ISSN:0975-9646
Kirti Raheja et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (4) , 2011, 1725-1728

To Study the Ways to Annotate Images Manual, Semi-Automatic and Fully Automatic in M2S and CAIR

Kirti Raheja[1], Deepti Gupta[2]
[1] M.D UNIVERSITY Rohtak, INDIA
[2] TITS, BHIWANI, INDIA

***Abstract* - Images are commonly used on a daily basis for research, information and entertainment. The introduction of digital cameras and especially the incorporation of cameras in mobile phones make people able to snap photos almost everywhere at any time since their mobile phone is almost always brought with them. The fast evolution in hardware enables users to store large image collection without high costs. Making use of these image collections requires efficient image retrieval techniques. Traditional image retrieval techniques like text-based image retrieval and content-based image retrieval techniques have shortcomings. New techniques or combination of existing techniques must be established to provide users with adequate image retrieval functionality.**

Keywords: API:-Application Programming Interface, CAIM:-Context-Aware Image Management CAIR:-Context Aware Image retrieval, CBIR:-Content-Based Image Retrieval, M2S:-MMS To Search, MMS:-Multimedia Messaging System

INTRODUCTION

In this we will describe two mobile services that enable users to retrieve images, videos, textual descriptions or other types of information from a data collection using MMS (Multimedia Messaging Service) or SMS (Short Message System). In this we will first present the motivation for it. Traditional techniques for image retrieval and context will then be introduced.

I. MOTIVATION

The number of images, both private and public increases every day due to the rapid growth in mobile technology the latest years. Mobile entities include PDA's (Personal Digital Assistant) and mobile phones. This has encouraged the use of digital images as one of the most important communication media for daily life. Images are commonly used on a daily basis for research, information and entertainment. The introduction of the digital camera and especially the incorporation of these into mobile devices such as PDA's and mobile telephones, enables people to take pictures almost anywhere and at any time. Several applications have been developed to support this trend, and more is yet to come. People can for instance distribute images quickly and easily by sending an MMS with their mobile phones. People have got used to retrieve desired information when and where they want it, since their mobile phone is almost always with them. The desired information is often only a call or SMS /MMS away.

A. Current image retrieval techniques

Text-based image retrieval finds images based on textual metadata associated with the image. The metadata can for instance be information like location, time, what the picture is about, who is on the picture and who captured it. Text-based image retrieval requires manual annotation of images. Annotating every image in a collection manually is a very time-consuming task that few users prioritize. Even if users annotate the images, this is a highly subjective task and personal phrases are chosen. A user can make use of synonyms and the annotation text can vary a lot from person to person.

B. *What is context?*

Context is a very wide concept and it is used in several fields. Giving a clear definition of context is not easy. However, there have been several attempts to make one. We believe Dey's definition suits our use:

"Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves".

C. Information retrieval in mobile environments

We live in an information society and people are overloaded with information. TV-and radio-commercials, posters and other types of advertisements are displayed and distributed on buses, airports, in shopping centers, to mailboxes and e-mailboxes.People are getting lots of information without asking for it. This might give rise to the expectation that when they need information about something they will get it right away.

D. *Goal*

The goal is to investigate context-aware image retrieval in a mobile environment. This has two sub-goals:

- An intermediate aim of this thesis is to design and develop a content-based information retrieval for mobile environment. We want to make a fun, easy-to-use, entertaining and useful service meant for mobile phones. We also want to look into content-based image retrieval techniques because this is and has been a debated technique in the literature and many commercial companies are now investigating the use of content-based image retrieval in their services.
- Another intermediate aim is to design a context-based image retrieval servicefor mobile environment. The use of

context is said to solve problems related to content and text-based information retrieval.

We wanted to experience context-based image retrieval by designing a system that utilizes this concept.

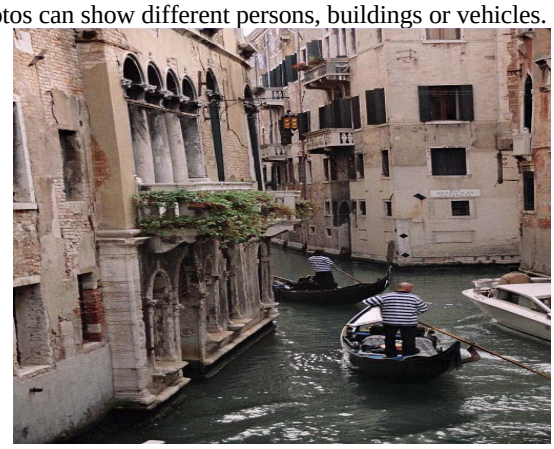
D. Method

We have studied the use of image retrieval techniques such as content-based and text-based image retrieval and investigated the use of context and how context can be usedfor information retrieval. Mobile environments have been considered and how imageretrieval can be applied to such environments. We have designed developed a service called M2S (MMS to Search) that enables users to retrieve information about a specific entity using MMS. CAIR is a mobile application that enables users to retrieve images from image collections based on context and using SMS.

II HOW TO COLLECT IMAGES

People collect images for archiving memories, storytelling and personal enjoyment. An event is what is happening in the picture, what the picture is about.Events are naturally associated with specific times and places such as a child's birthday, a vacation or a party.

However, events are difficult to define quantitatively or consistently. The photos associated with an event often exhibit little coherence in terms of both low-level image features and visual similarity. As an example of an event, consider pictures taken during a holiday trip to Venice. The photos can show different persons, buildings or vehicles.



A. About images

Information can be represented in different ways, as sound, text, symbols and images. Users can hold large amounts of information and it would be a great advantage to have efficient technique in order to retrieve the desired information. The complexity of the information retrieval depends on how the information is represented. To search for a specific piece of text is far more efficient than search for a specific piece of sound based on the text and/or the sound itself.

B. Image annotation

People seem to use very little time to annotate their personal images. How many amateur photographers are determined enough and have enough time and energy to go through developed pictures, and put them into albums, instead of just sticking thepictures in a shoebox? How many people go through their digital photos and give each one a unique file name in an appropriate directory instead of leaving them in the default directory created by the camera software? Not many. As a result, more and more people have thousands of digital photos with little or no organization, and they are resigned to gaining no more benefit or enjoyment from them than the photos stored in overfilled shoeboxes around the house. Well-performed annotation has the power to transform this almost random collection of images into a powerful, searchable and rich record of events in people's lives.

C. *What is annotation?*

The goal of annotation of images is to assign semantically meaningful information to images. Text is the most common and relevant way of annotation. An annotation can be for instance: "The celebration of Helen's 25 birthday". Other ways of the assigning information can be by drawings and sketches where these are used as basis/input in searches where the image searched for and the sketch/drawing looks alike. A recording of audio works the same way. An audio file is assigned the image and when the image is searched for, the input can be in form of singing, humming or whistling into the microphone .Metadata may be used in a number of ways:

• Embedding the metadata in the web page using META-tags in the HTML coding of the page.

• As a separate HTML document linked to the resource it describes.

• In a database linked to the resource. The records may either have been directly used.

2.2.3 Methods for image annotation

There are several ways to annotate image collections. This can be either done manually, semi-automatic or automatic.

D. Manual image annotation

This is the "old-fashioned" approach where people have non-digital paper-pictures in photo albums and write the associated text. Manual annotation is a completely human oriented task that deals with human oriented information. This type of metadata can be the event of the image, the photographer, the title and similar information. The **advantage** of manual annotation is the accuracy in extracting semantic information at several levels. It is the most precise way of annotation and for now, the only way of full value to add semantics to images. Another **drawback** is that the task of describing the content of images is highly subjective. The perspective of textual descriptions given by an annotator could be different from the perspective of a user. A picture can mean different things to different people. It can also mean different things to the same person at different times.

E. Semi-automatic image annotation

Semi-automatic annotation is manual annotation with machine extraction of some information. It depends on the user's interaction to some degree. The technical information (see automatic annotation) is added automatically from, for

instance, a camera, the user can then be prompted to add additional information to the image. The manually added information is typically semantic information. Semi-automatic annotation combines the efficiency of automatic annotation and the accuracy of manual annotation of images.

F. Automatic image annotation

Automatic annotation is machine annotation, where humans only verify the task. The information added by a camera is of a technical nature and is automatically added. This information is typically time, location, resolution of the image, camera model, which number the image has in the range of images taken, name of the image and other technical information. As we see from this type of information automatic annotation is limited due to computers lacking ability to extract semantic information from images.

III. TRADITIONAL TECHNIQUE

In the next subsections we will discuss traditional techniques for image retrieval. The techniques that we will describe are text-based image retrieval, content-based image retrieval and a variant of content-based image retrieval called region-based image retrieval.

A. *Text-based image retrieval (TBIR)*

The most common way to manage large image collections is to store text in the form of keywords together with the image. We separate between text based image retrieval techniques that use the surrounding text of the image and text based techniques where each image or image collection is annotated. The approach that deals with surrounding text searches the keywords that are physically close to the image.

B. Content-based image retrieval (CBIR)

Much of the motivation behind content-based image retrieval is for users to avoid annotation of images. In some systems text based annotation and retrieval of images is inadequate. It might include images from surveillance cameras, finger-print recognition or x-ray pictures. Motivated by the lack of an efficient image retrieval technique, content-based image retrieval was introduced. "Content-based" means that the technology makes direct use of content of the image rather than relying on human annotation of metadata with keywords.

There are three main approaches in content-based image retrieval

- Query by example. The user selects images and the image retrieval system returns a selection of images based on a set of matching points.
- Specifying colors.
- Sketches/drawings. The user creates a rough sketch to find matching images.

C. *Region Based Image Retrieval (RBIR)*

Region Based Image Retrieval (RBIR) is an extension of content-based imageretrieval techniques . Region-based image retrieval systems provide new querytypes to search for objects embedded in an arbitrary environment. An RBIR system automatically segments images into a variable number of regions, and uses a segmentation algorithm to extract of features (like color, shapes and sketches) for each region.

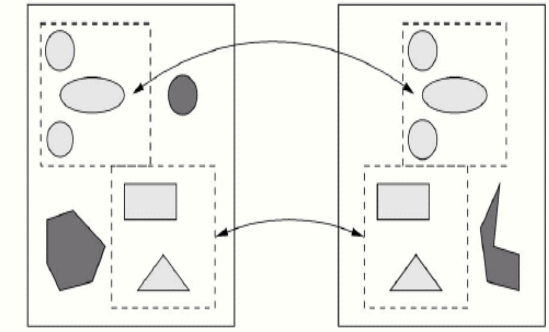


Figure 1: Images with similar regions

IV PRECISION AND RECALL

The next subsections will ways to measure the performance of information retrieval functions.

A. Image retrieval

Image retrieval is a subclass of information retrieval and inherits therefore many of the aspects that encompasses information retrieval. Image retrieval is concerned with retrieving images that are relevant to the user's request from collections of images. The essential aims of information retrieval are to be efficient and effective. Efficiency means delivering information quickly and without excessive demands on resources, even when there is a massive amount of information to be retrieved. Clearly efficiency is extremely relevant to information retrieval where late response is often useless information. Effectiveness is concerned with retrieving relevant documents. This implies that the user finds the information useful. If a user keeps retrieving information of low relevance it is natural to believe that the user quit using the system.

Consider an example of an information request I that is a collection of information. In this collection of information there is a set of relevant documents, R, where |R| is the number of documents in the set. Assuming that an information retrieval strategy processes the information request I, and generates an answer set A where |A| is the number of documents in the set.

Let |Ra| be the number of documents in the intersection if the sets R and A. The recall and precision are measured as follows: Recall is the fraction of the relevant documents (the set R) that has been retrieved.

Recall = |Ra|/|R|

Precision is the fraction of the retrieved documents (the set A) that is relevant.

Precision = Ra/A

Precision and recall are tested on a fixed and relatively small set of documents and has predefined queries to decide which documents goes with which queries. This is not a task that is done completely automatic.

The formula for F-score is:

F-score = (2. Recall.precision)/(precision+Recall)

The F-score is also known as the F1 measure, because precision and recall are equally weighted. There is also F05 measure F that weights recall twice as much as precision. The idea is to allow the user to specify whether he is more interested in recall or in precision. This value is often given in percentage and the higher the percentage, the better search strategies. All three measures consider documents as either relevant or irrelevant. In practice documents can have degrees of relevancy. In this thesis we refer to precision and recall to determine the performance of retrieval systems.

V CONCLUSION

we have discussed ways to annotate images both manual, semi-automatic and fully automatic and we discussed advantages and disadvantages with these approaches. A picture can mean different things to different people and the manual expression of metadata becomes problematic when the images are searched for. Traditional image retrieval techniques include text-based image retrieval and content-based image retrieval. There are several advantages and drawbacks to these approaches that we have discussed. Content-based image retrieval makes it possible for users to search for parts of an image that is not a part of the main activity of the image. Because content-based image retrieval techniques base their search only on the contents of the images, annotation is superfluous.

REFERENCE

1. Brown P.J, G.J.F. Jones, "Exploiting contextual change in context-aware retrieval." Proceedings of the 2002 ACM symposium on Applied computing, Madrid, Spain,
2. Brown P.J, G.J.F. Jones "Context-aware Retrieval: Exploring a New Environment for Information Retrieval and Information Filtering." Personal and Ubiquitous Computing, 2001.
3. Jones G.J.F., P.J. Brown, Context-Aware Retrieval for Ubiquitous Computing Environments. Lecture Notes in Computer Science. 2004, Berlin / Heidelberg: Springer.
4. Allan James (editor), Jay Aslam, Nicholas Belkin, Chris Buckley, Jamie Callan, Bruce Croft, Sue Dumais, Norbert Fuhr, Donna Harman, David J. Harper, Djoerd Hiemstra, Thomas Hofmann, Eduard Hovy, Wessel Kraaij, John Lafferty, Victor Lavrenko, David Lewis, Liz Liddy, R. Manmatha, Andrew McCallum, Jay Ponte, John Prager, Dragomir Radev, Philip Resnik, Stephen Robertson, Roni Rosenfeld, Salim Roukos, Mark Sanderson, Rich Schwartz, Amit Singhal, Alan Smeaton, Howard Turtle, Ellen Voorhees, Ralph Weischedel, Jinxi Xu, ChengXiang Zhai, "Challenges in Information Retrieval and Language Modeling." SIGIR Forum, 2003.
5. Baeza-Yates Ricardo, Berthier Ribeiro-Neto, Modern Information Retrieval. 1999, New York: ACM Press.
6. Pitoura Evaggelia, Bharat Bhargava, "Building information systems for mobile environments." Proceedings of the third international conference on Information and knowledge management, Gaithersburg, Maryland, United States.