Visual Object Detection from Mobile Phone Imagery for Context Awareness

Patrick Morris Luley
JOANNEUM RESEARCH
Wastiangasse 6,
A-8010 Graz, Austria
+43 316 876 1735
patrick.luley@joanneum.at

Lucas Paletta
JOANNEUM RESEARCH
Wastiangasse 6,
A-8010 Graz, Austria
+43 316 876 1735
lucas.paleta@joanneum.at

Alexander Almer
JOANNEUM RESEARCH
Wastiangasse 6,
A-8010 Graz, Austria
+43 316 876 1735
alexander.almer@joanneum.at

ABSTRACT
We describe a system which proposes a solution for multi-sensor object awareness and positioning to enable stable location awareness for a mobile service in urban areas. The system offers technology of outdoors vision based object recognition that will extend state-of-the-art location and context aware services towards object based awareness in urban environments. In the proposed application scenario, tourist pedestrians are equipped with a GPRS or UMTS capable camera-phone. They are interested whether their field of view contains tourist sights that would point to more detailed information. Multimedia type data about related history might be explored by a mobile user who is intending to learn within the urban environment. Ambient learning is in this way achieved by pointing the device towards an urban sight, capturing an image, and consequently getting information about the object on site and within the focus of attention, i.e., the user's current field of view. The described mobile system offers multiple opportunities for application in both mobile business and commerce, and is currently developed as an industrial prototype.

Categories and Subject Descriptors
H.4 [**Information Systems Applications**]: Miscellaneous
D.2.8 [**Software Engineering**]: Metrics – complexity measures, performance measures

General Terms
Algorithms, Human Factors

Keywords
Computer vision, object detection, mobile tourist information systems, location awareness

MobileHCI'05, September 19–22, 2005, Salzburg, Austria.
ACM 1-59593-089-2/05/0009.

1. INTRODUCTION
Location awareness for a mobile service in urban areas can, by the use of GPS only, not be assured everywhere and at any time, because of the known weaknesses of GPS signal availability in urban areas. Therefore, mobile systems operating in urban environments must take advantage of contexts arising from the spatial and situated information at a current location of the pedestrian user. Today, location based services are in principle able to provide access to rich sources of information and knowledge to the nomadic user. However, the kind of the location awareness that they do provide is not intuitive, requires reference to maps and addresses, i.e., the information is not directly mediated via the object of interest.

In contrast, the proposed system takes a decisive step towards getting in line with the user's current intention to relate information to its current sensorial experience, e.g., the object in its line of sight. In this way, the system can respond to the user's focus of attention, e.g., for the purpose of tourist information systems. A camera attached to the mobile system (PDA, or camera phone) pointing towards the object of interest (e.g. a building or a statue) will capture images on demand and would be capable of automatically finding objects in the tourist user's view. The images are then transmitted to a server that automatically extracts the object information, associates it to geo-referenced content, and sends the resulting data back to the mobile user. 'Mobile vision' is here referred to mobile visual data that are processed in an automated way to provide additional information to the nomadic client.

2. USER SCENARIO
The tourist moves completely free through an unknown area and if he is interested in any object (e.g. a historical building or a statue) he just has to take a picture of it with his camera-phone, with or without GPS device connected, and pressing the "Identify" button. As result he gets a detailed description of the object containing multi-media tourist information. As a second achievement he also gets the position of the identified object which can be used for navigation.

3. SYSTEM OVERVIEW
This Section will briefly describe how a common smart-phone with built-in digital camera can be used for image based object recognition (see Figure 1). A GPS device, built-in or connected to the phone via cable or Bluetooth can help to accelerate the

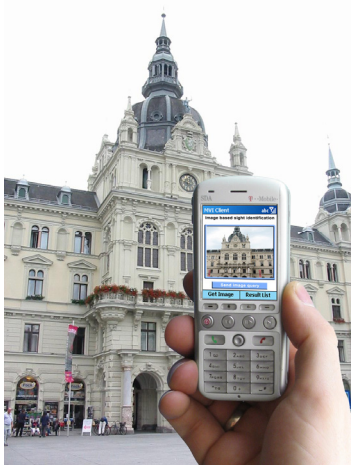

Figure 1. Taking a picture of the object to be identified.

recognition process considerably. The overall concept consists of three main phases.

In the first phase a software client is activated by a user on his personal smart phone. The software can be directly downloaded from a website with the internet enabled smart-phone and then installed on the device. The software-client offers functionality to take a picture of an object the user wants to identify. Next, if available, the smart-phone reads the actual position from the GPS device. If the GPS cannot obtain a position for any reason, the cell information of the phone-network provider can be used to approximate the user location instead. The picture of the object and the position of the snapshot-location are put together into a SOAP Message and send to the image-recognition web-service, running on a dedicated server, over a common wireless internet connection like GPRS or UMTS.

In the second phase the web-service reads the request from the client (smart-phone) and extracts the picture and the GPS position. The picture is then analysed by an image recognition algorithm to obtain representative features of the picture like edges and significant surfaces or colour transitions. Next, these features are compared with an object-database. The database contains pictures of different objects, with the pre-processed features and the position of the snapshot-point. The object is then identified by matching the features of the user-picture with the features of objects in the database Objects can be filtered by the users GPS position compared with possible object snapshot-points saved in the database. Once the object is identified the object-database is queried for some information containing text and pictures. This information is integrated into a SOAP-message and is then send to the smart-phone client as response.

In the third phase the web-service response is presented to the user on the smart-phone. Additional to the object quick information containing text and pictures there is an URL which can be used by the user to obtain detailed information about the object. The URL can be viewed in a common smart-phone internet browser like “Opera” or “Pocket Internet Explorer“. The URL contains, as a parameter, the unique identification number of the desired object and links to a dynamic website, which is generated at runtime on the server. The layout of the website is optimized for the requesting hardware platform – because different smart-phones have different display resolutions.

Figure 2 illustrates the technical concept and its corresponding three operational phases.

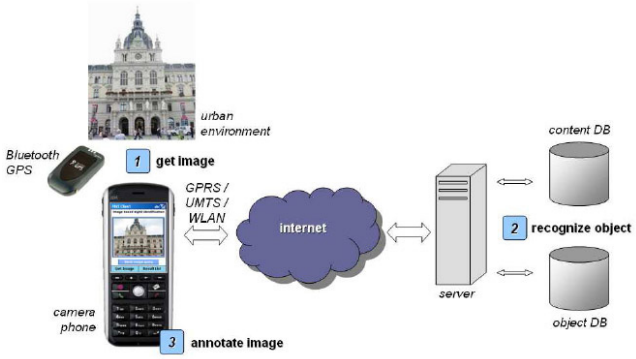

Figure 2. Overview of the mobile application systems concept

The aim of this client-server architecture is to bring the image based object recognition service to any person using a common camera-phone and to gain scalability in reference of the number of objects in the database and complexity of the image-recognition algorithms.

4. VISUAL OBJECT DETECTION AND IDENTIFICATION

The proposed image based service for object awareness requires both, robust and fast visual object recognition of typically low-quality outdoor images. Off-the-shelf camera technology in mobile devices provides up to 1.3 Megapixel resolution but often provides blurred, overtuned images with low contrast and bad color reconstruction. We therefore applied a methodology that seemingly offers several advantage for mobile vision applications, i.e., the Informative Features approach. By selecting highly discriminative local patterns for classification, it provides a distribution of true and false alarms that can be easily integrated for an efficient performance of a detection system.

5. CONLUSION AND OUTLOOK

In this paper, we provided an overview about a mobile system which includes GPS functionality and vision enhanced context awareness, with the goal to offer user friendly digital services for pedestrians in urban areas. The client-server architecture of the system allows bringing the image based object recognition service, which also provides location awareness, to any person using a common camera-phone. This mobile system enables its ubiquitous use in many business and commerce relevant situations.

For a successful market launch of a m-commerce application the quality of the service including a consistent availability of the service, an appropriate content, location awareness and target group oriented usability have to be realised.

These developments will be realized in frame of both the EU-project MOBVIS (FP6-511051, Vision Technologies and Intelligent Maps for Mobile Attentive Interfaces in Urban Scenarios) and the national, ASA funded project *Mobile City Explorer*.