

# Imaging Words – Wording Images

Adrian Popescu, Gregory Grefenstette, Cristophe Millet, Pierre-Alain Moëllic, Patrick Hède

**Abstract**—The rapid growth of the Internet information sources has led to organizing proposals, such as the Semantic Web initiative, with its ontological level providing a formal structuring for this disparate data. But given the amount of information to be treated even in a restricted domain, manual organization becomes rapidly unmanageable, and automatic methodologies for ontology building are required. Here we describe techniques for the automatic construction of a image ontology based on multimedia data (text and images) for a specific class of objects, manmade tools. Our approach combines modification of existing lexical resources and search engine querying in order to obtain raw images. These images are then clustered into representative concepts for the ontology. Our automated approach can be applied to any subset of physical objects.

**Index Terms**—Image, Ontology, OWL, Semantic Web, WordNet

## I. INTRODUCTION

As proven by initiatives like CYC [2], the manual construction of large scale ontologies is a costly effort and it is unrealistic to think that this approach can solve current needs for knowledge organization. This is especially true for highly dynamic resources like the WWW, where the increase in knowledge resources follows an exponential curve. The Semantic Web, with its description of content in ontologies has been presented as a potential solution to the information structuring problems. But, as underlined in [1], a vicious circle is created as the Semantic Web is dependent on the existence of metadata and these last rely themselves on the existence of a well populated Semantic Web. A way to cope with this problem is the development of automatic or semiautomatic methodologies for the ontology construction. Interesting results for automatic lexical ontology building are reported in [1].

In this paper, we describe our technique for automatically filling multimedia ontologies, grounding each concept in text and images. After a transposition of parts of WordNet [Miller] into OWL (Ontology Web Language) in order to create a taxonomical base, we have lexical information associated to concepts. For the image part of the grounding, we query the Web to gather pictures corresponding to objects in the taxonomy that are then clustered and filtered.

We structured the rest of this paper as follows: we discuss a translation of WordNet to OWL, we describe our image gathering and clustering tool and, before concluding, some

preliminary results of our method for image ontology construction.

## II. WORDNET TO OWL

### A. Automatic Ontology Construction

Our current work deals with the automatic construction of a grounded ontology. In order to automatically build such a formal structure, we need an associated taxonomy. There are two main possibilities that are offered to us: learning taxonomy from concepts found on the Web [1] or using one from an existing resource. We have chosen the second variant and used WordNet [5] as source for our taxonomy. Thus, we preserve the automatic character of our methodology and are able to exploit the richness of a resource that was manually constructed by lexicographers. We are aware of the criticisms raised by the transformation of WordNet into a formal ontology [4], but with the implementation choices we have made, we try to minimize their effects. There is notably the fact that our method only addresses picturable objects, which are ontologically less controversial than high level concepts. The approach we propose is domain independent. It depends uniquely of the knowledge contained in the resource we parsed. For exemplification purposes only, the examples furnished here are subconcepts of *tool* in WordNet.

The envisioned application, construction of a structured image catalogue, determined us to parse only parts of the information contained in WordNet to OWL. We transformed the sets of synonyms (synsets) in OWL classes, preserving the sense separation. Thus, *knife* from the lexical hierarchy becomes *knife\_1* in the ontology, while *garden tool*, *lawn tool* is transformed to *garden\_tool\_1*. *Lawn tool* is saved as an RDFS comment as another member of the *garden\_tool\_1* class. We equally parsed the terms definitions in the ontology.

Image clusters are associated exclusively to leave concepts in the OWL ontology. The rationale for this decision is that, with the use of hyponymy relation, we can propose image sets for all concepts in the ontology. Moreover, the leave terms generally are specialized concepts that point towards precise entities [6], which are less ambiguous both in language and the associated picture representation.

## III. IMAGE CLUSTERING MODULE

We propose a second structuring axis in our image catalogue. The use of an ontology allows inter-class organization, while an image clustering tool provides means for intra-class structure. A clustering process was run for each

leaf concept in the ontology. This process consists of two steps: image indexing and clustering following visual similarity.

#### A. Image indexing

We deal with pictures from broad domains and we need a general image indexing technique. Using an approach based on border/interior pixel classification [7]. We construct two histograms for each image, one for pixels on the image borders and pixels in interior regions. This indexing algorithm is fast, simple and provides information about colors in the image and, equally important, about sizes of image regions having a constant color (possibly objects). It leads to the construction of a vector containing 128 elements for each picture. We use the Riemann distance as similarity measure between two images. Distances are calculated between all pairs of images.

#### B. Image clustering

The indexed images are clustered using a k-SNN (Shared Nearest Neighbors) algorithm [3]. For each image, a neighborhood of k images is considered in the algorithm. The similarity of two images is assessed with respect to the degree of overlapping of their neighborhoods. Next, pictures that are most similar to their neighbors are considered as topic images and clusters are structured around them. A useful feature of the algorithm is that it does not impose the classification of all indexed images. Pictures considered weakly related to topics remain unclustered. This last feature is important in our application as we work in a noisy environment (there are a lot of images on the Web that are not annotated in direct relation to their visual content). We thus hope to isolate images that are irrelevant for the desired object and build highly coherent clusters of images containing it. Given that the classification is entirely automatic, there is noise that subsists in the clusters, but the obtained results seem more coherent than the set of images initially retrieved, though we have not yet performed extensive evaluation.

### IV. PRELIMINARY RESULTS

We already stated that our purpose here is to build a structured image catalogue using images from the Web. Instead of querying for images for all concepts in the ontology, we perform this operation for leaves only and, via hyponymy, propose picture sets for all other concepts in the hierarchy. This results in a structured presentation of results, while taking advantage of the fact that the image sets associated to leaves are less noisy (they correspond to well defined entities in the world[6]). An example of the obtained results is presented for *knife* in two situations. We use Google Image for the pictures in fig. 1 and our method (ontology for inter-class structure and clustering for intra-class organization).



Fig. 1. Selection of images for *knife* using Google Image.

#### Cluster for *carving knife*



#### Cluster for *bread knife*



#### Cluster for *clasp knife*



Fig. 2. Selection of images for *knife* using ontologies and image clustering

We observe that the images in fig. 2 illustrate better the notion of *knife* and are ontologically and visually organized, which is not the case for figure 1. Extensive evaluations are needed in order to assess if the proposed method performs better than existing ones in image retrieval tasks.

### REFERENCES

- [1] P. Cimiano, A. Hotho, and S. Staab, "Comparing Conceptual, Divisive and Agglomerative Clustering for Learning Taxonomies from Text", in Proc. of ECAI 2004, Valencia, Spain, 2004, pp.435–439
- [2] CYC, www.cyc.com
- [3] L. Ertoz, M. Steibach, and V. Kumar, "Finding Topics in Collections of Documents: A Shared Nearest Neighbor Approach", In: Wu, W., Xiong, H., Sheklar, S.(eds.): Clustering and Information Retrieval, Kluwer, 2003.
- [4] A. Gangemi, R. Navigli, and P. Velardi, "The OntoWordNet Project: Extension and Axiomatization of Conceptual Relations in WordNet", in Proc. of CoopIS/DOA/ODBASE 2003, Catania, Sicily, Italy, 2003, pp. 689–706.
- [5] G. A. Miller, "Nouns in WordNet: a Lexical Inheritance System. International Journal of Lexicography", 3,4, 1990, pp. 245–264.
- [6] E. Rosch, C. Mervis, C. B. Gray, D. M. Johnson, P. Boyes – Braem, "Basic Objects in Natural Categories", Cognitive Psychology, 8, 1976, pp. 382–439.
- [7] R. O. Stehling, M.A. Nascimento, and A.X. Falcao, "A Compact and Efficient Image Retrieval Approach Based on Border/Interior Pixel Classification", in Proc. of the Eleventh International Conference on Information and Knowledge Management. Mc Lean, Virginia, USA. ACM Press, 2002, pp. 102–109.