ENGINEERING AN AUGMENTED REALITY TOUR GUIDE

Panagiotis Ritsos, David J. Johnston, Christine Clark and Adrian F. Clark

University of Essex, UK

Abstract

This paper describes a mobile augmented reality system intended for *in situ* reconstructions of archæological sites. The evolution of the system from proof of concept to something approaching a satisfactory ergonomic design is described, as are the various approaches to achieving real-time rendering performance from the accompanying software. Finally, some comments are made concerning the accuracy of such systems.

1 Introduction

The UK has an embarrassment of rich archæological remains, dating back several thousand years. Sadly, the proportion of sites where there are substantial remains above ground is small. Indeed, the curators of archæological sites are faced by a dilemma: they wish to attract interested visitors but need to do so without disturbing any remaining archæology.

Visualizing the appearance of archæological sites is conventionally achieved through paintings and physical scale models. These are increasingly being replaced by computer graphic "walk-throughs." Both of these are unsatisfactory because it is difficult for a visitor to relate the appearance of the model (physical or virtual) to the archæological remains. A better solution is for the visitor to be able to visualize the appearance of the model as he or she walks around the site. This is normally done using labelled perspective views on plaques, as physical reconstructions are deprecated — but even plaques can be obtrusive; and many people still have trouble visualizing how the site would have looked in antiquity.

An exciting solution is to this dilemma is to exploit the simultaneous reduction in size and increase in performance of computer technology, and use virtual reality technology to visualize the ancient buildings as the visitor explores the site. Several research groups around the world have constructed proof-of-concept demonstrators for mobile augmented reality (see Section 2), though not necessarily with archæological reconstructions in mind. These combine a form of wearable computer with a head-mounted display and position-sensing technology to superimpose the reconstruction onto the visitor's own view of the surroundings. However, moving from proof of concept to something that could be used by the general public is a far from trivial task. Nevertheless, it is something we are attempting, even though we are pushing at hardware, software and system limits at every step.

The remainder of this paper is organized as follows. The next section gives an overview of previous work on wearable tour guides. Section 3 briefly introduces the archæological site that has formed the focus of the work, highlighting its important properties. The considerations that influence our design are described in Section 4. Section 5 then details the evolution of the hardware, and Section 6 the accompanying software. Section 7 describes the system in use and considers the sources of positioning errors and their severity. It also draws conclusions and outlines further work.

2 Previous and Related Work

One of the first mobile tour guide systems was Lancaster University's GUIDE [1]. GUIDE provides city visitors with location-related information using a tablet PC equipped with a radio-LAN (802.11b) interface. GUIDE does not attempt to produce 3D reconstructions of the surroundings but instead provides text and images related to the user's position. A number of base stations around the city provide each user with information relevant to their position. The authors argue that GPS, the *global positioning system* which they used in early versions of GUIDE, does not provide any advantages in such an environment when compared to network-based location mechanisms. In particular, GPS requires at least four satellites to be in view in order to obtain a moderately accurate position fix, yet this is rarely possible in the "urban canyons" formed by tall buildings.

One of the earliest research efforts in the field of mobile augmented reality is Columbia University's *Touring Machine* [2], which has evolved into the *Situated Documentaries* system [3]. This employs a GPS-equipped wearable computer to provide hypermedia presentations that are integrated with the actual outdoor locations to which they pertain. The prototype uses a tracked, see-through head-mounted display (HMD) and a hand-held pen computer to present 3D graphics, imagery and sound superimposed on the real world. A backpack wearable computer equipped with a GPS receiver provides the location information. The user roams within the university campus and is able to see information in the form of virtual tags, indicators and point-of-interest flags through the HMD.

A similar system, developed at Carnegie-Mellon University, is known as *Smart Site* [4]. This utilizes a wearable computer *and* a laptop to provide an intelligent tourist assistant system. The system architecture caters for multi-modal input and output. Location information is obtained from a GPS unit, while a camera provides visual input capabilities. A microphone and headphones provide audio input and output. Speech and gesture inputs may be used. The authors argue that for a tourist system it is important to use a number of input modalities to accommodate different scenarios. A user can roam in a site, derive location information from the GPS unit but also request for further details by speech or by gesture.

One of the most significant systems in this area is Tinmith-Metro, an interactive augmented reality 3D constructive solid geometry modeller [5]. This system is based around a wearable (backpack) computer with machine vision capabilities from a USB camera. User input is principally via a set of pinch gloves, allowing the user to execute commands using menus linked to finger gestures; indeed, the authors argue that desktop user interfaces should be avoided and replaced with speech recognition, camera input and hand-gesture tracking. Position is determined using (differential) GPS, while a digital compass measures head orientation. The system is capable of generating 3-D models of the external surfaces of buildings as the user roams, and can place pre-fabricated 3D objects within a scene. Objects are created by fitting infinite planes to surfaces and using their intersections to define buildings.

Archeoguide [6] is an EU-funded programme that aims to produce a virtual tourist guide. The system uses a centralized site information server, a mobile unit with GPS location capability, and a wireless LAN for communication. The site server includes an image database and a content creation mechanism. All rendering is performed on the information server using a VRML-based VR toolkit developed within the project, and distributed to the mobile stations over the wireless LAN.

3 The Archæological Site

Most readers will remember learning that armies under Julius Cæsar invaded Britain unsuccessfully in 55 and 54 BC. After Cæsar's departure, the capital was established at *Camulodunum* ("fortress of the war god"), to the south-west of modern Colchester at a place now known as the Gosbecks Archæological Park. Camulodunum reached its zenith under the reign of Cunobelin (Shakespeare's Cymbeline). Shortly after Cunobelin's death, Claudius decided to conquer Britain and, in 43 AD, his army landed in Kent and fought its way northwards, stopping just short of Camulodunum. In 49 AD, the first *colonia* (colony) of the new province of Britannia was established next to Camulodunum, on the site of modern Colchester. The Gosbecks area was also developed further with a 5,000-seat theatre, the largest in Britain, and a temple complex.

Following the end of Roman rule circa 450 AD, future development was centred within the city walls, in modern Colchester. The Gosbecks site proved an excellent source of building materials for local inhabitants — so much so that all that remains today are foundations, and these are below ground level. Indeed, the only way that a visitor to Gosbecks is aware of the scale and layout of the former buildings is by white lines painted on the ground and accompanying signage.

Our work to date has concentrated on reconstructing the Roman temple complex at the Gosbecks site; we shall extend it to encompass the theatre in the future. Gosbecks has characteristics that reduce some problems while introducing others. Firstly, it is situated in a reasonably flat area (modulo rabbit holes) at the edge of farmland near the brow of a low hill; the nearest buildings are over 500 m away. Hence, almost the entire hemisphere of sky is visible. There are no trees, pylons, or post-Roman remains on the site to interfere with our reconstructions. There are also no visible foundations that the 3-D reconstructions have to abut. On the other hand, much of the site has not yet been excavated, particularly the area surrounding the likely grave of Cunobelin; so English Heritage, the owner, is understandably protective of the undisturbed archæology and does not permit the ground to be disturbed, such as by erecting radio masts or burying cables.

4 Design Considerations

Our aim is to engineer an augmented reality tour guide system that could be used by the general public. This means it must be light, robust, accurate, have few external components and cables, exhibit long battery life — and yet be cheap to construct. This inevitably involves a number of trade-offs. Before considering them, however, the design of the overall system must be established.

System architecture The most obvious solution is to pre-render images from a large number of positions and orientations and simply display them as required, an adaption of the ideas underlying image-based rendering; but we consider this to be an inadequate solution for a site of the size of Gosbecks. Centralized rendering, as employed by Archeoguide, will not scale well to significant numbers of participants because of the need to transfer images over a wireless network to the body-mounted display system. Hence, our own system [7] is more in the spirit of the *Touring Machine*: each person has a dedicated system which carries out all the required computation. Rendering of a locally-stored 3D model is performed by the wearable computer and displayed on a head-mounted display with built-in orientation sensing. Location is determined using GPS. Certain locations are able to trigger the playing of a locally-stored audio clip. User interaction is restricted to a pair of buttons, one to replay audio clips and another to reset the unit; too many controls will lead to confusion. (Our prototypes all use more capable input devices.) Finally, each unit has an 802.11 interface, for use with differential GPS and for security purposes.

Accuracy Requirements Our ultimate aim is that a person should perceive the 3D reconstruction as being "in the right place," corresponding to an accuracy requirement for position of the order of 1 cm, clearly not possible with the kinds of systems currently being explored. A more achievable aim is that doorways in the model should be sufficiently positionally stable that the wearer can negotiate physical doorways without difficulty — say 0.2 m. Accuracy requirements for orientation are not, we believe, too stringent: 5° error seems to be acceptable to most users.

Determining Position and Orientation In the laboratory, the most common way of estimating position and orientation is via magnetic trackers; but their regions of operation is a few metres at best, making them unsuitable for this application. The ideal solution for an archæological site would probably be to mount radio beacons on masts and use triangulation; but we cannot do that on the Gosbecks site for the reason expounded above.

If triangulation using fixed beacons is not permitted, the obvious fall-back is to use the well-known Global Positioning System (GPS) — *i.e.,* triangulation with moving, remote satellite beacons. Single-receiver GPS, even after the removal in May 2001 of the "selective availability" random error superimposed on signals, does not provide sufficient accuracy for this application; instead, *differential* GPS is required. This involves the use of two receivers, a stationary *reference point* sited at a known position, and a roaming one situated on the wearable computer. As long as the two receivers are within a few kilometres of each other, the GPS signals that reach them travel through virtually the same atmospheric section and therefore have the virtually the same errors and delays. The reference point uses its known position to derive what the signal travel time should be, measures the actual travel time and, by subtracting these two values, calculates the error of the received signal. This error is then transmitted to the roaming receiver for correction purposes. For our work, the preferred approach is to use a laptop computer with attached GPS receiver at a well-defined position on the Gosbecks site, and to use the radio-LAN to broadcast the error signal to the wearable systems. Conveniently, each roaming computer can also broadcast its own position; in the future, we shall use this to superimpose toga-clad avatars on other visitors.

(a) First-generation system in use

(b) Second-generation system integrated into a jacket

Figure 1: First- and second-generation systems

GPS, even differential GPS, yields only position information. A separate sensor is required in the user's HMD to determine orientation. The particular device used for this work, Virtual i/O's *I-glasses*, incorporates an electronic compass and tilt sensor; while not particularly accurate, we have found (rather to our surprise) that it works well enough outdoors and meets our orientation accuracy requirement.

5 The Evolution of the Hardware

As with many other groups, the concept of mobile augmented reality was proven using a laptop in a rucksack. Our wish to develop a system dedicated to this task has led us through two specific generations of hardware to date.

The first generation hardware was based on the popular "Tin Lizzy" architecture (Figure 1a). It used a Digital Logic PC/104 motherboard equipped with a 266 MHz

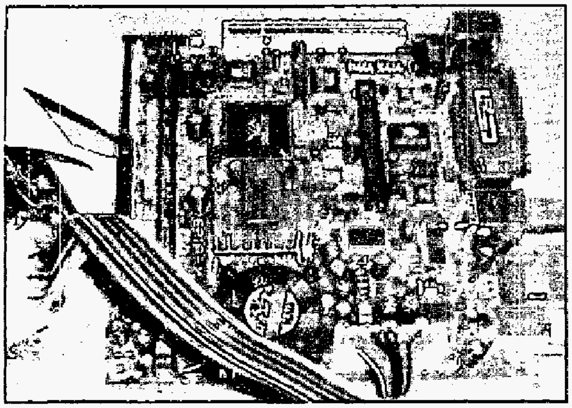


Figure 2: Mini-ITX motherboard used in the second wearable

Intel Pentium processor, 64 Mbyte of memory, four serial ports, one USB, VGA and sound. (The motherboard can now be upgraded to a 700 MHz processor.) A PCMCIA adaptor took a radio-LAN card. An IBM TravelStar hard disk was used, though we expect a production system to replace this with a compact flash card. An aluminium enclosure was used for protection, security, and to dissipate heat. The *I-glasses*' sensor attached to one serial port and the external GPS unit (a Garmin G12XL) to another. (An internal GPS receiver was tried but picked up too much electrical noise from the other components.) User input came from a HandyKey Twiddler2 [8], attached via USB. Power was supplied by two camcorder batteries, located in a separate unit for easy exchange; these also powered the HMD via a DC-to-DC converter.

The major problem with this configuration was power consumption: the two camcorder batteries provide a lifetime of only about forty-five minutes; we expect most users would spend about an hour on their tour, so we really need two hours' worth of power. However, increasing the number of cells adds significantly to the total weight and increases recharging time.

This configuration was designed to be worn on the belt. Its weight, including all cabling *etc.*, is 4.9 kg; we estimate that this figure can be reduced by about 20%. However, the system is bulky and becomes uncomfortable after about an hour's carriage.

The current, second-generation system has been designed specifically to reduce bulk and weight, while simultaneously providing for improved battery life. It is based around the increasingly popular mini-ITX series of motherboards. Mini-ITX boards, introduced by VIA,[1] are only 170 × 170 mm.

The EPIA M 9000 board (Figure 2) used in our second wearable includes a 933 MHz VIA C3 Processor, 256 MB DDR memory, a VIA CLE266 North Bridge with integrated AGP graphics processor, one PCI slot, two UltraDMA 133 connectors, on-board audio, Ethernet, Firewire, 2 USB and two serial ports. The motherboard is powered from a 55 W, 12 V Lex mini-ITX PSU, supplying power through a standard ATX connector.

The casing for the unit is larger in area but substantially thinner than the earlier PC/104-based device. This reduction in bulk has allowed us to integrate it into the large rear pocket of a photographer's jacket (Figure 1b). The GPS receiver and batteries then fit into the side pockets, and the GPS antenna is attached to an epaulet. The only remaining cables are then for the Twiddler (stored in a front pocket when not in use) and the HMD, which we have attached to the jacket where possible.

Trials demonstrate that this arrangement is much more comfortable, especially when the wearer is not in the first flush of youth. Equally important, it is much easier for visitors to the site to put on and take off the system when installed in a jacket than when attached to a belt.

6 Software Evolution

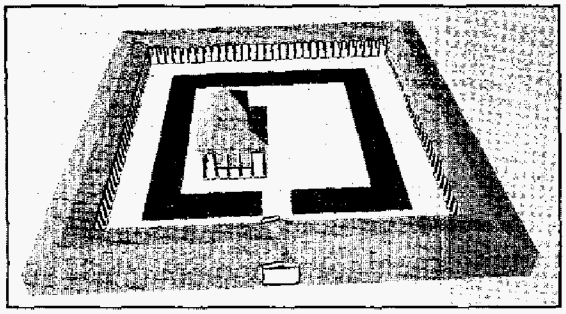
The design of hardware needs to go hand in hand with the application software that runs on it. In this case, the principal requirement is for rapid 3D rendering. The authors' preference is for a Linux-based system as this reduces cost, can run from flash, and allows us to have complete control over the user interface. However, drivers for accelerated hardware are not always available, so either one has to use software rendering or switch to Windows, where drivers are available but controlling the user interface is much more difficult.

Roman architecture was based on the principles of the earlier, more elegant Greek building styles; however, their approach was more formulaic to enable faster construction and employ less skilled artisans. A guide to building design due to Vetruvius [9] has survived from antiquity, which has enabled us, with advice from the Colchester Archæological Trust based on evidence they have uncovered, to reconstruct how the Gosbecks site probably looked.

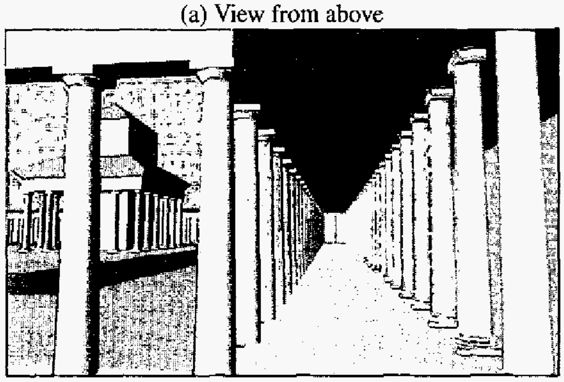
The temple complex comprises a square-shaped portico with an outer wall, an outer circuit of Doric columns and an inner circuit of Ionic columns, all covered by a tiled roof. The entrance faces roughly east. In the south-east of the inner square lies the main temple, surrounded by a ditch (Figure 3).

Software was written to generate VRML models from a few key measurements such as the number of columns and their diameters, and this formed the basis of the first software prototype. However, the resulting model was a large one, requiring accelerated graphics to render in

[1] `http://www.via.com/`



(a) View from above



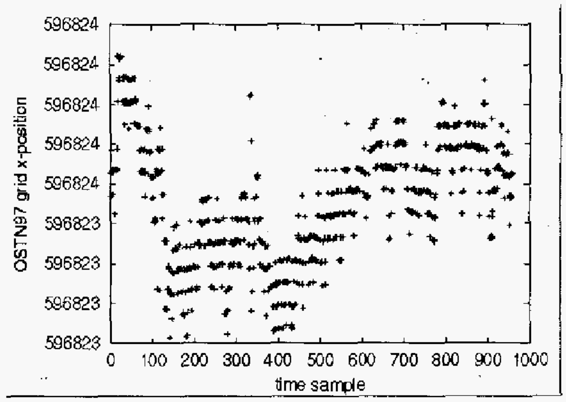
(b) Typical user's view

Figure 3: Views of the 3-D model of the Gosbecks temple complex

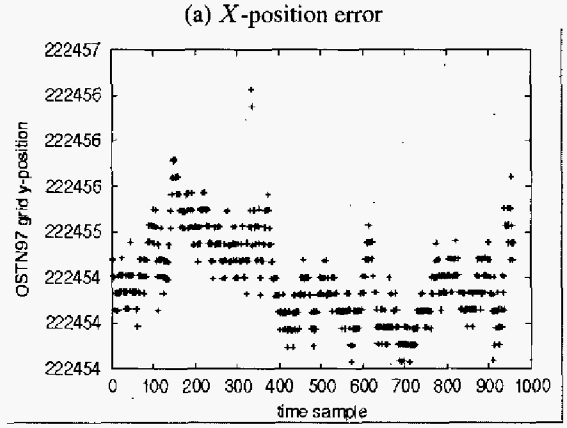
anything approximating real time even on desktop machines; sadly, suitable graphics cards were not available in PC/104 form. Furthermore, modifying existing open-source VRML viewers to accommodate GPS and HMD inputs proved surprisingly difficult.

Consequently, the model was converted to run with a custom OpenGL application, incorporating a number of optimizations, including rendering the columns, which are nominally cylindrical, as sets of eight octagonal prisms; avoiding textures; and using simple lighting. Nevertheless, rendering performance on the first-generation wearable remained too slow for practical use, under both Linux and Windows, because rendering lacked hardware acceleration.

The second-generation wearable has some hardware rendering capability and comes with appropriate drivers, though only for Windows. In developing the application software for the new hardware, we converted the VRML model to make use of a more structured scene graph. Experiments with the Cortona VRML viewer under Windows demonstrated that a 25 fps rendering rate is achievable. As Cortona itself is unsuitable for the eventual application, we anticipated using SGI's *OpenGL Performer* library, which has a VRML loader, as the basis of a custom application. Unfortunately, our experiments with it uncovered an incompatibility between Performer and the Windows driver wherein the screen would be re-



(a) X-position error



(b) Y-position error

Figure 4: The variation of GPS location with time

painted several times per second. As both components are closed-source, we are unable to ascertain precisely where the problem lies or correct it.

Consequently, our final solution is to re-code our original OpenGL code to employ a well-structured scene graph, and to cull those 3D objects that lie outside the view frustum within our own code rather than rely on an external library. In particular, we are aware that most of the rendering time is spent on the columns in the portico, so being able to cull them rapidly has a significant impact on the overall speed of the application. We are now able to render the model at about 35 fps under Windows with hardware acceleration and 20 fps under Linux without.

7 Disussion

Experiments with the latest generation of the tour guide system prove that it is considerately more comfortable to wear than its predecessor. Rendering performance is close to being adequate when the speed optimizations mentioned above are incorporated. Hence, we believe that truly viable solutions are close, perhaps as little as a few months away.

The remaining questions concern whether or not the overall system will be able to meet the design requirements outlined in Section 4 — in particular, whether the positional accuracy achievable with GPS is good enough.

A typical differential GPS reading will be within a metre of the right place when averaged over time, but it can wander violently by this sort of amount in the short term (Figure 4). The authors speculate that this is at least partly due to cycle slips and the rising and setting of GPS satellites; whatever the reason, the magnitude of the variations are problematic in practice. The effect can be ameliorated a little by maintaining a moving average of position (and orientation) values — but this also makes the system less responsive to the wearer's movements.

We now believe that better solutions for position estimation are required for this type of application. In the future, we intend to investigate other RF-based position estimation schemes, some of which exist but are awaiting licenses. However, even the best current outdoors position and orientation sensors will not be accurate enough to superimpose 3-D reconstructions onto existing archæology. In an attempt to achieve that, we are starting to investigate the use of a camera mounted on the HMD, using computer vision processing of its imagery to identify existing buildings and hence improve position and orientation accuracy. This increases the amount of on-body processing still further but is the only solution that will yield the ultimate accuracy required.

References

[1] Keith Cheverst, Nigel Davies, Keith Mitchell, Adrian Friday, and Christos Efstratiou. Developing a context-aware electronic tourist guide: Some issues and experiences. In *Computer-Human Interaction 2000*. IEEE Computer Society, 2000.

[2] Steven Feiner, Blair MacIntyre, and Tobias Höllerer. A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. In *The First International Symposium on Wearable Computers*. IEEE Computer Society, 1997.

[3] Tobias Höllerer and John Pavlik. Situated documentaries: Embedding multimedia presentations in the real world. In *The Third International Symposium on Wearable Computers*. IEEE Computer Society, 1999.

[4] Jie Yang, Weiyi Yang, and Matthias Debecke. Smart sight: A tourist assistant system. In *The Third International Symposium on Wearable Computers*. IEEE Computer Society, 1999.

[5] Wayne Piekarski and Bruce H. Thomas. Tinmith-metro: New outdoor techniques for creating city models with an augmented reality wearable computer. In *The Fifth International Symposium on Wearable Computers*. IEEE Computer Society, 2001.

[6] V. Vlahakis, J. Karigiannis, M. Tsotros, N. Ionnidis, and D. Stricker. Personalized augmented reallity touring of archaeological sites with wearable and mobile computers. In *Proceedings of the International Symposium on Wearable Computers*, 2002.

[7] P. Ritsos, D. J. Johnston, C. Clark, and A. F. Clark. Mobile augmented reality archæological reconstruction: Ergonomic and accuracy considerations. In *Proceedings of the Virtual World Heritage Conference*, October 2002.

[8] Handykey Corporation. Twiddler. `http://www.handykey.com/`.

[9] Morris Hicky Morgan. *Vitruvius, the Ten Books on Architecture*. Dover publications, 1967.