Interacting with Augmented Assets in Cultural Tourism

Maria Teresa Linaza[1], Ander García[1], Isabel Torre[1] and Jose Ignacio Torres[2]

[1]Department of Tourism, Heritage and Creativity, VICOMTech, Paseo Mikeletegi 57
20009 Donostia-San Sebastian, Spain
{mtlinaza, agarcia, itorre}@vicomtech.org
[2]The Movie Virtual, Paseo de los Olmos
2000 Donostia-San Sebastian, Spain
jitorres@themovie.org

Abstract. This paper presents the results of the PRISMA project, where a proposal to include a zoom-lens camera into an outdoor marker-less AR system for the cultural tourism sector has been developed and assessed. The key issue of PRISMA is the combination of the commonly known concept of tourist binoculars and AR technologies. PRISMA presents tourist information from a new point of view, allowing the user to interact with multimedia information. An panoramic view from one of the hills of Donostia-San Sebastian city in the North of Spain has been chosen as the real validation environment.

Keywords: Augmented Reality; multimedia management; cultural storytelling.

1 Introduction

Augmented Reality (AR) is not that much of innovation any more. Several number of projects and applications concerning AR using different types of registration techniques or recording cameras have been published. Regarding registration techniques, the use of marker-based tracking is widely extended due to its robustness. The position and rotation of the markers is extracted from the image recorded by a camera using image processing algorithms. Nevertheless, these techniques can not always provide the best solution, mainly in wide areas where the addition of markers is not always possible or can be cumbersome. Therefore, other technologies are used to track rotation and position of the user during the registration process. A variety of motion trackers can be found on the market, based on infrared, magnetic fields, inertial forces or other technologies.

However, the adquisiton of position and rotation data alone still is not enough to get an acceptable registration of an augmented scene. Further inaccuracies in the registration process can appear due to the specific parameters of the tracking cameras, like focus, opening-angle and distortions. To adjust these parameters, AR systems must include methods to calibrate the system according to the specifications of the selected camera. Thus, registration is still an up-to-date research topic with a wide potential for improvement, as synchronizing the virtual and the real worlds, and rendering the graphical interface from the point of view of the users is a crucial factor.

Up to now, few AR applications with zoom-lens cameras have been developed. Zoom-lens require an interactive adjustment of not just the position and orientation of the system, but also of the internal parameters of the camera like opening angle, distortions and focus. Marker-based tracking systems can overtake the difficulties associated to the usage of zoom-lens cameras implementing complex image processing algorithms for tracking. However, marker-less approaches including outdoor applications require additional developments to achieve proper registration.

This paper presents the results of the PRISMA project, where a proposal to include a zoom-lens camera into an outdoor marker-less AR system for the cultural tourism sector has been developed and assessed. The concept of binoculars on hills expecting tourists to put a coin in to see the surrounding area for a couple of minutes is widely known. They offer an overview over the buildings and streets of a city, nature and cultural sites as well as the chance to zoom tourist assets closer to the user. The view can increase the interest of tourists in visiting places and help them in choosing further targets to visit later on. However, it may often be hard to find anything except the nearby woods or the sky in the field of view of the binoculars. Even when a nice looking attraction is found, it turns out to be sparsely interesting because of the lack of sight and information about the resource.

Moreover, existing multimedia content presentations are distant from the real environment which means that users have to leave the tourist site to obtain additional information. If tourist organizations wished to reach wider audiences, they would have to build attractive multimedia content available on site. Therefore, new systems that support these innovative applications and provide added-value content are required.

The paper is organized as follows. Section 2 presents a brief state of the art of AR technologies and their application to cultural tourism applications. The following section includes a brief description of the PRISMA project and the validation scenario. Section 4 gives a technical overview of the prototype, including its components and performance. The assessment of the prototype in a real scenario is presented in Section 5. Finally, Section 6 presents some conclusions and further work.

2 State of the Art

Augmented Reality (AR) is a technique in which the view of the user is enhanced or augmented with additional information generated from a computer model [1]. In contrast to Virtual Reality, where the user is immersed in a completely computer-generated world, AR allows the user to interact with the real world in a natural way. In order to make AR systems effective, the computer-generated objects and the real scene must be combined seamlessly so that virtual objects align well with the real ones. It is therefore essential to determine accurately the location and the optical properties of the registration cameras. The registration task must be achieved with special care because the human visual system is very sensible to even small mis-registrations.

Many researchers believe that AR is an excellent user interface for cultural tourism applications, because it allows intuitive browsing of location-based information. Using AR, the perception of the user of the real world is enhanced by computer-generated entities such as 3D objects and spatialized audio. The interaction with these entities takes place in real time to provide convincing feedback to the user, giving the impression of natural interaction.

It can not be denied that the standard equipment used in the traditional tourism activity has remained unaffected by the common attitude to embrace the latest and greatest technological solutions. When visiting a tourist asset, the traditional visitor's guide and a trusty map remain as the standard equipment. Although everybody is familiar to such scenario, it shall shortly undergo a rapid metamorphosis. The standard equipment will change into Personal Digital Assistants (PDA); instead of a map, an electronic map enhanced with a reliable position sensing mechanism is envisaged; guides will change into online access to a data repository endowed with rich multimedia content, spatial data and other relevant up-to-date content.

As a key element in the standard equipment, maps are abstractions of the real world and are well suited to give an overview of an area. However, some cognitive effort is needed to determine the orientation of the user from looking at maps. On the other hand, AR allows matching the scene with reality immediately, although it could be difficult to provide location-based information without positioning and orientation errors.

Questions like "What can I see and do in he city?" or "What is important to know about it?" will be answered by future tourist guides or the so called Pedestrian Navigation Systems (PNS). When reviewing literature on PNS, it can be observed that most of current concepts have been developed in one of two following research approaches: location-based services and outdoor AR systems.

Some well-known systems that have been developed within the location-based services approach are Cyberguide [2], GUIDE [3], Deep Map [4], and Lol@ [5]. The presentation of the location of the tourist is mainly based on maps, which are displayed on mobile devices connected to wireless networks. Further multimedia information such as photos, videos and audio about landmarks and points of interest are linked to the map.

Concerning outdoor AR systems, different projects can be mentioned, such as the Touring Machine, ARREAL [6] and Studierstube Mobile AR [7]. In the Touring Machine project [8], users stand in the middle of the campus, wearing a prototype which includes a Head Mounted Display (HMD). The system overlays textual labels over campus buildings. However, the relative inaccurancy of the trackers is not a significant problem for the success and acceptance of the final application by the user. Although the approaches for PNS differ substantially, both of them mix reality and virtuality to support orientation and wayfinding; location-based services adding real information to virtual maps and Augmented Reality systems adding virtual objects to real scene views.

From the state of the art, it can be concluded that there is a shift towards more active, cultural and urban destinations. The attractiveness of urban culture and other aspects of the tourist resources of the city receive growing attention from tourists. How to deal with these opportunities is a challenge for the tourist industry as well as for local and regional governments. In order to manage quality tourism in tourist destinations, it is necessary to create new possibilities for services and dynamic ways of experiencing the destination.

3 The PRISMA Project

The aim of the PRISMA project is the design, development and implementation of a new 3D visualization device based on AR technologies and storytelling techniques for cultural tourism applications (Fig. 1). The emphasis within the project has been focused on developing experimental user interface software, not on designing hardware [9][10]. Therefore, commercially available hardware has been integrated in order to build the prototype.

Fig. 1. PRISMA prototype in the real environment.

The key issue of PRISMA is the combination of the commonly known concept of tourist binoculars and AR technologies. PRISMA presents tourist information from a new point of view, allowing the spectator to interact with multimedia information. By means of these technologies, the real scene can be enhanced by virtual information to increase the experience of the user and provide added value interactive personalized multimedia content about the current view. The proposed techniques are highly visual and interactive in order to access and understand tourism information.

A panoramic view from one of the hills of Donostia-San Sebastian city in the North of Spain has been chosen as the real validation environment. For the assessment of the prototype, the binoculars have been placed at the middle height of the hill to allow tourists and also local citizens to "visit" some of the interesting cultural and tourist attractions of the city, such as the Cathedral and Santa Clara Island in the middle of the bay, as shown in Fig. 2. PRISMA overlays graphical labels over cultural tourist attractions. The relative accuracy of the labels is crucial so that the spectator chooses the correct cultural asset and not another one in the nearby.

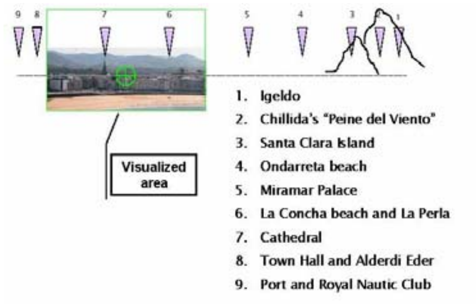

Fig. 2. Selected tourist attractions for the augmentation.

When the user discovers a tourist asset about which he/she is interested in getting additional information, it is enough to click on the graphical representation of the object and a menu with the available choices for the content (cultural and tourist information, historical data, interactive models) is displayed. PRISMA allows tourists to obtain multimedia personalized information about these cultural tourist assets, including textual information (timetables, directions, telephone numbers), 2D photographs and maps, or even videos and 3D reconstructions. The user interface is based on simple buttons placed on the system.

4 Technical Description of the Prototype

4.1 Components of the Prototype

As shown in Fig. 3, the prototype basically comprises a camera that records a real time image from the point of view of the user, and sends the image to a processing unit. The camera used in the prototype is the Sony SNC-Z20N, which incorporates a highly sensitive 1/4 type CCD. Equipped with auto-focus 18x optical zoom lens, this camera can zoom in on small or distant objects with exceptional resolution. Moreover, the camera can be controlled from an external control device, allowing local control of zoom parameteres of the camera over a serial RS-232C connection.

The prototype further includes a visualization device through which the user can view the augmented scene. Our current setup uses a 5DT HMD 800-40 Head Mounted Display fixed to the structure as an output device. It is basically a metaphor of conventional binoculars, including a video non see-through visualization system.

As an essential device involved in the registration process, the tracking of the system is achieved by an Intersense Inertiacube2 orientation sensor for inertial tracking as the prototype is mainly supposed to be moved in two dimensions (left/right, up/down). As seen in Fig. 3, the visualization device and the camera are mounted on the same mechanical axis so that they move in a synchronous way. Therefore, the tracking system calculates the position of the visualization device and thus, the position of the camera.

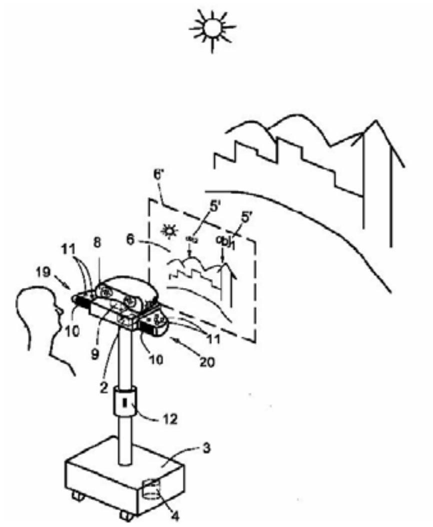

Fig. 3. Technical components of the PRISMA system.

The prototype also includes seven buttons distributed among the left and right sides of the prototype for simple and ergonomic interaction. On the one hand, the two buttons on the left side allow the user zooming in or out over the augmented scene. On the other hand, there are five buttons on the right side: one button in the middle which can be used to click on the augmented graphical objects and also to choose among the menus of the additional multimedia content; and four buttons distributed around the middle button, allow browsing through the menus for selecting the additional multimedia content.

Additionally, some sound speakers are included so that the user can listen to presentations or other audio information included in the additional multimedia content.

Content is stored in two different databases in the processing unit: an augmentation database and an auto-administrable content database. The augmentation database stores the 3D graphical objects, including the name of the main tourist attractions and other graphical objects that are superimposed to the points of interest for the spectator through applicable AR methods. These methods convert 3D graphical objects into 2D graphical representations, which are superimposed to the real scene, resulting in an augmented scene. The auto-administrable content database stores the additional multimedia content, including videos, movie clips, interactive 3D panoramas or even 3D models of existing and reconstructed cultural tourist attractions.

We have used ARToolKit as the software platform for developing the applications. ARToolkit is a software library for building AR applications, which offers the possibility of augmenting a video stream with graphical objects. The graphics supported are principally OpenGL, but also a VRML import PlugIn exists which allows importing 2D as well as 3D objects and animations. The ARToolkit is based on marker-based tracking that detects special markers in the video stream by means of computer vision algorithms and calculates the real camera position and orientation relative to the physical markers in real time. As the tracking approach used in

PRISMA depends on inertial sensors fixed to the binoculars, the ARToolKit has been modified.

Finally, an Authoring Tool has been implemented in order to simplify the manipulation of the 3D graphical objects, their placement in the augmented scene and the associated additional multimedia content they are supposed to show. The programme can load XML files to change them or to create new ones in the proper style for the main programme. The prototype includes up to ten points of interest with up to five subobjects to choose within the menu. 3D graphical objects include tourist points of interest and their corresponding markers in the augmented scene, while subobjects stand for different types of additional multimedia content shown when choosing 2D graphical representations.

4.2 Performance of the Prototype

Fig. 4 shows the functional schema of the PRISMA prototype. As explained previously, the real camera captures the Field of View (FoV) of the user. The tracking system calculates and sends certain positioning information to the processing unit, informing about the current location and orientation of the visualization device. The processing unit then runs a graphics adaptation process which converts 3D graphical objects into 2D graphical representations depending on an orientation vector obtained from the positioning information. In other words, the processing unit adapts a shot of the 3D graphical objects to obtain a virtual scene, controlling a "virtual camera" using the orientation vector of the real camera. Therefore, the real scene and the 2D graphical representations are synchronized according to the position of the camera. This synchronization allows composing the real and the virtual scenes of the augmented scene.

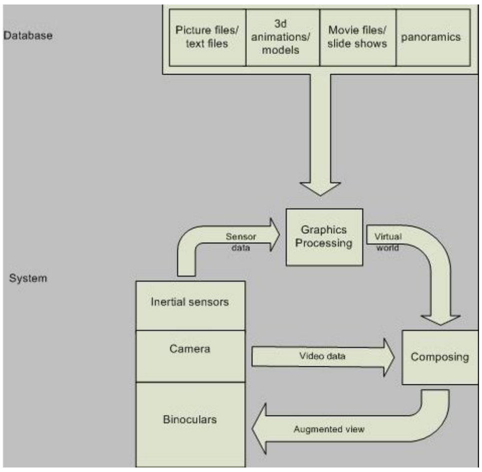

Fig. 4. Functional scheme of the PRISMA system.

If the user moves the prototype to the left or to the right, the tracking system informs the processing unit about this change. The graphics adaptation process then updates the 2D graphical representations of the 3D graphical objects so that they match the new real scene. As a result, when the user changes the position of the prototype, the augmented view changes as a whole.

In order to clarify the above mentioned concepts, Fig. 5 depicts an example of the performance of the PRISMA prototype. In the real validation scenario, the augmented scene viewed by the user is formed by the real view of the Bay of San Sebastian and some hills, some of the buildings and one of the hills being pointed out by corresponding 2D graphical representations.

As the user sees an interesting object about which he/she would like to get more information - for example, the high building in the middle of the screen pointed out by the text "obj1"-, he/she presses on the buttons on the right side of the prototype until selecting the 2D graphical representation of that interesting object. Then, the prototype displays the additional multimedia content regarding the selected 2D graphical representation.

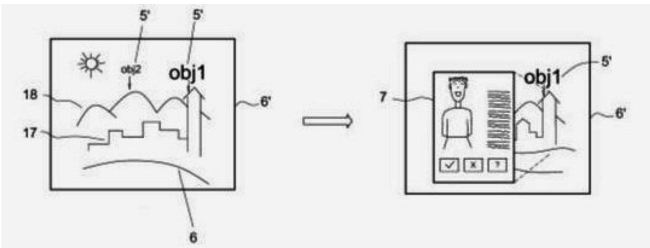

Fig. 5. Example of the performance of the PRISMA system.

For example, as shown in Fig. 5, the prototype pops up a screen showing some text and playing a video of a guide explaining certain characteristics of the selected tourist asset, and/or offering extra navigation options. The system provides personalized information, where different contents are displayed for several user profiles, including aspects such as multilingualism. For example, an English-speaking tourist with a cultural background receives additional in-depth information in English about the history of the interesting building.

5 Assesment of the Prototype in the Real Scenario

The overall aim of the usability analysis of PRISMA has been the development of typologies of users and standards, so that new interactive tourist experiences could be personalized in the future taking into account concepts such as usefulness and ease of use. Therefore, the behaviour of the user when dealing with new technologies as a means of communication of tourist contents has been assessed. In order to achieve this aim, the quality of the communication between the prototype and the user has also been measured. Therefore, it is crucial to assess the impact of PRISMA within the added-value experience with the innovative 3D visualization prototype.

Concerning the sample, about 100 people took part in the usability study, from which 47 answered in-depth questionnaires, at the real scenario at the Monte Urgull in San Sebastian (Spain) in March 2006. The average profile of the participant is a man between 20 and 35 years old (61% of the sample) and quite familiar with the use of new technologies (mobile devices, laptops, digital cameras, Internet). Most of the participants worked in the tourism sector, including destination managers and content providers.

In order to measure the scope of the results and conclusions, it must be mentioned that the profile of the sample has been biased, as most of the selected participants worked in the tourism sector. They had a double role within the usability study: a common tourist role for the quantitative analysis and a professional of the tourist sector role for the qualitative one. The participants provided a critical vision taking into account not only their personal point of view, but also an extended discussion about further developments of the prototype. The conclusions have been grouped on the basis of the issues that have been highlighted within the evaluation process.

Most of the participants agreed that the approach of using advanced visualization technologies such as Augmented Reality enhances the interactive experience with tourist content. The interviews concluded that the device provides added-value multimedia contents in an interactive way. This way, information is more accessible with new visualization means and innovative interaction metaphors. Moreover, more than 80% of the participants were willing to pay up to three euros for the experience. As a conclusion, the participants have positively assessed the usefulness of the prototype as a tourist information point for urban environments due to its capability of providing Location-Based contents to allow better route planning.

Concerning the ease of use, the results of the quantitative analysis were mostly positive, although issues related to interaction, comfort and simplicity in the access to contents have to be improved. Many limitations or usability problems were related to the fixed height and the restricted physical adaptability of the prototype to the physical conditions of the user. Position of legs, arms and hands of the user was influenced by the binoculars, the interaction devices, the mechanical rotation of the visualization device and the design of the device as a whole. Further improvements in the physical design of the device will have a great impact not only on a more comfortable experience for the user, but also on a smaller damage when using the device. A final recommendation was the need of the contextualization of the prototype in order to facilitate its usage.

Regarding the contents, more than half of the sample found them usable, interesting and innovative. Moreover, two thirds of the users were impressed by the historical images that were displayed. Interactive contents based on QuickTime panoramas had the greatest impact, specially the navigation and zooming inside the Cathedral of the city. However, the potential of the prototype could be extended by panoramic approaches, holograms or non-linear story-telling techniques. As a final conclusion, it can be said that PRISMA can be considered as a prototype with a great potential impact over the tourist sector. Further application scenarios such as environmental impact of architecture planning projects have been suggested.

6 Conclusions

This paper presents the results of the PRISMA project, where a proposal to include a zoom-lens camera into an outdoor marker-less AR system for the cultural tourism sector has been developed and assessed. The main objective of the project was the implementation of user interfaces and personalized contents, and not the hardware development. Therefore, commercially state-of-the-art equipment has been integrated to build the prototype.

Regarding the hardware, the inertial tracker works quite accurately although the answer is quite slow in some cases. Therefore, when the system was moved too fast left to right, the delay of the tracking system got very obvious, limiting the immersion effect of the 3D graphical objects. Further refinements of the prototype should include complementary hybrid tracking systems depending on further outdoor tracking possibilities.

The main technical challenge in the programming of the prototype has been the implementation of the augmented zoom. Although many different alternatives were tested to achieve appropriate results, it can be concluded that there are two possible ways of programming the zoom algorithms: implementing the linear interpolation of the changes in the FoV with slow change rates, or using absolute zoom positions with the most correct overlay but minor delays of the real behind the change of the virtual position. As a general statement, it can be affirmed that a higher precision of the information about the position of the zoom provided by the real camera as well as a higher communication speed could be a possible solution.

As the merging of the 3D graphical objects with the real data is crucial in AR, the overlay of the graphics on the video-stream in PRISMA is also a key issue. Due to several reasons, the graphics merged with the real environment suffer from some slight inaccuracies. Main shortcomings are the synchronisation of the zoom between the real and virtual cameras, the change of the rotation centre as the camera lens move from the centre while zooming and the delay of the inertial sensor on fast movements. These problems may be solved using another type of camera or some kind of mechanical mechanism to move the camera, as the lens must be kept in the centre of the rotation.

Other optional features for further refinements of the prototype are the option to see the surrounding area during different seasons of the year, i.e. with snow, or in summer, or without clouds, or the possibility to combine the AR aspecta with a virtual immersive enhancement. By selecting an object of interest such as a cultural building in the area, a virtual walkthrough is started. This application "moves" the user from the real position into the building, letting him/her move around and discover the virtual interior of the building.

This paper can be considered as an example of how an AR system can annotate the environment of a user to visualize related information. Unlike desktop visualization environments in which the system can control what the user sees, in tourism applications, the tourist determines the viewing specification and the physical environment determines which tourist assets are visible.

As cultural tourism is dominated by visual experiences to a large extent, it can be regarded as a rich site for both the 'creation' and analysis of visual evidence, by both researchers and practitioners. Despite the importance of the visual sight in tourism, image-based research methods are simply not on the agenda for many tourism researchers. When visual images are used, it is often only to support the use of written text. PRISMA aims at developing specialised language or standardized methodologies and modes of using and analysing image-based sites.

As concurrence gets harder, it will become increasingly important that tourist destinations develop products tailored to their main consumer groups. Thus, it becomes ever more important to understand visitors' perceptions of destinations. Urban environments serve multiple functions and often become ideal destinations for both business and pleasure travellers as they increase in size and importance.

Acknowledgements

This paper is part of the project "PRISMA- Augmented vision for Cultural Tourism Applications", financed by The Movie Virtual and Ereiten Kultur Zerbitzuak within the INTEK program of the Department of Industry of the Basque Government.

References

1. Azuma, R.: A Survey of Augmented Reality. Presence: Teleoperators and Virtual Environment 6(4), 355--385 (1997)
2. Abowd, G.D., Atkeson, C.G., Hong, J., Long, S., Kooper, R., Pinkerton, M.: CYBERGUIDE: A Mobile Context-Aware Tour Guide. ACM Wireless Networks 3, 421--433 (1997)
3. Cheverst, K., Davies, N., Mitchell, K., Friday, A., Efstratiou, C.: Developing a Context-aware Electronic Tourist Guide: Some issues and experiences. In: 2000 Conference on Human Factors in Computing System, pp. 17--24 (2000)
4. Malaka, R., Zinf, A.: DEEP-MAP- Challenging IT Research in the Framework of a Tourist Information System. In: Information and Communication Technologies in Tourism ENTER 2000 (2000)
5. Gartner, G., Uhrlitz, S.: Cartographic Concepts for Realizing a Location-based UMTS Service: Vienna City Guide LOL@. In: 20th Int. Cartographic Conference ICC (2001)
6. Baus, J., Krüger, A., Wahlster, W..: A Resource-Adaptative Mobile Navigation System. In: Int. Conference on Intelligent User Interfaces IUI, San Francisco (2002)
7. Reitmayer, G., Schmalstieg, D.: Location-based Applications for Mobile Augmented Reality. In: 4th Australasian User Interface Conference AUIC'03 (2003)
8. Feiner, S., MacIntyre, B., Höllerer, T., Webster, A.: A Touring Machine: Prototype 3D Mobile Augmented Reality Systems for Exploring the Urban Environment. In: Int. Symp. On Wearable Computing ISWC, IEEE Press (1997)
9. Alzua-Sorzabal, A., Linaza, M.T., Susperregui, A.: Providing On-Site Augmented Information to Tourists. In: M. Hitz, M. Sigala and J. Murphy (eds.) Information and Communication Technologies in Tourism 2006, SpringlerWien New-York (2006)
10. Fritz, F., Susperregui, A., Linaza, M.T.: Enhancing Cultural Tourism Experiences with Augmented Reality Technologies. In: Proc. of the 6th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage VAST2006 (2006)