

# Augmented Reality for the Real World

Brett Wilkinson, Paul Calder

School of Informatics and Engineering, Flinders University of South Australia

{ Brett.Wilkinson@flinders.edu.au, Paul.Calder@flinders.edu.au }

## Abstract

*Augmented Reality (AR) systems suitable for use in everyday situations should not intrude unduly on their user's lifestyle (or that of others around them). This paper argues that such systems should make use of technologies and devices such as mobile phones, PDAs, and portable entertainment units that people are likely to be already using for other purposes. However, such devices have limited capabilities, which present challenges when designing a system for interaction with a 3D world.*

*This paper describes a prototype lightweight AR system built from off-the-shelf components and discusses some of the techniques we are developing to enable effective interaction with the system.*

## 1. Introduction

Picture this: You're attending a conference in an unfamiliar city and have arranged to meet a colleague at a café. The conference organizers have already uploaded places of interest, including the address of the café, onto your iPod, so when you reach the city the location of the café is displayed on your "e-shades" as a beacon superimposed on your view of the surroundings. You head for the beacon and are soon enjoying a cappuccino with your friend.

Or this: You're at a football match with your son but you get separated in the crowd. Becoming a little anxious, your son uses his mobile to contact yours and display your location on his "video-pod" eyewear. As he heads towards you, you're doing the same thing to display his location, so you're soon reunited.

Or perhaps this: You've signed up with a tour group, been given a pair of e-shades and have been taken to a popular lookout; the whole city is laid out before you. Your tour operator is talking about the various historical sites of the city and triggers virtual details with a PDA, which are displayed on the group's glasses. You ask where the statue of a local hero is, so the tour guide displays a map of the city on his PDA, locates the statue, and clicks the "show me" icon to display its position on your e-

shades. Having spotted the statue, you ask what that large building is a little further away. The tour guide hands you the PDA which you use to move a superimposed cursor over your view of the building and click a "what's this" button on the PDA screen. The PDA scrolls the map to the corresponding location and indicates that the building in question is the central library.

This is the future of Augmented Reality (AR); lightweight consumer level systems with minimal to no social impact on their users. AR systems have been demonstrated in domains as diverse as architecture [7], defense [5], surgery [14], and gaming [9]. However, most current systems require the user to carry bulky backpacks, wear cumbersome head-mounted displays, and use unfamiliar input devices such as data gloves. Although appropriate for specialized applications, we believe such systems are unlikely to be popular for the everyday uses we envision. Instead, we are investigating how lightweight and non-intrusive AR systems can be built using the kind of technologies that people are likely to be already using for other purposes: phones, music players, PDAs and the like.

The scenarios above are not yet a reality, but much of the technology needed is either already available or likely to be available in the near future. Lightweight see-through head-mounted displays, rather like a pair of sunglasses although still somewhat more bulky, are becoming cheaper. Orientation trackers to monitor head movements are already small enough to be built into the frame of the display. GPS units that can report positions to within a few metres are increasingly turning up in PDAs and mobile phones. And the current generation of PDAs, phones, and portable media players boast enough CPU power to generate graphical displays rapidly, and they frequently include Bluetooth capability to tie the components together and Wi-Fi to connect the system to local networks.

Although these technologies are unlikely to be good enough for demanding medical or military applications, we believe that they are more than adequate for everyday outdoors applications where targets are relatively far away (so that GPS position accuracy is good enough) and where

the graphics are relatively simple (so that portable devices with low-powered CPUs are fast enough).

Indeed, we believe that the main missing ingredient is not hardware, but rather the software that will enable people to interact with the system using the limited input capabilities of familiar devices. How does one build an effective interface to the 3D world using the small touch screen of a PDA, or the keypad of a mobile phone, or even the click wheel of an iPod?

To investigate these questions and try out our ideas, we have built a prototype system with off-the-shelf hardware and are experimenting with ways to interact with the system. This paper briefly examines some of the related work on which we are building, gives an overview of our system hardware and software, describes some of the interaction techniques we are investigating, and suggests which directions we are planning to head in the future.

## 2. Related Work

Current AR research ranges from stationary to mobile, collaborative to isolated, indoor to outdoor, and head-mounted to portal-based displays. This review is focused on some of the issues relevant for mobile AR technologies and the interaction techniques used.

### 2.1 Mobile Outdoor AR

The Tinmith system [11], built by Bruce Thomas and Wayne Piekarski [8] at the University of South Australia, allows extensive mobility and exceptional interaction with AR objects and the virtual environment. The system features a small and well designed backpack, but is still too cumbersome and awkward for everyday use.

The GEIST system [4] of Holweg is another example of an accurate and relatively compact system. The backpack containing the computational equipment is larger than the Tinmith apparatus but the GEIST system employs a PDA for its input device. Users of the GEIST system are able to input queries into the PDA [3] while reliving a historical story. The use of the PDA for user input is a logical choice in this scenario as the user is dealing with textual queries. However an issue arises here with the concept of separate displays, the user's main view through the HMD and then their view of the possible input data on the PDA.

Other outdoor, mobile systems such as the BARS [5] and MARS [2] projects also rely on specialised and bulky equipment. Military applications such as BARS require highly accurate position information. However, many everyday applications can tolerate the positional errors of up to 10 metres that are characteristic of standard consumer GPS receivers.

### 2.2 Portal Window AR

Portal window applications such as HandheldAR [16] promise mobility and a small system size. These systems require the user to hold a display device at arm's length and to view the augmented world as if through a window. Although simple, portal systems are not appropriate for many applications, especially those that require the user to move around while viewing the display.

Typical examples of portal-based systems include SitePack [7] and MagicLens [6]. SitePack provides specialized services for architecture design and display. The equipment is quite compact consisting of a TabletPC, a web cam, and a GPS tracker. MagicLens is intended for collaborative indoor AR and presents an augmented view of the world as if viewed through a magnifying glass.

### 2.3 Image Capture, Processing, and Recognition

Many AR systems use some form of image capture to facilitate their AR environment, often to help determine the user's position or the position of key objects in the world. For example the ARToolkit [1], an extensively used AR library, provides support to determine the position and orientation of fiducial markers attached to real-world object. These markers can then serve as visual anchors for virtual objects. Although appropriate for controlled environments, where pre-preparation and detailed mapping of the surroundings is possible, this approach would not work for wide-area use of personal AR because the world is not conveniently decorated with easy-to-recognise features.

Other systems rely on image processing to infer the position of objects from features such as edges, shadows, and landmarks. Although this approach doesn't require any pre-preparation of the real world, it requires substantial computational capability and is beyond the capacity of low-powered portable platforms.

### 2.4 Current Interaction Techniques

The Tinmith system provides user input via a data glove with a small attached fiducial target pad. The system tracks the location of the glove and can provide intricate levels of interaction with virtual objects [10, 15]. However, the approach requires that the glove is within view of the camera, which limits its use for everyday applications where people are often not looking at their hands. We see our solution, by allowing a user to discreetly tap at a PDA or manipulate a mobile phone's buttons without having to overtly announce their actions, as a more socially acceptable approach.

Reitmayr's [13] approach to AR with the Studierstube system investigates similar principles to our own research evaluating the features of a 2D interaction tool when manipulating 3D worlds. Where their research differs

though is in the use of a virtual tablet anchored to fiducial markers to enable interaction. We see the solidness and physical presence of an input device as important in the accessibility of the system. Without a tangible apparatus to manipulate many users would be hesitant to use the system.

Reitmayr's iOrb [12] is a compact interaction device utilizing an orientation tracker held in the hand. Coupled with the interfaces they are testing, the iOrb seems a successful solution to unobtrusive interaction with AR systems, although it requires the user to operate the device in an unfamiliar manner and there is still a level of hand movement that would attract unwanted attention.

## 2.5 Our Interaction Philosophy

In trying to develop a real world AR system, various interaction techniques needed to be considered. The idea of adopting data gloves, twiddlers, arm-mounted keyboards, and other current 'techie-friendly' input devices didn't present a realistic solution to consumer level AR interaction. The dismissal of these tools led discussion to the current technology available to expected users and the type of interactions that they are employing with these devices. Most expected users are comfortable and literate with portable technologies such as mobile phones, PDAs, and portable media players. All of these devices incorporate some form of point-and-click interaction technique, from the navigation keys of a mobile phone or media player to the stylus of a PDA. This project is seeking to leverage the user's current understanding and comfort with this style of interface and provide a technique of AR interaction that is private and familiar

The aim of our research is to develop a mobile, unobtrusive AR system using compact low-cost equipment that would enable an average user to interact effectively with virtual targets. We see one of the main hurdles for mass acceptance and usability of AR as the bulk of the equipment. Social issues would prohibit most users donning a bulky HMD, data gloves, and a body mounted computing device to view virtual information as they navigate about their daily lives.

As the power of computing technology continues to grow and its bulk diminishes AR, will become a very attractive information delivery and communication system. However, the mass acceptance of AR will rest on its interaction techniques; how usable are the devices and what sort of user training will be required to interact with the systems? If we can provide a familiar set of tools for users to engage with, in a manner that is private and contained within an acceptable personal space, consumer level AR would be achievable and commercially attractive.

The primary vision that has driven this research is the notion of a busy city street populated by AR-equipped pedestrians. If all these users were stretching their gloved hands out and gesticulating to complete their interactions

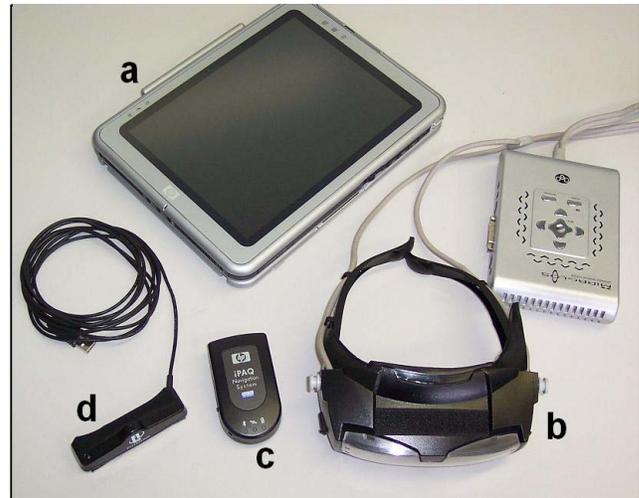


Figure 1: System Hardware

there would be constant collisions and interference. We see our solution as providing a discrete, confined and socially responsible interaction process.

To investigate the assumptions about familiar input techniques and their effectiveness at manipulating an AR environment we constructed a simple test platform. The following sections will outline the hardware that comprises our system and the style of interaction techniques that are employed to emulate many of the tasks that the existing unfamiliar AR interaction tools currently implement.

## 3. Equipment

The system comprises four items of off-the-shelf hardware (excluding the battery pack for the HMD); the complete system hardware is depicted in Figure 1.

The test system runs on a HP tc1100 TabletPC (Item a in Figure 1) enabling touch screen interaction similar to a PDA but providing a more stable system for testing interface developments.

The system uses an optical see-through HMD from Deocom, the Viotor SX-2, (Item b in Figure 1). An optical see-through HMD was chosen, as opposed to a video see-through device, because low-powered platforms would struggle to accurately render the scene based on the video image captured by an attached camera. For the kinds of applications we envisage, we believe that the limitations of see-through displays (such as not being able to have virtual objects occlude real-world objects) would not detract from the expectations of the *general* user and therefore argued that optical see-through was the best mode of display delivery.

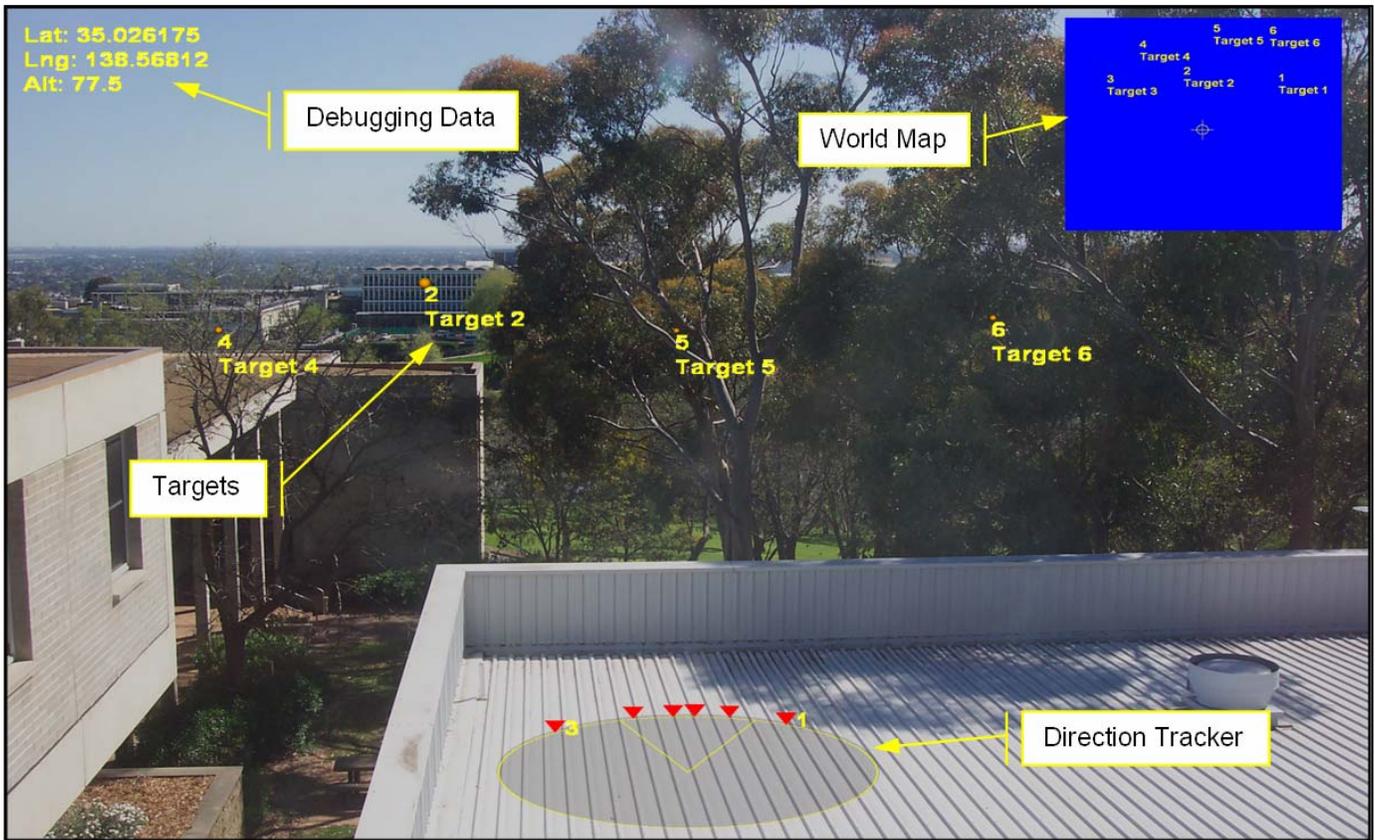


Figure 2: User Interface

For location tracking the system uses a Bluetooth GPS receiver, the HP Navigation System GPS Receiver from Navteq (Item c in Figure 1). Future platforms for the AR system will likely include PDAs or mobile phones with built-in GPS receivers; the predominance and low cost of these devices make them a natural selection for the expected system platform.

Finally, to track the movements of the user's head the system uses the Intersense Intertrax2 USB model (Item d in Figure 1) to provide the required 3 degrees of freedom (3DOF) tracking while maintaining the low cost of the complete system. However, when field testing with the unit began the limited capabilities of the device and its sensitivity to lateral movements of the user were discovered. Approaches to counter this limitation have been adopted with the test system; a stop-and-survey procedure is used to avoid the drift caused by the user's lateral movements.

#### 4. Construction of the Interface

To test our system we created a test application that provides a head-mounted, see-through navigation and target-finding scenario for the user. The application presents the user with a series of target real-world objects on a university campus, identified only by their position in the augmented view. Users are asked to use the system to

walk towards and name each target in turn. Targets are familiar objects around the campus, such as rubbish bins, garden seats, and notice boards. We are using the application to provide data on the performance of the system's components, the usability of the system, and the social impact on the user.

The application software is implemented in C# on the .NET platform. Orientation-tracking input, matrix calculations, and rendering are supported by Microsoft's DirectX libraries. DirectInput allows access to the raw orientation tracker data, while the Direct3D library provides the various transformations and scene rendering technologies. These libraries were used for their ease of development; in addition, future releases of the Windows Mobile operating system are expected to incorporate DirectX functionality.

#### 4.1 User interface

Figure 2 shows the application interface overlaid on a real-world scene. The display shows the targets in the main window, a direction tracker that indicates the direction of targets not currently within the user's field of view, and a top-down world view indicating the user's current location relative to the targets. Annotations about the user's current location, viewing direction, speed, and current time can be displayed for debugging purposes.

Targets are represented as simple labeled spheres, sized to indicate relative distance. Extra information about selected targets can be displayed as needed (not shown in Figure 2). The simple representation of targets preserves as much of the user's view of the real world as possible and reduces the processing power needed to render the scene.

The direction tracker represents the direction of the targets from the user's current location and bearing. An ellipse simulates the horizon of the user's world, with a pie wedge indicating their current field of view. The display is aligned with the user's view; turning to face a new direction causes the target indicators to be repositioned around the perimeter of the ellipse.

The world view details the user's location relative to the targets in the world. This view is represented as a map, oriented with north at the top of the screen and centered on the user's current position. As the user moves throughout the world they will see the targets move within the world view.

The interface also includes components for performing various application functions, such as calibrating the orientation tracker, turning optional display elements on and off, and displaying instructions and status messages. These functions use interface components such as pop-up menus and message boxes that are modeled on familiar desktop widgets.

## 4.2 Interaction Techniques

The current system uses the TabletPC stylus as the main interaction device. The user operates the interface by holding the tablet in one hand and the stylus in the other; in effect, the tablet and stylus act as a pointing device that allows the user to move a cursor over their view of the augmented world. Since the user's view is generated by the HMD, the tablet's display screen is not needed once the application is launched. (In the current prototype, the HMD is simply slaved off the tablet display; in future, we plan to investigate using the tablet display to provide additional control and information capabilities.)

The target-finding application requires users to navigate their way to real-world locations identified by on-screen targets. The application allows them to select a target in order to find out additional information about it. Selected targets are highlighted in the main window (and in the direction tracker and world view if displayed).

Since the TabletPC tracks the stylus location while it is on or near to the tablet surface, selection is easily achieved by simply moving the stylus until the virtual cursor is over the desired target, then tapping the tablet. Where multiple targets overlap, subsequent taps deselect the selected object and select the one behind it, thus allowing the user to "drill down" to distant targets. Where a selected target would be obscured by nearer targets, it is repainted in a separate layer; the effect is as if the selected target were visible through a hole in the nearer target.

Since targets might be outside the user's current field of view (because the user is facing away from them), they can sometimes be difficult to locate. Often such targets can be located using the direction finder or world view, which display all targets. However, where targets are close together they can be hard to distinguish in these small-scale views. To help locate such "lost" targets, the application also allows users to select targets by choosing them from a pop-up menu system.

For an application implemented on another portable device, different input techniques would be needed. For example, most PDAs use a passive stylus for input, where the position of the stylus is not known (until of course it touches the display screen). On such systems, a different strategy would be needed to implement the virtual cursor. For example, the cursor could be controlled by gestures: a stroke might move it a specified distance in the indicated direction, and a flick might start it moving with a specified velocity. On a mobile phone, yet another strategy would be needed. Perhaps the phone's "arrow" keys could be used to move the cursor around, and the "enter" key to select a target. So far, we haven't been able to devise a simple strategy to control a cursor using an iPod touch wheel, but we're working on it!

## 5. Future Development

We are currently testing the prototype system to evaluate its performance and usability, and to find out how users feel about its "social" impact. Further development of the interaction tools will include specialized functionality such as improved depth selection, rotation of targets through all three axes, and context-specific selection criteria. Effective screen shortcuts to enable selection and subsequent manipulation of the virtual targets will be considered. We will also investigate ways to input new targets or textual information while immersed within the virtual world.

Finally we will investigate porting the system to other platforms such as a PDA or a mobile phone. Such platforms will present additional issues such as display refresh speed and resolution, compatibility of hardware, and usability of the interfaces.

Currently the capabilities of the system are demonstrated using a navigation-style application. Continuing development of other applications that are usable within this system could include education services (such as applications that allow students to view historical information about sites), entertainment services (such as a stargazing application that provide details about the stars viewable from the user's position), and military services (such as a stripped-down AR system updatable in real time by a central command environment).

Part of the issue involved with developing effective consumer level AR is the creation of useful applications; this investigation of interaction techniques will present

other applications with a context and location dependence to provide more information to the average user.

## 6. Conclusions

We expect that the use of data gloves and other specialized interaction devices will remain an important aspect of AR applications but that they will be used within specific domains. For most people, familiar interaction techniques using familiar input devices will be good enough, and will have significantly less physical, financial, and social impact on their lives. Our research into AR input methods has identified some useful techniques and confirmed our assumptions that users can effectively interact within a 3D head-mounted, AR system with a simple, unobtrusive interface.

## References

- [1] ARToolkit 2003, <http://www.hitl.washington.edu/artoolkit/>
- [2] T. Hollerer, et al. Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System. *Computers and Graphics*, 23(6), pp. 779-785, 1999.
- [3] D. Holweg and O. Schneider. GEIST: Mobile Outdoor AR-Information System for Historical Education with Digital Storytelling. In Proceedings *Federal Ministry of Education and Research: Virtual and Augmented Reality Status Conference*, Leipzig, Germany, 2004.
- [4] U. Holweg, et al. GEIST Outdoor Augmented Reality in an Urban Environment. *Computer Graphics Topics*, (6), pp. 5-6, 2002.
- [5] S. Julier, et al. BARS: Battlefield Augmented Reality System. In *NATO Symposium on Information Processing Techniques for Military Systems*, Istanbul, Turkey, October 2000.
- [6] J. Looser, et al. Through the Looking Glass: The Use of Lenses as an Interface Tool for Augmented Reality Interfaces. In Proceedings *Second International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia*, Singapore, June 2004.
- [7] M. Nielsen, et al. Mobile Augmented Reality Support for Architects Based on Feature Tracking Techniques. In Proceedings *International Conference on Computational Science*, Krakow, Poland, June 2004.
- [8] W. Piekarski, et al. A Mobile Augmented Reality User Interface for Terrestrial Navigation. In Proceedings *22<sup>nd</sup> Australasian Computer Science Conference*, Auckland, NZ, 1999.
- [9] W. Piekarski and B. Thomas. ARQuake: The Outdoor Augmented Reality Gaming System. In *Communications of the ACM*, Vol. 45 (1), pp. 36-38, 2002.
- [10] W. Piekarski and B. Thomas. Using ARToolkit for 3D Hand Position Tracking in Mobile Outdoor Environments. In Proceedings *1<sup>st</sup> International Augmented Reality Toolkit Workshop*, Darmstadt, Germany, 2002.
- [11] W. Piekarski and B. Thomas. Tinmith-evo5 – An Architecture for Supporting Mobile Augmented Reality Environments. In Proceedings *2<sup>nd</sup> International Symposium on Augmented Reality*, New York, USA, 2001
- [12] G. Reitmayr, et al. iOrb – Unifying Command and 3D Input for Mobile Augmented Reality. In Proceedings *IEEE VR 2005 Workshop on New Directions in 3D User Interfaces*, Bonn, Germany, 2005.
- [13] G. Reitmayr and D. Schmalstieg. Mobile Collaborative Augmented Reality. In Proceedings *Second International Symposium on Augmented Reality*, New York, USA, 2001.
- [14] M. Schnaider, et al. Medarpa – A Medical Augmented Reality System for Minimal-Invasive Interventions. In Proceedings *11<sup>th</sup> Annual Medicine Meets Virtual Reality Conference*, Newport Beach, USA, January 2003.
- [15] B. Thomas and W. Piekarski. Glove Based User Interaction Techniques for Augmented Reality in an Outdoor Environment. In *Virtual Reality: Research, Development, and Applications*, (6), 2002.
- [16] D. Wagner and D. Schmalstieg. First Steps Towards Handheld Augmented Reality. In Proceedings *Seventh IEEE International Symposium on Wearable Computing*, New York, USA, 2003