In-Situ Visualization for Cultural Heritage Sites using Novel Augmented Reality Technologies

Didier Stricker[1], Alain Pagani[1], Michael Zoellner[2]

[1] DFKI – TU Kaiserslautern – Augmented Vision Lab, Germany
[2] Fraunhofer IGD, Department Virtual and Augmented Reality, Germany

Abstract

Mobile Augmented Reality is an ideal technology for presenting information in an attractive, comprehensive and personalized way to visitors of cultural heritage sites. One of the pioneer projects in this area was certainly the European project ArcheoGuide (IST-1999-11306) which developed and evaluated Augmented Reality (AR) at a very early stage. Many progresses have been done since then, and novel devices and algorithms offer novel possibilities and functionalities. In this paper we present current research work and discuss different approaches of Mobile AR for cultural heritage. Since this area is very large we focus on the visual aspects of such technologies, namely tracking and computer vision, as well as visualization.

Key words: *Mobile Augmented Reality, Tracking, In-Situ Visualization*

1 Introduction

Mobile Augmented Reality for cultural applications represents a very challenging area. Highly complex and highly advanced systems and technologies need to work in a balanced way in order to produce the apparently simple and intuitive overlay of virtual information onto images of the reality. In such applications many technological areas - mobile computing, energy management, localization and tracking, human-computer interaction, and visualization – converge together. In that sense, a mobile augmented reality system can be compared to complex mobile robotic systems.
The paper delivers insight into the current state of the art technology, which has been developed and evaluated in practical European projects, namely ArcheoGuide (IST-1999-11306), Matris (IST-002013), and iTACITUS (IST 2.5.10 – 034520). The paper presents two major contributions in localization and visualization areas, respectively novel optical tracking algorithms and the concept of reality filtering for adapted visualization (Zoellner, 2008).

The paper is structured as follows: in section 2 we present the components and architecture of mobile AR-systems. Section 3 describes novel tracking solutions which are robust and cope with jitter. Section 4 is about visualization of available or reconstructed data and artifacts of cultural sites. We finally present real applications (section 5) and conclude in section 6.

2 Deploying mobile AR

Site preparation
Before deploying a mobile augmented reality system, a preparation step is mandatory. It consists in a site survey to collect the necessary data and plan the hardware installation. In ArcheoGuide, we collect aerial photographs and surveying data and enter them in a geographical information system (GIS), which is also used to construct a digital elevation map. This 3D site representation enables us also to identify major monuments and corresponding viewpoints with unobstructed views. The digital information and virtual 3D models is then attached to geographical points and thus defines suitable tours for the visit. Moreover high-definition photographs of the ruins are captured from the predefined viewpoints along the

tour paths. For each viewpoint, we took a set of tiled photographs to simulate user movement around it. These pictures are also calibrated and localized in 3D and build the reference for online localization of the user on the site.

System overview

A generic and complete architecture and infrastructure for augmented reality for archeological sites has been proposed within ArcheoGuide. It can be divided into three basic sub-systems: the Site Information Server (SIS), the Mobile Units (MU), and the Network Infrastructure (refer to Figure 1).

The server is built on a high-end PC with sufficient storage space to implement a GIS and multimedia database. It is used as the central repository of the system for archiving the multimedia information used in the construction of augmented reality tours. The SIS communicates this information to the clients via a Wireless Local Area Network (WLAN). The clients (MUs) are portable devices carried by touring users in the archaeological site.

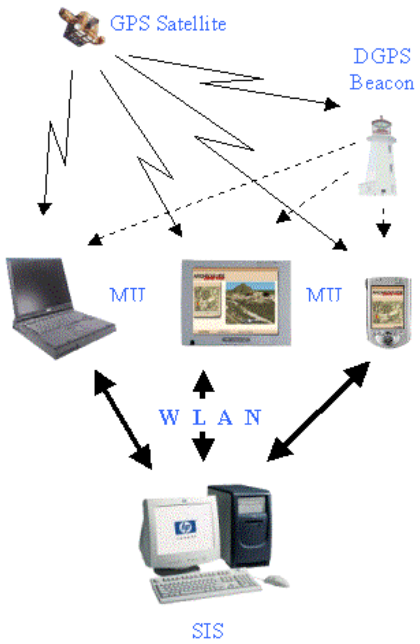


Figure 1. ArcheoGuide system architecture

The Mobile Unit (MU)

The MU system consists of a mobile computer unit, and a binocular display with a camera attached in front of it. Internally, a Global Positioning System (GPS) and orientation sensor keep track of the visitor's current location on the site to provide background information as well as the respective overlays. Additionally, an optical tracking system determines the exact position and view direction of the visitor in order to exactly place the virtual augmentations into the visitor's view (See Figure 2 (a)).

(a) (b)

Figure 2. (a) Mobile unit (MU) with augmented reality binocular (b) Augmented reality view

3 Tracking

The localization of the user on the site represents a key issue of augmented reality. GPS and compass provides only coarse positioning information, which are sufficient for delivering context-related media to the user. However, such sensors are to slow and not accurate enough to allow proper superposition of virtual views onto images of the surrounding in real-time as presented e.g. in figure 2 (b).

Optical systems provide the required data and many approaches have been developed , the main issue being reliability and accuracy (no jitter) and this in spite of the difficult conditions of outdoor environments.

2D tracking with reference images

In many cases, visitors look at the ruins at a given point of view. Thus, the virtual objects are only perceived from the same side, i.e. in 2D. The tracking can then be reduced to find in the current image video the right position and orientation of the virtual graphic objects in 2D.

For this reason, we propose the use of a set of calibrated images captured from the user's viewpoint as basic references for the optical tracking. The user's view, i.e. the current live video image, is compared to all reference images and a correlation score is computed. The best score is retained and the 2D transformation between the current video-image and the reference image is evaluated. This is passed to the rendering system for creating the augmented world.

The reliability of the algorithm represents the most important aspect of our choice of the registration technique. A lot of changes may appear between the live and reference images as new objects or visitors may be present in the scene and new lighting conditions, due for example to changing sun direction or clouds, may create new shadows or highlights in the images. We therefore prefer for these kind of applications to orient our choice to algorithms that exploit rather global than local properties of the image. Basically, this corresponds to algorithms working directly on the pixel intensity of the whole image or in the frequency space. We opted for a Fourier-based approach due to its robustness and fast calculation (Stricker 2001). This approach allows the recovery of only rotation, scaling, and translation, effectively limiting its application to pre-defined viewpoints and restricted camera motion.

3D Tracking

The 3D tracking approach is similar to the one presented above. However, the full 3D position and orientation of the user view is recovered. Here we rely on identification and tracking of local image features and compute back the rotation out of the motion of the points in the sub-sequent video images. As presented in (Zoellner 2008), we learn a representation of each image by using the Randomized Trees framework introduced by Lepetit et al. (Lepetit 2005). Following preprocessing steps are done for each reference image: first, a set of *P* robust salient points are selected from the image. This is done by rendering thousands of randomly warped versions of the image (by warping, we mean a realistic projective transformation of the image plane), and running a corner detector on every version. The found corners are back-projected in the original view, resulting in clusters of found corners growing with the number of tries. After a given number of tries (typically several thousand), the *P* clusters accumulating the most votes are kept as robust corners. For each of these robust corners, the 3D coordinates of the point is computed by making use of the camera pose *[Rn; tn]* and of the known 3D model of the object of interest. At runtime, the Randomized Trees select then the points of the live video image.

Once the tracking system has been initialized, we start to follow the position of the object by using feature tracking techniques. Feature tracking describes the process of following 2D points in subsequent images that correspond to the same physical point in the environment explored.

By 2D points, we mean salient points that can be found by classical corner detectors. These points are described by small image patches around them, and the set point and patch is called a feature. In order to follow 2D points, correspondences between points from two consecutive frames of the image stream must be built and maintained as long as possible. The longer features are being successfully tracked the more stable is the camera pose estimation.

After the KLT features have been successfully tracked in 2D, we end up again with 2D-3D correspondences. It is then possible to compute the camera position with usual minimization techniques as in the initialization step.

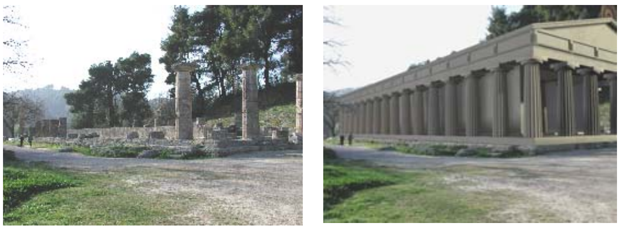
4 Visualization

The visualization of additional virtual content onto existing objects and scenes represents a sensitive element of the design of the augmented reality application. We present in the following two paragraphs two different approaches which show the actually large spectrum of possibilities offers by state of the art mobile graphics.

Seamless Integration: Realistic AR-Visualization

When 3D virtual reconstructions of buildings already exist, as for example for the site of Olympia, it is of great interest for the visitor to contemplate them directly on the site, and thus get a realistic impression of their size, color and global appearance in the environment.

This impression is unique and one realizes only then for example how tall and colorful were the Hera, Philipeon (see figure 1(b) and 3) or Zeus temples at Olympia.



(a) (b)

Figure 3: Real and augmented view of the Hera Temple (Olympia, Greece)

In order to get the highest quality, we pre-rendered the virtual objects using commercial tools and off-line processing. We create a set of high-quality 2D views (textures) which were then inserted at the right place in the image, using the data of the optical tracking.

Adapting the context: Reality Filters

In many cases, the appearance of the buildings is only documented with drawings (see figure 5) and any realistic representation is a pure interpretation. In this case we developed another techniques for presenting this information as they are in it real context, again in a seamless way, but this time, we adapt the context (i,e. the background image) to the virtual information.

The following approach is a solution to adapt the real view of a scene to blend well with black and white drawings in Augmented Reality applications by applying an inverted Sobel filter on the video stream. In contrast to the methods described in (Fischer 2005) we are concentrating on 2D textures instead of 3D models. This avoids the expensive and longsome process of producing high quality 3D models and the realistic integration into the video stream. The result is a reduced aesthetic defined by the original material and an affordable application. Furthermore it needs less computing power and thus saves battery life of the mobile device. This applies not only to cultural heritage sites with only historic material available for visualization but also for example for architectural visualizations of 2D plans. In cultural heritage this visual effect results in a visual time machine via Augmented Reality because the whole scene is rendered like a real time drawing. A drawing that is controlled by the user's movement and is displaying real buildings and people like a sketch (Figure 8).

Due to the reduced black and white style of the environment accentuations are much stronger than on a real colorful background. That enables a more efficient visualization of points of interest and drives a viewer's attention on them.

5 Applications

ArcheoGuide: Olympia, Greece

The user positions him/herself at a viewpoint and stares at the place of interest. In essence, the system treats him as an active pointing device, and the MU identifies his/her desire to view the augmentation of this specific monument. It transmits a request to the SIS, which mines the corresponding audio-visual data from its database and transmits it back to the MU. The reconstruction model or animation (see figure 4) is matched to the live video stream from the web camera, transformed accordingly, and rendered.

Figure 4: Visualization of virtual athletes on the stadium of Olympia

At the same time, the audio narration is synchronized to the visual presentation and are both presented to the user via the binocular and a pair of earphones. The image seen by the user is illustrated in figure 3(b). He/she can interrupt or alter the flow of information by moving away from the viewpoint or turning to another direction.

iTACITUS: Reggia Venaria Reale

Within iTACITUS, one of the field test areas of the Augmented Reality applications is Reggia Venaria Reale, an UNESCO World Heritage site in Italy close to Turin. The former residence of the Royal House is comparable to the French Versailles. The site has been restored over the last years and was opened to the public in fall 2007. While there are only a few 3D reconstructions of some buildings there is a vast archive of historic drawings and paintings. There are frontal drawings of facades of complete streets and the main palace's buildings.

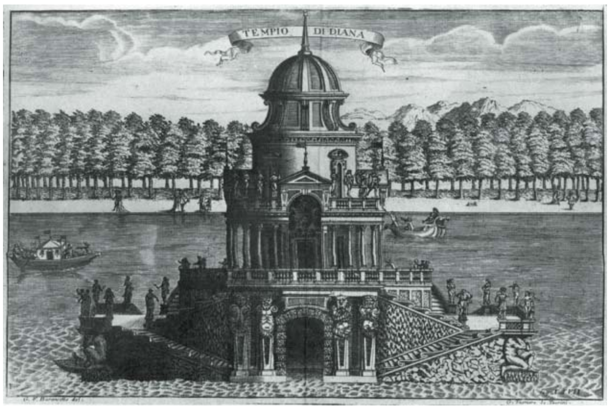


Figure 5: Drawing of Tempio di Diana at Reggia Venaria Reale

This vast archive is the basis of our visual time machine. With the Reality Filters we are able to create this effect through the display of a Sony UX Ultra Mobile PC. In order to match the source material we are rendering the video stream like a black and white drawing. Thus overlays of frontal drawings of buildings onto the real field of vision are seamlessly integrated. The whole scene looks like a real time ancient drawing. Even visitors, guides and guards are rendered in black and white. The application consists of three spots at the site. The first Tempio di Diana (Figure 5) was located at the end of a long path along a small artificial creek. It was surrounded by water and only accessible by boat. Only ruins of the fundament are left.

Figure 6: Tempio di Diana in a reality filtered scene of the gardens

At the horizon we are displaying the drawing of Tempio di Diana (Figure 5) on top of these ruins. Standing at a viewing platform visitors can look around and watch through the display of a Sony UX via video see through (Figure 6). As soon as they are looking at the direction of the temple the video turns black and white and the drawing of the temple is superimposed. The Fountain d'Hercule in front of the viewing platform will be superimposed the same way.

Figure 7: Historian watching Palazzo di Diana on Sony UX UMPC

Reggia Venaria Reale's Palazzo di Diana's architecture was modified several times over the years. Each state of the buildings was documented on drawings. We are superimposing these drawings of the modifications of the architects Castellamonte (1674) and Garove (1700-1713) on the facade of the main building. Visitors standing in the large courtyard are seeing the fountain and the current restoration of the palazzo rendered like a drawing. While listening to the audio guide's story about the palazzo the buildings appearance is switching through the centuries while seamlessly integrated into the environment (Figure 8). Finally, the whole scene looks like a real time ancient drawing (Figure 7).

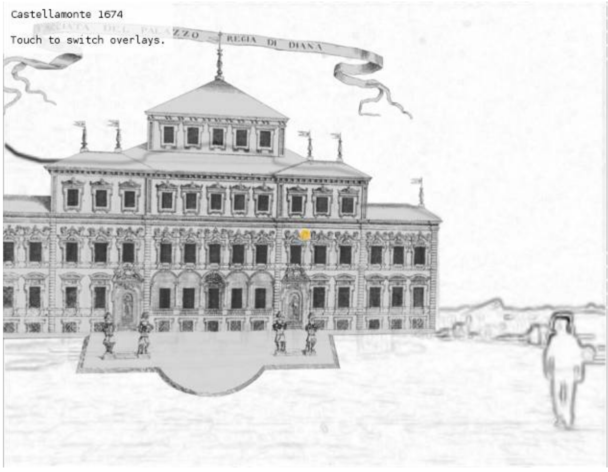


Figure 8: Historical drawing of Palazzo di Diana and reality filtered tourist

6 Conclusion

In this paper, we have presented our Augmented Reality concept for cultural heritage sites. Its main features are the possibility to show 2D content in very different manners, either as high realistic or as superimposed on lookalike (modified) reality. The adaptation of the reality is made possible by the use of interchangeable filters that enable a better integration of the ancient content in the reality. We showed results of our application in the area of cultural heritage, where the system runs on an Ultra Mobile PC (Sony Vaio UX) with 15 frames/sec.

In a future version of our system, we will investigate the automatic detection of the right filter for the best integration of the content in the real image.

Acknowledgements

The work discussed in this article was supported by the European Union IST framework (IST 1999-11306) project ArcheoGuide and is continued in the current project iTACITUS (IST 2.5.10 – 034520).

References

LEPETIT V., LAGGER P., FUA P.: *Randomized trees for real-time keypoint recognition*. Conference on Computer Vision and Pattern Recognition 2 (2005), 775–781.

VLAHAKIS, Vassilios; IOaNNIDIS, Nikos; KARIGIANNIS, John; TSOTROS, Manolis; GOUNARIS, Michael; STRICKER, Didier; GLEUE, Tim; DAEHNE, Patrick; ALMEIDA, Luis *Archeoguide: Challenges and Solutions of a Personalized Augmented Reality Guide for archaeological sites* IEEE Computer Graphics and Applications 22 (2002), 5, pp. 52-60

STRICKER, Didier *Tracking with Reference Images: A Real-Time and Markerless Tracking Solution for Out-Door Augmented Reality Applications* International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST), Glyfada, Greece, 2001, pp. 91-96.

ZOELLNER, Michael, PAGANI Alain, STRICKER, Didier: "*Reality filtering: A visual time machine in augmented reality*", in VAST 2008. Proceedings (December).

ITACITUS, 2008. "*Intelligent tourism and cultural information through ubiquitous services*." http://www.itacitus.org.

FISCHER, J. 2005. *Stylized augmented reality for improved immersion*. Proceedings of IEEE Virtual Reality, 195–202.