differentiate between them. The scene ontology could guide a manual annotation process and suggest classes of interest according to the currently observed objects in a given image. Then, according to the new annotations, a new (interactive) learning process based on image examples can be applied or rules can be created. Finally, new automatic annotations can be produced under the human interpreter's control.

New rules construction The construction of new rule is a new way the ontology can evolve in time. Of course, we also include the possibility to exploit active learning approaches to produce new classifiers ; but this way has already been exploited in Costache PhD thesis [3] and is under study in Blanchart undergoing PhD thesis.

These rules are logical rules, they can pilot automatic extractors (like region size computation, ...). Rules are defined as logical combinations of concepts and relations. They are close to the notion of semantic query, and should be formally described (using SPARQL language for example). The construction of rules is not obvious ; it is related to the way how ontology's objects were constructed : it necessitates to go through terminologic and syntactic analysis.

It is difficult to know when rules should be defined and when machine learning should be used. We have the intuition that complex objects, like ports and airports, could benefit from rules definition. We already know that theses classes are difficult to detect with automatic classifier because these classifiers necessitate some examples which are not always available. But we have the feeling that airport can semantically be defined as a spatial configuration of specific roads, buildings, fields and planes, which could be automatically identified.

Rules creation necessitates a specific editor. Once a rule is created, a corresponding query is defined and occurrences of such query can be obtained. These instances are used to check the validity of the rule and to refine it. It is similar to a feedback approach as used in active machine learning approaches.

Reasoner tool

Until recent years, reasoner tools were not considered as sufficiently powerfull to be helpfull ; hence engineers were more interested in relational databases than formal ontologies. Nowadays, ontologies can be exploited in similar ways as relational databases (as descriptive tools) and also in conjunction with reasoner, allowing access to never stored but inferred information.

There are several well-accepted reasoner like Pellet, Racer. Most of them are linked to ontology editors, and particularly Protege in order to ensure the development of correct and rigorous ontologies.

5.4.4 Conclusion

SISA is the name given to the annotation application under development in DAFOE project. It will put in practice the mentioned applications, with the help of Mondeca engineers. A scenario is under definition allowing multi-view navigation for the same scene (SPOT5, QuickBird images and CLC map), Marseille for example. The goal will be to demonstrate the interest of combining automatically extracted information with knowledge modeling for :

- Classification checking
- Annotation help and particularly annotation's specialization
- Navigation using visual and semantic queries
- Ontology evolution with active learning and online query definition and storage

Chapitre 6

Conclusion and perspectives

6.1 Results summary

This project has been dedicated to knowledge representation tools for satellite imagery. It was supposed to benefit from Soϕia project but it has not been the case and this project can be considered as standalone.

The main contributions are :

- Study of the notion of knowledge : many things had already been produced in COC around the notion of classification and relations between visual classes, but knowledge notion was not clearly identified at this time. Positioning our work in the knowledge engineering domain is opening COC to already unexploited approaches that could be combined to our actual competences.
- Interrogation about human interpreters work : during past projects (EXITER 2008) we observed difficulties to communicate with interpreters and we needed to take time to better understand them. It seems now clear that automatic systems cannot replace human interpreters but they could be designed to help them while extending the human cognitive process involved in the interpretation task. We applied Mental Gestures theory to derive such an helping system.
- Development of ontologies : we explored how knowledge-based applications usually model knowledge and discovered formal ontologies. Based on state-of-the-art approaches, we developed a visual ontology and a scene ontology that can be linked. A complete satellite image annotation application based on these ontologies and automatic extractors will be developed until July 2010 (end of DAFOE project).

6.2 Perspectives

6.2.1 Methodological perspectives

From a methodological point of view, we suggest to focus on systems able to assist interpreters and not to replace them. The interpretation task supposes to give a meaning to an image and computers are not able to do that. But computers have better memories and computation abilities than humans ; hence they can be used as support to cognition activities in order to help interpreters extract a meaning and translate it on a map.

The system we would like to develop now should of course contain communication abilities, but also interaction means. Interaction is facilitated when semantics can be used : knowledge-based representation are developed for that purpose. Interaction can be offered at different levels in the cognitive process of the interpreter as presented in chapter 4.2.

A knowledge-based system for satellite image browsing and annotation is usefull and can be considered as an extension to GIS. They should be able to integrate information (and tools) coming from GIS as well as images and to navigate and query among all the views.

6.2.2 Collaboration perspectives

As mentioned during this report, this project has been the initial point for 3 different collaborations :

- DAFOE : in the context of ANR DAFOE, a new and effective collaboration is born with Mondeca (and Inserm). This collaboration will lead to the development of an interactive annotation tool based on the two ontologies in order to exploit both image processing and reasoning.
- XediX : XediX platform is identified as promising to exploit large quantities of XML files associated to satellite images. Financial support is not found yet.
- A. Bretto : an old contact as been reactivated to deal with hypergraph representation. Professor A. Bretto will work at Telecom ParisTech, as invited researcher, from January 2010 to September 2010 in order to extend the collaboration. The relation with this current project is the will to discuss about the ability of hypergraphes to represent knowledge and navigate inside, particularly in the context of satellite image interpretation.

6.2.3 Scientific perspectives

From a scientific point of view, many questions are still open.

- Ontologies : methodologically ontologies are difficult to create. DAFOE platform should reduce this difficulty but there are still problems to handle their complete life cycle (and particularly their capacity to evolve in time).
- Hypergraphs : we demonstrated that hypergraphs are powerfull to represent image content and then to exploit this representation in order to structure low-level information. Could hypergraphs be considered as knowledge representation tool ? could they simply be a visual representation of ontologies ? this would be usefull since ontologies are semantically defined but when dealing with images and enormous content, visual representations are precious. This subject is proposed in ANR proposed in 2010 with A. Bretto.
- Knowledge and expertise : on a more philosophical point of view, the notion of *expert* seems related to a certain degree of knowledge that is not completely sharable. It would be interesting to reexamine this aspect considering Mental Gestures Theory as well as other cognitive theories and to better study what so-called expert systems provide until now.

Chapitre 7

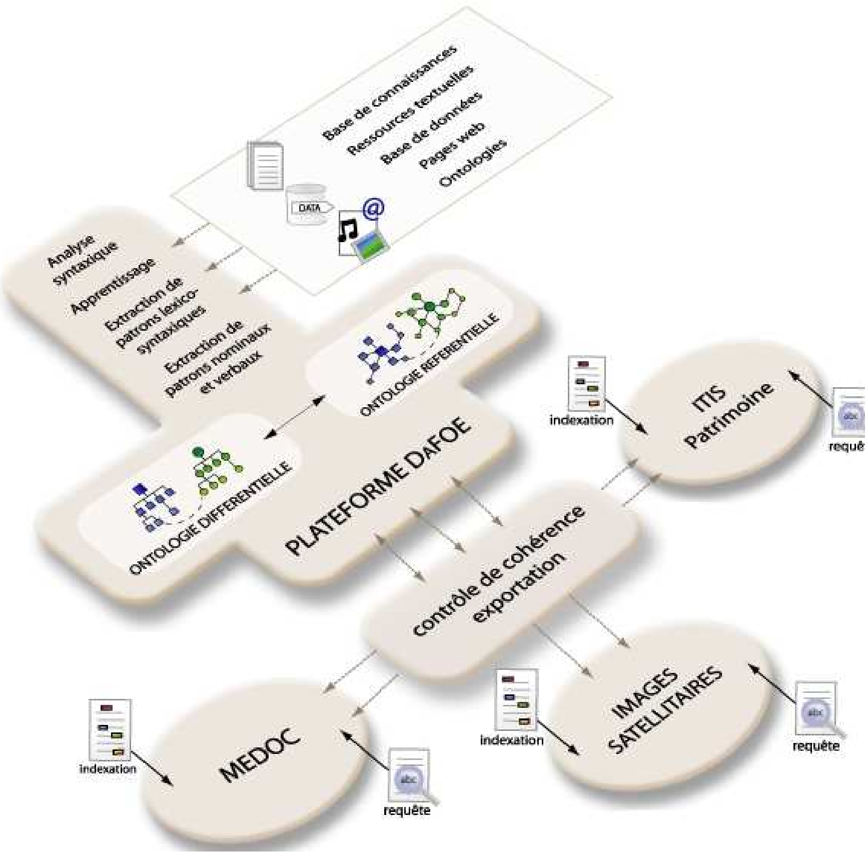
Appendices

7.1 DAFOE project

ANR project Le projet DAFOE (Differential and Formal Ontology Editor) est un projet ANR débuté en janvier 2007, sous le pilotage de Jean Charlet (INSERM). Les partenaires du projet sont L'UMRS 872 équipe 20 de L'INSERM, Télécom ParisTech, Heudiasyc (HEUristique et DIAgnostic des SYstèmes Complexes), l'entreprise Mondeca, l'équipe Ingénierie des Connaissances, de la Cognition et de la Coopération (IC3) de l'IRIT, Le Laboratoire d'Informatique de l'Université Paris-Nord (LIPN), Le LISI (Laboratoire d'Informatique Scientifique et Industrielle) et SUPELEC.

Platform to create and manage ontologies La plupart des outils développés autour des ontologies permettent de les construire en précisant comment représenter les concepts et formaliser leur sémantique, ils ne précisent pas comment trouver les concepts ni comment expliciter leur signification. L'objectif du projet est de proposer une méthode complète associée à une plateforme technique pour concevoir des ontologies, de la modélisation à partir du domaine à leur évolution en passant par leur formalisation et exploitation. S'appuyant sur les acquis de travaux antérieurs, à la fois issus des partenaires et de la littérature du domaine, le projet a pour but de prendre en charge la modélisation sémantique des concepts ontologiques pour motiver et justifier les représentations formelles qui seront utilisées et en faciliter la révision.

La plateforme technique DAFOE est un ensemble d'outils dont un éditeur d'ontologies qui prend en charge toute la question de la sémantique de ces ontologies, à travers des questions épistémologiques liées aux concepts formels de haut niveau et, vis-à-vis de la composante métier, à travers des travaux sur les corpus textuels. On obtient ainsi une ontologie formalisée qui pourra être traitée dans un éditeur d'ontologie respectant les standards des langages d'ontologies du W3C (OWL).



Final evaluation using application L'évaluation finale de cette plateforme se fera via des applications pour lesquelles une ontologie est nécessaire et ce à travers sa dimension sémantique et donc son interaction avec l'utilisateur. Les applications développées correspondent à des tâches d'indexation de documents puis de recherche d'informations à leur sujet. Ils ont donc proposé de mettre en oeuvre ces applications dans trois domaines

- l'aide au codage médical ;
- l'indexation patrimoniale ;
- l'indexation d'images satellitaires.

Nous participons donc à ce projet en tant qu'utilisateurs désireux de créer une ontologie avec pour objectif de faciliter et améliorer notre processus d'annotation automatique des images satellitaires dans le cadre d'une application nommé SISA (Satellite Image Semantic Annotation). Les résultats produits serviront à valider la plate-forme développée dans le cadre du projet DAFOE. Ce processus nécessite la définition des ontologies nécessaires à la tâche, et c'est ce que nous proposons dans le cadre de ce projet.

Bibliographie

[1] K. Aouini. La construction d'une ontologie appliquée au traitement d'images satellitaires. Master's thesis, Université René Descartes - PARIS 5 - UFR de Mathématiques et Informatique, 2009.

[2] B. Bachimont. *Engagement sémantique et engagement ontologique : conception et réalisation d'ontologies en Ingénierie des connaissances.* Eyrolles, 2000.

[3] M. Costache. *SVM and Bayesian methods for category based semantic learning : a search engine for satellite images*. PhD thesis, Telecom ParisTech, September 2008.

[4] A. de la Garanderie. *Comprendre les chemins de la connaissance - Une pédagogie du sens*. Chronique Sociale - 2eme édition, 2009.

[5] N. Durand, S. Derivaux, G. Forestier, C. Wemmert, P. Gançarski, O. Boussaïd, and A. Puissant. Ontology-based object recognition for remote sensing image interpretation. In *IEEE International Conference on Tools with Artificial Intelligence*, volume 1, pages 472–479, Patras, Greece, october 2007. IEEE Computer Society.

[6] J.P. Gaté and T. Payen de la Garanderie. *Introduction à Antoine de La Garanderie*. Chronique Sociale, 2007.

[7] T. R. Gruber. Towards principles for the design of ontologies used for knowledge sharing. In N. Guarino and R. Poli, editors, *Formal Ontology in Conceptual Analysis and Knowledge Representation*. 1993.

[8] C. Hudelot. Ontologies, un outil pour l'analyse et l'interprétation d'image, 2006.

[9] C. Hudelot, J. Atif, and I. Bloch. Ontologie de relations spatiales floues pour l'interprétation d'images. In *Rencontres francophones sur la Logique Floue et ses Applications, LFA 2006*, pages 363–370, Toulouse, France, October 2006.

[10] I. Kyrgyzov. *MINING SATELLITE IMAGE DATABASE OF LANDSCAPES AND APPLICATION FOR URBAN ZONES : CLUSTERING, CONSENSUS AND CATEGORISATION.* PhD thesis, École Nationale Supérieure des Télécommunications ENST2008E011, May 2008.

[11] I.O. Kyrgyzov, H. Maître, and M. Campedel. A method of clustering combination applied to satellite image analysis. In *IEEE - International Conference on Image Analysis and Processing ICIAP 2007*, pages 81–86, Modena, Italy, sep 2007.

[12] M. Lienou. *Apprentissage automatique de la production des cartes d'occupation du sol et représentation en mots visuels des images satellitaires*. PhD thesis, Ecole Nationale Supérieure des Télécommunications, March 2009.

[13] E. Tomai M. Kavouras, M. Kokla. Comparing categories among geographic ontologies. *Computers and Geosciences*, 31 :145–154, 2005.

[14] N. Maillot. *Ontology Based Object Learning and Recognition*. PhD thesis, Université de Nice Sophia Antipolis, dec 2005.

[15] D. Marr. *Vision : a computational investigation into the human representation and processing of visual information*. W. H. Freeman, San Francisco, 1982.

[16] V. Mezaris, I. Kompatsiaris, and M.G. Strintzis. Region-based image retrieval using an object ontology and relevance feedback. *EURASIP Journal on Applied Signal Processing, Special Issue on Object-Based and Semantic Image and Video Analysis*, 6 :886–901, 2004.

[17] Commission of the European Communities. Corine land cover, 2001.

[18] K. Petridis, D. Anastasopoulos, C. Saathoff, N. Timmermann, Y. Kompatsiaris, and S. Staab. M-ontomat-annotizer : Image annotation linking ontologies and multimedia low-level features. In *KES'06*, 2006.

[19] P. Plesel. Clés d'intepprétation/de compréhension de la scène satellitaire. Master's thesis, Université Paris Sorbonne - Ingénierie de la Langue pour la Gestion Intelligente de l'Information, 2009.

[20] P. Plesel. Rapport de stage de master ii, conception d'une ontologie pour l'annotation d'images satellitaires. Master's thesis, Université Paris Sorbonne - Ingénierie de la Langue pour la Gestion Intelligente de l'Information, 2009.

[21] A. Th. (Guus) Schreiber, B. Dubbeldam, J. Wielemaker, and B. Wielinga. Ontology-based photo annotation. *IEEE Intelligent Systems*, pages 2–10, May-June 2001.

Dépôt légal : 2010 – 4ème trimestre
Imprimé à Télécom ParisTech – Paris
ISSN 0751-1345 ENST D (Paris) (France 1983-9999)

Télécom ParisTech

Institut TELECOM - membre de ParisTech

46, rue Barrault - 75634 Paris Cedex 13 - Tél. + 33 (0)1 45 81 77 77 - www.telecom-paristech.frfr

Département TSI