

Capturing and Reusing Case-Based Context for Image Retrieval*

Dympna O’Sullivan¹, Eoin McLoughlin¹, Michela Bertolotto¹ & David C. Wilson²

¹University College Dublin
Belfield, Dublin 4, Ireland

²University of North Carolina at Charlotte
Charlotte, NC 28223-0001 USA

dymphna.osullivan@ucd.ie, eoin.a.mcloughlin, michela.bertolotto.ucd.ie, davils@uncc.edu

Abstract

Like many other application areas, task-based domains that employ digital imagery are faced with the problem of information overload. Modeling the relationship between images and the tasks being performed is an important step in addressing this problem. We have developed an interactive approach for the capture and reuse of image context information that leverages a measure of a user’s intentions with regard to tasks that they address. We analyze aspects of human-computer interaction information that enables us to infer why image contents are important in a particular context and how specific images have been used to address particular domain goals.

1 Introduction

Recent advances in digital image capture and storage technologies have increased the problem of information overload in imagery task domains. As a consequence, intelligent application support is needed to help manage imagery tasks. The majority of current retrieval techniques search for images based on similarity of appearance, using low-level features such as color and shape or by natural language textual querying where similarity is determined by comparing words in a query against words in semantic image metadata tags. The biggest problem arising from these techniques is the so-called semantic gap [Hollink *et al.*, 2004] — the mismatch between the capabilities of the retrieval system and user needs.

In order to address this problem we have developed a decision support system that can integrate information about underlying visual data with more high-level concepts provided by users as they complete specified tasks. Capturing human expertise and proficiency allows us to understand why relevant information was selected and also how it was employed in the context of a specific user task [Leake *et al.*, 1999]. From a case-based standpoint our research focuses on case knowledge acquisition and case knowledge reuse, where a case is represented by a complete user task including all system interactions in the course of carrying out the task. The approach allows us to capture and reuse best practice techniques

*The support of the Proof of Concept Fund of Enterprise Ireland is gratefully acknowledged

by automatically constructing a knowledge base of previous user experiences. This knowledge base can then be employed to improve future context-based query processing. Similar tasks are no longer interpreted as simply a collection of natural language terms: rather image retrieval requests take into account the context of the specific user domain goals. This approach allows for a dramatic reduction in both the time and effort required to carry out new tasks as amassed contextual knowledge is reused in support of the similar tasks. Furthermore, there are benefits from a knowledge management standpoint as contextual knowledge pertaining to particular tasks may now be stored and reused as an additional resource for support, training and preserving organizational knowledge assets.

2 System Overview

Our system enables a user to search directly for imagery corresponding to their current task needs. A typical task-based query to our image repository may consist of any combination of specified metadata, semantic task information and a sketched configuration of image-objects. The user may search an image database for imagery corresponding to their current goals or a knowledge base of previous user experiences (called sessions) for other user’s tasks that may bear resemblance to their own. In either case the imagery/sessions are ranked according to a percentage matching score. The user may then interact with the returned information as part of their task by annotating and highlighting relevant image aspects.

3 Capturing Case-Based Context

Our research draws on work in capturing task knowledge with the ultimate goal of performing more effective image retrieval using annotations [Lieberman and Liu, 2002]. To this end we have developed tools for direct image manipulation to assist the user in organizing information about relevant imagery and their task to capture important contextual information about specific user goals. These insights are then distilled into a form the system may use to perform similarity matching. The tools for direct image manipulation include filters, transformation, highlighting, sketching, post-it type and multimedia annotations. They allow the user to identify regions of interest that can be linked to clarifications and rationale. All con-



Figure 1: Capturing User Context through Image Annotation

textual knowledge is gathered by the system using implicit analysis so that users are shielded from the burden of knowledge engineering [Claypool *et al.*, 2001]. We unobtrusively monitor, interpret and respond to user actions as they interact with the imagery via the annotational tools and record this as user context.

Figure 1 shows how a user can make use of the image annotation tools as part of a geo-spatial image task. The particular task that this user had in mind was the construction of an airport on the outskirts of an urban area. The user has uploaded some multimedia and textual annotations; these textual annotations are represented by the icons on the image. The notebook icon replaces the text box where the user has typed comments and prevents the image from becoming cluttered. The camera icon replaces the video uploaded by the user, and when clicked on will play back the recording. The dark rectangles around the icons are the areas associated with these annotations and are emphasized once the icons are moused over. The area emphasized by the rectangle surrounding the notebook icon represents an undeveloped area outlying a residential part of the urban region. The user comment states that this is an area of low elevation, away from residential areas and drained naturally by a large river (i.e. suitable to proceed with the new development). The user has also highlighted a group of buildings by sketching an oval shape around them (they may have concerns that these buildings are quite close to the area selected to build the airport). Once the user has finished interacting with the imagery their entire task process is stored as an encapsulated session case in the knowledge base.

4 Reusing Similar User Context for Retrieval

The application allows for the reuse of similar user context for retrieval. More specifically, if a user working in a similar task domain to a previous user queries the system, the application reuses the previously captured context to compute similarity between the new task and the goals of previous users in order to recommend images and user sessions that may be of help to the current user. A more complete description of how

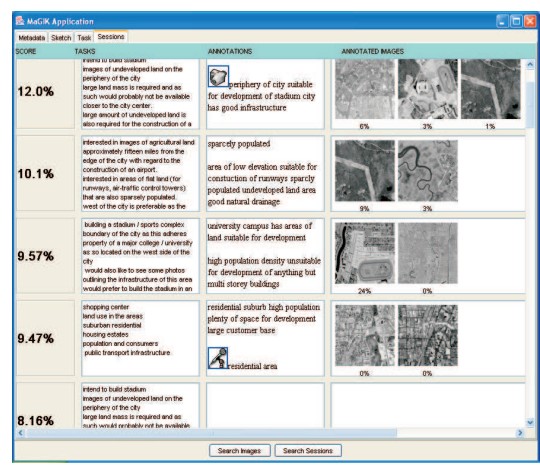


Figure 2: Session Retrieval

this similarity is calculated is described in [O’Sullivan *et al.*, 2004]. The current user can compare their task to previous work both to find relevant imagery and also to examine the decisions and rationale involved in addressing the previous task. Figure 2 shows the interface for displaying retrieved sessions.

Each row in the interface represents a similar user session and is summarized to include the percentage matching score, the most similar queries, the most important annotations added by the similar user, media buttons to play multimedia annotations and thumbnails of any annotated imagery. The user can reuse these similar sessions in carrying out their task by incorporating all or part of a relevant session into their own task context. Because the task-based retrieval system is tightly coupled with the activities that the user is performing, the system has the capacity to make proactive recommendations in an unobtrusive manner by monitoring the current task context. Information requests, increments in the information accessed and annotations provided are no longer taking place in an isolated environment. Rather they can be grounded in the context of the activities that the user is performing so the system can correspondingly anticipate and update what previous knowledge would be relevant, making it available to the user. This knowledge is provided unobtrusively, and need only be accessed when required. Thus the process does not distract from the task at hand, yet makes relevant knowledge available just-in-time.

5 Experimental Evaluation

The system was originally designed for the retrieval of geo-spatial imagery and an earlier evaluation of session retrieval indicated that sessions were useful as a basis for task-based retrieval [O’Sullivan *et al.*, 2004]. In this paper we show that task-based annotation context has advantages for individual image retrieval by employing the application as an aid for travel planning. We evaluated the system using images from the Corel Image Dataset [Wang and Li, 2003]. We began by analyzing a subset of 500 images (corresponding to 5 countries - Australia, England, France, Ireland and Thailand) and

manually annotating each of them with four/five keywords describing image objects and scenes depicted in the images. In order to extract contextual information for each image, the set of keywords for each image was entered as a query to web sites offering information on holidays in the outlined countries, for example <http://www.visiteurope.com/france.html>.

The resulting web pages were analyzed and the most relevant paragraphs from these pages were applied as annotations to the images. These annotations were between 2 and 5 sentences in length. An example of an annotation for a photograph featuring a part of the English Lake District was "The Lake District National Park in the north-west of England is the largest of Englands National Parks. Its 2,292 square kilometres cover high fells, lush green dales, still lakes, vibrant villages and quiet hamlets". The purpose of these annotations was to act a baseline for comparison with task-based queries relating to holidays in the outlined countries.

In order to demonstrate the task-based retrieval capabilities of the system, several tasks were outlined relating to holidays in the countries and user behavior was simulated by annotating relevant imagery with task-specific information. An example of a task description for a holiday in England was "Planning a trip to England, spend a few days in London, visit the markets, followed by a few days in the countryside". Once the images had been annotated according to this first set of tasks another set of tasks corresponding to more specific types of holidays in the countries were outlined. An example of such a task for planning a holiday in England was "Interested in an adventure holiday in the English countryside - I am particularly interested in hill-walking and water sports". The purpose of this was to see if the contextual annotations added to the images for the first set of tasks could be re-used to find imagery relevant to the second set of tasks. The 500 images were then categorized by hand as either relevant or not relevant to each outlined task in the second category of tasks. Each task was entered as a query to the system and the returned images were analyzed. Figure 3 demonstrates the average scores associated with the top five most relevant images (as judged by a real user) returned for the task-based queries. The left/right columns give the similarity respectively for keyword-based retrieval and annotation-based retrieval. For the most part the annotation-based retrieval vastly outperforms the keyword-based retrieval, demonstrating that task-based annotations can provide useful additional context that can improve image retrieval over a dataset. We plan to expand the library both in terms of the number of images and in terms of the number of applied task-based queries. This experiment is a starting point for automating the construction of datasets from where information can be retrieved based on current context rather than just on keyword or content-based search.

6 Conclusions

We have introduced our approach to developing a context aware system with a task-based focus for retrieving imagery and knowledge. Our intention is to use the presented experimental results as a baseline for future evaluation frameworks where we will examine the relative contributions of differ-

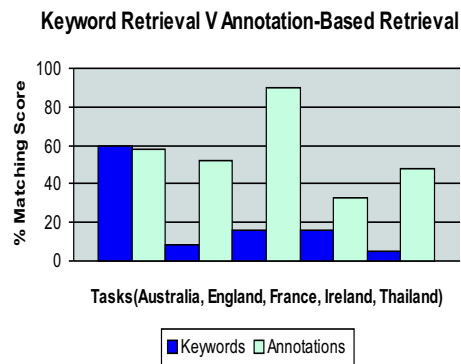


Figure 3: Keyword Retrieval V Annotation-Based Retrieval

ent levels of user, task and annotation contexts. We hope to perform a comprehensive user trial in the near future. We plan to extend our implicit knowledge acquisition techniques by extracting information from a greater variety of user actions. We also intend to supplement our primarily text-based retrieval system by including multimedia annotational information for retrieval.

References

- [Claypool *et al.*, 2001] Mark Claypool, Phong Le, Makoto Waseda, and David Brown. Implicit interest indicators. In *Proceedings of ACM Intelligent User Interfaces Conference, IUI-01*, pages 33–40. ACM Press, 2001.
- [Hollink *et al.*, 2004] Laura Hollink, Guus Schreiber, Bob Wielinga, and Marcel Worring. Classification of user image descriptions. *International Journal of Human Computer Studies*, 66(5):601–626, 2004.
- [Leake *et al.*, 1999] David B. Leake, Larry Birnbaum, Cameron Marlow, and Hao Yang. Task-based knowledge management. In *Exploring Synergies of Knowledge Management and Case-Based Reasoning, Proceedings of The American Association of Artificial Intelligence (AAAI-99) Workshop*, pages 35–39. AAAI Press, 1999.
- [Lieberman and Liu, 2002] Henry Lieberman and Hugo Liu. Adaptive linking between text and photos using common sense reasoning. In *Proceedings of the 2nd International Conference on Adaptive Hypermedia and Adaptive Web Based Systems*, pages 2–11. Springer-Verlag, 2002.
- [O’Sullivan *et al.*, 2004] Dympna O’Sullivan, Eoin McLoughlin, Michela Bertolotto, and David C. Wilson. A case-based approach to managing geo-spatial imagery tasks. In *Proceedings of ECCBR 2004, Seventh European Conference on Case-Based Retrieval*, pages 702–716. Springer-Verlag, 2004.
- [Wang and Li, 2003] James Wang and Jia Li. Automatic linguistic indexing of pictures by a statistical modeling approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9):1075–1088, 2003.