Semantic Enrichment of Multimedia Document

Ajay Chakravarthy [1], Vitaveska Lanfranchi[1], Daniela Petrelli[2], Fabio Ciravegna[1]

[1] Department of Computer Science, University of Sheffield
[2] Department of Information Studies, University of Sheffield
Regent Court, 211 Portobello Street,
Sheffield, S1 4P, United Kingdom
{a.chakravarthy, v.lanfranchi, f.ciravegna}@dcs.shef.ac.uk,
{d.petrelli}@sheffield.ac.uk

1 Introduction

Nowadays a large and growing amount of information is stored in various multimedia formats, such as images, video, audio.. While much research has been undertaken into the efficient storage, access, usage and retrieval of textual information, this is still under developed for multimedia information.
In order for multimedia information to be efficiently utilised it is very important to add suitable annotations.

Background

The Web Intelligence Technologies group is part of the NLP Research Group within the Department of Computer Science at the University of Sheffield. Our research activity is focused on Intelligent Technologies for tomorrow's Web, with a particular focus on Technologies for Knowledge Sharing and Reuse, which includes tools for annotating information.

In our current projects (IPAS, X-Media) we are cooperating with industrial organisations that already annotate a vast quantity of information but in an unstructured way. Semantic Web Technologies can have a great impact in these evironments, helping them to better annotate and utilise their knowledge.

Research activity

Our research activity on annotations started with the development of a tool for annotating textual document, called Melita [1]. Melita is an ontology-based annotation interface that uses Adaptive Information Extraction from texts (IE) for reducing the burden of text annotation. While users annotate texts, the IE system learns in the background inducing rules on how to recognise annotations. When a rule reaches a (user-defined) level of accuracy, Melita presents new texts with a preliminary annotation derived by the rule application.

Annotation technologies play a very important role in our new projects (IPAS and X-Media) however the scenario is more advanced than the Melita tool, as the documents we are dealing with are Multimedia documents, i.e. that contain both text and images, interlinked and referring to each other.
Many other tools exist for annotating text for the Semantic Web that do not provide any integration with images; in recent years tools for annotating images have been developed, but again they isolate the type of content to annotate without providing textual annotation.
In this scenario we need to formalise a new method for adding annotations, that reduces the boundaries between free-text and ontology-based annotation. This enables a user to recognise instances inside images and text, relate them, insert comments and annotations.
Annotated multimedia documents will then be available in RDF format for intelligent retrieval, manipulation, allowing the user to formulate queries like "Which are the main causes of failure of an engine?".

AKTiveMedia

Our group has started the development for a tool for this scenario, called AKTiveMedia, a User Centric system for Multimedia Document Enrichment, that uses Semantic Web and language technologies for acquiring, storing and reusing knowledge.

AKTiveMedia has three modalities of use, namely, Image Annotation mode, Text annotation mode and Multimedia annotation mode (where both text and images can be annotated at the same time). For all the modalities, it is possible to annotate using a mix of ontology-based and free-text annotations, and is possible to add comments and relations. Moreover AktiveMedia embeds an EXIF metadata extractor for enriching images with sophisticated metadata as shutter speed, exposure and so on.
For what regards specifically image annotation, AKTiveMedia allows the user to perform batch annotations (for annotating whole images inside a collection) or to annotate regions of the image.
The aim is to provide a seamless interface that guides the user through the process, reducing the complexity of the task. Languages technologies and a web service architecture are adopted to provide a context specific annotation mechanism that uses suggestions inferred from both the ontology and from the previously stored annotations to help the user: the ontology is pre-filtered to present only the top-level concepts (the most generic ones); when the users identify an instance of a top-level entity the system suggests the possible relevant sub-entities, e.g. when annotating a region of an image as a "part", the system suggests all the possible parts present in the ontology for that engine and the user can select the right one. The same happens for relations, again inferred from the ontology and suggested to the user on the base of the concept selected: for example, when the part has been chosen, the user can select a "has_fault" relation and drag and drop the text in the document that describes the fault (see Figure 1).

Formatted: p1a

Deleted: Figure 1

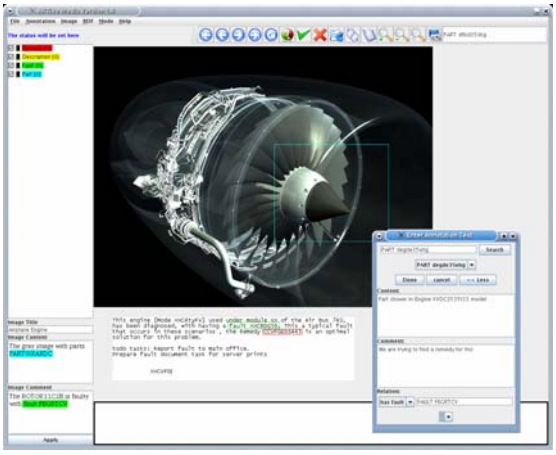

Figure 1 - AKTiveMedia interface

When an instance has been identified it is possible to add free-text comments, to highlights findings or opinions, or to state some generic or contextual information about the instance, e.g. findings on why the failure happened.

Formatted: p1a

The produced knowledge is then used as a way to establish connections with and to navigate the information space: when the user annotates a part of an image as a "sand-damage" on a "turbine" the system uses those annotations to retrieve other related images and documents. New relationships can then be established with the found knowledge, e.g. the damage can be related to other previous cases, and through free-text comments the relationship may be made explicit (e.g. this type of failure happens constantly on this blade in hot conditions, and this is proved by document x).

AKTiveMedia architecture uses a two steps persistence model to save the annotations: 1) the annotations are fist saved in a local repository, 2) then they are imported by a web service into a RDF central repository.

This operation is repeated at regular intervals (see Figure 2).

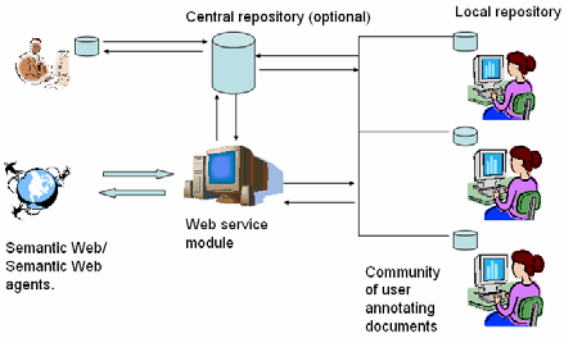

Figure 2 - AktiveMedia two step persistent architecture

Our future research activity will be focusing upon how to improve user experience when annotating multimedia documents and how to offer a satisfying way to query the resulting semantic information. In a future perspective we are also planning to integrate a semi-automatic annotation strategy that will be complementary to the user role, learning from past annotations and relations and offering suggestions to the user, which will then be able to accept or reject them.

References

1 F. Ciravegna, A.i Dingli, Y. Wilks and D. Petrelli: Using Adaptive Information Extraction for Effective Human-centred Document Annotation. In J.Franke ,G.Nakhaeizadeh ,I.Renz (ed.): "Text Mining, Theoretical Aspects and Applications", Lecture Notes in Computer Science, Springer Verlag, 2003