

AHRC ICT Methods Network Workshop

ANNOTATING IMAGE ARCHIVES TO SUPPORT LITERARY RESEARCH

WELSH E-SCIENCE CENTRE, CARDIFF UNIVERSITY, 10 – 11 OCTOBER 2007

Report by Omer Rana, Paul Rosin, David Skilton, Julia Thomas

Introduction

This report summarizes the discussion at the AHRC ICT Methods Network funded workshop on 'Annotating Image Archives to Support Literary Research', which took place at the Welsh eScience Centre (Cardiff) on 10 and 11 October, 2007 (in collaboration with the Centre for Editorial and Intertextual Research).

The Centre for Intertextual Research (CEIR) at Cardiff has completed an AHRC funded project to create a Database of Mid Victorian Wood Engraved Illustrations (DMVI). Taking 1862 as a representative sample year, it contains 868 Victorian literary illustrations, drawing in the main on two major collections: the periodical illustrations of the 1860s and 1870s in the School of Art Museum and Gallery, Aberystwyth University, and the Forrest Reid collection in the Ashmolean Museum, Oxford. Each image in the database is accompanied by iconographic annotations, enabling searches to be performed not only on the title, artist, author, and other material and bibliographic properties, but also by the pictorial content of the illustrations themselves. Successful iconographic description requires a standardized vocabulary and a set of keywords (divided into a set of predefined "facets"), which allow the database user to consistently locate illustrations matching their search terms. Identifying the best representation scheme to provide these annotations has been the basis of discussion between researchers within CEIR, the Welsh eScience Centre and the School of Computer Science.

Based on research within computer science, for instance, such keywords and their relationships may be expressed using Resource Description Framework (RDF) and the Ontology Web Language (OWL). In addition, there could also be potential benefit in adopting techniques from Content Based Image Retrieval (CBIR) research – which has primarily focused on using the content of an image to match against similar types of images within a database. In this approach, the focus is on comparing a template image with those in an archive, rather than providing explicit annotations/keywords that may be associated with the image.

Given this context, the aim of this workshop was to bring together researchers within literary studies and cultural history (particularly those focusing on nineteenth-century literature and illustrations), and members of the computer science community focusing on Semantic Web and CBIR technologies:

1. To better understand how individuals from different user groups (historians, cultural critics, journalists, researchers in literature and arts, social scientists) interpret mid-Victorian illustrations;
2. Whether common themes emerge in the use of such systems by these individuals, and whether such analysis of use can be utilized to provide a recommendation system.

The intention was to better understand how a community of users interprets image content based on their own expertise and interests, and whether these annotations can enhance those already available within DMVI. The workshop also attempted to determine whether annotations added through a "community of usage" could add value to existing collections in DMVI, and how useful/acceptable this could be for potential users.

The workshop was divided into two sessions:

- A morning session that was dedicated to understanding the requirements of the arts and humanities (A&H) community. The talks in this session focused on the types of research questions that were of interest to the A&H community, and the types of interpretations of images that were of interest to them;
- An afternoon session that was dedicated to understanding computer science techniques for supporting image annotation. The talks in this session focused on the automated techniques for image analysis and annotation.

The full schedule can be found at:

<http://www.methodsnetwork.ac.uk/activities/act27prog.html>.

The rest of this report outlines particular issues that were discussed in the workshop.

Based on the requirements in the morning session, it was clear that the types of annotation that the A&H community had in mind were extremely hard to automate with the current technology. It was also generally agreed that although some automated image annotations were currently possible (such as being able to distinguish between indoor and outdoor scenes with a high probability of success, identifying images with buildings or those containing water, etc), these were generally seen to be too restrictive in scope.

Annotation Standards

It was generally agreed that standards for annotation of images were necessary to enable their wider usage in the research community. Although agreed standards were currently being used for cataloguing books, there was limited application of de facto standards for cataloguing illustrations. For instance, existing standards would only provide support for encoding metadata about an image – such as its size (measured in number of pixels) and the image format. It was also generally agreed that there were limited standards for encoding multimedia content (such as images and videos). Hence, support for metadata standards for illustrations would be a useful next step – and would be of interest to a very large research community.

Although standards such as Dublin Core were considered to be useful in this context, they did not provide support for managing illustration metadata. Similarly, other approaches such as LSCOM (<http://www.ee.columbia.edu/ln/dvmm/lscm/>) followed a 'silo'd' approach, and were hard to fully integrate within existing systems.

Automating Annotation

Annotation may be:

1. Provided by a domain expert – this is the current practice;
2. Automated by undertaking an analysis on the illustration/image.

In both cases, it was necessary to ensure that adequate editorial control was applied. In case (1) above, annotations may be provided by a non-expert but validated by an expert (in the same way as the addition of content to "Wikipedia"). In the second case, where automated annotation was being used, editorial control would become even more significant. Although computer scientists were generally enthusiastic about automating annotations, researchers from the literature community were skeptical, and did not feel that this would be useful for them. Automated annotation was therefore considered to be useful, but when used in combination with other information.

A variation on the above themes was to consider how to ‘fine tune’ a system based on usage and computer analysis – where input from a “group” of users (using a ‘learning’ based approach) could be utilized.

It was generally felt that current annotation techniques were inflexible, and often based on one particular categorization scheme. There was clearly a need for a search mechanism that would allow different types of representation standards to coexist.

Hence, two approaches could be considered in this context:

1. To create a ‘superset’ of all terms. In this case, it would be possible to map from a given annotation specification to this superset – enabling various annotation schemes to coexist. This technique would also enable cosearchability of multiple databases;
2. To translate a search request (query) into a format that the underlying search engine/mechanism could understand.

The Precis system provides a subject approach

(<http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=ED176803>), whereas Google provides an alternative to a library system, and enabled aggregation across different types of metadata. A point raised during the discussion was the need to better understand how Google would index the illustrations within a book, for example how it would deal with metadata associated with an illustration. The participants felt that current indexing was primarily based on the use of text placed around the image, and did not utilize approaches that took image content into account. Similarly, it was felt that identifying the relationship of such approaches with the indexing adopted by Amazon would also be useful, especially to determine how the format used by Amazon would compare with indexing used by Google?

An observation made by the participants was that major market players used indexing approaches for marketing, and not necessarily as an archival standard. Such a motivation for indexing therefore limited their use for academic research. It was, however, necessary to identify how a rating system (such as that used by Amazon) could be used to support better indexing – and make use of a ‘learning’/community based approach. Issues of pre- or postcoordinated indexing were also discussed in this context.

Based on the discussion provided by the computer scientists, it was pointed out that only a limited type of objects could be automatically detected within an image. This limited the usefulness of the approach to very specific contexts (for instance, finding an image that contained an outdoor scene vs. one that did not; an image that contained a large body of water, or green fields, etc). Professor Rueger emphasized the use of feature detectors within an image, enabling the response from these feature detectors to be aggregated to help identify more complex objects. With reference to the images contained in DMVI, it was felt that low level image features (such as crosshatched lines, or the massing of dark tones in one area of a plate) may be something that could be picked up by such feature detectors. Therefore, a useful way to integrate some of the computer science image analysis with DMVI would be to enable a user to show an ‘example’ image, and ask the system to find images that were ‘similar’ (in this instance, similarity being measured by comparing the response from the feature detector(s) when shown the query image and other images in the archive).

It was recognized that the automated analysis of images may be useful in identifying new research questions that were not initially considered by A&H researchers. For instance, finding associations and co-occurrence between objects between images could also provide a useful finding. The key lesson was to ‘extend’ the existing approach based on analysis by human experts, rather than ‘replace’ it.

The need to support *browsing* in addition to *searching* was considered to be an important requirement. It was also recognized that there was very limited support for *browsing* when using popular search engines, such as Google. Browsing was also a limitation with existing Digital Libraries. The ACM Digital Library and the online bibliography search engine CiteSeer were given as examples of edited online content – although there was limited browsing support provided by these also.

Precision vs. Coverage

According to the developers of the DMVI system, a key requirement they had was to ensure that the database was 'precise', and contained annotations that were provided by experts. For them, data quality remained an important objective, and therefore delegating annotations to a social network was generally not a useful option. A counter example to this, would be data collection by a family historian – where access to a large quantity of information that enabled wider access was more significant.

One could therefore consider two types of queries that could be made to such systems:

- A query to find a precise, specific object;
- A query to find all classes of objects.

Both types of queries were useful, and would necessitate different data capture techniques. It was also generally agreed, that given current development of archives, a high recall search was much more difficult to implement than a high precision search. Hence, there was a need to identify a mechanism for a 'comprehensive' search, and the social network supported annotation techniques could prove useful in this context.

Limits to Digitization

The participants also discussed the requirement to maintain the original artefact in addition to its digitized equivalent. Participants noted that the digitization process cannot record all aspects of an object, such as the 'aesthetics' associated with an object. They discussed how assumptions must be made during the digitization process about the attributes of an object that it is essential to record. This may be limiting for some researchers, especially where information about an object that is not considered to be of key importance could be used to answer particular research questions. However, due to limited space availability at museum and libraries, and the rapid decay of images printed on low-grade paper, it was inevitable that digitization would become increasingly important in the future.

Dissemination

The following activities have already started to disseminate the outcome of this workshop:

- Representation at the Open Grid Forum research group on 'Arts and Humanities' - led by David de Roure (University of Southampton) and Alexander Voss (Edinburgh University)
Representation was made at the session of this working group in Seattle (at OGF21).
- Three additional workshops are planned under the AHRC Museums, Libraries and Galleries Network Scheme – the first of these will be at the Victoria and Albert Museum in London (in December 2007). A summary of the major points from this workshop will be presented.

It was also noted that the AHDS image archive (with limited online access available at: <http://www.ahds.ac.uk/>) could provide a useful resource for computer science research.