

Towards Cross-Media Document Annotation

Ajay Chakravarthy

Department of Computer Science, Regent Court, Portobello Street,
Sheffield S1 4DP
A.Chakravarthy@dcs.shef.ac.uk

Collecting and aggregating multimedia knowledge is of fundamental importance for every organisation in order to gain competitiveness and to reduce costs. It is possible that knowledge contained in just one medium E.G. text documents, does not carry the full evidence looked for. Therefore connecting information stored in more than one medium is often required. It is clear that current knowledge management technologies and practises cannot cope with such situations, as they mainly provide simple mechanisms (E.G. keyword searching). Currently knowledge workers manually pierce together the information from different sources. In this report we focus and envisage research methodologies that will enable the semantic enrichment of multimedia documents, both on multiple media and across media through annotation.

Annotation of a document is a complex and labour intensive task. So far, research has focused [1][2][4] on supporting the annotation of single media. Much less attention has been paid to the issue of annotating material across media. For this reason there is a growing interest in developing methodologies able to capture the content and the context of multimedia documents, in order to enable effective searching (and document-based knowledge management in general). Previous research in personal image management [6] and text annotation [3] demonstrated how annotating images or documents could be a way to organise information and transform it into knowledge that can be used easily later. Metadata enables the creation of a knowledge base which can then be queried as a way both to retrieve documents (via content and context) and to query the structured data (E.G. creating charts illustrating trends). We address many of these problems with AKTive Media¹ which is a system implemented during the PhD. AKTive Media is a user centric ontology based cross-media annotation system. The goal is to automate the process of annotation by means of knowledge sharing and reuse, thereby reducing user effort during the annotation process. The system actively queries web services and central annotational triple stores as a background service to look for context specific knowledge. The aim is to provide a seamless interface that guides the user through the process, reducing the complexity of the task. Language 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); The produced knowledge is then used as a way to establish connections with and to

¹ <http://www.dcs.shef.ac.uk/~ajay/html/cresearch.html>

navigate the information space: Example when the user annotates a part of an image of a car engine as “abrasion-damage” on a “crank-shaft” 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). AKTive Media has information extraction (IE) plug-ins built in (T-Rex)[5] which automate the annotation of textual documents. The main difference between our approach when compared to other state of the art annotation approaches [4][6] is that we use knowledge across media for annotation and also to further relate these annotation instances. This helps in greatly reducing user effort during manual annotation of documents, by providing intelligent suggestions derived from across media, the IE engine and from the central annotation server. The other major difference is that in AKTive Media, an effort is made to bridge the semantic gap between low level image features and semantically annotated metadata provided by users during annotation. We achieve this by providing means to index image collections and enable the user to query over the index using the visual content of the source image that is being annotated, the user can then use free hand mark-up over regions of the images to perform semi-automatic image segmentation and map high level ontology concepts to these segmented regions.

This research methodology has been deployed in various research projects including AKT, Memories for Life, X-Media and we have scheduled a detailed user evaluation in Rolls Royce UK at the end of year, for the annotation of strip reports.

References

1. Ciravegna F., Dingli A., Petrelli D. and Wilks Y.: User-System Cooperation in Document Annotation based on Information Extraction. In Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02), 1-4 October 2002 - Siguenza (Spain)
2. Dzbor M., Domingue J., Motta E. Towards a semantic web browser. Knowledge Media Institute, The Open University, Milton Keynes. UK. 2002
3. Handschuh S., Staab S., and Ciravegna F. S-CREAM- Semi-automatic CREAtion of Metadata. In Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02), 1-4 October 2002 - Siguenza (Spain), Lecture Notes in Artificial Intelligence 2473, Springer Verlag
4. Hendler J., Bijan P., Grovel M., Schain A., Golbek J., Halaschek-Wiener C., PhotoStuff – An Image Annotation Tool for the Semantic Web, 2003. University of Maryland, MIND Lab, 8400 Baltimore Ave., College Park, MD 20742, USA 2NASA Headquarters, Washington, DC 20546, USA.
5. Iria, J. T-Rex: A Flexible Relation Extraction Framework. In Proceedings of the 8th Annual Colloquium for the UK Special Interest Group for Computational Linguistics (CLUK'05), Manchester, January 2005.
6. Kuchinsky A., Pering C., Creech M.L., Freeze D., Serra B., Gwizdka J., FotoFile: A Consumer Multimedia Organisation and Retrieval System, Proceedings of ACM CHI99 Conference on Human Factors in Computing Systems, 496-503, 1999.