

PERSONALIZED AUGMENTED REALITY TOURING OF ARCHAEOLOGICAL SITES WITH WEARABLE AND MOBILE COMPUTERS

Vassilios Vlahakis[†], John Karigiannis[†], Manolis Tsotros[†], Nikolaos Ioannidis[†], Didier Stricker^{*},

[†]Intracom S.A., ^{*}Fraunhofer Institut für Graphische Datenverarbeitung
E-mail: vvla@intracom.gr

Abstract: This paper presents ARCHEOGUIDE, a novel system offering augmented reality tours in archaeological sites. The system is based on wearable and mobile computers, networking technology and real-time computer graphics and 3D animation and visualization techniques. The user can participate tours adapted to his profile and automatically receive information based on his position and orientation as calculated by a hybrid technique making use of GPS, compass and image-based tracking. The user can interact with the device via multi-modal interaction techniques and request navigation and other information. ARCHEOGUIDE has been tested at the archaeological site of Olympia in Greece.

KEYWORDS

Wearable, Augmented Reality, Image Tracking, Personalized Tours.

PROJECT URL

<http://Archeoguide.intranet.gr>.

1. INTRODUCTION

ARCHEOGUIDE is a research project pursued by a consortium of European companies and institutions under the support of the European Union IST framework. It stands for Augmented Reality-based Cultural Heritage On-site GUIDE.

Recent years have seen attempts to introduce mobile computing and image processing in archaeology and cultural heritage. The aim of these attempts is to provide visitors with a more realistic and informative way to access their cultural heritage.

From the early audiotape and CD guides on hire at many museums worldwide, a lot has changed with the introduction of portable computers with enough processing power to handle audio, image and video information. Advances in computer graphics allowed 3D modeling of artifacts and monuments and the creation of reconstruction and visualization tools like 3D-MURALE [1]. Projects like TOSCA, MUSE and SHAPE [1] provide their views on the use of pen-tablets and palmtops for guiding and informing tourists on archaeological sites.

Similarly, “cave” installations like CVR and IHW give an immersive experience in the exploration of ancient sites. Augmented Reality kiosks have been implemented and installed both at outdoor sites and museums like Ename [1].

However, all these applications provide either a “pre-fabricated” solution allowing no user interaction, or require excessive post-processing and cannot be used in real-time touring. This gap is filled by the ARCHEOGUIDE project [2], [3], which is described in this paper. It provides personalized AR tours by means of mobile computers, accurate position and orientation tracking algorithms and image processing techniques.

The paper focuses on the system architecture and the description of the mobile units. It presents example of available features and briefly presents the optical tracking algorithm. Finally, conclusions are drawn and references point to sources of further information.

2. ARCHEOGUIDE ARCHITECTURE

The ARCHEOGUIDE system [3] is implemented as a set of mobile devices offering Augmented Reality [4] tours in cultural heritage sites. It is built on a client-server architecture, where a local server acts as the central repository of cultural heritage information and is connected to the mobile units via a Wireless Local Area Network (WLAN) as shown in Figure 1.

The system allows the user to participate personalized tours inside the cultural heritage site and provides him audio-visual information on the ruins of the site (e.g. monument reconstructions, audio narration, navigation information, avatar animations, access to related information). This material is stored on the server and is downloaded to the mobile computers according to the user’s profile, position and orientation.

The data transfer is achieved over the Wireless LAN (WLAN) installed in the periphery of the archaeological site. The WLAN is built on the IEEE 802.11 WLAN and operates at 11Mbps per second, resulting in reasonable response times of the order of a few seconds. The device is equipped with a Lucent Orinoco PCMCIA wireless LAN adaptor card for connecting to the network.

The network is installed in the periphery of the archaeological site for eliminating physical damage to the

sensitive area and for minimizing visual disturbance. Its topology was based on the morphology of the site and the areas accessible by visitors, resulting in a network of three access points (AP). Two wired and a third one located at a remote corner and connected to the network hub via a point-to-point wireless link. Due to the unavailability of power at its location, it runs on solar panels, which charge two 12-volt rechargeable batteries. A DC-AC converter is added between the batteries and the AP for providing the necessary 220V AC power.

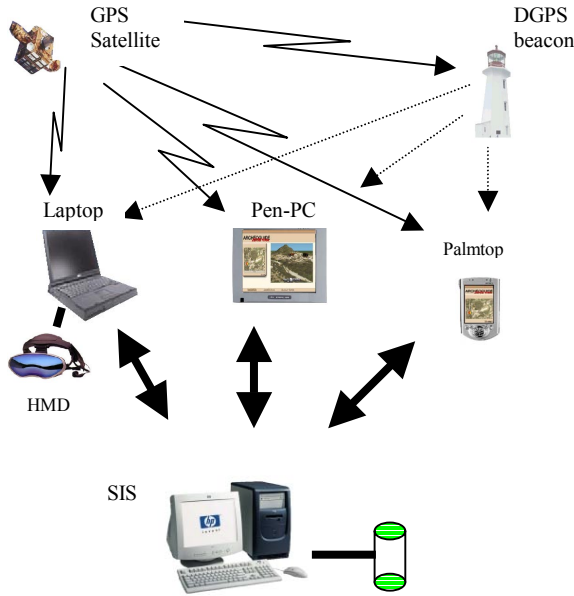


Figure 1: System Architecture.

3. THE MOBILE DEVICES

The mobile devices are designed to provide user-friendly operation to all possible users irrespective of their computer skills. They come in three versions, providing different levels of processing power and capabilities.

All devices provide personalized information according to the user's profile entered prior to the beginning of the tour and support position tracking.

ARCHEOGUIDE uses a hybrid technique for the calculation of the user's position and orientation. An initial estimate of the user's position is made via a Global Positioning System (GPS) receiver. This calculation is based on the signals received from a constellation of GPS satellites, orbiting at geo-stationary orbits. A minimum of four (up to a maximum of twelve) satellite signals are used to triangulate the user's position with an accuracy of the order of 2-4 meters. This accuracy depends on the number of satellites in the line-of-sight of the receiver and the corresponding signal strength.

In regions where signal strength is not sufficient, a Differential version of the GPS (DGPS) system can be employed. In this case, a DGPS beacon is installed at a pre-selected position whose geographical coordinates are known with very high accuracy and which is at close range to the mobile units. The beacon transmits a *correction* signal to the receivers on the mobile devices, which can provide accuracies of less than 1 meter to the calculations.

The user's orientation is calculated with a digital compass. The selected device is based on a magnetic transducer and can record angular changes with an accuracy of 0.5° when held flat. Its accuracy goes down if held at an angle or if in the presence of electromagnetic interference and can reach 5°.

Higher accuracy is needed in real-time AR presentations. This is achieved by refining the initial estimates by the GPS and compass with an optical tracking algorithm described later in this paper.

ARCHEOGUIDE introduces multi-modal interaction techniques in uncontrolled outdoor environments and eliminates the use of traditional, in computer terms, and cumbersome devices like keyboards and mice. It treats its user as an active pointing device and responds by automatically presenting him with the audio-visual information corresponding to the object he is staring at.

Nevertheless, the user is always in control of the operation and the flow of information. He can intervene and alter it simply by moving or turning his head, or if he prefers with a special pen or gamepad.

The following sections present the mobile units.

3.1. Mobile AR devices

The top-of-the-range device is built on a laptop computer. The initial prototype features an 800 MHz Pentium III processor with 256 MB RAM, 20GB hard disk, and an nVidia graphics card. It is integrated with a Garmin 35-HVS GPS receiver and a Precision Navigation TCM2 digital compass for position and orientation estimation (Figure 2). The laptop is carried in a small backpack with the GPS receiver attached upright on one strap for better signal reception. A pair of Sony Glasstron AR glasses are integrated with the compass and a pair of earphones, all mounted on a lightweight helmet allowing easy carrying and fitting all heads. The helmet also features a Philips PCVC 690K USB web camera with a maximum video resolution of 640X480 at 30 frames per second for capturing the user's view.

This design allows the user to see augmented views in 360° from pre-defined viewpoints. The reason for these viewpoints being the limitation of the processing requirements since it facilitates the optical tracking (described later) by the selection of the appropriate reference image, therefore allowing real-time operation.



Figure 2: Mobile AR Device



Figure 3: AR view of the reconstructed temple of Zeus in Ancient Olympia, Greece.

This design is currently under modification, with the replacement of the laptop with a more compact and easy to carry wearable computer of similar processing power. A second step is the removal of the helmet and the further integration of its sensors in a single device similar to the AR glasses currently used but with a mounted mini camera and compass.

Up to this point, a modified design has been produced where the glasses and helmet have been replaced by an AR binocular unit (nVision VB-30). The device features no compass as its proposed use (pick up and place in front of your eyes when looking from a predefined point and direction) do not necessitate it. However, this limits the presentation of augmented views to one per viewpoint as otherwise it is not easy to differentiate between different directions in real time.

The final product is expected to be produced when new compact and lightweight hardware has been tested.

In its current state, the laptop-based device offers its user real-time Augmented Reality in uncontrolled outdoor environments. The user can automatically:

- see realistic reconstructions of ruined monuments based on archaeological information (Figure 3)
- listen to audio narration giving historical and other information in the user's preferred language
- view ancient life revival in its original settings (Figure 4)

based on his position and the particular monument he is staring at. The life animation features avatar humans modeled with the H-ANIM standard and gives a new meaning to archaeological site visiting as he can better understand the use of particular monuments in the ancient time, or understand ancient customs and culture. A typical example includes the revival of sports disciplines of the ancient Olympic Games in the very place where they originally took place, the ancient stadium. This allows ARCHEOGUIDE users to see avatar athletes compete among real visitors of the site.



Figure 4: Avatar athletes competing in the stadium of ancient Olympia.

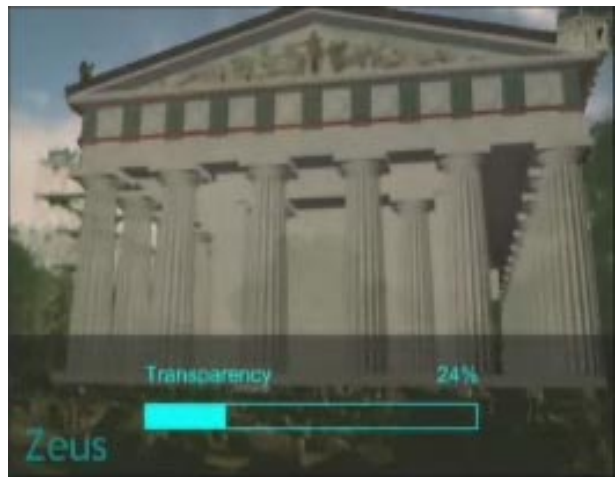


Figure 5: Menu selection on HMD.

All this information is automatically presented to the system user. He can also request additional information at any time during his tour with the use of the optional gamepad or the mouse buttons on the binocular display. Both interaction means result in the projection of a scrollable menu of actions (Figure 5) on the two displays which the user can navigate and select the appropriate choice:

- receive navigation information in the form of site maps indicating the major monuments, the path he follows and his current position and orientation updated in real time (Figure 6)
- related information (e.g. photographs of artifacts found in the reconstructed monument and currently stored in the museum or architectural details and other information)
- visualization and manipulation of 3D objects (e.g. statues)
- Audio conveying additional information on the visualized material
- Change of language
- Change animation content at a viewpoint
- Change transparency of reconstruction model
- Enter or change his profile



Figure 6: Navigation information

Figure 5 was shot through the AR display and the image is dimmed for as long as the menu is displayed to make it more visible under all lighting conditions. This opposes the very high quality image of Figure 3, which has been pre-computed so as to incorporate obstructions from trees etc. This is the image displayed to the other ARCHEOGUIDE devices (see below). The same image, however, without the obstructions, is presented through the AR visualization devices.

This information is tailored to the user's preferences. Prior to the beginning of his tour, he can enter his profile and the device will automatically adapt the tour and its contents. For example an archaeologist will receive more

detailed information than an ordinary visitor, or a sports fan will get more sports-related information than general facts. This profile can also be altered during the augmented tour.

The mobile devices automatically download the relevant multimedia information that comprises the tour. This is done based on the user's profile and in response to his current position and orientation. The data transfer is achieved over the Wireless LAN (WLAN).



Figure 7: Pen-tablet User.

3.2. Pen-tablet device

A more compact and lightweight version of the mobile device is built on a Fujitsu Stylistic 3500R pen-tablet computer, equipped with a 500 MHz Celeron Processor, 256 MB SDRAM, 15 GB hard disk, and 10.4" reflective SVGA TFT monitor.

The device is integrated with a Teletype PCMCIA GPS receiver card and a Precision Navigation TCM2 digital compass. It features a compact and easy to carry solution aimed for use as the electronic equivalent of a paper guidebook. The user holds it with his hands and can write or select items on the device's pressure-sensitive screen with a special pen (Figure 7). The screen displays a GUI based on Internet Explorer and features clickable bookmarks like those used in filofaxes (Figure 8). They enable the user to select one of the following:

- Navigation through a 3D model of the archaeological site augmented with reconstructed monuments and avatar animations. This model is navigated automatically as the user moves or turns and is available at every point on-site
- Augmented panoramic views available at pre-selected viewpoints. 360° natural views augmented with reconstruction models and animations. Synchronized with the user's natural view as he turns around the viewpoint

- Very high-quality images augmented with reconstructed monuments, available at pre-selected viewpoints
- Related topics (e.g. artifacts, avatar animations)
- Audio narration in the user's preferred language
- Navigation aid in the form of site map or plan, indicating main monuments, user's trace and current position and orientation
- Preview facility of available multimedia material at viewpoints indicated on the site map

The user can select the type of presentation (listed above) by clicking on the bookmark tabs with the special pen. The information corresponding to his profile (entered prior to the beginning of the tour), position and orientation is automatically presented. These multimedia objects are downloaded from the central server over the wireless network.

The software supporting these applications is based on Java applets and servlet components for maximum portability.



Figure 8: Pen-tablet user interface.

3.2. Palmtop device

The lightest version of the ARCHEOGUIDE mobile unit is built on a Compaq iPAQ PDA with 64 MB RAM and a touch sensitive screen. The device is equipped with a dual PCMCIA card expansion jacket where a Teletype GPS receiver and a WLAN adapter are plugged. It can be carried in one hand like a small notepad and operated with a special pen (Figure 9). The application is written in Java and runs under with Pocket Internet Explorer allowing the user to:

- view 2D images augmented with reconstructed models. They are presented automatically based on the users position and profile
- listen to audio narration in his language of choice
- read historical and other information in his language of choice

- navigate inside a 3D model of the site featuring reconstructed monuments. The navigation is performed either automatically based on the user's position as given by the GPS receiver, or with a special pen operating on the device's screen
- augmented panoramic views at pre-selected viewpoints. Scrolled with the special pen.
- get navigation information in the form of site maps marked with the main monument, pre-selected viewpoints and corresponding viewing angles, and user's trail and position in real time
- get navigation information in the form of a 3D site map augmented with clickable models of the main monuments

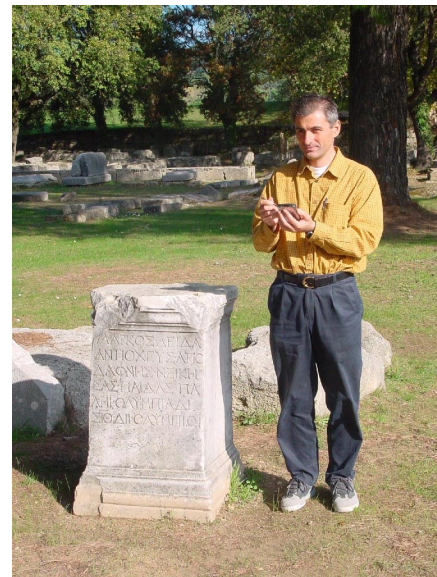


Figure 9: PDA user.

4. THE OPTICAL TRACKING ALGORITHM

This accuracy in the position and orientation tracking, provided by the GPS and the digital compass, is sufficient for navigation and controlling the flow of audiovisual information. However, ARCHEOGUIDE needs better performance for the calculation of augmented scenes.

In its scenario, the system user should be able to wear a special HMD in the form of a pair of optical see-through AR glasses, or video see-through AR binoculars and visualize reconstructed monument ruins etc. placed at their exact position in the natural scene. Therefore, even small misalignments will result in augmented images where the reconstruction models will fly in the air and create an unrealistic effect.

For this reason we have implemented an optical (image) tracking technique that uses the initial estimate of the user's position and orientation to match the user's view (captured by a web camera worn during the site visit) with predefined calibrated images or 3D models.

The technique is based on Fourier analysis to calculate parameters such as translation, rotation, and scaling between the live video and the reference views. These parameters are then used to transform reconstruction models of the ruined temples for rendering them on the user's natural view on the HMD or binocular devices. The same technique is also applied in rendering avatar animations on the natural scene.

The following sections give a brief introduction to the optical tracking algorithm. The interested reader may refer to [5], [6], [8] and [9] for more detailed information.

The algorithm caters for inaccuracies from the GPS and the compass and results in aligned graphics and natural images within 2-3 pixels and occlusions of up to 60%. These results may vary if large variations in the shading occur, therefore, completely altering the image content as it is perceived from the algorithm. In these cases tracking still performs well but the misalignment may increase by a factor of 5-10 approximately. The algorithm allows the user to be situated within a circle of radius of approximately 2 meters whose center is indicated in the site by a small bronze plate. The user can see it for verifying the position of the viewpoint as indicated on a digital map of the site, which is marked with his position and orientation in real time and which he can display on the HMD. The algorithm can support any number of viewpoints. In its current installation at ancient Olympia, a total of 10 viewpoints are used. These can be increased according to the requirements of the curator and the response of the users.

4.1. Translation

Let f_1 and f_2 be two images differing only in a 2D translation $t(t_x, t_y)$. The images are related by:

$$f_2(x, y) = f_1(x - t_x, y - t_y) \quad (1)$$

The Fourier shift-theorem gives their relation in the frequency domain:

$$F_2(\xi, \eta) = e^{-j2\pi(\xi t_x + \eta t_y)} F_1(\xi, \eta) \quad (2)$$

where F_2 and F_1 are two arrays of complex numbers.

The power spectrum of the Fourier transform F_1 does not change if the function f_1 is shifted by an amount of (t_x, t_y) . The power spectrum is translation invariant.

The translation vector (t_x, t_y) can be easily isolated by computing the cross-power spectrum of F_1 and F_2 :

$$\frac{F_1(\xi, \eta) F_2^*(\xi, \eta)}{|F_1(\xi, \eta) F_2^*(\xi, \eta)|} = e^{j2\pi(\xi t_x + \eta t_y)} \quad (3)$$

where $F_2^*(\xi, \eta)$ is the conjugate-complex of $F_2(\xi, \eta)$.

Recalling that the inverse Fourier transform (IFT) of an exponential function is a Dirac function, we can estimate

the maximum of the IFT of equation (3) which provides the image shift (t_x, t_y) .

4.2. Rotation

If image $f_1(x, y)$ is transformed into image $f_2(x, y)$ by a translation $t(t_x, t_y)$ and a rotation with angle ϕ , then the relation between f_1 and f_2 is defined as:

$$f_2(x, y) = f_1(x \cos \phi_0 + y \sin \phi_0 - t_x, -x \sin \phi_0 + y \cos \phi_0 - t_y) \quad (4)$$

which in the frequency domain corresponds to:

$$F_2(\xi, \eta) = e^{-j2\pi(\xi t_x + \eta t_y)} F_1(\xi \cos \phi_0 + \eta \sin \phi_0, -\xi \sin \phi_0 + \eta \cos \phi_0) \quad (5)$$

A rotation in the spatial domain generates a similar rotation in the frequency domain. The magnitude spectra M_1 and M_2 of F_1 and F_2 are related as follows:

$$M_2(\xi, \eta) = M_1(\xi \cos \phi_0 + \eta \sin \phi_0, -\xi \sin \phi_0 + \eta \cos \phi_0) \quad (6)$$

Alternatively, point $P(\xi, \eta)$ in the magnitude spectra is represented using polar coordinates by a point $P(r, \phi)$. Both magnitude spectra in polar coordinates are defined:

$$M_2(r, \phi) = M_1(r, \phi - \phi_0) \quad (7)$$

A rotation is represented as a translation of value ϕ_0 in the polar-transformed magnitude images and can be found by phase-correlation, and thus the rotation angle.

4.3. Scaling

Let $f_2(x, y)$ be the scaled image of the image $f_1(x, y)$ with the factors (a, b) , so that:

$$f_2(x, y) = f_1(ax, by) \quad (8)$$

Then, the Fourier spectra of both images are related by:

$$F_2(\xi, \eta) = \frac{1}{|ab|} F_1\left(\frac{\xi}{a}, \frac{\eta}{b}\right) \quad (9)$$

If the horizontal and vertical axes of the frequency domain are scaled in a logarithmic way, the scaling parameters can be found as a translation in the frequency domain. This can be written as:

$$F_2(\log \xi, \log \eta) = \frac{1}{|ab|} F_1(\log \xi - \log a, \log \eta - \log b) \quad (10)$$

Using phase-correlation, the translation $(\log a, \log b)$ can be found and thus the scaling factor (a, b) .

In conclusion, the power spectra are invariant for translation but variant for scaling and rotation.

4.4. Rotation and Scale

In most cases, the horizontal and the vertical scale factors are equal. A rotated and scaled copy of one image can be found by a log-polar transformation of the magnitude images (see equation 7).

$$M_2(\log r, \phi) = M_1(\log r - \log a, \phi - \phi_0) \quad (11)$$

4.5. 3D Optical Tracking

3D optical tracking can be used to enable free touring and eliminate pre-defined viewpoints by exploiting standard VRLM features. The 2D warping described by the matrix H , is related to the 3D rotation R as [7]:

$$H = ARA^{-1} \quad (13)$$

where A is the matrix of the intrinsic camera parameters:

$$A = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & & 1 \end{pmatrix} \quad (14)$$

f_x : focal length along the X-Axis of the camera

f_y : focal length of the Y-Axis of the camera

c_x, c_y : coordinate of the principal point

The Fourier-based registration algorithm computes H . Knowing A the 3D rotation can be easily deduced:

$$R = A^{-1}HA \quad (15)$$

The position of the 3D viewpoint is known and defined as the translation vector T (t_x, t_y, t_z).

Finally, the orientation R and the position T are passed to the rendering system.

8. CONCLUSIONS

The ARCHEOGUIDE project has tackled the problem of unsupervised, markerless, real-time position and orientation tracking in outdoor environments, and produced a prototype offering personalized augmented reality tours in archaeological sites. The prototype is based on mobile-wearable computers and advanced visualization and multi-modal interaction techniques. The system was warmly accepted by both ordinary visitors and experts in archaeology and computing. The initial installation at the archaeological site of Olympia in Greece proved the effectiveness and user-friendly operation of the system. The AR units with the special AR visualization devices were judged as the most realistic and promising for future use. However, many people preferred the pen-tablet, as it is lighter and more compact. Its user interface appearance facilitated its use, as it resembled the paper agendas they were accustomed to use. In contrast, these users felt less comfortable with seeing the world through the AR glasses

and binoculars even though they agreed that they offered the most realistic representation of the ancient world. Women and children were interested in using the PDA device, as it was very handy to carry with them. Overall, the AR devices were judged as the users' favorite solutions provided they become more compact and lightweight. The other two devices should probably target visitors who participate in group tours and those who feel less at ease with computers.

Currently, work is in progress to come up with a final prototype based on more compact, wearable hardware, offering more freedom to its user and incorporating full 3D tracking.

Upon the successful upgrading of the existing hardware, the consortium aims at marketing and installing the system at several sites across Europe and beyond.

9. ACKNOWLEDGMENTS

We would like to acknowledge the EU who selected ARCHEOGUIDE to be financially supported within the IST program (contract number IST-1999-11306).

10. ARCHEOGUIDE CONSORTIUM

The ARCHEOGUIDE consortium is made up of INTRACOM (Greece), IGD (Germany), ZGDV (Germany), CCG (Portugal), A&C2000 (Italy), Post Reality (Greece), Hellenic Ministry of Culture (Greece).

REFERENCES

- [1] Proc. Int. Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST01), Glyfada, Nr Athens, Greece, 28-30 Nov. 2001.
- [2] T.Gleue, P.Daehne, "Design and Implementation of a Mobile Device for Outdoor Augmented Reality in the ARCHEOGUIDE Project", Proc. VAST01, Glyfada, Greece, 28-30 Nov. 2001.
- [3] V. Vlahakis, N.Ioannidis, J. Karigiannis, et al., "ARCHEOGUIDE: Challenges and Solutions of a Personalized Augmented Reality Guide for Archaeological sites", IEEE CG&A Magazine, Special Issue "Computer Graphics in Art, History and Archaeology", late 2002.
- [4] R. Azuma, "A Survey of Augmented Reality", Proc. SIGGRAPH 1995, Developing Advanced VR Applications Course, ACM Press, 1995.
- [5] R. Holloway, "Registration Errors in Augmented Reality Systems", Ph.D. Thesis, Univ. North Carolina at Chapel Hill, 1995.
- [6] A. State, G. Hirota, D.T. Chen, et al., "Superior Augmented Reality Registration by Integrating Landmark Tracking and Magnetic Tracking", Proc. SIGGRAPH 1996, ACM Press, 1996.
- [7] Davis, J., "Mosaics of Scenes with Moving Objects", IEEE Comp. Soc. Conf. on Computer Vision & Pattern Recognition (CVPR98), 1998

- [8] L.G. Brown, "A Survey of Image Registration Techniques", *ACM Computing Surveys*, 24(4) pages 325-376, Dec. 1992
- [9] F. Seibert, "Augmented Reality by Using Uncalibrated Optical Tracking", *Computers and*

Graphics 23 (1999), Special Issue on Augmented Reality, pp 801-804.