Image Retrieval Using a Multilingual Ontology

Adrian Popescu
Commissariat à l'Energie Atomique - LIST
18 route du Panorama
92260 Fontenay aux Roses, France
adrian.popescu@cea.fr

Abstract

Search engines are among the most useful Internet applications. There exist several media types on the Web and, given the particularities of each of them, adapted search solutions are required. We limit our discussion to image search engines. While rapid and robust, existing image search engines offer results that respond only partially to the user's queries. An improvement of image search results might be obtained with the introduction of semantics in the dedicated systems. Here, we discuss the construction and the utilization of a multilingual lexical resources (WordNets in several languages) to improve image retrieval on the Internet. Given the initial nouns hierarchies in the WordNets, we build a multilingual OWL ontology including knowledge in English, Italian, and Spanish. A pictured representation of a dog remains a representation of a dog in spite of the associated name (would this be *dog*, *perro* or *cane*). The use of a large scale multilingual ontology allows us to offer the consequent sets of responses for the concepts in the hierarchy irrespective to the initial language the query was formulated in. With the use of an ontology to structure an image database, we can solve problems related to the ambiguity of a query content and we obtain an important gain in precision in the image sets rendered to the user compared to state of the art system.

Introduction

There exist an important number of image search solutions on the Web. All major actors on the market, like Google, Yahoo, MSN or AOL, propose image related facilities and there exist image dedicated solutions like Picsearch. All these applications are impressively rapid and offer, most of the times, a large number of images results in response to a query. A problem with current image search systems is the precision in the retrieved picture sets. Namely, a good part of the image responses are irrelevant to the formulated query.

We showed elsewhere (Popescu et al., 2006) that the use of semantic resources in image retrieval (IR) can improve the precision of the obtained results and that the use of an ontology to organize an image database provides an economic and efficient way of rendering an important number of results in response to the user's queries. Here, we propose an extension of the ontology described in (Popescu et al., 2006), which provides the possibility to query the system in several languages and obtain consistent sets of results irrespective to the language of the query. Moreover, with the introduction of other languages in ontology it is possible to detect and treat inter-lingual ambiguities. We propose a partial translation of several WordNets to a multilingual OWL ontology. For the moment, our ontology includes nouns in English, Spanish and Italian. The noun hierarchies for the two last languages are aligned to the English version of WordNet and this alignment allows an integration of the three resources into a unique OWL ontology. The English WordNet is by far the most developed and we used this resource to gather images and render them to the user. The other WordNets are used in the image rendering process and for disambiguation. Given the sets of images associated to leaf terms in a hierarchy and using the hypernymy relation, we can provide answers for all concepts included in the hierarchy. The construction of a multilingual ontology provides the possibility to offer similar image responses to queries that point towards the same type of entity in the world when the system is queried with similar concepts in different languages. The premise for offering unitary

picture sets for irrespective to the language the query is expressed in is that a *dog* covers similar entities whether we call it *dog*, *perro* or *cane*. Linguistic particularities can be taken into consideration with the presentation of the results in different orders, in accordance with the existence and the frequency of the subconcepts of the queried concepts in the language of the query. A secondary use of a multilingual ontology in image retrieval is that it allows disambiguation for terms that are homonyms in several languages.

We structured the remainder of this paper as follows: in the next section, we situate our work in the context of relevant research in computational linguistics and Semantic Web. Further, we discuss the construction of a multilingual OWL ontology using the information in different WordNets. The relevance of this ontology in image retrieval constitutes the subject of another section. Before concluding, an evaluation of the image retrieval results we obtain using an ontology compared to a state of the art image search system.

Related Work

The WordNet project (Miller, 1990) has generated derived related work in an important number of fields. The three dedicated conferences organized since 2002 are a proof for its success. Here, it is interesting to situate three types of applications: the creation of lexical hierarchies in other languages, the transformation of WordNet into a formal ontology and the utilization of WordNet in image retrieval applications.

The initial WordNet was created in American English, but currently there exist versions in dozens of other languages. There are some large scale projects like EuroWordNet (EuroWordNet, 1999), which proposes lexical hierarchies in 8 languages, as well as projects concerning individual languages like Hebrew (Ordan & Wintner, 2005). The noun hierarchies for the languages other than English have variable sizes but none approaches that of the English WordNet. In this paper we are mainly concerned with the Spanish (Daudé et al., 2000) and Italian (Pianta et al., 2002) WordNets, which are lexical databases that are strictly aligned with the English version 1.6. The main advantage of a strict alignment is that the databases are easy to manipulate in computer application, while the principal drawback is that the structures of the languages are not always similar and there are gaps that appear in the hierarchies. We include these three languages in a multilingual ontology and use it in image retrieval. For this task, the advantage offered by the alignment prevails over the gaps that appear in the hierarchies because we do not perform any linguistic analysis, but simply employ the ontology to provide multilingual access to our application and to collect images in an improved manner.

WordNet has the same basic shape as formal ontology, that of a hierarchy. The transformation of the lexical hierarchy into ontology generated a lot of discussions. Propositions were made to modify WordNet following principles of formal ontologies. Gangemi et al. (2003) propose a rearrangement of a part of the WordNet nouns hierarchy into a formal inheritance system. But, given the enormous volume of required work, a complete transformation of WordNet following the rules in (Gangemi et al., 2003) does not exist and it is improbable to be generated. Nonetheless, the discussion about the alignment of WordNet to formal ontologies principle resulted in the refinement of the ontological structure of lexical hierarchy. Its current version, 2.1, a difference between classes and instances appears whereas it does not exist in previous versions. Given the size of included knowledge and its similarity to ontologies, WordNet is an interesting resource for people in the Semantic Web community. A translation of the lexical hierarchy to OWL (Ontology Web Language) is proposed in (Van Assem et al., 2006). This transformed version of the lexical hierarchy exists uniquely for English. The authors propose a

complete translation of the WordNet structure to an ontological form. For the application envisioned here, image retrieval using ontologies, only a part of the information in the lexical hierarchy is interesting (e.g.: we do not use meronymic relations among WordNet concepts).

Concepts in WordNet have been used in image related applications. Wang et al. (2004) propose the constitution of an image thesaurus using images from the Internet. They employ taxonomic relations in the lexical hierarchy to filter word senses and to expand queries for images. This approach is close to ours but there are some key differences that appear. First, in (Wang and al.), the hypernymy relation is exploited differently, as they use concepts on several levels in the hierarchy to build their database. We propose to attach images only to leaf synsets in WordNet and to use hypernymy in order to offer images for the other concepts in the hierarchy. Second, we preserve polysemy, while Wang et al. (2004) use only the first WordNet sense for polysemic terms. Third, the number of images in (Wang et al., 2004) is smaller than 20000, while we include more than 1 million images in our system. The size limitation for the former work is partly due to the fact that detailed image analysis (region extraction and naming among others) is performed and this is a time costing procedure.

The dominant Internet search paradigm is syntax based. This is true for all kinds of media existing on the Web. Recent work like Squiggle (Celino et al., 2006), related to the Semantic Web initiative, propose the introduction of a semantic layer in the information retrieval architectures. The semantics is mainly encoded in ontologies, which contain structured information about a domain of application. Celino et al. (2006) propose an application somewhat similar to the one described in this paper, namely the use of a multilingual ontology for image retrieval in a given domain (ski) is described. The differences come from the size of the ontology, the construction method and the principle of utilization. We propose the automatic building of a large scale resource, departing from existing sources (WordNets in three languages), while in (Celino et al., 2006) a small scale hand-built ontology is employed. When a query is launched in Squiggle, the system translates the term in the other languages in the ontology and proposes answers related to the translations. This is mainly an effect of the fact that in their domain of application it is often the case to have too few answers to a query. In our application, there are typically to much responses and the problem that arises is the efficient presentation of results.

Multilingual Ontology

In this section, we present a methodology for creating a multilingual ontology employing existing aligned resources. We already mentioned that an official translation of WordNet to OWL exists (Van Assem et al., 2006)] but, in order to better accommodate the purposes of the image retrieval application, we proposed an alleged version of the WordNet representation in OWL (Popescu et al., 2006). We are primarily interested in obtaining an ontology whose structure and size will allow real-time processing when a user queries for images.

Ontology Construction

The total number of synsets in the English WordNet 2.1 is 117595, while the current versions for Spanish and Italian include 105494 and 32700 synsets respectively. These statistics include all types of synsets (nouns, adjectives, adverbs, and verbs). We note that the Italian and Spanish versions are not as rich as the English one. There are two explanations for this situation:
-the English WordNet 1.6 version is not as developed as the 2.1 variant. The WordNets discussed here are strictly aligned to the 1.6 version and there are no translations for the synsets that were added afterwards.

-not all the synsets in the English WordNet 1.6 were translated in Italian and Spanish. There are two main reasons for these differences between the three languages: first, the Italian and Spanish hierarchies are not as developed as the English one and they contain a smaller number of specialized terms. Second, there are English words that do not have translations in the other languages. For example, *neckwear* has no equivalent in Italian, while its immediate hyponym, *necktie* is translated as *cravatta.* The pair *necktie – cravatta* illustrates well the difference in the detail level between the hierarchies in the two languages. In English, necktie has a total of subconcepts for *necktie* is 8, while in the Italian hierarchy there is only one hyponym.

The inclusion of Italian and Spanish in the ontology is not straightforward because they are aligned to an old version of WordNet (1.6). It is necessary to create a passage between versions and this task is realized using sense-key mapping files, available at (WordNet, 2006). We obtain a correspondence between the versions of WordNet we employed. After the mapping step, we create a raw data file that contains English, Spanish and Italian synsets. A previously created ontology for English is then used so as to add Spanish and Italian translations for the English terms.

The ontology we propose is to be used in image related applications and we limit our discussion to nouns, which are, in majority, picturable entities. The multilingual ontology is constructed employing the following transformation rules:

1. Each term in WordNet synset (set of synonyms) becomes a class in the OWL version. If several terms exist in a synset, they are considered equivalent OWL classes. The rationale for this design choice is that all members of a synset correspond to the same entity in the world (abstract or physical). A naming convention for the translation is established in that each class name includes the concept and its associated sense number. This way we preserve the sense separation for polysemic concepts, a central structural property of WordNet, equally important in image retrieval tasks. For ambiguous terms, different meanings of the same term cover separate entities in the World and it is suitable to provide individual image sets for each sense of a concept. The ontology includes three languages; the difference between the terms in these languages is stated using a suffix that individualizes the classes. For English, we add *EN* at the end of the class name. For example *dog* has seven senses, the most prominent one being that of member of *mammals*. This meaning of the term will be translated to an OWL class containing *dog__1__EN, dog__1__EN*, *domestic_dog__1__EN*, *Canis_familiaris__1__EN,* while the others will range between *dog__2_EN* and *dog__7_EN*.

2. We extend the English ontology to Italian and Spanish using the same design rule described in 1. Each member of a synset in the two languages (with its associated sense) is represented as an OWL class. The suffixes that individualize the languages are respectively IT and SP for Italian and Spanish. The equivalent classes that represent dog in the multilingual ontology are: *dog__1__EN, domestic_dog__1__EN, Canis_familiaris__1__EN, cane__1__IT, Canis_familiaris__1__IT, can__1__SP, perro__1__SP, perro_doméstico__1__SP.*

The obtained result is a multilingual hierarchy, where one or several OWL classes point towards the same depicted entity and it is possible to render images associated to these concepts in a structured manner. We stress that, for ambiguous terms, an image set is associated to each meaning. Disambiguation is an important gain in semantics driven information retrieval systems (Celino et al., 2006) because the accuracy of the obtained results is improved. A detailed discussion on disambiguation is given in the section describing image collection.

Utilization of the Ontology

In recent years, we note an impetus towards the development of tools meant to exploit ontologies in practical situations. Ontology editors like Protégé (Protégé, 2006), meant to create and visualize ontologies and reasoners (Racer, 2006), meant to exploit them, are proposed but there are still a lot of problems that subsist. When one wants to exploit large scale ontologies, like the OWL translation of the English WordNet, simple actions like its visualization become difficult (for example, it was impossible to visualize the whole WordNet hierarchy on a P IV PC, with 512Mb RAM. Nonetheless, it was possible to open the artifact hierarchy which represents about a third of the whole lexical hierarchy). When reasoning over the hierarchy, inconsistencies are detected when encountering multiple inheritances (allowed in WordNet) because this relation is not allowed in OWL (based on description logic). These problems are not essential in our application because we only use hypernymy and class equivalence and it is easy to parse the OWL and to extract the necessary information without the use of a dedicated reasoner.

Multilingual Ontology in Image Retrieval Tasks

We propose an image retrieval framework and it is necessary to create an image database and to propose a structure for its utilization. If an ontology is employed in IR tasks, it proves useful in both stages. It can be used to query an existing image search engine in order to populate the image ontology. A concept hierarchy can equally be used in the utilization phase, to propose structured presentation of results to the users. In the following subsections, we discuss the role of a multilingual ontology for the constitution of a structured image database and its utilization.

Image gathering

A picture database is automatically constituted using an existing image search engine (Yahoo!). Leaf synsets in the English WordNet were selected and inserted into a list and queries for pictures corresponding to each item in the list were launched with Yahoo! and stored into separate directories. The queries for pictures are formulated using only English concepts. There are three main reasons for making this choice:

-the English hierarchy provides better coverage than its Italian and Spanish coverage. This allows us to obtain a broader set of images. Even if a part of the leaf synsets does not exist in Italian and Spanish WordNets, the hypernymy relation makes them relevant for the upper level terms in these two languages. We illustrated the differences between the hierarchies with the example of *necktie*. Another example is that of *guard dog*, which is translated as *cane da guardia* (Italian) and *perro guardian* (Spanish). In English, this concept has 7 leaf synsets attached while in the other two languages, only 1 such synset is attached to the translations of *guard dog*.

-the image retrieval in English is more precise than those in other languages (see the Evaluation section) and the obtained image database is of a better quality when using queries in this language.

-the English leaf synsets have broader images sets associated than their translations. When using the Yahoo! Search engine, there are about 55000 pictures associated to *Malinois*, a dog breed in the English hierarchy and only 3 images associated to its Italian correspondent , *pastore belga di Malines*. This is an anecdotic example that confirms the fact that image retrieval in more comprehensive in English than in other languages.

For each leaf synset in WordNet, we request images for all included terms (if several exist). For example, the class images for *cotton rat* and *Sigmodon hispidus* will be associated to the *cotton_rat__1* class because the 2 terms are synonyms. The utilization of all the members in a

synset increases the number of pictures associated to a concept. This is particularly important for the cases when the number of image responses to a query with a specialized concept is small. A class of images is obtained for each leaf synset (if there are images that are indexed with those concepts by the Yahoo! search engine). The rationale for using leaf concepts in WordNet for querying the Web is that they point towards specific entities, while more general terms represent bigger and more ambiguous parts of the world.

Figure 1: Images for *Alsatian* from Yahoo!.

There are two types of leaf terms in WordNet and different query types are launched with regard to this separation:
-non-ambiguous terms – for these leaf concepts, queries containing only that term are launched. For example, *German shepherd* has only one sense in WordNet and the image query is formulated using the concept alone.
-ambiguous leaf concepts – when a word has several senses in the hierarchy, queries with the word and its immediate hypernym (or hypernyms if several exist) are launched. Disambiguation is discussed in more detail in the next subsection.

Disambiguation

There are two types of ambiguity, intra-lingual and inter-lingual. Hereafter, we present an example of intra-lingual polysemy. *Alsatian*, a synonym of *German shepherd* has two senses in English. The queries for images of *Alsatian* as *dog* will be *Alsatian sheep dog*, *Alsatian shepherd dog* and *Alsatian sheepdog,* where *sheep dog*, *shepherd dog*, and *sheepdog* are the immediate parents of queried leaf concept in WordNet. We present, in figure 1, the first 20 picture responses from Yahoo! obtained for *Alsatian* using uniquely the term. In figure 2, we the

images rendered for *Alsatian shepherd dog*. The pictures in figures 1 and 2 are presented in the order proposed by the search engine.

Figure 2: Images for *Alsatian shepherd dog* from Yahoo!.

We observe that for an ambiguous term like *Alsatian*, with the use of a hypernym to expand the query, the obtained images are more appropriate for the meaning *dog* than those rendered when using the word alone. For the latter case, images of people appear, corresponding to the other meaning of the term, inhabitant of Alsace. Other images in the same set correspond to typical Alsatian dishes, like *flammkuchen* and *sauerkraut*. The same disambiguation technique is used for all polysemic leaf terms in the English WordNet.
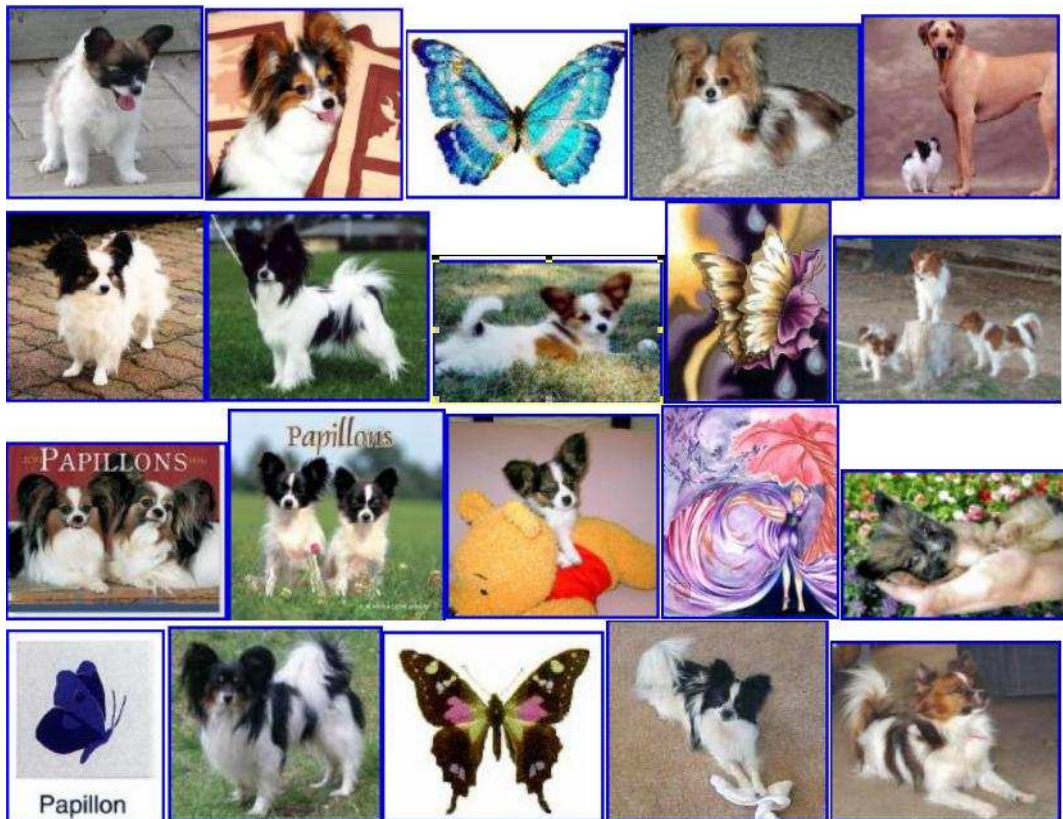
Figure 3: Images for *papillon* from Yahoo!.

Figure 4: Images for *papillon toy spaniel* from Yahoo!.

A second type of ambiguity which appears is the inter-lingual one, when the same term points to different entities in each language. This kind of ambiguity cannot be solved with the use of a monolingual ontology. We present here the example of *papillon*, a term which has a unique entry in the English WordNet and designates a type of *dog*. With the introduction of Italian in

the ontology, we find a second sense, that of *tie* (*cravatta*) and the picture expanded query associated to this concept will be *papillon toy spaniel*. We present, in figures 3 and 4, the first 20 results obtained with the initial query and with the expanded one using the Yahoo search engine.

When a simple query is formulated, there are two main senses of *papillon* that appear: *dog* and *butterfly*. The second sense comes from French, a language that is not yet included in our ontology but for which a WordNet hierarchy exists (Catherin, 1999) and we currently assess its inclusion in the multilingual hierarchy. This example shows that the utility multilingual resources increases with the number of included languages. There is a drastic improvement of precision when using an expanded query for *papillon* compared to the case when the word was used alone. All the images in figure 4 correspond to the desired meaning of the word, a similar situation to that presented in figure 2.

Querying the database

The constitution of a multilingual ontology provides the possibility to query the picture repository attached to the hierarchy. As we described above, the images are gathered for English leaf synsets, but we can offer responses to queries corresponding to terms in English, Spanish or Italian. With the use of English for picture gathering, the image sets for the other two languages are broader than if we would have used the leaf synsets in Italian and Spanish because the English WordNet is by far more developed than the other two.
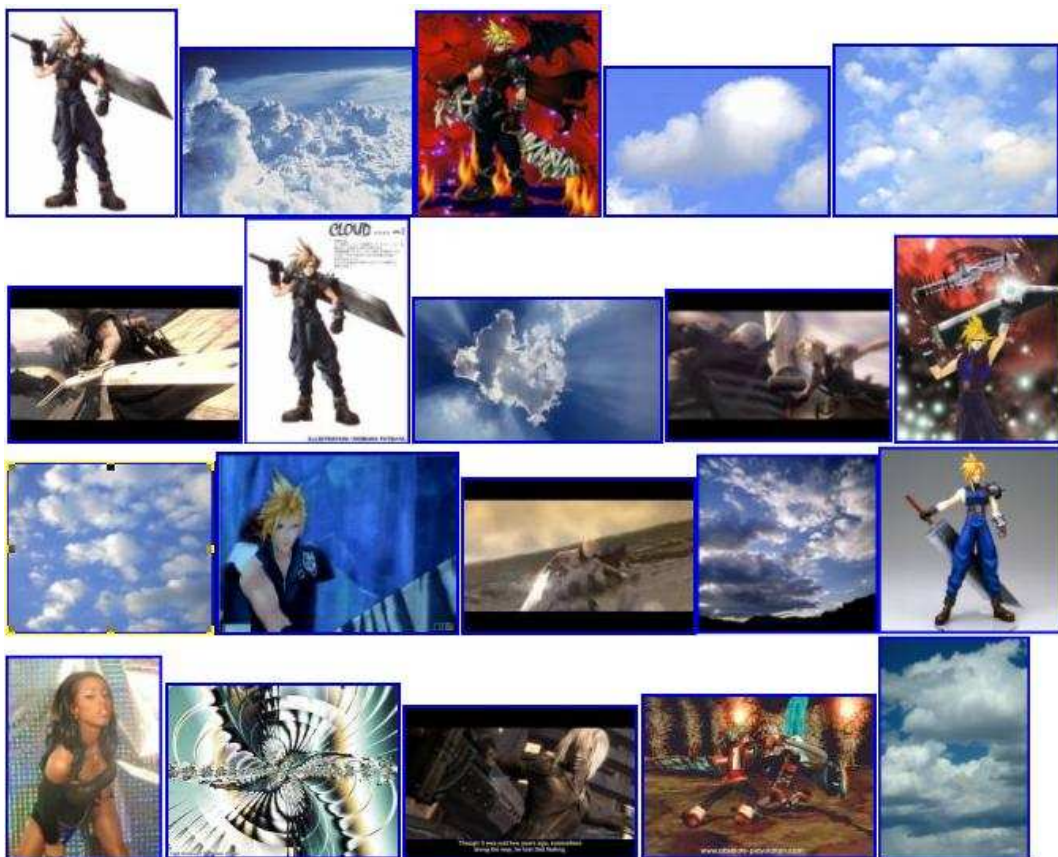

Figure 5: Images for *cloud* from Yahoo!.

We illustrate the differences between current image retrieval paradigm and the use of an ontology for the same task with the case of a query for *cloud* (as *"a visible mass of water or ice particles suspended at a considerable altitude"* (WordNet, 2006)). The Yahoo! Search engine

offers the responses in figure 5 to a query with *cloud* (we present the first page of responses only). Several senses of cloud appear in the answer set. The majority of images depict scenes from cartoons or video games and only 7 responses out of 20 are in accordance with the formulated request. Currently, there is no sense separation for the requests we formulate when using an Internet search engine. The responses are rendered using string matching, without any semantic treatment of the content of the query.

When the images in a database are structured employing an ontology, the responses are organized following the hypernymy relation in the hierarchy. In figure 6, we present some results for cloud using our structured database, with images associated to leaf concepts in the hierarchy. We present results for three leaf subconcepts of cloud: *Cirrocumulus*, *Cumulonimbus* and *Altostratus*. These results in figure 6 also stand for *nube* and *nuvola*, the synonyms for cloud in Spanish and Italian.

Figure 6: Images for cloud from the structured database (Yahoo! is used as raw image source).

When compared to the unstructured picture set for *cloud* (figure 5), the images in figure 6 are semantically organized and the precision of the answers is greatly improved. All the pictures in figure 6 are representative for *cloud*.

Results organization

The inter-categorical organization of image results for a concept was discussed in the above section. This structure is provided using the hypernymy relation in WordNet and it results in improved interactivity options for the user. Once a query is formulated, the proposition of pictures for subsumed concepts using ontological relations is straightforward.

Another structuring dimension is an intra-categorical organization which is obtained using an image clustering algorithm. It is impossible to perform real-time clustering for large image databases and this is an important reason for proposing this facility at the level of leaf concepts. Moreover, with the increase of coverage of a concept, content based grouping of images

becomes irrelevant because pictures from different domains are clustered together. Several picture clusters are proposed for each leaf concept in WordNet that has associated images on the Web. The images are grouped in clusters following low-level visual similarities (color and texture). A presentation of the clustering module is to be found in (Popescu et al., 2006). The grouping of similar images may provide a possibility to have faster access to interesting instances, as they are organized in similar sets.

Up to the moment, we cannot provide an assessment of the effect of organizing the results for image queries but we expect positive evaluation from the users because the system proposes intuitive organization of images and increased possibilities for interaction. We currently design a user test meant to provide information about the effect of structuring and increased interaction.

Evaluation

The assessment of performances for image retrieval applications is not a trivial task. First, the choice of test parameters is not straightforward and second, when it is necessary to evaluate the performances manually, assessment becomes a time costing effort. Two measures are usually employed for quantitative evaluation: precision and recall. For a given query, the former stands for the number of correct answers in the set of retrieved images, while the later accounts for the number of retrieved images out of the total number of images which are relevant for the request. Recall is important for small scale databases, where there are not a lot of images representing a given concept, but its importance decreases for large bases, where each concept is represented by a big number of images. Search precision is an important parameter in both situations.

The Internet is the largest image database available and, for simple queries, the number of image responses is usually enormous (Yahoo! indexes more the 6 million images for *dog* and nearly 1 million for *cloud*). It is highly improbable that all relevant images for a term are indexed by the search engines (e.g.: Google indexes around 3 millions images for *dog*, number which represents less than a half of the images indexed by Yahoo! for the same concept) and, consequently, it is impossible to accurately evaluate recall for Internet retrieval systems. Our application is close to Web search and we propose a precision test in order to assess the performances of our approach. The evaluation framework is meant to asses:
-the variation of precision when querying the Web for images in different languages. We compare the performances of the Yahoo! search engine when querying for images in English, Spanish and Italian.
-the eventual improvement of performances when using an ontology in image retrieval compared to the performances of an existing system. The image database attached to the ontology is created using Yahoo! to gather images and this system is chosen as comparison term.

The pictures in the picture database were not previously evaluated by a user and we are obliged to perform a manual assessment of precision. This is a time costing effort but the results are reliable. We assessed the performances of Yahoo! for 10 terms (table 1). Two criteria guided the choice of the concepts we used in this test:
-we propose categories that are familiar to most of people using a search engine. This roughly corresponds to the basic level of representation for categories established by Rosch and al. (1975).
-coverage of both natural categories and artifacts (manmade objects) is intended

The utilization of terms from different domains gives a fair idea about the generality of the proposed image retrieval method. The above conditions provide the conditions for an easy evaluation of images, as they all represent commonly known concepts. The list of the evaluated concepts, in English, Italian and Spanish, is presented in table 1.

English	Spanish	Italian
Apple	Manzana	Manzana
Car	Coche	Coche
Cloud	Nuvole	Nube
Dog	Cane	Perro
Dolphin	Delfino	Delfin
Eagle	Aquila	Aguila
Flower	Fiore	Flor
Hammer	Martello	Martillo
Rock	Roccia	Roca
Toy	Giocatollo	Juguete

Table 1: List of the evaluated concepts.

When constituting the list of concepts to be evaluated, we employ a classical separation for concepts, that between natural categories and artifacts (Keil, 1992) and propose 7 terms for the former type and 3 for the latter. The natural categories subset includes: 3 concepts for animals which typically live in different environments (water, ground, air), 2 concepts for plants, and two for inanimate entities (cloud and rock). The artifacts contain *car*, *hammer* and *toy*. It is important to note that the choice of familiar concepts is in accordance with the tendency people have to name objects in pictures. Rosch and al. (1975) show that basic level names for categories are preferred over more general or more specific ones. It is probable that the queries in general public applications follow the same trend.

For ambiguous concepts like *dog*, *rock* or *cloud*, only one sense of the term was evaluated as correct. The translations in Spanish and Italian are based on this sense of the English word. If several translations (synonyms) existed for one meaning of an English term, we have chosen the best known word in the two other languages. An interesting case was that of *nube*, term that designates the same entity in Italian and Spanish. We chose to use a synonym, *nuvole*, for Italian.

We present, in table 2, an evaluation of the precision results in four situations: querying the Yahoo! search engine with the ten concepts presented in table 1 in English, Italian and Spanish and the results from a structured image database for the same set of concepts. We remind the reader that leaf terms in the English WordNet are employed to form the image classes that stand for the 10 evaluated concepts when an ontology is used in image retrieval. For each of the four cases, 50 images per concept were proposed for testing. The tester was thus presented with a total of 2000 images and was asked to decide if the picture he saw is representative for the concept it is meant to depict. The evaluator had no information about the way the picture classes were obtained or structured. The only information she had was the English name of the category associated to each photograph. The results for the compared approaches to image retrieval on the Internet are in percents and the last line contains the mean precision for each method.

Concept	Precision[%]			
	English	Italian	Spanish	Ontology
Apple	20	10	36	**60**
Car	84	50	66	**100**
Cloud	46	78	36	**96**
Dog	66	32	90	**100**
Dolphin	72	14	64	**78**
Eagle	46	12	34	**88**
Flower	**90**	56	38	88
Hammer	20	14	20	**48**
Rock	58	**74**	52	72
Toy	38	**64**	54	58
Mean	*54*	*40,4*	*51,4*	*78,8*

Table 2: Number of relevant images out of 50 for each concept. Precision results in 3 languages compared to the use of an ontology in image retrieval.

The results in table 2 show that the precision of the image search varies from one language to another. The mean precision for English is 54%, 2.6% better than that for Spanish. The results for Italian are the worst, with only 40.4% representative images in the evaluated set. These results confirm the fact that the search in English is the best in mean and sustain the choice of this language for populating the picture ontology. It is to be noted that the use of an ontology in image research significantly improves the quality of the image search compared to state of the art systems. When comparing the results of our method to those of Yahoo! for English the difference is of 24.8%, with even greater precision gains over Spanish (27.4%) or Italian (38.8%). In table 2, we stressed the best results obtained for every category. With the use of the ontology the obtained results are superior for 7 concepts out of a total of 10 when compared to queries in one of the three evaluated languages. We note that very good results are obtained for *car*, *cloud*, *dog* or *eagle* when using a concept hierarchy. For the three other terms, our method does not obtain the best score, but it is second best with a small difference compared to the best results. The obtained results show that introduction of a multilingual ontology ameliorated image retrieval. The improvement of precision is important when compared to any of the three tested languages and we think it is worthwhile to propose unique sets of images for queries in all languages in the ontology.

Conclusions and perspectives

In this paper we discussed the utility of using multilingual semantic resources in image retrieval tasks. First, we presented techniques for automatically constituting a multilingual OWL ontology using existing linguistic hierarchies. This integration allows inter-lingual disambiguation and enables the system to respond to queries in all languages includes in the ontology and to perform the necessary reasoning in order structure the images attached to the hierarchy.

Second, we proposed a comparison between state of the art image search engines and the case when an ontology is used. A precision test for 10 concepts proved that, when employing concept hierarchy in image retrieval, the precision in the response sets is significantly improved. We equally discussed the picture structuring advantages obtained with the introduction of an ontology in image retrieval tasks.

We presented an exploratory work that provides encouraging results, but there are a number of issues that are to be treated in the future. For example, when presenting the results for a class

that has many associated leaf concepts, it is impossible to present excerpts for all subcategories. One potential solution is to present for the leaves that appear the most often in the language. A second problem is that of the leaf concepts that exist exclusively in English. When querying in another language than English, it would be possible to propose, at least in a first time, only responses for leaves of the hierarchy that exist in the respective language.

References

Catherin, L. The French WordNet.*Technical report*, 1999.

Celino, I., Della Valle, E, Cerizza, D., Turati, A. (2006) Squiggle : a Semantic Search Engine for indexing and retrieval of multimedia content. In *Proc. of First International Workshop on Semantic-enhanced Multimedia Presentation Systems*, Athens, Greece.

Daudé, J., Padró, L. & Rigau, G. (2000). Mapping WordNets Using Structural Information**.** In *Proc. of the 38th Annual Meeting of the Association for Computational Linguistics,* Hong Kong.

EuroWordNet project (1999). *http://www.illc.uva.nl/EuroWordNet/* (consulted 25/11/2006)

Gangemi, A., Navigli, R. & Velardi, P (2003). The OntoWordNet Project: Extension and Axiomatisation of Conceptual Relations in WordNet. In *Proc of. CoopIS/DOA/ODBASE 2003*, (pp. 689–706), Catania, Italy.

Keil, F. C. (1992). Concepts, kinds, and cognitive development. *MIT Press*.

Miller, G. A. (1990). Nouns in WordNet: a Lexical Inheritance System. *International Journal of Lexicography* (pp. 245-264), 3(4).

Ordan, N. & Wintner, S. (2005). Representing natural gender in multilingual lexical databases. *International Journal of Lexicography* (pp. 357-370), 18(3).

Pianta, E., Bentivogli, L., Girardi, C. (2002). MultiWordNet: developing an aligned multilingual database. In *Proc. of the First Int. Conference on Global WordNet*, Mysore, India.

Popescu, A., Grefenstette, G. & Moëllic P.-A.(2006b). Using Semantic Commonsense Resources in Image Retrieval. In *Proc. of the First Semantic Multimedia Adaptation and Presentation Workshop 2006*, Athens, Greece.

Protégé project (2006). *http://protege.stanford.edu/* (consulted 13/12/2006).

Racer (2006). *http://www.sts.tu-harburg.de/~r.f.moeller/racer/* (consulted 13/12/2006).

Rosch, E., Mervis, C.B., Gray W.D., Johnson D.M. and Boyes-Braem, P.(1976). "Basic Objects in Natural Categories," *Cognitive Psychology* (pp. 382-439)*, 8(3).

Van Assem, M., Gangemi, A. and Schreiber, G. RDF/OWL Representation of WordNet. *http://www.w3.org/TR2006/WDwordnet-rdf-20060619* (consulted 25/11/2006).

Wang, X. J., Ma, W. Y., and Li, X., (2004). Data-driven Approach for Bridging the Cognitive Gap in Image Retrieval", In *Proc. of ICME 2004* (pp. 2231-2234), Taipei, Taiwan.

WordNet project (2006). *http://wordnet.princeton.edu/* (consulted 25/11/2006).