IDiexis: Mobile Image-Based Search on World Wide Web
- A Picture is Worth a Thousand Keywords

Konrad Tollmar, Tom Yeh, Trevor Darrell

MIT - Computer Science and Artificial Intelligence Laboratory
LU - Department of Informatics
konrad.tollmar@ics.lu.se, tomyeh@mit.edu, trevor@csail.mit.edu

ABSTRACT

Images of objects as queries is a new approach to search for information on the web. Image-based information retrieval goes beyond only matching images, as information in other modalities also can be extracted from data collections using image search. We demonstrate a new system that uses images to search for web-based information. We introduce a point-by-photograph paradigm, where users can specify an object simply by taking pictures. Our technique uses content-based image retrieval methods to search the web or other databases for matching images and their source pages to find relevant location-based information. We have developed a prototype on a camera phone and conducted user studies to demonstrate the efficacy of our approach compared to other alternatives.

Keywords

Content-based Image Retrieval, Object Recognition, Mobile Interface, Mobile Information Retrieval

INTRODUCTION

Mobile Internet applications have been in public use for a few years. However compared with a desktop computer, mobile devices such as smart phones and PDAs with ability to connect to the Internet differ quite considerably. First, the size of the display is small. Even if a PDA has larger display than a mobile phone it is still far from the size of an average desktop computer. Second, text input is slower than with a full desktop keyboard and in many mobile situations attention is divided which makes text input even harder. Third, the data transfer on mobile Internet is still far slower than wired Internet. In short, handheld web environments are still severely limited, and it has been suggested that mobile internet applications should be designed to support more task-specific uses [Maarek2002].

We have developed a mobile content-based search system which matches cell phone camera images against the attributes of other images found by Web-crawling servers [Yeh 2004, Tollmar 2004]. What makes our system interesting is the context of wireless connectivity and cheap digital imaging hardware; what makes it potentially useful, and not just a nifty demo, is the Web's vast database of imagery for comparison against what the camera sees.

The concept of a general mobile image-based Web search brings many questions to life. How would users perceive such a search system and what would they use it for? Would its availability change users Web searching behavior? How will users interact with such a system? What will it look like?

Creating a general image-based Web search engine and opening it to the general public on a larger scale is not yet feasible. Increasing the number of indexed images would make a search engine less accurate, as it is more likely that images will be similar without containing the same object. The challenges are clear when one compare a database of a hundred thousand images with Google's eight billion indexed Web pages [Biever 2006]. While we and other researchers continue the work of refining and optimizing image-matching algorithms this demonstration explores instead the interaction sides of mobile image-based Web searches. This area is ripe for exploration as we approach the inevitable availability of the technology.

PREVIOUS WORK

Kim, L. and Albers study of search strategies in small displays found a high data spread across all search times and conditions and claimed that designing information differently doesn't play a major role for some users. They suggested that future work should focus on other factors such as user motivation for finding the information, user knowledge of the types of information, and different types of information [Kim 2001]. Studies of various techniques that streamline web content to small screens indicate also that this causes confusion and that models that are more closely related to the mental model of full screen are preferred [Buyukkokten 2000].

The waste amount of systems based on image-based search has been developed for desktop searching of media libraries, and such content-based queries have been the subject of much research - e.g. ImageRover [Sclaroff 1997] and QBIC [Smith 1997]. But for the majority of these applications image search has turned out to be less popular than traditional keyword search methods. Rather than matching based on appearance, images are typically pre-labeled with keywords or matching is performed based on image captions or filenames. The uses for these fixed systems are however rather different from the mobile case, and, the utility of mobile images to search for information has been demonstrated for specific domains, e.g. location-based information [Tollmar 2004].

Figure 1, Example of an image search with the IDeixis system

The research field of image retrieval is about finding images that match in different ways. In this study, we're interested in finding information related to an image. It is not the matching image itself that is sought (as is the main purpose of the image retrieval field) it is the related information found in its vicinity, its context, which is the goal. More recently work [Fan 2005] has also considered the more specific case of mobile image-based web search.

DEMO - THE IDEIXIS SYSTEM

In this demonstration, we present the IDeixis system, Image-based Deixis (IDeixis) is an image-based approach to specifying queries for information retrieval. The key idea is that with a camera phone, users can point at things by taking images, send images wirelessly to a remote server and retrieve useful information.

IDeixis consists of two major components: a client-side application running on the mobile device, responsible for acquiring query images and displaying search results, and a server-side search engine, equipped with a computer-vision module, specifically, a content-based image retrieval system (CBIR), to facilitate image-based information retrieval.

Our overall system has several steps. First, an image-based (as opposed to keyword-based) URL index needs to be constructed to allow searching. A webcrawler crawls through the web, looks for images, and records the URLs containing these images. Appropriate features are extracted from each image and stored in the database. After the indexing is complete a mobile user can take photos of the object he/she is interested in. These photos are sent to the image database through a wireless internet services. A search engine looks for a set of images most similar to the query image using an image distance metric. The result will consist of a list of (candidate image, source URL) pairs.

The platform we chose to build the prototype system is a J2ME enabled camera phone, and image matching techniques and web-services has been developed in C++, Matlab and Java.

REFERENCES

1. Biever, C. See it, snap it, and let your phone find it on the web. New Scientist, p. 20-21. (2006, January, 7)

2. Buyukkokten, Orkut; Molina, Hector Garcia; Paepcke, Andreas; and Terry Winograd. Power Browser: Efficient Web Browsing for PDAs. Proceedings of CHI 2000. (2000)

3. Kim, L. and Albers, M.: Web Design Issues when Searching for Information in a Small Screen Display, Proceedings of SIGDOC'01. (2001).

4. Maarek, Y., Soffer, A. and Chang, B. (Eds), Working Notes of the WWW2002 Workshop On Mobile Search. (2002)

5. McDonald, S.; Lai, T. & Tait, J. Evaluating a content based image retrieval system. SIGIR '01. (2001).

6. Sclaroff, S., Taycher, L., and La Cascia, M. 1997. Image rover: A content-based image browser for theWorldWideWeb. In Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Libraries (1997).

7. Smith, J. R. and Chang, S.-F. 1997. An image and video search engine for the World-Wide Web. In Proceedings of the SPIE Conference on Storage and Retrieval for Image and Video Databases V (1997).

8. Tollmar K., Yeh T., and Darrell T., IDeixis - Image-Based Deixis for Finding Location-Based Information, In Proc. Mobile HCI 2004. (2004).

9. Yeh T., Tollmar K., and Darrell T., Searching the Web with Mobile Images for Location Recognition. In Proc. IEEE Conf. on Comp. Vision and Pattern Recognition. (2004).

10. Fan, X.; Xie, X.; Li, Z.; Li, M. & Ma, W. Photo-to-search: using multimodal queries to search the web from mobile devices, MIR '05: Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval. (2005).