

MobiAR: Tourist Experiences through Mobile Augmented Reality

David Marimon¹, Cristina Sarasua², Paula Carrasco³, Roberto Álvarez⁴, Javier Montesa⁵,
Tomasz Adamek¹, Idoia Romero², Mario Ortega³, Pablo Gascó⁵

¹Telefonica Research and Development, Barcelona, Spain; ²Visual Interaction Communication Technologies Vicomtech-IK4, San Sebastian, Spain; ³LabHuman, Valencia, Spain, ⁴Indra Software Labs, Madrid, Spain, ⁵Brainstorm, Valencia, Spain;

E-mail: ¹{marimon, tomasz}@tid.es, ²{csarasua, iromero}@vicomtech.org, ³{pcarrasco, mortega}@labhuman.i3bh.es, ⁴rceballos@indra.es, ⁵jmontesa@brainstorm.es
⁵pablogasco@gmail.com

Abstract: Tourists usually request precise and tailored information while exploring a destination. They need accurate information about accommodation, restaurants and tourist attractions, amongst others, in order to make the most of their experience. Mobile Augmented Reality (AR) can help tourists in the process of obtaining such information in a very simple way. MobiAR is an Android service platform for tourist information based on AR, which allows users to browse information and multimedia content about a city through their own mobile devices. Not only does the platform handle location-based information and user preferences, but it also takes advantage of computer vision technologies, so as to determine the tourist resource the user is interested in. This paper presents the main use cases covered by MobiAR and describes the main technological components. Furthermore, the validation scenarios that have been defined for San Sebastian and Valencia are also explained.

Keywords: Mobile Augmented Reality, interactive experiences, context-based multimedia, tourism.

1 INTRODUCTION

Augmented Reality (AR) is considered to be a new medium of creative expression that can enrich the manner in which human beings experience reality [1]. Combined with the information located on the Web, social media and streaming techniques, AR can enhance the way users interact with the physical world, adding additional information about people, buildings or places in order to evoke previous memories or complement present stories.

Recent advances on mobile computing provide new opportunities to AR systems and applications, which can employ wireless networking and location-based technologies on small, lightweight and useful devices that present high processing capabilities. Moreover, the development of mobile AR systems has been improved due to the launch of smartphones, such as the iPhone 3GS and the devices supporting the Android operating system, which incorporate crucial hardware components for AR (i.e. global positioning system, compass, accelerometers and camera).

Regarding application sectors, fielder developments are expected to have a great impact on tourism [2]. Tourists are

likely to see a world of mobile applications and gadgets designed to inform about the location, possible activities and history of the places. Travel guides will come to life in real time, so that tourists can visualise historical places and beautiful landscapes using 3D objects in real time locations.

In such context, MobiAR aims at developing a service platform based on AR technology on smartphones. Furthermore, MobiAR is able to provide users with location-based and personalised information while wandering around an urban environment. The system targets both tourists that are interested in further information during their visit, and citizens that want to discover hidden secrets of their own city.

This paper is organised as follows. Section 2 briefly describes the current status of AR mobile technologies, including existing technologies and main drawbacks, with special attention to the tourism sector. Section 3 introduces the specific tourist arena targeted by the MobiAR project, providing several use cases covered by the platform. Section 4 focuses on the technical description of the platform. The validation scenarios that have been selected for the evaluation of the project are dealt with in Section 5. A discussion and the proposed future work conclude this paper.

2 RELATED WORK

Augmented Reality (AR) is a technology which allows computer generated virtual imagery to exactly overlay physical objects in real time [3,4]. AR creates the illusion that virtual, computer-generated objects exist in the real world, going beyond the static graphics technology where the graphics imposed do not change with the perspective. Development of the needed technology for AR systems, however, is still underway within the research community.

Nowadays, there are several so-called Mobile AR browser AR applications. Layar, Acrossair and Wikitude are probably the most well known.

Layar¹, is an Android and iPhone based mobile AR browser that was launched in 2009. Users can explore their physical surroundings, call up geo-tagged information from the web and superimpose it on the video captured by the camera of the device. The platform has an application programming interface that allows developers to contribute with different

¹ <http://www.layar.com/>

“layers” to the browser. Hundreds of new data layers are available to view on top of the camera viewer of the mobile device, from Wikipedia entries when one is looking at geographic Points Of Interest (POI) to real estate listings that are viewable when pointing at homes for sale. Acrossair [REF] has a similar interaction with “layers” of content. The application is only available on the iPhone and those “layers” are close-managed by Acrossair developers.

Wikitude² is an Android, iPhone and Symbian application launched originally on Android in the fall of 2008. It pulls information from Wikipedia and Qype, the European user-generated review service, and overlays that geo-located data onto the display. Version 3 of Wikitude is integrated with the proprietary user-generated geo-tagging application Wikitude.me. Users can create their own POIs and location-based, hyper-linked digital content that can be viewed through the Wikitude browser application.

Two further examples of more user-centric world browsers are Junaio³ and Tagwhat⁴. They allow users to tag and upload content from the physical world and to share and discover the content that other users have uploaded.

Junaio provides information about POIs and the ability to add 3D animations and share the edited images via the usual social networking sites. Each user generated geo-tagged POI is then visible by all the other users.

Mobile AR browsers currently work by using a combination of the compass, accelerometers, and GPS data to identify the user’s location and field of view, retrieve data based on the geographical coordinates, and overlay that data over the camera view. This combination has a number of limitations which impoverish the AR experience.

Usually, the GPS on a mobile phone only gives a position within around 20 meters, while the smartphones’ compass orientation is only accurate to around 20 degrees. This can lead to problems in determining exactly the field of view of the camera. Real and augmented objects may also be poorly aligned with each other, so the virtual AR objects end up “floating” in the view rather than being solidly anchored to real objects. This becomes a big problem if, for example, two restaurants with different review ratings located beside each other.

The MobiAR system goes one step beyond and adds a competitive advantage on top of the systems previously mentioned. We have developed an augmented reality mobile system which uses computer vision techniques to improve accuracy in identifying landmarks. The location of the user can then be identified up to an order of magnitude more accurately than when using GPS and compass data alone. Details about the developed solution are provided in Section 4.3.

² <http://www.wikitude.org/>

³ <http://www.junaio.com/>

⁴ <http://www.tagwhat.com/>

Additionally, it is worth mentioning that tourism is one of the most productive economic activities in the world. It has allowed the spread in the world of regional and local cultures worldwide, generating diverse business opportunities. Tourism is an economic activity which has a big influence on the Spanish economy.

Spain is the second country in the world according to the number of international arrivals (7% of total) [5]. Tourism industry represents an 11% of Spain’s Gross National Product of Spain [5]. Therefore, MobiAR targets one of the most representative sectors of the country.

Other projects outside of Spain have developed tourist contents for non-specific Augmented Reality systems. We have developed a tourism-specific AR system that gathers needs and expectations of the Spanish tourism industry.

An identified problem of most Mobile AR browsers is that users are often overloaded with tagged information and it turns out to be difficult to wade through the information. Therefore, MobiAR selects the content to be shown according to the preferences stated by the user in order to simplify the AR experience.

3 AUGMENTED REALITY FOR TOURISM IN THE MOBIAR PROJECT

The goal of the MobiAR project is to develop a mobile service platform for tourist information based on the Augmented Reality technology. One of the specific objectives of MobiAR is to employ off-the-shelf smartphones instead of specific equipment.

MobiAR targets users who are willing to discover or better know a destination with enhanced experiences by enriching the visit with multimedia and location-based information. Users can be either tourists who need a guided tour, or local citizens eager to discover more about emblematic places in a city. In both cases, users can generate their comments and reviews, so as to share them with others.

3.1 Example Use Cases

This section describes a couple of typical use cases within the MobiAR project. First, a user browsing her environment to check for interesting Points Of Interest. Second, a user accessing multimedia information related to a certain tourist spot.

In order to discover the surroundings, the user can simply point at any direction with the camera of the mobile device. In the display, the user can observe the video captured by the camera (the reality) and the POIs overlaid according to the distance and direction from the standing location. Different types of multimedia content are associated with different types of POIs. See Figure 1(left) for an example screenshot.

Thanks to the tactile display of the mobile device, users can select a POI just touching the screen. Immediately, the numbers of available multimedia elements (description, comments, images and video) are displayed on the bottom bar. This enables quick awareness of the existing information and

hence helps the user to decide whether to check for more content or not.

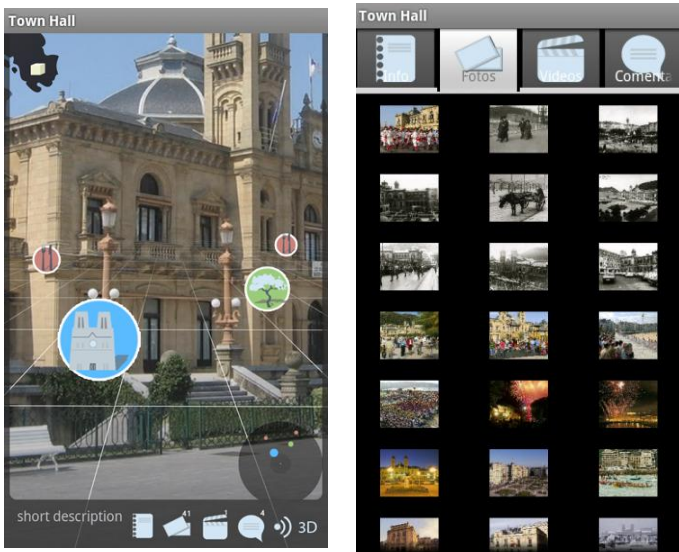


Figure 1: Snapshots of the MobiAR application. (Left) AR view of the environment. (Right) Access to multimedia content related to that POI.

A possible outcome of such awareness is depicted in the second use case described in the following. When the user wants to access more detailed information, just by touching the desired element (e.g. pictures), the application switches to another view. This view has different tabs (one per element type) and directly activated is the one corresponding to the user's choice. An example screenshot of a list of related pictures is shown in Figure 1 (right).

4 THE MOBIAR PLATFORM

The MobiAR platform is based on a mobile client-server architecture. The system is composed of several modules (see Figure 2), namely, the mobile application to visualise and interact with the multimedia content (including the 3D visualisation engine); the content server, which manages users and multimedia content; and the visual recognition engine to determine the localisation of the user. This section provides details about all those modules.

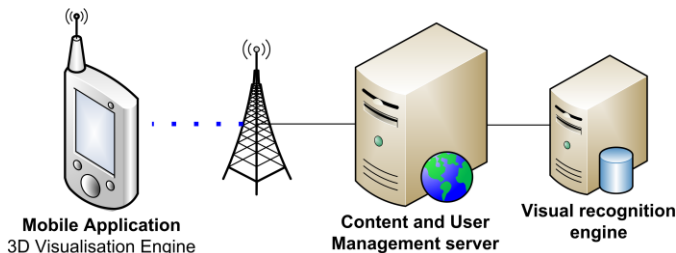


Figure 2: Block diagram of the MobiAR platform.

4.1 Mobile application

Android has been selected as the development platform, although the final prototype may be ported to other platforms that fulfil the basic requirements (management of 3D content,

multilingualism, usability). The MobiAR project bets heavily on 3D graphics content, making it one of the main requirements that were taken into account when choosing the platform, but not the only one. Additionally, the platform also needs a digital compass, accelerometers and a GPS antenna. The two platforms that best meet all these needs are iPhone OS and Android.

Android was born not more than two years ago but is growing at a blistering pace [6]. Several technological consultants such as Gartner [7] predict that "the growth of the Android platform in the coming years will be enormous and may even exceed sales and market share of iPhone in 2012 [8].

In terms of software implementation, the MobiAR mobile application is an Android activity that encompasses the view of augmented reality, offering the user the ability to choose content in both 2D and 3D. The information about the POIs, received periodically from the server, is relative to the position of the user. The user's position is discovered through the GPS built into the terminal and the triangulation of phone masts.

Taking into account context-based data and user profiles, the MobiAR application queries the content server for multimedia items that have been location-tagged (categorized by latitude, longitude and altitude data). When the appropriate contents are retrieved, the AR view is composed with the real-time images captured from the camera of the mobile device and the digital information (menus and POIs) overlaid. There are two possible modes to handle augmentation: 2D and 3D modes. The 2D mode shows multimedia content POIs enriching the real images captured by the camera. This mode is very suitable to discover interesting places nearby. The 3D mode has both the content and the user interface in 3D. Therefore, this mode is tailored to an enriched and leisure-oriented experience.

All of those POI representations would be completely static, if it was not due to information that is acquired through the sensors, namely the digital compass and the accelerometers. Thanks to those sensor readings, the information shown on the screen is dynamically positioned on the screen at the right coordinates.

4.1.1 3D Visualisation Engine

This section describes the technology that enables the 3D AR visualisation mode previously mentioned. Although this 3D view mainly consists of interactive 3D models placed over the POIs registered on the map, it also enables displaying and animating complex 3D interactive menus.

One of the main technological challenges of the project is to place interactive dynamic 3D models on the augmented view. Thus, and in order to maintain coherence between the real and virtual spaces, it is compulsory to reproduce the movement of the real camera in the computer-generated world. Further data from the sensors of the mobile device (compass, accelerometers) is required in order to correctly fulfill this requirement.

A general-purpose 3D engine has been designed taking those requirements into account, and another additional purpose: it is a first step towards the porting of Brainstorm eStudio engine to mobile phones. The 3D engine on Java programming language and the OpenGL ES graphics library have been selected in order to display 3D objects in real time.

The 3D engine is composed of two modules: a core library with the basic 3D functions, and a utilities library where more sophisticated functions satisfy special needs for the project. Among its capabilities, the implemented engine can display images and 3D objects from the memory card, 3D texts and animated menus; pick objects clicked by the user; set the optical parameters of the virtual camera; animate 2D and 3D multimedia contents; animate the camera position; and convert geographical coordinates to UTM.

4.2 Content and user management server

The MobiAR content and user server manages all the multimedia information related to a tourist destination and its resources. More specifically, it provides users with personalised information about the tourist resources close to their location. Furthermore, it manages user information (i.e. a set of preferences that are used by the system to perform information retrieval), and it is responsible for storing and managing the information generated by users (images and user comments).

The communication between the mobile device and the server is performed when the client requests a list containing all the POIs placed around the user or specific information about one of these resources. Another situation that requires the communication with the server is that in which the user sets his or her preferences, which include the language in which the user understands the content, the radius that establishes the distance between the user and the resources to be selected, the maximum number and the type of resources to be filtered. Due to distributed system architecture, the communication is carried out using HTTP POST requests through the Internet.

The server consists of three elements: multimedia content about the destinations (i.e. images, audios, video files and 3D objects); a database implemented in MySQL; and the Web service to manage the multimedia content and the users of the system. The MobiAR Web service is based on Representational State Transfer (REST), a software architecture for distributed hypermedia systems. It offers an easy and efficient way of building client-server architectures, where clients initiate requests to servers that are answered by appropriate responses. Moreover, the MobiAR server is deployed through the open source servlet container Apache Tomcat.

Another key role played by the server is the correction of the user position. When the client asks the server to correct the position of the user, the server stores the reference image captured by the client, and makes an HTTP request to the visual recognition engine, in order to send the location information. This information includes the GPS coordinates and the reference image itself. When the image recognition

module analyses the information sent, it returns a corrected position, which is sent back to the client via the server. Details about this process are given hereafter.

4.3 Visual recognition engine

The visual recognition engine is the core technology that permits identifying the environment in front of the user. This technology, as identified in Section 2, is one of the technological improvements w.r.t. the state of the art Mobile AR applications.

The visual recognition engine matches a query image with a set of reference images stored in a database of geo-located images. By identifying the reference images and their spatial relation with respect to the image captured by the mobile device, the MobiAR platform can recognise the location of the user. It is worth emphasising that this is performed taking the advantage of the GPS information, the recognition is constrained to reference images that were captured close to the query image.

The MobiAR approach, similarly to other state of the art solutions, relies on SIFT features [9], hierarchical dictionaries of visual words [10], inverted file structures, and a spatial verification stage of the top ranked initial results [11][11]. However, in contrast to the above-mentioned methods, the commonly used TF-IDF scoring mechanism [11] is replaced by an early clustering of matches in the pose space (limited to orientation and scale), in a way similar to the voting in the well-known Hough transform. The inclusion of even such a rudimentary spatial verification mechanism in the very initial stage of the recognition is very helpful in cases of small objects buried within complex scenes. For example, when tested with the Oxford dataset (5k ref. images of outdoor locations) this initial stage alone obtains recognition results (Mean Average Precision: 0.47) that are similar to other methods using a more complex spatial verification stage [11].

Once the list of most similar reference images is obtained, data acquired from the device, namely, GPS and orientation information is fused with the spatial relation calculated between the reference images and the query. This fusion establishes the most probable location of the mobile device, assisting and correcting the GPS data.

5 VALIDATION SCENARIOS

Two main application scenarios have been defined for the validation of the MobiAR platform, so that the functionalities of the prototype can be evaluated. Specific areas in two different Spanish cities have been chosen due to their tourist attractiveness.

5.1 Tourists in San Sebastian

This scenario will evaluate the benefits of the MobiAR services for tourists in order to improve their experience while navigating a city. San Sebastian will provide personalised and context-aware experiences, so that users will be able to define their preferences with regard to the types of resources they are

interested in and the maximum distance between them and the resources to be visualised. This customisation will enable the display of only suitable content. Furthermore, this scenario will be exploited in a social environment, so that multiple users can interact with each other. The features related to user-generated content (c.f. Figure 2 (right)) will be put into practice, so as to allow users to add comments to the system (about tourist resources) as well as to upload their own pictures.



Figure 3: Snapshot of the MobiAR application (left) in San Sebastian (right) user-generated content related to the place.

In San Sebastian, a set of 20 POIs have been selected, which include accommodation, restaurants, places of interest and monuments. Some of the most outstanding POIs are the Town Hall, the Boulevard and La Concha Bay. Users will be able to feel the drum festival called Tamborrada that takes place in San Sebastian every year on January 20th, by means of pictures that depict children playing the drums, the popular songs that are played, as well as a video of the official start of the festival.

The selection of the multimedia content for the MobiAR experience has taken into account the promotion strategy of the city. Fomento de San Sebastian, one of the partners of the project, has signed an agreement with several public and private institutions and archives in order to access their content.

On the one hand, film-related contents have been selected, as there is a very close connection between the city and the audiovisual sector. It must be mentioned that San Sebastian hosts the International Film Festival and other representative festivals. On the other hand, buildings and monuments that have a key role in the current life and collective memory of the city have been selected. Thus, both citizens and tourists can better understand the main changes suffered by the city during the last century.

5.2 One day in the City of Arts and Science of Valencia

The MobiAR platform will be also validated in an enclosed environment such as the Arts and Science Museum of Valencia to assess the enhancement of the experience in

cultural and tourist organisations. One of the main objectives of this scenario is the validation of the multilingual capacity of the platform, so that a user can customise the application in his/her own language. Moreover, interactive 3D contents will be integrated in the real view taken by the mobile device.



Figure 4: Snapshot of 3D content for the Valencia scenario.

The selection of the multimedia content for evaluating MobiAR experience in Valencia has been centred on this architectonic spot. It is one of the most relevant architectonic modern sets of buildings of Spain which includes the Science Museum, the Oceanográfico, an IMAX cinema and the Opera Hall and it attracts nearly 4 million people per year to the city [ref].

In this experience, 3D content is integrated in the real view of the Science Museum, so that tourists can better understand some of the concepts explained. Furthermore, citizens can view architectonic 3D models of all the buildings to appreciate the whole set of modern buildings from another point of view.

15 POI of the Valencia city including monuments, places of interest, hotels and restaurants have been selected for the evaluation trial. Some of the most outstanding POIs of Valencia, besides the City of Arts and Sciences are the Serrano's Towers, the Lonja and Colon's Market.

6 CONCLUSIONS AND FUTURE WORK

MobiAR is a mobile Augmented Reality service platform, orientated towards the tourist sector. It allows users to discover destinations through an enriched experience that combines reality with multimedia content and information about tourist resources. The platform, which is based on Android, selects the content to be displayed taking into account location and user preferences. Once the content is correctly filtered, it can be visualised both in 2D and 3D modes.

Even though several related mobile AR applications have recently emerged, one of the major improvements included in MobiAR entails the exploitation of computer vision techniques to improve accuracy in the identification of landmarks. As a result, the position of the user can be calculated more precisely, which can lead to the provision of

more appropriate contents, especially in places that are highly populated by POIs.

Future work will focus on the definition and execution of field trials, with real users in both scenarios (i.e. San Sebastian and Valencia) in order to assess the performance and the usability of the system, whether the users are satisfied with the application, as well as their main motivations to use MobiAR.

Mobile Augmented Reality can help users to make the most of their tourist activities, as users can discover much more accurate information while walking around a city, using only one device. One of the key aspects of mobile AR is to deploy the technology on the devices that users already hold and are familiar with. The more competitive platforms are developed, the more improved the state of the art will be. As there are still challenges to be handled, MobiAR will attempt to contribute to this improvement.

Acknowledgement

This project has been done within the MobiAR project with the funding support of the Ministry of Industry, Tourism and Commerce inside the Avanza program.

Telefónica I+D participates in Torres Quevedo subprogram (MICINN), cofinanced by the European Social Fund, for Researchers recruitment.

References

- [1] Becker, G. (2010). Challenge, Drama & Social Engagement: Designing Mobile Augmented Reality Experiences. Web 2.0 Expo, San Francisco.
- [2] Gretzel, U., Law, R., Fuchs, M. (2010). Information and Communication Technologies in Tourism, ENTER 2010, Proceedings of the International Conference in Lugano, Switzerland, February 10-12, 2010.
- [3] Azuma, R. (1997). A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environment* 6(4): 355-385.
- [4] Azuma, R. (2001). Augmented Reality: Approaches and Technical Challenges. In *Fundamentals of Wearable Computers and Augmented Reality*, Woodrow Barfield and Thomas Caudell, editors. Lawrence Erlbaum Associates, Chapter 2, pp. 27-63.
- [5] INE (Instituto Nacional de Estadística). (2010). La demanda final turística representa el 10,7% del PIB de España en el año 2007. <http://www.ine.es/prensa/np533.pdf>
- [6] Open Handset Alliance. (2010). Android. http://www.openhandsetalliance.com/android_overview.html
- [7] AndroidTapp. (2010). List of Android devices. <http://www.androidtapp.com/list-of-android-devices/>
- [8] AdMob Mobile Metrics. (2010). Mertics Highlights. <http://metrics.admob.com/wp-content/uploads/2010/06/May-2010-AdMob-Mobile-Metrics-Highlights.pdf>
- [9] Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *Intl. Journal of Computer Vision*, 60(2):91-110.
- [10] Nister, D. and Stewenius, H. (2006). Scalable recognition with a vocabulary tree. In *Proc. Computer Vision and Pattern Recognition (CVPR)*.
- [11] Philbin, J., Chum, O., Isard, M., Sivic, J. and Zisserman, A. (2007). Object retrieval with large vocabularies and fast spatial matching. In *Proc. Computer Vision and Pattern Recognition (CVPR)*.