A Framework for Content-Based Image Retrieval Fully Exploiting the Semantics of Annotation

Tommaso Di Noia, Eugenio Di Sciascio, Francesco di Cugno, Eufemia Tinelli
Politecnico di Bari, 70125 BARI, Italy
{t.dinoia,disciascio,f.dicugno,e.tinelli}@poliba.it

Francesco M. Donini
Università della Tuscia, 01100 Viterbo, Italy
donini@unitus.it

ABSTRACT

We present a framework and an application for semantic-based retrieval of images. Our approach adopts a two-level ontology structure in a subset of OWL-DL. In the core ontology only generic spatial relations are represented, while domain ontologies are specific for the image collection. The approach allows semantic-based relevence ranking and results explanation for query refinement, by exploiting standard and non-standard inferences in Description Logics.

1. INTRODUCTION

We present an innovative Description Logics (DLs) framework for semantic-based image retrieval that introduces non-standard inferences for conceptual query and query refinement. DL-based engines usually provide at least two basic reasoning services, namely subsumption and satisfiability. Satisfiability accounts for the internal coherency of the description of a concept (no contradictory properties are present), and subsumption accounts for the more general/more specific relation among concepts, that forms the basis of a taxonomy. Such services provide basically a yes/no result. This might be not enough when such services have to provide query answering. For content-based image retrieval we believe that at least a ranking function should be provided, and this should be based on logical criteria. To this aim we propose using non-standard inference services, namely Concept Abduction and Concept Contraction [1, 2], and show here their rationale in semantic-based image retrieval. *Concept Abduction* captures the reasoning mechanism – namely, making hypotheses – involved when some of the features required in a query Q are not specified within available image descriptions. *Concept Contraction* captures the possibility to relax some of the features requested in Q when they are in conflict with those of an available image description. We use such inferences in our approach both to provide interaction and refinement in the query / retrieval process and to rank images based on their semantic-based distance from the query.

2. FRAMEWORK AND PROTOTYPE

Let us consider a simple conceptual query:

Q: *I'm looking for images showing a skyscraper, the sun and the sea as background.*

Let us suppose that, among various images, there are three in the knowledge base whose –simplified– descriptions are as follows:

Im1 *Landscape with a Person and the sun.*

Im2 *Person near a building and a garden as background.*

Im3 *Building in background with a garden, and the sun.*

Expressing both the query and image descriptions with respect to a DL we obtain:

Q: $Skyscraper \sqcap Sun \sqcap \exists background \sqcap \forall background.Sea$

Im1: $Person \sqcap Sun$

Im2: $Building \sqcap Person \sqcap \exists background \sqcap \forall background.Garden$

Im3: $Building \sqcap Sun \sqcap \exists background \sqcap \forall background.Sea$

Now suppose to have a simple ontology containing only two axioms, *i.e.*, $\mathcal{T} = \{Skyscraper \sqsubseteq Building; Sea \sqsubseteq \neg Garden\}$.

With respect to the previous ontology we can evaluate the match degree between **Q** and each image description as well as some semantic-based explanations, exploiting the relations among concepts as formalized within the ontology axioms. Notice that the approach is based on an Open World Assumption (OWA), *i.e.*, the absence, in an image description, of a feature requested in the query is not interpreted as a constraint of absence, or, in other words, as if the feature was negated. The OWA allows to handle incomplete information. It is hence possible to deal with underspecified descriptions.

SIM(Im1,Q): **Im1** can potentially satisfy **Q**. No feature in **Im1** is conflicting with those represented in **Q**, but characteristics depicted in **Im1** do not completely fulfill **Q**. Using DLs syntax we say $Q \sqcap Im1 \not\equiv \bot$ and $Im1 \not\sqsubseteq Q$. Explanation hypothesis are needed for the non-subsumption relation. Solving the related *Concept Abduction Problem*, a solution is $H_{Im1} = Skyscraper \sqcap \exists background \sqcap \forall background.Sea$. Considering **Im1**, **Q** and the ontology $\mathcal{T}$, a semantic-based score for H_{Im1} has to be computed.

SIM(Im2,Q): Features in **Im2** are conflicting with what is requested in **Q** *i.e.*, $Im2 \sqcap Q \equiv \bot$. A revision of the query may be needed if we are interested in images similar to **Im2**. Solving the related *Concept Contraction Problem* a solution is $Q = G_Q \sqcap K_Q$, where $G_Q = \forall background.Sea$ and $K_Q = Skyscraper \sqcap Sun \sqcap \exists background$. Now the relation $K_Q \sqcap Im2 \not\equiv \bot$ holds, but we have $Im2 \not\sqsubseteq K_Q$. Again, solving the related Concept Abduction Problem a solution $H_{Im2} = Skyscraper \sqcap Sun$ results.

The ranking function, computing the overall score, in this case must consider G_Q, K_Q, H_{Im2} and/or the relative scores for each of them.

SIM(Im3,Q): $Q \sqcap Im3 \not\equiv \perp$ and $Im3 \not\sqsubseteq Q$. $H_{Im3} = Skyscraper$.

Notice that because of the axiom $Skyscraper \sqsubseteq Building$ in $\mathcal{T}$, the concept $Skyscraper$ in H_{Im2} must be evaluated differently from $Skyscraper$ in H_{Im2}. To cope with this property, both monotonicity and anti-monotonicity behavior [3, 4] has to be respected by the penalty function computing the overall score. Ranking $Im1$, $Im2$ and $Im3$ w.r.t. Q, the first image in the list is $Im3$. Obviously, if there were one or more images ImX, such that $ImX \sqsubseteq Q$, then ImX would be the highest ranked.

We would like to point out we are not expressing here an explicit ranking function. Various ranking functions can be in fact defined, having as arguments concepts provided by G, K and H, which can be used to determine the match degree. The structure of the annotation, at an abstract level, can be exploited to perform a matching process which takes into account the nature of the annotated resources – the images – and the annotation itself. To this aim we modeled a core ontology considering only spatial relations. Such ontology is strongly property oriented, that is the notions of `background`, `up`, `down`, `on the left side`, `all around` and so on, are modeled as properties – from now on R^i_{core} – whose restrictions describe the content of a particular region in the picture. Both the image annotations and the queries are expressed as a conjunction of quantified properties. Such properties, R^i_{core}, are related with each other. For instance the semantics of `around` is related to the ones of `left` and `right`. If something is all around the picture, then it is on the `left` side and on the `right` one. The core ontology contains the set $\mathcal{R} = \{R^i_{core}\}$ and axioms on such properties stating their mutual relation. To model the actual content of images we consider domain-specific ontologies. Then, using the $< owl : imports/ >$ OWL TAG, not only we make possible to use different sets of knowledge domains but we allow to introduce the content within an image annotation. Using $< owl : imports/ >$ we import the domain ontologies within the image core ontology, reusing such ontologies for image content description. Since the matching process involves basically the image content, domain ontologies were developed using the $\mathcal{ALN}$ subset of OWL-DL, for which algorithms to solve Concept Abduction and Contraction problems exist [1, 2]. Importing in the core image ontology the domain ontologies, we can annotate images with reference both to spatial relations and actual image content. The semantic-based image retrieval process is performed as explained hereafter, with the aid of the following simple image annotation Im and query Q :

Im: $\forall \texttt{background} .(b\_annot) \sqcap \forall \texttt{around}.(a\_annot)$

Q: $\forall \texttt{within}.(w\_query) \sqcap \forall \texttt{left}.(lx\_query) \sqcap \forall \texttt{right}.(rx\_query)$.

1.The user selects the content domain (or a set of content domains). This selection corresponds to the identification of the domain ontology (or a set of domain ontologies) to be imported in the core image ontology.

2.The user selects the properties she is interested in, with respect to the core image ontology, and composes restriction of such properties (*i.e.*, the content of selected image regions). Considering Q qnnotation the user first selects the property `within` and then composes its restriction – *w_query* – using domain ontologies, then `left` with its *left_side_request* and finally `right` and the related *rx_query*. **3.**For each R^i_{core} in the query, the corresponding restriction is selected both in the query and in the image annotation and an extended matchmaking process is performed (See [4] for details on the concept of extended matchmaking). In the image annotation, also the restriction related to properties R^j_{core} such that $R^j_{core} \sqsubseteq R^i_{core}$ are selected and put in conjunction with the one of R^i_{core}.

With reference to our example query Q and image annotation Im the following steps are executed:

a. *w_query* is selected as the restriction of `within` in Q.

b. in Im, due to the sub-property relations in the image core ontology, the restriction of both `background` – *b_annot* – and `around`– *a_annot* – are selected.

c. an extended matchmaking process is performed between $a\_annot \sqcap b\_annot$ and *w_query* returning $\langle G_{\texttt{within}}, K_{\texttt{within}}, H_{\texttt{within}}\rangle$.

d. *lx_query* is selected as the restriction of `left` in Q.

e. in $Im$the restriction of `around`– *a_annot* – is selected.

f. an extended matchmaking process is performed between *a_annot* and *lx_query* returning $\langle G_{\texttt{left}}, K_{\texttt{left}}, H_{\texttt{left}}\rangle$.

g. a similar process is performed for *rx_query* as the restriction of `right`in Q returning $\langle G_{\texttt{right}}, K_{\texttt{right}}, H_{\texttt{right}}\rangle$.

Notice that for R^i_{core} in Q, if neither R^i_{core} nor any $R^j_{core} \sqsubseteq R^i_{core}$ are in Im, then $\top$ is considered as the restriction of R^i_{core} in Im.

4. a score is computed by means of a penalty function using $\langle G_{\texttt{within}}, K_{\texttt{within}}, H_{\texttt{within}}\rangle$, $\langle G_{\texttt{left}}, K_{\texttt{left}}, H_{\texttt{left}}\rangle$ and $\langle G_{\texttt{right}}, K_{\texttt{right}}, H_{\texttt{right}}\rangle$.

In the query the user is also able to express a strict constraint, that is the restriction of a property R^i_{core} in the query must not be in conflict with the corresponding restriction in the retrieved image. The above user specification is modeled in our framework as a condition on $G_{R^i_{core}}$. If the user states R^i_{core} as a strict property, then $G_{R^i_{core}} \equiv \top$ must occur. No Concept Contraction has to be performed on the strict property restriction. For instance, with respect to the above example, if the user states that $R^i_{core} = \texttt{left}$ is a strict constraint in the query, then the result of the step 3.f. **must** be $\langle \top, lx\_query, H_{\texttt{left}}\rangle$ otherwise the image is not selected and is discarded. That means that if the `left`specifications are not in conflict with the related ones in the image annotation then no contraction is needed – $G_{\texttt{left}} = \top$, nothing as to be contracted, and $K_{\texttt{left}} = lx\_query$, all the `left`specification is kept – and then the strict constraint is respected, making the image selected and presented within the ranked result list to the user.

3. REFERENCES

[1] S. Colucci, T. Di Noia, E. Di Sciascio, F.M. Donini, and M. Mongiello. Concept Abduction and Contraction in Description Logics. In *DL'03*, 2003.

[2] S. Colucci, T. Di Noia, E. Di Sciascio, F.M. Donini, and M. Mongiello. A Uniform Tableaux-Based Method for Concept Abduction and Contraction in Description Logics. In *ECAI '04)*, pages 975–976, 2004.

[3] T. Di Noia, E. Di Sciascio, F.M. Donini, and M. Mongiello. A system for Principled Matchmaking in an Electronic Marketplace. *IJEC*, 8(4):9–37, 2004.

[4] T. Di Noia, E. Di Sciascio, and F.M. Donini. Extending Semantic-Based Matchmaking via Concept Abduction and Contraction. In *EKAW 2004*, pages 307–320. 2004.