Campus Navigation System Based on Mobile Augmented Reality

QIN Yong-xu, LIU Jia-min, QIU Hui, YAN Bo, JIN Chang-xu
School of Information Science and Technology
Shenyang University of Technology
Shenyang, China
E-mail: qinyongxu008@gmail.com

***Abstract*—Due to the limited of precision, navigation technology in the local area cannot accurately locate each building and obtain the functions related to the building. This paper investigates augmented reality and navigation technologies and realizes the campus internal navigation on a mobile phone. In the campus, the exact position and detail data of every school can be obtained by using the augmented reality technology. The navigation system merged with augmented reality makes navigation technique in the local application more exact, humanized and convenient.**

Keywords- mobile device; augmented reality; campus navigation

I. INTRODUCTION

As computers increase in power and decrease in size, intelligent mobile phone and wearable devices are rapidly becoming feasible, providing people access to online resources always and everywhere. Augmented reality (AR) is a new rising research field in recent years. AR system integrates virtual information into the real world and enhances the users' visual experience and makes users more intuitive and more real understanding of the real world [1]. Mobile augmented reality systems (MARS) provide this service without constraining the individual's whereabouts to a specially equipped area. Ideally, they work virtually anywhere, adding a palpable layer of information to any environment whenever desired [2].

According to the identification and information acquisition, mobile AR technology is classified into location-based AR and vision-based AR. Location-based AR first positions an area by GPS in the mobile device, or using current location information achieved by network, and then obtains the virtual information corresponding to the current position [3]. The first successful location-based AR application is called iButterfly [4]. By using GPS, the iButterfly obtains the users' position and then users can capture the butterflies which are set beforehand to achieve business objectives. Other than location-based AR, vision-based AR must first obtain the images by the camera; then matches the captured images with a given template; and the last registers the corresponding virtual information on the images [5]. For example, a mobile augmented reality application for books described in [6]. By capturing the images in the books, it will provide more information to make the books more vivid. Compared with location-based AR, vision-based AR needs more compute and more powerful processors.

Vision-based AR provides more information for the specified target in real scenes by virtually adding texts, images and videos. However, it cannot accurately point out the target position. Conversely, location-based AR can acquire the target position in the real scene. However, its navigation is not able to generate accurate position data when the destinations are within a small area, e.g. the buildings in a university campus.

Considering the special scenario where the locating information in a university campus needs to be acquired, this paper investigates mobile location-based and vision-based AR technologies, and implements the campus navigation system by using these technologies. This system provides the related information for each building, e.g. teacher, research and teaching information of a faculty. Moreover, it aids navigation by directly pointing out locations in order to improve imprecise location in a local area and supplements latitude and longitude data without navigation. Hence, the system becomes more perfect and easy for user to operate.

II. SYSTEM DESIGN

A. Overall design

This navigation system is developed under Android system of intelligent mobile phones. According to the requirement, the navigation system will automatically obtain the current location wherever and whenever the users execute this system in campus. But the location data cannot always direct the right place. During the campus tour, the user can find the artificial markers pasted on each building at which the school, library or classroom is located. The marks are all different from each other and each building in the campus has a different marker to be represented. Markers carry on much information of the building which the marker represented. When user scans the marker by the phone camera, the detailed functions and position data of this building are presented on navigation interface. The data related to the building are provided in the way of images, texts and videos. At the same time, the virtue model of this building is also to be shown on the interface. In addition, the route guide can be generated according to the current position of the marker under the situation where the user cannot ensure the current position in the campus. The overall flowchart of the system is shown in Fig. 1.

978-1-4799-2809-5/13 $31.00 © 2013 IEEE
DOI 10.1109/ICINIS.2013.42

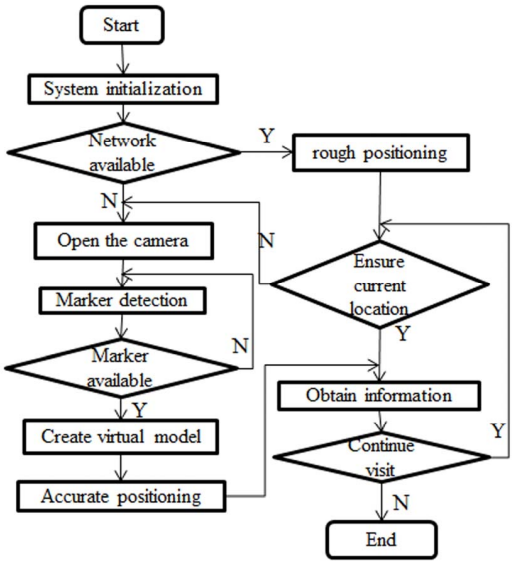

Figure 1. System overall flowchart

Mobile AR navigation system combines the augmented reality with the navigation technology. By using the function of obtaining detailed information, it not only makes the user to find him/her way while visiting the campus, but also shows the user a correct route to the target location. To get the goals, this system realizes the modules of user location, marker detection, three-dimensional (3D) display and campus guide. The function modules of the system are shown in Fig. 2.

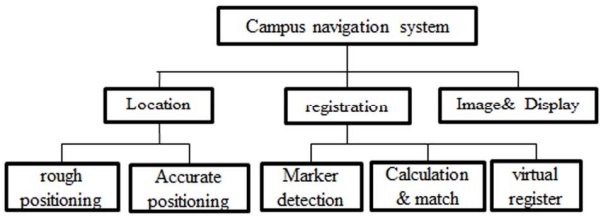

Figure 2. Function modules

B. Implementation

1) Location

This system exploits Map Android API provided by AutoNavi Company to realize the navigation. The API supports the digital map display and operation, interest point searching, geocoding, offline map and other functions [7]. Based on the precision of positioning, both rough location and precise location are designed in this system. The building's location provided by the API has a deviation from the corresponding real location, this leads to pointing to a wrong position where you would not go. Therefore, the rough location related to latitude and longitude of the building is first achieved using the satellite map. The precise location to the building is then adjusted by identifying the given marker.

In the case of a given current position, the user directly selects a destination on the system interface. According to the location of a guide marker, the system searches the vector map in the corresponding position, and then provides the user an optimal route in the navigation interface.

2) Registration

To achieve the data and precisely locate a building, the marker on the given positions is first recognized. It is then tracked in real time. The virtual data which involves in the building functions is registered on the marker. For that reason, the registration by vision-based AR technology is implemented in the following steps.

Step 1: Make sure the whole guide marker is involved in the video of camera. The guide marker in the video image is correctly detected by using edge detection.

Step 2: By using the image that is generated from Step 1, the guide marker image is matched with the template image which is built in the system beforehand.

Step 3: To register the virtual data on the guide marker, the marker in the image is identified by using AR technology. The camera external parameters need to be figured out. The mapping of world coordinate system to image plane coordinate is then calculated by the camera parameters [8, 9]. The transforming process is implemented by (1).

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \lambda M [r_1 \quad r_2 \quad r_3 \quad t] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}. \qquad (1)$$

Where (u, v) is a point on the imaging plane; (x, y, z) is a point in the world coordinate system; M and [r1 r2 r3 t] are intrinsic and extrinsic parameters of camera matrix respectively; λ is a scaling factor.

Assume that each marker has a coordinate system, so that the camera parameters can be solved by using homography matrix H which is created in the process of marker plane mapped to the imaging plane. Calculating homography needs four or more corresponding point pairs. The more to the points, the higher registering accuracy is. This paper, according to the characteristics of the markers, uses 4 point pairs consisting of the vertex of the square on the marker to deduce the homography. Hence, Equation (1) is transformed into (2).

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = [h_1 \quad h_2 \quad h_3] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}. \qquad (2)$$

According to (1) and (2), Equations (3) and (4) are deduced, that is

$$H = [h_1 \quad h_2 \quad h_3] = \lambda M [r_1 \quad r_2 \quad r_3 \quad t]. \tag{3}$$

$$\begin{aligned} r_1 &= \lambda M^{-1} h_1 . \\ r_2 &= \lambda M^{-1} h_2 . \\ r_3 &= r_1 \times r_2 . \\ t &= \lambda M^{-1} h_3 . \end{aligned} \tag{4}$$

Scale factor λ is determined by the orthogonal condition, i.e. $\lambda = 1/|M^{-1}h|$. Thus the camera pose is obtained by homography matrix.

Step 4: Since the process of generating the target scene by OpenGL is similar to the camera imaging processes, the virtual objects will be drawn by using OpenGL API and the camera external parameters.

After the target is successfully tracked and registered by visual-based technology, location-based tracking will be further adopted. The exact data of the marker on the target are given and showed on the navigation interface.

3) Image and display

The images in this system contain guide markers and data of each building. The marker designed is a black square within a pattern inside [10]. The pattern on the marker refers to the building. Each marker has its own feature for each building. These markers are stored as the template images and used for consequential image recognition. The images with data are given an identifier linked to data of the buildings. When the user utilizes the camera to capture a marker image, the user enters the image capture interface and directs the camera to the marker image. If the whole marker is not included in the image, the system will not provide the information for users.

A built-in camera and display which can meet the requirements have been completely equipped. The preferable interfaces are designed for the management of data display. To easily operate various functions provided by the system, the operation interface are designed by the button and drop-down list controls.

III. EXPERIMENTAL RESULTS

The system is developed by AutoNavi MAP API, OpenGL API and AndAR API, and tested on an android mobile phone. AutoNavi MAP API provides the system GPS support; OpenGL API provides the system display support and AndAR API provides the system identification support. The user executes this application if the network is available, and directly access to the interface. The rough positioning and navigation is presented as shown in Figure 3, i.e. the dark area in Fig. 3.

Figure 3. Rough positioning

Under the circumstances of getting lost, the accurate location can be obtained according to the guide markers pasted on each building. The user can observe a 3D virtual model of the building with the building's name. This virtual model is created by the designer beforehand.The virtual model (i.e. a cube with literal characters) is registered on the marker in the real scene (Fig. 4), and then presice position is presented on the map (Fig. 5).

Figure 4. Virtual registration

Figure 5. Precise positioning

Through accurately positioning to each school, the system will provide detailed information about the school, as shown in Fig. 6 and Fig. 7

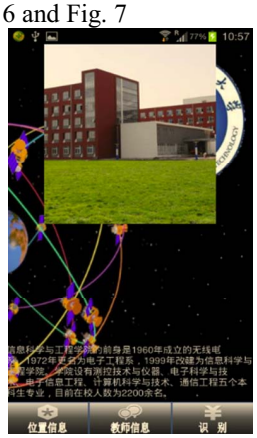

Figure 6. School information

Figure 7. Teacher information

Given the current position, the user can select the drop-down list box to the other destination; the system will provide an optimal path, as shown in Fig. 8.

Figure 8. Route selection

IV. CONCLUSIONS

The paper investigates augmented reality, navigation and image processing technology, and integrates them to develop a campus navigation system. Through the experiments, the system can guide users to visit the buildings within the campus, acquire more accurate information about the school. This contributes to better understanding school to some extent. This system can be further extended to the similar applications, e.g. large factories, the World Expo and large commercial areas.

ACNOWLEDGMENT

The authors would like to thank the participants and the anonymous reviewers for their comments and feedback. This research was supported by the Science and Technology Department of Liaoning Province in China (Grant No. 201102164) and by Liaoning BaiQianWan Talents Program (Grant No: 2012921041).

REFERENCES

[1] Chen Jing, Wang Yongtian, Lin Liang. Implementation of the augmented reality on PDA [J]. OPTICAL TECHNIQUE, 2007, 33(1): 52-55.

[2] Yue L. Key Tech of Mobile Augumented Reality & Its Research Progress[J]. Office Informatization, 2013, 2: 003.

[3] Paucher R, Turk M. Location-based augmented reality on mobile phones[C]//Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on. IEEE, 2010: 9-16.

[4] DENTSU, iButterfly [OL]. http://www.ibutterfly.hk/chi/index.html, 2011.

[5] Beier D, Billert R, Bruderlin B. Marker-less vision based tracking for mobile augmented reality[C]//Mixed and Augmented Reality, 2003. Proceedings. The Second IEEE and ACM International Symposium on. IEEE, 2003: 258-259.

[6] Zou Youjia. A Mobile Augmented Reality Application For Books[J]. Computer Technology and Development, 2013,8.

[7] AutoNavi, AMAP [OL]. http://api.amap.com/Android/index.html,2013.

[8] Liu Jiamin, An Lexiang, Chang Yan, Zhu Shitie, YangDan. Three-dimensional registration method based on triangular marker in augmented reality[J]. Journal of Shenyang University of Technology, 2013,35(1):79-84.

[9] Jiang Qinyun, Wang Cheng, Li Lijun, Guan Tao. Study on Registration Algorithm in Augmented Reality[J]. Computer & Digital Engineering, 2006,34(9):38-40.

[10] Yang Dan, Liu Jiamin. Research on Augmented Reality Interactive Scene Based on ARToolKit[D]. Shenyang University of Technology, 2012.

2016 IEEE International Symposium on Mixed and Augmented Reality Adjunct Proceedings

Mobile Guide to Augmented Reality for Campus of the Autonomous University of Nayarit

Adalberto Iriarte-Solis* Palmira González-Villegas Ricardo Fuentes-Covarrubias Gerardo Fuentes-Covarrubias

Universidad Autónoma de Nayarit Universidad de Colima

ABSTRACT

Mobile applications have become very important not only as a mean to disseminate information, but also to serve as guide in unfamiliar contexts. This guide developed with mobile Augmented Reality (AR) aims to give information on the location of points of interest within the UAN campus. The tools used were through markers with Unity3D and Vuforia Software.

Keywords: localization, augmented reality, murals.

Index Terms: H.5.1 [Information Technology and Systems]: Information Interfaces and Representation (HCI)—Artificial, augmented, and virtual realities; K.3.1 [Computing Milineux]: Computers and Education—Computer Uses in Education

1 INTRODUCTION

Communication and information exchange is not only for the use with computers. Communication is at hand through mobile devices. In recent years, mobile applications have become very important not only as a mean of to disseminate information, but also to serve as guide in unfamiliar contexts [6]. People are increasingly demanding accurate information at the time and place where they are. The devices and mobile programming have contributed largely to meet this need; however, the location and context are important factors to provide the required information. Consider, for example, technologies that can detect the location and characteristics of a place, such as a building, a laboratory, or room; a person or a teacher; or course, which are very important and necessary [12].

Augmented Reality (AR) is one of the most used emerging technologies. Through it, the user is allowed to see the actual information or compounds overlapping with the real world virtual objects. The virtual objects can be manipulated by individuals, coordinating to get the view they wish [9]. The AR is in development, but at the same time there has been a positive impact on the market, thanks to the several functional projects that have been and continue to be made using this technology. The potential of mobile devices associated with the AR, which enables to interact with the stored data and the actually captured with the geographic position of the user itself [2].

1.1 Problematic situation identified

Having the location somewhere with their respective facilities is of great help and comfort to people when they are not there. However, the Autonomous University of Nayarit (UAN) does not have a concrete and specific orientation of each building for new students, and due to its large extension could present conflicts or delays to reach the desired destination.

* email: adalberto.iriarte@uan.edu.mx

Similarly it is of great interest and a cultural part, students engage with representative wall paintings (knows as murals) of the institution that are in some of its buildings, are often unaware of the significance and their great importance, affecting the identity of students with their institution.

2 THEORY

For the development of this work, several articles were read in English and Spanish for over 5 years, in order to analyze the latest trends have had with augmented reality technology were reviewed. We found that AR works based on overlapping information about reality from three basic technological resources sometimes complementing each other: the firing patterns of software also known as markers, geo-localization and interaction with Internet [5].

2.1 Augmented Reality Applications

The AR is used in military and technological systems. In the assembly and construction of aircraft, for example, operators connecting cables do wear special glasses that overlay the actual image, the image of the wiring diagrams and wiring diagrams. By superimposing the image of the wiring diagram on the actual items where the operator must intervene the operation much easier, and above all safe [3].

Systems have been developed that attempt to connect texts and published in the press, or ads with mobile phone messages. In experimental settings, with visible marks AR has borne fruit in many systems. They are based on marks in black and white, a kind of two-dimensional barcode that is read by a camera, a computer system that uses that information to superimpose a meta-reality on the captured image [1].

In the work of Delail, Weruaga and Zemerly, a new system that can be used efficiently in a campus of a university to make members aware of their location, and allows it to be easier to find places is presented. The use of image detection campus, allows the system to run on virtually any part of the campus, provided the database for recognition markers, obtaining the location information through AR. Furthermore, the inertial navigation system provides a suitable approach for tracking user location indoors. Thus, the system allows students locate their classrooms, teachers and friends. The user profile should contain information about the student, such as calendar, which can be used to notify students of their next class. The image detection process is carried out with Qualcomm Vuforia SDK that delivers distance and orientation of the marker 3D image. The goal, orientation and posture are used by the RA layer to overlay virtual 3D objects that merge seamlessly with the scene of the video capture. Meanwhile, the location is obtained from the local database of the information related to the detected marker such as location, therefore, may be implicitly obtain the current location of the mobile user [4].

Explore new artistic practices that use mobile consumer devices with GPS technology works to create geo-located artistic works. These technologies allow to create virtual parts that can be accessed via phone or GPS equipped tablets and the appropriate software. An example is The Museum of London app. This mobile phone application allows the user to view photographs of

978-1-5090-3740-7/16 $31.00 © 2016 IEEE
DOI 10.1109/ISMAR-Adjunct.2016.18

the city of London on the current locations. Other artists in the recent Venice Biennale with augmented reality geographic parts are rare Vogels, Don Pigeon, giant pigeon at the center of the Plaza de San Marcos, and James Franco, posing a pervasive, ubiquitous game in we have to find its different appearances in public spaces in the city [10].

Another example is at the University of the Andes, where an article titled "Viewing urban geo-referenced information through augmented reality", is about the development of a mobile application described. It could manipulate and display 3D data in geo-referenced iPhone, to enable the users compare information of edifications geo-referenced stored in a database, in order to compare the information of what should be built in a sector city against what is actually built [14].

Research on techniques to evaluate and improve performance in the use of navigation and route finding with mobile augmented reality, held at the University of Florida, was to carry out a series of experiments to investigate the performance in search tasks with augmented reality [7]. Another highlight research was published in the University of Toronto by Tran, which aimed to develop a platform that records real-time position of mobile devices and laptops [13].

While at the University of Granada, a mobile application called "Healthy Routes UGR", offers several routes within the metropolitan area of Granada you can tour was created walking and cycling, which include both iconic and significant places of the city as the teaching institution. It is based on Android system operating and is available for smartphones and tablet [10].

3 Description of the Application Developed

It was felt that it was needed to develop a mobile application intuitive, lightweight, upgradeable, and scalable. Research on the technologies required for the development of the application, and the extent that they anticipate was performed. The Android operating system was selected because it is one of the most popular for mobile applications. Also because it has a variety of tools and languages.

The main features of the mobile guide application are: (1) share basic information and important points of the university; (2) Know the user´s location and address; (3) Inform the degrees and curricula on campus;(4) Show the location of the main buildings on campus; and (5) Explain the characteristics of each building from the point of view of the new students.

All information is static, i.e. it does not interact with any database to be retrieved, because is not a large amount of data, they are simply stored within the application. A database is useful when large amounts of data handled or want to save information from the user [11].

Thanks to technological advances, was installed on a prototype implementation for mobile devices with Android OS 2.3 and later updates of the same system, integrating augmented reality, allowing the user to have interaction, appreciation and understanding shown in the murals of campus used as markers, including the audio with the explanation of each. As part of the geo-localization it was also included the locations of the main buildings of the university.

In the main menu, the main buildings are divided into three areas: administrative, academic, and athletic. In addition, options of virtual tour using geo-location and the murals were added, as shown in Figure 1. If the user selects any area of the buildings, the list of each is given and information is available to more added detail and location, as shown in Figure 2. About the option of murals of the content of the menu is displayed all murals as markers, taking 17 of them to be viewed by the user.

Figure 1: Application Main Menu.

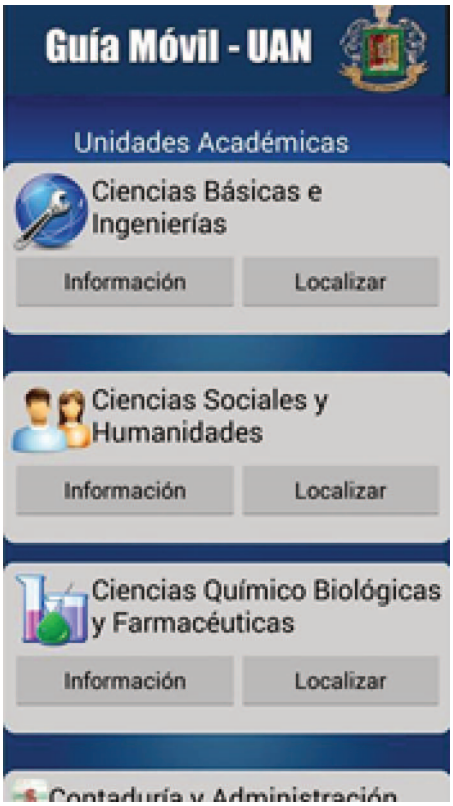

Figure 2: Overview of Academic Units.

To begin the design of the 3D models of the murals, it was necessary to install Sketchup and Cinema 4D software. The first was used to develop 3D models, while the second was to provide animations. These same will added audio with a short story about the meaning of each mural. Audacity software was used for this purpose. Later with Unity3D software models superimposed on the augmented reality and with the help of Vuforia could be linked to the application loaded.

Were taken points of altitude, latitude, and longitude system for geo-location of the main university facilities. During the investigation made on the application design, it has been programmed and finally all the components are added, integrated as is shown in Figures 3, 4, 5 and 6.

Figure 3: The screen shows geolocation information.

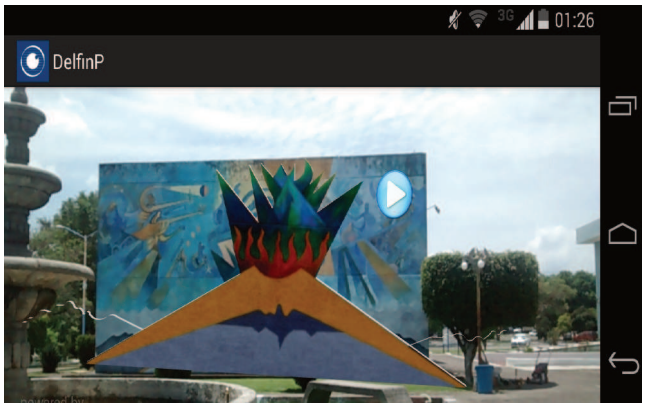

Figure 4: Mural called “The Origin of Our Culture” with AR.

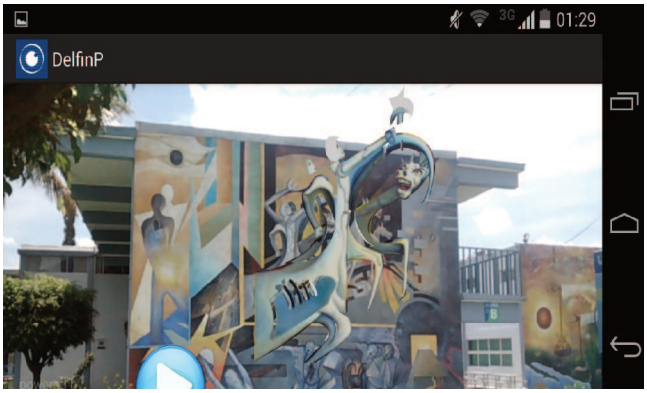

Figure 5: Mural called “Apocalypse” with AR.

Figure 6: Mural called “By as ours to universal” with AR.

After the implementation, form of distribution to end-users was discussed. It was uploaded in the normal way applications to the store device software for which it is built. Also the application was placed directly on a server, in this case as the application is a mobile guide for new students, and was placed on the website of the UAN to be downloaded from their website.

4 Evaluation of the Application

An empirical evaluation was made on a sample of potential users of the application (students and general public). They were provided a tablet with the application and were allowed to use it to evaluate the application in the university context (UAN). After the shoot is over, we applied a questionnaire to assess the application.

The questionnaire basically was designed with twelve questions organized into 3 topics, usability of the application, descriptive knowledge of the site, and user preferences.

The questionnaire was administered to a sample of thirty people divided as follows; twenty students, five teachers and five outsiders providing the following results: 80% say that application allowed a better understanding of the UAN; most users (90%) prefer to use their cell phone to manage application, however it was considered intuitive and fast. For people outside the university, application seemed helpful to them to locate themselves in the surroundings of the UAN.

The application responded to the needs of search corresponding to the location and description of each building information, or a mural in particular. It should be noted that people of the sample considered these applications useful for location in the vicinity of the university, as it allows them to have a wealth of accessible information in real time, without connecting to a remote server.

5 Conclusion

Having tested the application and submitted to an evaluation by potential users, it was concluded that it’s a stable application and was obtained with acceptable performance. Despite being an early version, meets the expectations set at the beginning of development.

As future work, we consider formalizing quality tests corresponding to the application software. Also developing a connection to a server with the required information and updated in real time, minimizing the application size. This would seek to achieve a better response and avoid the need to update the application completely.

Finally, features to consider in future versions of the application are: a tablet version and to other operating systems, and use other input data as RFID (Radio Frequency Identification).

REFERENCES

[1] R. Azuma, M. Billinghurst, and G. Klinker. Special section on mobile augmented reality. *Computers & Graphics,* 35(4), 7-8, 2011.

[2] G.A. Carrasquilla, H. Pinilla, and L.C. Carrasquilla. Aplicación de la realidad aumentada en la enseñanza de la simetría molecular para lograr un aprendizaje significativo. In *Informe de investigación, Universidad de Cartagena, Programa de Ingeniería de Sistemas*. Cartagena, Colombia, 2011.

[3] A.B. Craig. *Understanding augmented reality: Concepts and applications*. Morgan Kaufmann. USA, 2013.

[4] B.A Delail, L. Weruaga, and M.J. Zemerly. CAViAR: ContextAware visual indoor augmented reality for a university campus. *IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, 3, 286-290, 2012.

[5] J. De la Torre, N. Martin-Dorta, J. Saorín, C. Carbonell, and M. Contero. Entorno de aprendizaje ubicuo con realidad aumentada y tabletas para estimular la comprensión del espacio tridimensional. *Revista de Educación a Distancia*, 37, 1-17, 2013.

[6] G. Ghiani, C. Santoro, and L. Davide Spano. A location-aware, multidevice museum guide. *Interacting with Computers*, 21(4), 288-303, 2009.

[7] G. Jaramillo, J.E. Quiroz, C.A. Cartagena, C.A. Vivares, and J.W. Branch. Mobile augmented reality applications in daily environments. *EIA*, 14(1), 125-134, 2010.

[8] D. Lorett, and T. Marrugo. INSITU, Sistema de guía turístico insitu, basado en realidad aumentada para cartagena de indias. In *Universidad de Cartagena, Programa de Ingeniería de Sistemas*. Cartagena, Colombia, 2011.

[9] J. Martín-Gutiérrez, J. Saorín, M. Contero, M. Alcañiz, D. Pérez-López, and M. Ortega. Design and validation of an augmented reality for spatial abilities development in engineering students. *Computer & Graphics*, 34(1), 7-91, 2010.

[10] J. Munárriz. Arte geo-localizado. ASRI - Arte y Sociedad. *Revista de Investigación,* 4, 2013.

[11] A. Silberschatz, H. Korth, and S. Sudarshan. *Database System Concepts*. 6th Ed. New York:McGraw-Hill, 2011.

[12] C. Shih-Chun, H. Wen-Tai, F.L. Gandon, and N. M. Sadeh. Semantic Web technologies for context-aware museum tour guide applications. *Advanced Information Networking and Applications*, 2, 709-714, 2005.

[13] E. Tran. A real-time mediated reality platform for outdoor navigation on mobile devices and wearable computers, In *University of Toronto, Department of Electrical and Computer Engineering*, 2010.

[14] D. Wilches, and P. Figuero. Visualización de información urbana georeferenciada por medio de realidad aumentada. In *XXXVI Conferencia Latinoamericana de Informática*, Bogotá, Colombia, 2010.

CAViAR: Context Aware Visual indoor Augmented Reality for a University Campus

Buti Al Delail, Luis Weruaga and M. Jamal Zemerly
Dept. Electrical and Computer Engineering
Khalifa University of Science, Technology and Research, Sharjah, UAE
{*100037712, luis.weruaga, jamal*}*@kustar.ac.ae*

***Abstract*—CAViAR is a mobile software system for indoor environments that provides to the mobile user equipped with a smartphone indoor localization, augmented reality (AR), visual interaction, and indoor navigation. These capabilities are possible with the availability of state of the art AR technologies. The mobile application includes additional features, such as indoor maps, shortest path, inertial navigation, places of interest, location sharing and voice-commanded search. CAViAR was tested in a University Campus as one of the technologies to be used later in an intelligent Campus environment.**

***Keywords*-Context aware, indoor localization, visual computing, image detection, augmented reality.**

I. Introduction

The outstanding computing capabilities in today's high-end phones or "smartphones" in combination with its wealth of sensors, such as GPS (Global Positioning System), inertial sensors, camera, wireless receivers, are powering the mobile application sector to the extent of becoming the fastest growing one in information communication technologies. Knowledge about the orientation and location delivered by current handheld devices allows for context-aware information services tailored to the user. In that regard, the excellent performance in mobile visual computing [1] permits the detection and tracking of markers and planar objects in real time. This information can be used for mobile visual location recognition [2] in dense outdoor or urban areas where GPS measurements are either unavailable or inaccurate.

On the other hand, localization in *indoor* scenarios is mainly dominated by wireless-based positioning techniques in combination with inertial sensors, and others [3], [4]. The cooperation of computer vision and inertial measurements has been also explored in indoor environments [5], but this particular solution requires unfortunately uncomfortable "wearable" computing equipment. A breakthrough application that derives somewhat from mobile visual computing is augmented reality (AR), namely, the display of virtual 3D objects that merge seamlessly with the actual video scene captured by the device camera [6]. As AR is founded on detection of a known marker, as well as its tilt in 3D coordinates [7], when the marker is in a permanent static location, *de facto* localization is accomplished.

Following the previous concept, this paper proposes a scalable software architecture for a mobile application that provides indoor location detection and tracking based on a combination of image marker recognition and inertial measurements. This localization service is core to deliver a context-aware information system built around an augmented reality software layer. The AR layer informs the user of a close point of interest, overlaying self-explanatory 3D virtual objects related to the location on the real-time video capture. The points of interest nearby are also shown in a 360 degree fashion supported on the device compass readings. Any AR virtual object is "clickable" or "touchable", such that the user obtains information related thereupon. CAViAR's system was initially developed for the University campus where the authors work. Nevertheless, the idea can be extended to any scenario without major architectural changes. The structure of this paper follows: Section II presents the architecture of the global system, which includes a mobile application and the information Internet server; Section III describes with sufficient detail the main implementation aspects of each module; Section IV addresses the evaluation of the system; Finally, Section V brings the future prospects and conclusions.

II. System Overview

The overview of the system architecture is shown in Fig. 1. The user localization process starts with the video capture from the device camera. The video frames are analyzed in real-time to detect and track well-known markers. These markers must correspond to high-contrast planar objects that are part of the landscape, and are thus assumed to be static. The image detection process is accomplished with Qualcomm's Vuforia SDK [8], which delivers the distance and 3D orientation of the image marker (as in [7]). The target orientation and posture are further used by the AR layer to overlay 3D virtual objects that merge seamlessly with the scene in the video capture. Simultaneously, the *location manager* fetches from the local database the information related to the detected marker, such as its location, hence implicitly obtaining the current location of the mobile user. That location can be displayed on an indoor map. Indispensable in this process is the communication module, which

978-0-7695-4880-7/12 $26.00 © 2012 IEEE
DOI 10.1109/WI-IAT.2012.99

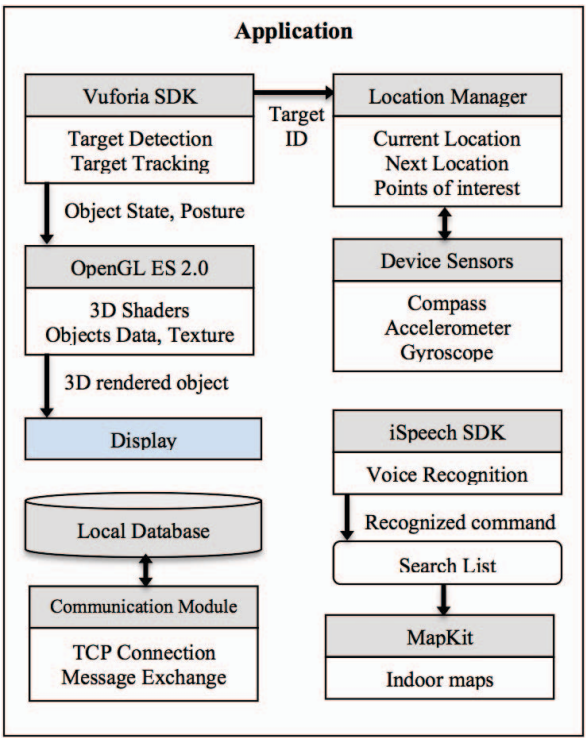

Figure 1. CAViAR's Software Architecture

connects to the service server via a TCP socket, so insuring that the local database is up to date. Above this core layer, the AR 360-degree visualization of points of interest in the vicinity of the user is an important system feature. Other features, such as voice-commanded search of locations have been explored and implemented.

The system is scalable, and it can be extended to include multiple buildings, new locations and detectable markers without major architectural software modifications. In summary, the universal system would consist of different dynamic sub-systems whose input data, such as maps, coordinates, locations, markers, 3D objects, and contextual information would be stored in the server. It is thus clear that since the server provides the data, the application code need not be modified. The iPhone application has been published on Apple's AppStore for members of the University campus, the process of data update being fully transparent to the user.

III. Implementation

The software implementation was modular, based on an object-oriented software architecture. The mobile App was implemented in the iPhone with the official iOS SDK (Software Development Kit). It could be implemented for other mobile devices (Vuforia SDK [8] is available for Android), but given this task would imply major software development efforts because of programming language disparity.

Regarding the programming languages used in the implementation, native iOS applications must be written in Objective-C in the most part. However, the common interface of the communication messages was written in C++ to provide conventional and portable messaging.

A. Mobile Application

1) Indoor maps: The indoor maps were implemented as an overlay on the Google maps. This technique allows the application to show multiple indoor maps of buildings. The map sets are downloaded from the server and stored in the application cache storage. Fig. 2 shows the University campus indoor maps overlaying Google maps.

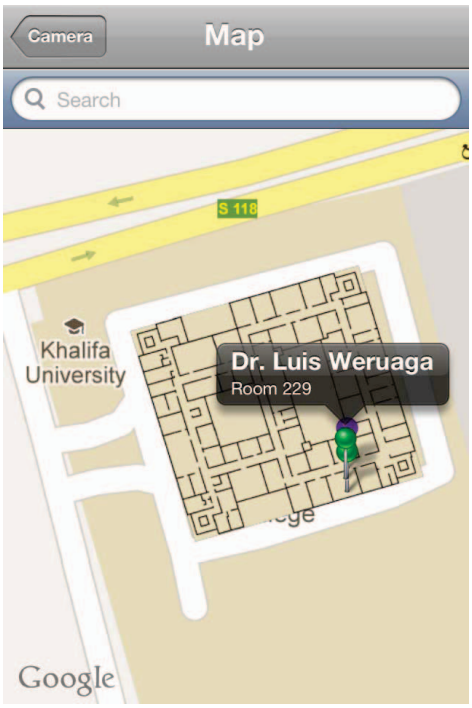

Figure 2. Custom indoor map overlaid on Google maps iOS API.

2) Indoor Localization: The localization system combines several features including image-based localization, different types of navigation, and augmented reality to show informative 3D objects and the surrounding points of interests. Three core processes build this module:

- Getting the user location: In order to get the initial location of the user, the mobile user must find an image marker with the camera. Each marker has a permanent indoor location that is stored in the database. When a marker is detected (see Fig. 3 for an illustrative explanation of the feature or fiducial points of a certain marker), its unique identifier is used to search the database for the location of that marker. Since the user is close to the marker, this location can be assumed to be the user location.

Figure 3. Marker image features and fiducial points.

- Updating the user location: Two different methods to update the user location have been implemented. The detection of a new marker will obviously trigger the location update. However, if the user moves away from that location, his/her location is updated with the device sensors that register motion (accelerometer) and direction (compass). This technique is called *inertial navigation*: the method detects the user footsteps based on the accelerometer readings (see Fig. 4), and updates the user location 2 ft towards the direction read from the compass.

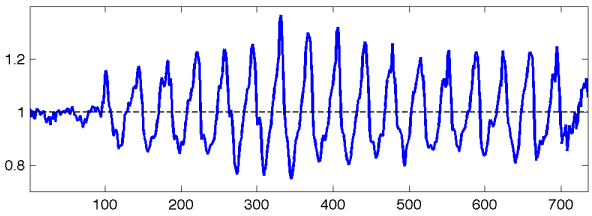

Figure 4. Magnitude of the accelerometer readings with walking user: every cycle in the signal corresponds to a user footstep.

3) Augmented Reality: two types of augmented reality were implemented:

- The 3D Virtual Object AR, which is based on Vuforia SDK [8] and OpenGL ES [9]. Fig. 5 shows a snapshot of the 3D AR objects overlaid on the actual scene captured by the camera. The information of the detected marker (shown in Fig. 3) delivered by Vuforia corresponds to the distance and tilt of the scene, information used to render the 3D object in the "right" place. Moreover, "clickable" virtual buttons, provided also by the Vuforia SDK, were implemented as they allow for visual interaction with the detected image, and used for information retrieval. The 3D object data can be downloadable from the server, this makes the system portable and easily reconfigurable, as the 3D data is obtained from the server on the fly.
- Points of interest AR, shown in Fig. 6. This display method relies on the device sensors, the compass and the gyroscope, to provide a 360-degree view with the points of interest (PoI) in the vicinity of the current mobile user location. As proof of concept and for the sake of simplicity, the PoIs are displayed with a simple clickable button.

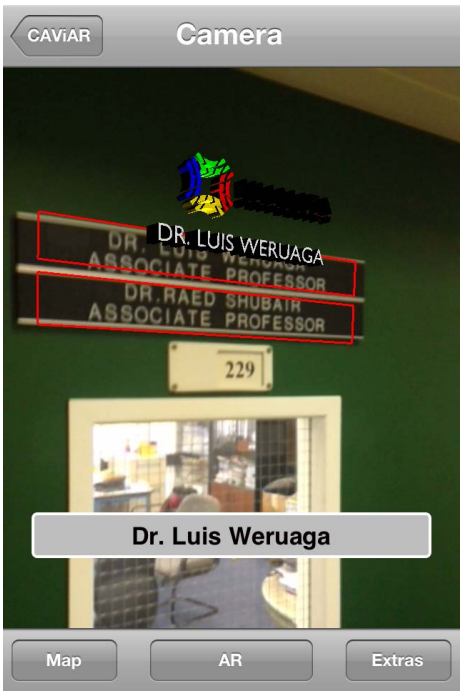

Figure 5. Augmented reality: 3D virtual object displayed in real-time on the video scene; 'clickable" virtual buttons (in red line); when a virtual button is pressed, the related information (faculty name) is displayed.

Figure 6. Indoor 360-degree points-of-interest view.

4) Database: The database is automatically created and managed on the application using the data supplied by the Internet server. Allowing the iOS to create on the fly the database on the device solves compatibility problems between different iOS releases, and satisfies the application scalability by allowing the addition of new buildings and new databases without needing to modify the code.

5) Search and Finding Directions:

- Locations, points of interest, and users can be searched by typing in the search field (see Fig. 2). Search requests can also be performed with a voice recognition engine, detailed in the next point below.
- Voice-commanded Search. The speech recognition module is optimized to recognize a set of voice commands issued by the user, the recognition is done using the API of iSpeech SDK [10]. It offers the ability to provide the speech recognition with a list of words, allowing the recognition of text that is not available in the dictionary nor the English language, such as Arabic names. It also achieves very good accuracy at recognizing different accents. An example command can be "Where is Dr. Luis Weruaga?".
- Providing Directions: The localization system can be used effectively to obtain the user location and search for other locations. Showing the path from the user location to the destination is the link to providing indoor navigation. In order for the system to search for the path, intersections within the building are added to the location database with links that identify which locations can be directly connected. This data can be used to search for the shortest distance path between any two locations and A^* algorithm is the best and most efficient algorithm to use for this type of search. The algorithm guarantees finding the shortest path in the fastest possible way without the need of visiting or testing all locations.

6) Client-Server communication: a lightweight communication protocol was implemented over TCP. The messages are coded in C++, which allows interfacing the same objects in the application and the server code. In order to simplify the designing, coding and sending messages, each message is derived by inheritance from the Message parent class that defines the common header fields. The object of the child class is sent over the network, on the receiving side, the header is used to identify the message type and load the object to its specific class.

B. Server Application

The server application is software running on a PC or workstation. It was implemented using Boost cross-platform C++ libraries, which makes the software to be compiled and run on different operating systems either Windows or various Linux distributions. The server relies on a main database stored in a MySQL server. Fig. 7 shows the server architecture.

The application server provides control information to update the client database, process multiple client requests, and acts as a medium between clients to share their information. Therefore, to make the system intelligent, we can decide the type of information that should be sent to a specific user or group. For example, in a University, students need to know their classrooms, supervisors and public locations available such as the canteen, while the staff may need to know additional information. By making the decision and data customization on the server, which may be hosted on the Internet or locally on the Intranet, enables an intelligent behaviour of the campus.

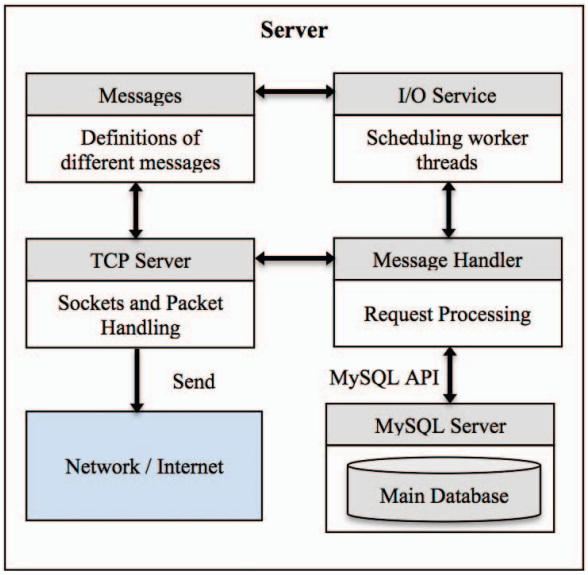

Figure 7. CAViAR's Server Architecture

The clients database can be looked at as cache of the main database, the cached database update mechanism uses timestamps. When the client connects to the server it sends the database timestamp, and as a response, the server checks for new update or entries in the main database and sends only the new data, which saves bandwidth. Furthermore, the server allows pushing information, such that new updates like friends locations can be automatically pushed to the user. However, files like target database, AR objects and the indoor map images are transferred via ftp or http.

IV. Evaluation

The system was evaluated in the University campus of the authors. The inertial navigation system is reasonably accurate. Fig. 8 shows the estimated path between two points distant more than 300 ft, such that the resulting localization error is less than 2 % while holding the telephone in front of the user. A current weakness though is that the precision of the inertial tracking gets influenced by the way the user is holding the phone. We are currently addressing this problem with machine learning techniques such as principal component analysis (PCA). On the other hand, the precision of the trackable detection depends on the features of each trackable image and their uniqueness. Although Vuforia's image detection algorithm is quite robust agains occlusions, we are exploring the effect of removing from the image

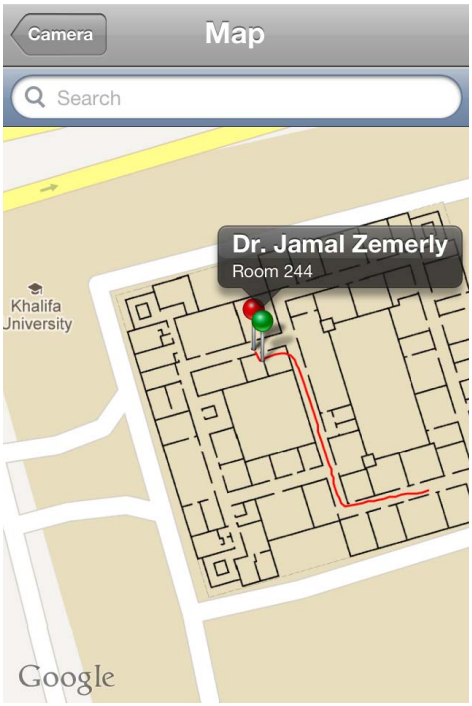

Figure 8. Trace (red) obtained from the inertial location tracking module for a walking path length of about 300 ft.

trackable common features present in different trackables, such as the title of the faculty staff member.

Finally, Table I shows a comparison of features available in CAViAR to other mobile applications that provide augmented reality and indoor maps. The combination of features that are included in CAViAR, such as image detection localization with inertial navigation tracking, and AR with indoor localization distinguishes it from other applications.

Feature	*CAViAR*	*accrossair*	*junaio*	*micello*
Augmented reality	YES	YES	YES	no
3D AR Objects	YES	no	YES	no
Points of interest	YES	YES	YES	no
Search for location	YES	YES	no	YES
Indoor maps	YES	no	no	YES
Outdoor maps	YES	YES	no	no
Indoor localization	YES	no	no	no
Information sharing	YES	no	no	no
User Profile	YES	no	YES	no
Friends	YES	no	no	no
Locate friends	YES	no	no	no
Voice command	YES	no	no	no
Virtual navigation	YES	no	no	no
Indoor tracking	YES	no	no	no
Shortest path	YES	no	no	no

Table I
COMPARISON AMONG CAVIAR AND OTHER MOBILE APPLICATIONS.

V. CONCLUSION AND FUTURE WORK

This paper presents a new system that can be used efficiently in a University campus to make members aware of their locations, and makes it easier to find places. The use of image detection within the campus enables the system to virtually run on any campus, provided the database for recognizing markers, obtaining the location and displaying the information in AR exists. In addition, inertial navigation provides a suitable approach to track the user location indoors. Furthermore, the system allows students to locate their classrooms, teachers and friends. The user profile should contain data about the student, such as timetable, which can be used to notify students of their next classroom.

In the future, we plan to address the following issues for enhancing the system and improving user experience:

- support for multi locations/multiple floors,
- more accurate tracking, including the combination of wireless localization such as WiFi,
- location detection for multiple identical markers,
- enriching the user profile with more information,
- using profile data to provide user-specific customizable information and services,
- advanced representation of points of interests using graphics, and
- supporting more voice commands.

REFERENCES

[1] B. Girod, et al., "Mobile visual search," *IEEE Signal Processing Mag.*, vol. 28, no. 4, pp. 61–76, Jul. 2011.

[2] G. Schroth, et al., "Mobile visual location recognition," *IEEE Signal Processing Mag.*, vol. 28, no. 4, pp. 77–89, Jul. 2011.

[3] L. Klingbeil, et al., "A modular and mobile system for indoor localization," *Intl. Conf. Indoor Positioning & Navigation*, 2010, pp. 1–10.

[4] S. Lee, B. Kim, H. Kim, R. Ha, and H. Cha, "Inertial sensor-based indoor pedestrian localization with minimum 802.15.4a configuration," *IEEE Trans. Industrial Informatics*, vol. 7, no. 3, pp. 455–466, Aug. 2011.

[5] D. Chdid, R. Oueis, H. Khoury, D. Asmar, and I. Elhajj, "Inertial-vision sensor fusion for pedestrian localization," *IEEE Intl. Conf Robotics Biomimetics*, 2011, pp. 1695–1701.

[6] K. Roche, *Pro iOS 5 augmented reality*, Springer, 2011.

[7] D. Wagner, et al., "Real-time detection and tracking for augmented reality on mobile phones," *IEEE Trans. Visual. Comp. Graphics*, vol. 16, no. 3, pp. 355–368, May/Jun. 2010.

[8] Vuforia SDK, Qualcomm Inc., https://ar.qualcomm.at/, Feb. 2012.

[9] A. Munshi, D. Ginsburg, and D. Shreiner, *OpenGL ES 2.0 Programming Guide*. Addison-Wesley, 2008.

[10] iSpeech SDK, http://ispeech.org/, Apr. 2012.

A Novel Campus Navigation APP with Augmented Reality and Deep Learning

Cheng-Hung Lin, Yang Chung, Bo-Yung Chou, Hsin-Yi Chen, Chen-Yang Tsai
Department of Electrical Engineering, National Taiwan Normal University
Taipei, Taiwan, R.O.C
Email: brucelin@ntnu.edu.tw, Tel: +886-2-7734-3432
Email: potasdichrome@gmail.com, Tel: +886-978-035-671

Abstract

Augmented reality has been widely used in many applications because of its ability to offer an amazing way to overlay computer-generated images over the user's real-world view, creating a composite view rooted in real and virtual worlds. Augmented Reality is a realistic, direct or indirect view of the physical reality environment whose elements are "enhanced" through computer-generated or sensory input such as sound, video, graphics, tactile, or GPS data. In this paper, we present a novel campus navigation APP that uses augmented reality to provide users with a new and interesting way to meet our campus. With advanced augmented reality technologies such as computer vision and object recognition, the information about the campus environment and its objects is overlaid on the real world and becomes interactive. In order to improve the APP efficiency, this paper presents a virtual terrain modeling interface with deep learning to improve the object recognition ability.

Keywords: augmented reality, deep learning, computer vision, computer vision

I. Introduction

Due to the emergence of powerful mobile devices and bandwidth-improved broadband wireless networks, multi-media mobile applications (such as streaming or check-in places anywhere) can be executed anywhere. Among the multimedia mobile applications, augmented reality (AR) attracts a lot of attention because of its ability to offer an amazing way to overlay computer-generated images over the user's real-world view, creating a composite view rooted in real and virtual worlds. AR provides a realistic, direct or indirect view of the physical reality environment whose elements are "enhanced" through computer-generated or sensory input such as sound, video, graphics, tactile, or GPS data. Since augmented reality technology allows people to interact with the world around them, AR has become another new aspect embedded in modern life.

In this paper, we propose a campus tour APP with AR technology. The campus tour APP allows users to learn about the campus by quickly taking photographs of campus attractions. In order to identify the large number of tourist attractions on campus, the proposed APP combines positioning techniques to reduce search delays. In addition, we adopt machine learning to improve image recognition. Section II proposes the key components of the campus tour APP. Section III discuss the experimental results and Section IV presents the conclusion.

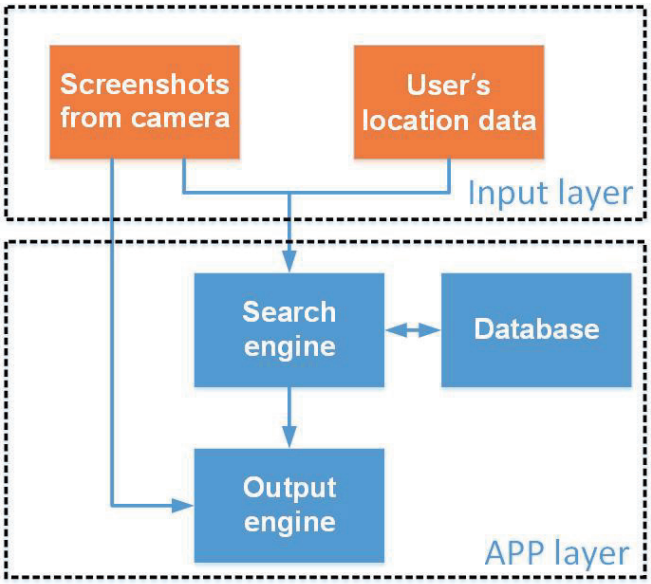

Fig. 1: Architecture of the AR-based campus tour APP

II. Architecture of the AR-based campus tour APP

The key technology of the proposed campus tour APP is based on image recognition. In our project, the features of the campus attractions are extracted by Scale-Invariant Feature Transform (SIFT)[1] and then stored in a database. We also determine the 3D model placement and direction by neural network. Furthermore, we utilize GPS information for assisting recognition performance. In the following, we will discuss the architecture of the proposed campus tour APP.

Fig. 1 shows the architecture of the AR-based campus tour APP that contains three major components, including

ISBN 978-1-5386-4342-6

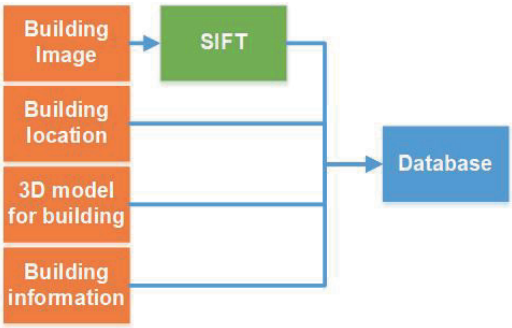

Fig. 2: Database structure

search engine, database, and output engine. The APP imports screenshots and user's location data into the search engine to compare with the database, and then the output engine outputs the match results, including the augmented reality elements, 3D model, and attraction information, on the screen of mobile devices.

A. Database construction

As shown in Fig.2, the database stores the following attraction data: the features of campus attractions, GPS coordinates information, attraction 3D models, and their corresponding information. The features of the campus attractions are extracted by SIFT and saved as a one-dimensional vector. With GPS coordinates information, we can use these as a basis for attraction recognition. When an attraction is recognized, its 3D model and information will be output through the output engine.

SIFT uses the edge and corner positions of an image as its features to identify images in different situations, including rotated or distorted images. Fig. 3 shows an example where SIFT can successfully distinguish the building from three different perspectives. Even if these perspectives are very different, SIFT can still distinguish that they are the same building.

B. Search engine

As shown in Fig. 4, after taking a picture of a target point, the image and the user's GPS coordinates information are sent to the search engine to compare with the database. The search engine first uses SIFT to extract the features of the input image and then compares the SIFT features and GPS data with the database to find the most similar match.

We also use GPS coordinates information as an assistance for data search. Fig. 5 shows that the APP only loads the blue cube database near the user through the user's location. Therefore, an input image does not have to be compared with every attraction in the database. The search speed can be significantly increased.

Fig. 3: Building recognition with SIFT

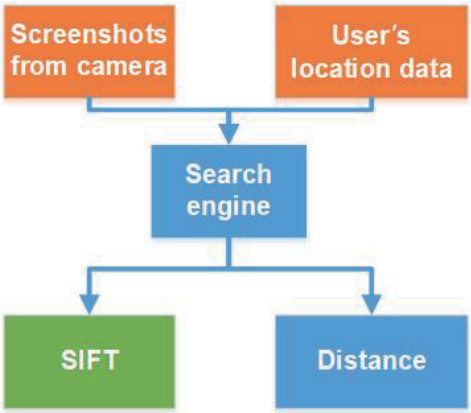

Fig. 4: Search engine structure

C. Output engine with augmented reality

Fig. 6 shows the Output engine structure. According to match results, the Output engine will show the attraction data using augmented reality. The APP achieves augmented reality by placing a 3D model at a specific location in front of the attraction and facing a specific direction. When the user interacts with the 3D model, the attraction's information is also displayed.

The Output engine places a 3D model in a specific place base on a CNN model. We utilize CNN[2] to determine

ISBN 978-1-5386-4342-6

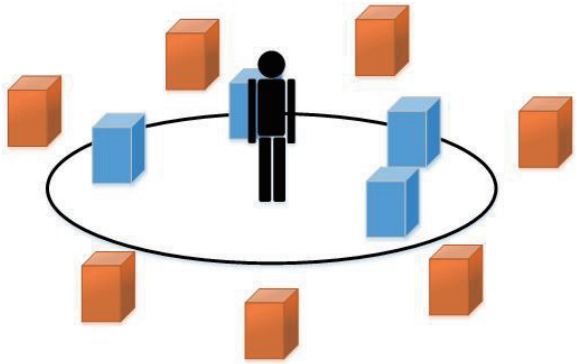

Fig. 5: GPS assistance diagrammatic sketch

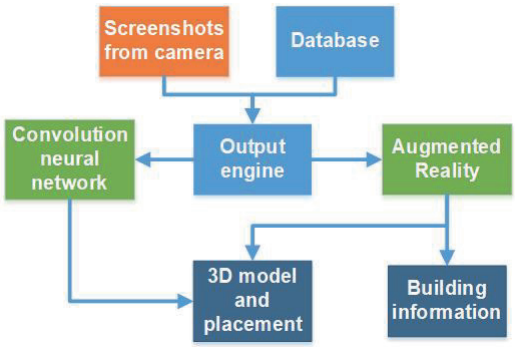

Fig. 6: Output engine structure

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \frac{1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

Fig. 7: Convolution kernel

the specific placement of the 3D models in each angle of attractions. After enhancing specific attraction's features with convolution, we make them as inputs of neural network. The output will be the placement of the 3D model and its facing direction. Fig. 7 is the convolution kernel that we used to find the feature of every attraction. Fig. 8 shows the proposed Neural Network which has 324 input nodes, 2 output nodes, and 2 hidden layers. After performing the convolution, we reduce the image to an 18x18 matrix as input to the neural network, and the output will be the position of the 3D model, including the coordinates X and Y and their rotation.

III. Experimental results

We implement the proposed APP on an ASUS_Z012DA mobile phone, which is equipped with a Qualcomm® Snapdragon™ 625 2.0GHz CPU, an Adreno™ 506 GPU, and a 16 million resolution camera. The APP achieves attraction recognition accuracy rate

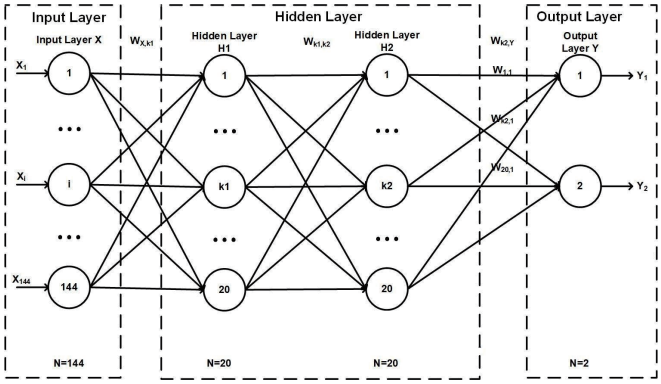

Fig. 8: Back-propagation neural network concept diagram

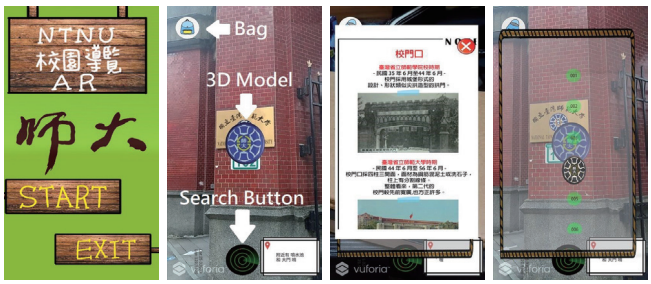

Fig. 9: Application screenshots

of 90%. The attractions can be easily identified by their appearance and characteristics. However, the use of 3D models with CNN was significantly delayed when running on mobile devices. The convolution process takes too much time, even if we have removed the training process from the phone.

IV. Conclusion

We have proposed a novel campus navigation APP that uses augmented reality to provide users with a new and interesting way to meet our campus. The experimental results show that the APP successfully identified the attractions we stored in the database in advance. We will study how to improve the effectiveness of CNN and make it run efficiently on mobile devices.

References

[1] D.G. Lowe. Object recognition from local scale-invariant features, 1999.

[2] Michael Taylor. Make Your Own Neural Network: An In-depth Visual Introduction For Beginners, 2017.

[3] Aurlien Gron. Hands-On Machine Learning with Scikit-Learn and TensorFlow, 2017.

[4] Jin-De Wang. Neural Network and Fuzzy Control Theory, 2007.

ISBN 978-1-5386-4342-6