Camera Based Navigation System with Augmented Reality

M. Marcu*, D. Cuc* and D. Tulvan*
* "Politehnica" University of Timisoara/ Faculty of Automation and Computers,
Timisoara, Romania

***Abstract*—Nowadays smart mobile devices have enough processing power, memory, storage and always connected wireless communication bandwidth that makes them available for any type of application. Augmented reality (AR) proposes a new type of applications that tries to enhance the real world by superimposing or combining virtual objects or computer generated information with it. In this paper we present a camera based navigation system with augmented reality integration. The proposed system aims to the following: the user points the camera of the smartphone towards a point of interest, like a building or any other place, and the application searches for relevant information about that specific place and superimposes the data over the video feed on the display. When the user moves the camera away, changing its orientation, the data changes as well, in real-time, with the proper information about the place that is now in the camera view.**

I. Introduction

What if you could find out in real-time when you are walking on the street any contextual information you would like just by pointing your mobile phone's camera towards one of the surrounding buildings? You take out your phone, you point the camera towards the building, and a dialog box pops up on the screen showing you the name of said building. You select the option "Additional Information" from your smartphone's touchscreen and you can see information, such as the name of the architect that built it, the surface the building occupies, how tall it is, along with other useful data gathered from the web.

This is made possible by the aforementioned Augmented Reality, which enriches the user's environment with additional layers of information. The first products of this kind, although primitive and with room for improvement, can already be found on the market but are not very common yet. In the meanwhile, scientists are constantly researching and developing new solutions that might help in the advancement of this field.

The basic principle of Video Augmentation is the following: the user points the camera of the smartphone towards a point of interest, like a building or any other place, and the application searches for relevant information about that specific place and superimposes the data over the video feed on the display. When the user moves the camera away, changing its orientation, the data changes as well, in real-time, with the proper information about the place that is now in the camera view.

The smartphone uses the GPS receiver to compute the geographical coordinates of the user, the digital compass and the accelerometer to obtain the orientation and angle of the device. The data received from these sensors is then used to look up information about the target on the Internet, on different websites.

The goal of this project is to develop an Augmented Reality solution for Windows Mobile devices based on Wikimapia and different sensors available on the device, such as the GPS receiver, compass and accelerometer.

The advantages of this system over a normal map are that it is much simpler to use, the device is just pointed at the building/place you want recognized and the application will show its name, calculate how far away you are from the building, show additional information and pictures of that building, recognize different parts of the building and will also allow you to see a list of all the other buildings in the vicinity.

II. State-of-the Art

Paper [1] presents an approach for a mobile semantic collaboration platform based on the OntoWiki framework. It allows users to collect instance data, refine the structure of knowledge bases and browse data using hierarchical or faceted navigation on-the-go even without a present data connection. The critical part of the application is the advanced replication and conflict resolution for RDF content. The application is designed using MVC (Model-View-Controller) architecture; where the user interface is build using HTML 5 and JQuery Mobile Framework which includes the core jQuerylibrary in an improved version to ensure compatibility across all of the major mobile platforms. To access the device's hardware (e. g. camera, GPS sensor) OntoWiki Mobile uses the extended HTML5 API. Local storage is a part of the HTML5 application caches and is persistent data storage of key-value pair data in Web clients. Seeing as one of critical requirements of the application was to work without internet connection the issue of replication has occurred. When dealing with non-conflicting change-sets the merge is done asynchronously for all parties that are trying to perform the synchronization. In the cases of conflicting change-sets the merges are synchronized and a new version is created for each merge, as opposed to the single version created when the merge is asynchronous.

The authors of [2] describe an application used for displaying information about pieces of art inside a museum. The information displayed changes for every room and the direction in which the view is pointed in that room. The position and orientation of the actor and he's view is obtained from a GIS (Geographical Information

ISSN 1844 – 9689

System). The system lacks in accuracy, that's why subtracting data from a GPS and inertial navigation system would greatly improve the accuracy.

The authors of [3] developed an analytical model of the overall data collection process in sparse sensor networks with mobile relays (MRs). The model is flexible enough to incorporate different discovery and data transfer protocols. The discussion is limited to a simple discovery algorithm where the MR sends periodic advertisements and the sensors follow an asynchronous scheme based on a low duty cycle. An ARQ communication protocol with selective retransmission was used for the data transfer. Findings show that low duty cycles can be actually used for a large class of environmental monitoring applications. Surprisingly, a low duty cycle may not always be the most energy efficient option, depending on a number of different factors such as the speed and the mobility of the MR.

Paper [4] describes a series of context aware applications developed in the 90's that show the evolution of mobile computing and the context awareness capabilities over the course of a decade. It splits the up into three major categories: outdoor – where the most used system for positioning is GPS; indoor – where correctly pinpointing the position is more difficult because GPS doesn't penetrate most of the building. Here alternative positioning systems are used, some developed around infrared and some around radio signal strength; hybrid - location-tracking systems that are based on network domains, and they are not specifically targeted for indoor or outdoor use, using a wireless LAN technology deployed on a metropolitan area.

This paper [5] focuses on the benefits that context awareness could bring to the industry, by developing an application with context awareness that helps maintenance workers. The system uses an RFD sensor to send information to the worker. The paper introduces an architecture based on document knowledge base, including context-aware, knowledge service manager, end user components, combining the perverse environment and context. At last, it applies the architecture into the application of production device maintenance in enterprises. The context -aware is the basis of individual service, which can identify end users and tasks greatly. The application presented reflects the feasibility and availability of the architecture and improves the efficiency of maintenance worker.

In the paper [6] the authors presented a tool called SymPA, developed for the Symbian Platform that correlates traffic information, radio-access-technology measurements, and location data to help developers evaluate mobile applications in the field. The main feature offered by SymPA, that enable mobile application analysis are: Analyzing Data Connections -traffic capture, received-signal-strength tracking, and radio-access-technology identification; Checking Connectivity - lets developers verify mobile-to-Internet connectivity by establishing a data connection and then obtaining the IP address associated with the connection; Localizing Measurements - two features that allow this: the GPS-tracking function activates monitoring of the mobile terminal's speed and position and the Google Static Maps API gives developers a map of the area directly on a mobile device; Tracking Energy Consumption - lets developers fuse battery-consumption information with all the other available information so that they can associate results with the actual conditions during the measurement; Correlating Data - Correlating data from different levels lets developers find the source of errors at the application level and associate them with locations.

This paper [7] presents the design, implementation and evaluation of several techniques used for optimizing the information uploading process for continuous sensing on mobile phones. The two main optimization models the paper describes are the Optimizing User State and Location-based State Uploading. The first one is split into two cases Online Stream Analysis Strategies (techniques that can be used when network connectivity is always available based on the analysis of the data streams) and Off-line Strategies: Markov Chain based Prediction (an alternative strategy that tries to forecast the next state during a disconnection).The first one maximizes accuracy and the second the best performance in terms of overhead, making a combined solution the best choice. The second one requires a GPS receiver and it implies to associate a state transition matrix to each location.

SocioScope [8] is a urban information system specially developed to support the needs of researchers, practitioners, local authorities, and enterprises interested in studying through digital traces. SocioScope streamlines the collection of spatio-temporal urban data, controls access to the conventional data based on roles and associated privileges, and provides tools to query and visualize the information. SocioScope can facilitate and foster the research activity in digital sociology by offering tools that can be used to connect data providers with researchers and practitioners and by defining common set of methodologies to collect, store, analyze, and visualize urban data.

The paper [9] presents context-aware mobile knowledge management architecture. The key component of the architecture is the "Contextualizer". This middleware implements the knowledge management activities (e.g., knowledge storage, retrieval, presentation, and acquisition) by taking into account the context elements, namely the user, the technical environment, the situation (i.e. time and place) and the specific task at hand. Moreover, in contrast to existing approaches, the "Contextualizer" comprises a meta-repository, which decides dynamically where the requested knowledge resides. The architecture still lacks thrall testing and it hasn't yet been applied to real-life situations.

[10] presents an AR System developed in de City of Vienna that was created with the purpose of replacing tourist guides. The system used consists of: a notebook computer (2GHz processor, an NVidia Quadro4Go graphics accelerator and a Wireless LAN network adaptor to enable communication with other mobile units), a Trimble Pathfinder Pocket differential GPS receiver, a Sony Glasstron optical-see-through stereoscopic color HMD fixed to a helmet as an output device, an InterSense InertiaCube2 orientation sensor and a PointGrey Research Firefly camera for fiducially tracking are mounted on the helmet. The system is designed around a three tier architecture, which allows the use of different databases, through a well-defined interface. The problems the developers faced are mostly related to low accuracy of the sensors, this would cause the image to shutter, this leaving a lot of room for further improvements.

III. Solution Architecture

The system in question represents an Augmented Reality solution for smartphones with an integrated camera and a high speed wireless Internet connection. Based on the contextual data received from different sensors of the device, the application looks up the information on the Wikimapia website and superimposes it over the video feed from the camera.

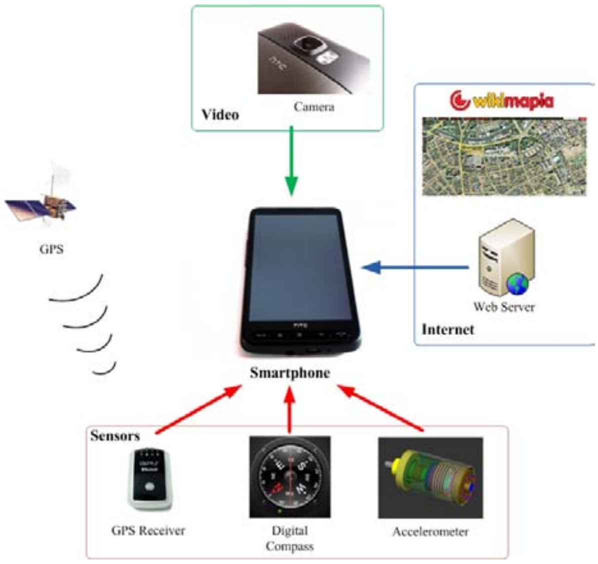

Figure 1. Overall solution architecture

As can be seen in the block diagram (Figure 1), the system consists of three main components that send data to the smartphone:

- Sensors - Three sensors that provide location context: GPS receiver, digital compass, accelerometer;
- Web Server - The HTTP server of the Wikimapia website the application connects with, in order to download information regarding the location context;
- Video - The smartphone camera, which provides video feed;

The GPS receiver acquires geodata from the GPS satellite network and computes the geographical coordinates of the user – latitude and longitude, as well as altitude if necessary. Both the digital compass and the accelerometer are used to determine the orientation and angle of the device. All the data obtained from these sensors comprises the information required to establish the context of the user and is sent further to a web server.

Using a wireless Internet connection the smartphone logs on to the HTTP web server of the Wikimapia website through an API that allows free access to its database. Once connected, the application looks the database up and extracts the data that corresponds to the context information provided by the three sensors.

Wikimapia is an online editable interactive map that combines Google Maps with a wiki collaborative system, allowing users to annotate location names anywhere on the map. Its aim, as stated on their website, www.wikimapia.org/wiki/Main_page, is "to create and maintain a free, complete, multilingual, up-to-date map of the whole world" with the collaborative effort of its users. Collaborative web maps, such as Wikimapia and OpenStreetMap, allow users to contribute with content, like annotating places, adding, descriptions, pictures, videos and animations about them. Because all this is done on a shared surface, problems related to concurrent access to the map, versioning or cluttering can appear. One solution to these issues is the use of overlays. An overlay is a grouping together of items, whose visibility can be toggled on or off, allowing for easier revision control and eliminating cluttering.

Having received the information about the user's context, the application can now overlay it upon the video feed provided by the camera of the smartphone or it can use it in any other way to facilitate the user's access to the relevant information

The main function of this context-aware mobile system is to provide real-time information about the points of interest found in the user's field of view. Thus, the user points the camera of the smartphone towards a point of interest, like a building or any other place, and the application searches for relevant information about that specific place and superimposes the data over the video feed on the display. When the user moves the camera away, changing its orientation, the data changes as well, in real-time, with the proper information about the place that is now in the camera view.

Some of the most important features included in the system are:

- It displays the name of the buildings or places found in the center of the smartphone's camera view;
- It displays the distance, in meters, to the building or place found in direct view of the camera;
- It displays the name of the area surrounding the user, or if the user is inside a building, the name of the building he finds himself in;
- It enhances the association of buildings in a larger context;
- It provides a list of all the buildings found in the proximity of the user, in a range chosen by him or her;
- It lets the user choose the searching range of the application. The user can choose a range between 200, 400, 600, 800 and 1000 meters.
- It supplies a web link for every building or place within range that points towards the appropriate building's page on the Wikimapia website where additional information is provided, and even a map centered on the building's location;
- The user can choose to turn camera view on or off;
- Shows the orientation of the user, either by displaying cardinal points, or by displaying the angle relative to the North;
- The user can choose which language will be used to display the names of the buildings or places. English is the default language, but there are several languages available to choose from, such as Romanian, French, German, Hungarian

All the features described above can be accessed through the user interface, which was conceived for ease of access with the immediate needs of the user in mind.

The application has three distinct interfaces:

- Camera, which is basically the Augmented Reality mode;

- Camera with Options, camera mode with a layer of buttons/text fields that allows the user to change working parameters;
- Place List, which displays nearby places and additional information about them, but with the camera turned off

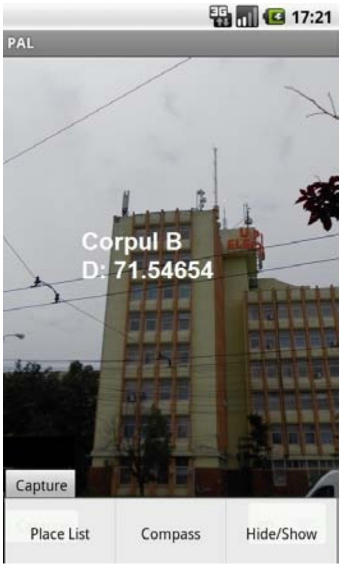

Figure 2. Camera based GUI

The first interface (Figure 2) displays the video feed taken from camera of the mobile device, in the background, and in the center of the screen we have details about the building that the camera is focused on (name, distance to its center and the container place, if the place is included in a larger context). The interface also displays additional information about GPS positioning GPS status when it is enabled/ disabled, and if a picture is taken and sent to the server it displays the progress.

IV. Solution Design

The application was designed with an object-oriented mindset and having a three-tier architecture. Three-tier is a client-server architecture in which the user interface, the application processing and the data management are developed as independent modules. It consists of:

- The Presentation Tier
- The Logic Tier
- The Data Tier

This architecture was specifically chosen to take advantage of the benefits of modularity so that the system could be adapted to changes in technology or requirements. This is made possible by well-defined interfaces which can easily be modified if necessary.

The presentation tier is the top-most level of the application. It represents the part of the code through which the user interacts with the system, giving control over it and allowing easy access to its features. It translates the information so it can be understood by the user without much difficulty.

The business tier (logic tier) contains the logic of the application, making logical decisions and performing computations, but also manages the data access, acquiring and analyzing data from the sensors and the data tier. The sensors (GPS, Digital Compass, Accelerometer and Camera) are accessed through various programming interfaces (APIs)

The third tier, the data tier consists of a database web server, which stores information, keeping it independent from the business and presentation tier. The data is retrieved by the business tier through a HTTP request from the Wikimapia web server, which sends the information queried in the JSON data interchange file format. An active Internet connection is required, but it gives an improved performance and increased scalability

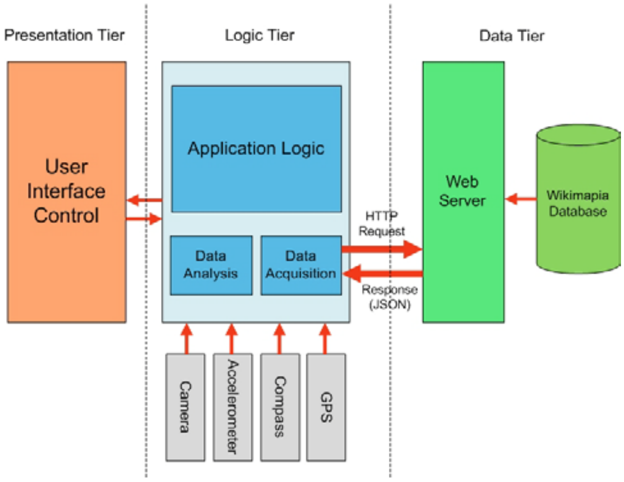

Figure 3. System design

The main function of the application is to detect the building located in front of the smartphone. Thus, an incremental approach was taken, and such, before implementing the method that actually returns the place which is in front of the device, two other additional algorithms were required:

- An algorithm that checks if a given point is inside a polygon;
- An algorithm that checks if a building is included in a larger context.

To determine the orientation of the user, the digital compass of the smartphone was used.

The data analysis performed in a place consists of computing the distance from the user to the respective place, calculating the number of vertexes that make up the polygon's outline, as well as the center for such a polygon. A geographical coordinate specified through latitude and longitude is mapped through a Point class. An important aspect of the implementation was the calculation of the distance between two such points. Because of the near-spherical shape of the Earth (technically, an oblate spheroid), calculating an accurate distance between two points requires the use of spherical geometry and trigonometric math functions.

The core of this feature is the data extraction logic which uses the Wikimapia API. The request sent to the server has to contain the API key, the file format of the output, the language, and most importantly the four coordinates, representing the four corners of the square that make up the corners of the minimap, which has the current position at its center.

The server responds with a JSON file, chosen as the default file format, containing information about all the buildings found in the proximity of the user. This contextual information is represented by the name, URL, extreme points of the location, and the coordinates of the edges of the polygon that makes up the contour of each place. All this data had to be parsed and put into a *Place* object, each one corresponding to a building or region.

All of the *Place* objects were then added and grouped together into a *Layer* object, where further data analysis regarding the relationship between places was performed.

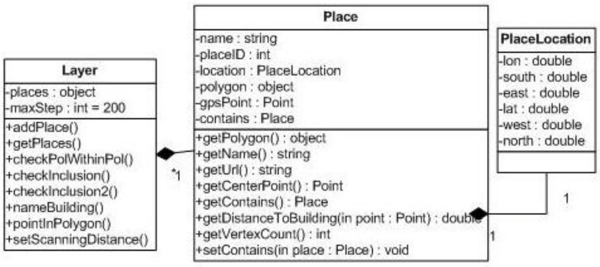

Figure 4. Data Analysis component design

To obtain a sufficient level of accuracy the following formula was used:

$$A = \left[\sin\left(\frac{dlat}{2.0}\right)\right]^2 + \left[cos\left(\frac{\pi lat1}{180}\right)\right] * \left[cos\left(\frac{\pi lat2}{180}\right)\right] * \left[sin\left(\frac{dlon}{2.0}\right)\right]^2$$

$$B = 2 * arctan\left(\frac{\sqrt{A}}{\sqrt{1-A}}\right)$$

$$C = r * B$$

Where dlat =lat2-lat1 and dlon =lon2-lon1, with (lat1, lon1) and (lat2, lon2) being the geographical coordinates between which the distance calculation needs to be performed. And r is the radius of the Earth in the type of units is required:

- Nautical miles: 3437.74677
- Kilometers: 6378.7
- Statute miles: 3963.0

To obtain the distance in meters, after replacing r with 6378, C is multiplied by 1000 [11].

To compute the distance between the current position of the device and a building, the center of every building was needed. Using the most northern, southern, western and eastern points of each building, points provided by the Wikimapia API, the center point was obtained.

$$lat_{center} = \min(north, \quad south) + \frac{north - south}{2}$$

$$lon_{center} = \min(east, west) + \frac{east - west}{2}$$

Every building contains a field where the distance to the current GPS position is kept. This value is recomputed for every building contained by the minimap.

After integration, tests were done for every feature of the application separately. One of the more interesting tests, depicted in Figure 33, involved a 360° rotation of the smartphone; the test was run successfully. All the problems revealed by testing where fixed, and in the end all the tests passed.

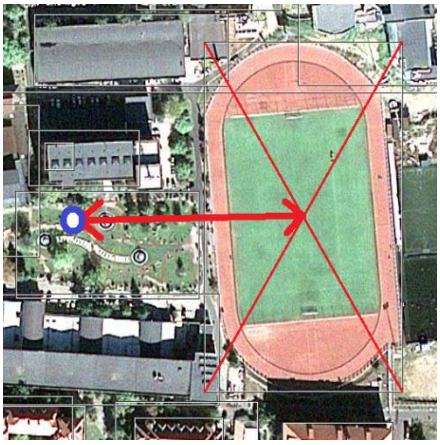

Figure 5. Distance to the centre of a place

Figure 6. 360° View Test

V. Conclusions

The goal of this work was to develop an augmented reality application, that could be used anywhere in the world, especially in metropolitan areas, for orientation and place recognition in a user friendly manner.

The advantages of this system over a normal map are that it is much simpler to use, the device is just pointed at the building/place you want recognized and the application will show its name, calculate how far away you are from the building, show additional information and pictures of that building and will also allow you to see a list of all the other buildings in the vicinity.

Acknowledgment

This work was supported by Hungary-Romania Cross-Border Co-operation Programme 2007-2013 project “Cross border academic development of an image-based recommendation system for regional educational purposes” - HURO/1001/283/2.3.1 (www.huro-cbc.eu and www.hungaryromania-cbc.eu). The content of this paper does not necessarily represent the official position of the European Union.

REFERENCES

[1] T. Ermilov, N. Heino, and S. Auer, "OntoWiki mobile: knowledge management in your pocket", Proceedings of the 20th International Conference Companion on World wide web, WWW'11, New York, NY, USA, 2011.

[2] L. Roffia, G. Raffa, M. Pettinari, and G. Gaviani, "Context Awareness in Mobile Cultural Heritage Applications", Proceedings of 7th International Conference of Ubiquitous Computing, UbiComp'05, September 11–14, Tokyo, Japan, 2005.

[3] G. Anastasi, M. Conti, and M. Di Francesco, "An Analytical Study of Reliable and Energy-Efficient Data Collection in Sparse Sensor Networks with Mobile Relays", Proceedings of the 6th European Conference on Wireless Sensor Networks, EWSN'09, Heidelberg, Berlin, 2009.

[4] G. Chen and D. Kotz, "A Survey of Context-Aware Mobile Computing Research", Technical Report, Dartmouth College Hanover, USA, 2000.

[5] K. Liao and Q. Liu, "The Architecture of Mobile Knowledge Service Based on Document Knowledge Base & Its Application", International Conference on Management of e-Commerce and e-Government, ICMECG '09, 16-19 Sept. 2009.

[6] A. Díaz, P. Merino and F. J. Rivas "Mobile Application Profiling for Connected Mobile", IEEE Pervasive Computing, Vol. 9, Iss. 1, Jan.-Mar. 2010.

[7] M. Musolesi, M. Piraccini, K. Fodor, A. Corradi, and A. T. Campbell, "Supporting Energy-Efficient Uploading Strategies for Continuous Sensing Applications on Mobile Phones", Proceedings of the 8th International Conference of Pervasive Computing, Pervasive 2010, Helsinki, Finland, May 17-20, 2010.

[8] A. Vaccari, F. Calabrese, B. Liu, and C. Ratti, "Towards the SocioScope: an Information System for the Study of Social Dynamics through Digital Traces", Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS'09, New York, NY, USA 2009.

[9] O. Thiele, H. Knapp and M. Schader, "A Context-aware architecture for mobile knowledge management", Int. Conf. on Wireless Information Networks and Systems, WINSYS'06, Setubal, Portugal, January 2006.

[10] G. Reitmayr and D. Schmalstieg, "Collaborative Augmented Reality for Outdoor Navigation and Information Browsing", n Proc. Symposium Location Based Services and TeleCartography 2004, Vienna, Austria.

[11] Distance Calculation, How to calculate the distance between two points on the Earthhttp://www.meridianworlddata.com/Distance-Calculation.asp, 2010.