Mobile Solutions for Image Retrieval and Meta-Data Extraction

Project Proposal Report

ZEPHYR

Tolga AKIN
1502038
tolga.akin@ceng.metu.edu.tr
0554 738 34 80

Çiğdem AVCI
1448398
cgdmavci@gmail.com
0505 758 64 28

Gülin Seray TUNCAY
1635267
gulizseray@gmail.com
0554 333 69 68

Hüseyin ULUSOY
e1502798@ceng.metu.edu.tr
1502798
0505 672 48 67
Mobile Solutions for Image Retrieval and Meta-data Extraction

Alias: MOS-IRME  
Sponsor: Department (Intelligent Systems Lab.)  
Aspects: Mobile software, web services, Image Processing, Artificial Intelligence & Algorithms, GPS

**Initial Thoughts**

Mobile technologies have gained popularity lately throughout the world. Using Internet technologies, combined with artificial intelligence and image processing techniques, it is possible to develop innovative software solutions. Taking these issues into consideration, together with the help of the Intelligent Systems Lab, we are planning to develop a software system which will allow users to retrieve information from the pictures they take with their mobile phones.

**Problem Definition:**

In the system, the user takes a photo of a landmark and wishes to get information about the place. The GPS information will be extracted from the image (exchangeable image file format – EXIF). The software will use an API of a virtual globe application like Google Earth and using GPS information the name of the place will be obtained. Next, related data will be retrieved from sources like Wikipedia. Finally, the user will be informed about the picture. Additionally, there may be a live camera support of this application.

Furthermore, detailed information can be sent to the user if desired. This detailed data can be about the characteristics of a famous landmark or it can be about the interior objects of a building. To make this possible, a wide database of images will be compiled. By means of some image processing and machine learning techniques, the system will try to match the user-supplied image with the images in the database. If the system succeeds in matching the image, the user will be supplied extra information about the interior details.

Since our project is both a commercial and a research project, it will be better to apply spiral model during software development process.

**Main Components of the Project**

1) **GPS**

According to the device on which our software system runs, the API which will going to be used varies, therefore the determination of the technology related with GPS will be clarified after the mobile device is selected.

For MOS-IRME software system's GPS component to be capable of functioning on a diversity of devices, each device's specific API should be handled.
2) Obtaining the Relevant Information

GPS information will be used to identify the location and the direction of the place. Using this information, the system will extract the data about the place with the help of virtual globe technologies such as Google Earth or NASA World Wind.

3) Image Processing

We plan to put emphasis on image retrieval methods which are based on similarity. According to many articles related to image retrieval, feature (content) extraction is the basis of content-based image retrieval visual features (color, texture, shape, faces). The color feature is one of the most widely used visual features in image retrieval. It is independent of image size and orientation. In image retrieval, the color histogram is the most commonly used color feature representation. In addition to the color histogram, several other color feature representations have been applied in image retrieval, including color moments and color sets. Many research results suggested that using color layout (both color feature and spatial relations) is a better solution to image retrieval. To extend the global color feature to a local one, a natural approach is to divide the whole image into sub-blocks and extract color features from each of the sub-blocks. With the help of those leading informations, different algorithms will be searched. If the main parts of the project are completed as required, machine learning methods will also be applied to image retrieval part to get better results. Images retrieved from the database and the image we currently have will be compared using those methods which are defined above.

4) Information Retrieval

By using web services and the tags of the most similar images collected as defined in image processing part more textual information will be extracted from the web.

5) Graphical User Interface

The graphical user interface will run on the mobile device such as a mobile phone or a PDA. Target architecture (i.e. processor, operating system) is yet to be determined. However, currently the best candidate is the Android Operating System.

Plan of Activities

For this semester, we plan to determine the requirements and create a design for the project. By the end of the semester, we intend to provide a demo to present the basic features of the software. In the second semester, we will complete our detailed research on the image processing part of the project.

Market Search

When we made a brief search through the Internet, we faced with an application similar to ours named Wikitude. WIKITUDE World Browser presents the user with data about their surroundings, nearby landmarks, and other points of interest by overlaying information on the real-time camera view of a smart-phone. This product proves the feasibility of our project. In some aspects Wikitude has some
similarities to our project. It uses the GPS, compass tools and web sources. In our project we intend to include these features as well. We will start by implementing the features similar to those of Wikitude. Then, we are planning to apply new techniques that aren't implemented in Wikitude such as “Image Retrieval”. Plus, we plan to narrow down our field of interest so as to give more detailed information about a place such as METU. GPS technology is more than enough to identify the location of somebody in METU; however it would provide very restricted information about the details of different locations such as being at the Computer Engineering Building or the dormitories.

For further information:

- http://wikitude.org

Literature Survey

Our project can be classified as “Augmented Reality”. In augmented reality, view of a physical real-world environment elements are merged with (or augmented by) computer-generated imagery. The input is usually a live video stream generated by a video recorder; however, we are planning to exclude live video streams and use user-generated still images instead. To realize augmented reality with images, we will need to employ image processing. For the image processing part of our project, we will be utilizing some algorithms based on image retrieval techniques. After a considerable literature review, we decided that content based image retrieval is suitable for our work.

Related Articles:

- Smeulders, Arnold W.M.; Worring, Marcel; Santini, Simone; Gupta, Amarnath; Jain, Ramesh. “Content-Based Image Retrieval at the End of the Early Years” (2000)
- Nastar, Chahab; LiitscMe, Llatthias; Nleilhac, Christopher; Boujemaa, Nozha “Surf image: a Flexible Content-Based Image Retrieval System” (1998)