

1

"We all agree that your theory is crazy, but is it crazy enough?"

Niels Bohr (1885-1962)

Chapter 1 - Introduction

Augmented reality (AR) is the registration of projected computer-generated images over a user's view of the physical world. With this extra information presented to the user, the physical world can be enhanced or augmented beyond the user's normal experience. The addition of information that is spatially located relative to the user can help to improve their understanding of it. In 1965, Sutherland described his vision for the Ultimate Display [SUTH65], with the goal of developing systems that can generate artificial stimulus and give a human the impression that the experience is actually real. Sutherland designed and built the first optical head mounted display (HMD) that was used to project computer-generated imagery over the physical world. This was the first example of an augmented reality display [SUTH68]. Virtual reality (VR) was developed later using opaque display technology to immerse the user into a fully synthetic environment. One of the first integrated environments was by Fisher et al., combining tracking of the head for VR with the use of tracked gloves as an input device [FISH86].

Augmented reality and virtual reality share common features in that they present computer-generated images for a user to experience, with information anchored to 3D locations relative to the user's display, their body, or the world [FEIN93b]. The physical world seen when using AR can be thought of as a fourth kind of information that the user can experience, similar to world-relative display but not artificially generated. A typical example of a head mounted display is shown in Figure 1-1, and an example AR scene with both physical and virtual worlds is depicted in Figure 1-2. The schematic diagram in Figure 1-3 depicts how a see-

Chapter 1 - Introduction

through HMD (used to produce AR images for the user) can be conceptualised, and the next chapter discusses in depth their implementation. While other forms of sensory stimulation such as haptics and audio are also available to convey information to the user, these will not be discussed since the focus of this dissertation is HMD-based AR.

Although the first HMD was implemented in 1968 and performed augmented reality, the first main area of research with these devices was for virtual reality. VR has a number of similar research problems with AR, but does not rely on the physical world to provide images and can also be viewed on display monitors or projectors. Since the user is normally tethered to the VR system, they are not able to walk large distances and virtual movement techniques such as flying are required to move beyond these limits. These same problems restricted initial AR research because, by its very nature, the user would like to walk around and explore the physical world overlaid with AR information without being tethered to a fixed point. While there are some important uses for AR in fixed locations, such as assisted surgery with overlaid medical imagery [STAT96] or assistance with assembly tasks [CURT98], the ability to move around freely is important. When discussing the challenges of working outdoors Azuma states that the ultimate goal of AR research is to develop systems that “can operate anywhere, in any environment” [AZUM97b].

A pioneering piece of work in mobile augmented reality was the Touring Machine [FEIN97], the first example of a mobile outdoor AR system. Using technology that was small and light enough to be worn, a whole new area of mobile AR research (both indoor and outdoor) was created. While many research problems are similar to indoor VR, there are unsolved domain specific problems that prevent mainstream AR usage. Older survey papers



Figure 1-1 Example Sony Glasstron HMD with video camera and head tracker

Chapter 1 - Introduction

(such as by Azuma [AZUM97a]) cover many technological problems such as tracking and registration. As this technology has improved, newer research is focusing on higher-level problems such as user interfaces, as discussed by Azuma et al. [AZUM01].

Many outdoor AR systems produced to date rely only on the position and orientation of the user's head in the (sometimes limited) physical world as the user interface, with user interfaces that can indirectly adjust the virtual environment. Without a rich user interface capable of interacting with the virtual environment directly, AR systems are limited to changing only simple attributes, and rely on another computer (usually an indoor desktop machine) to actually create and edit 3D models. To date, no one has produced an outdoor AR system with an interface that allows the user to leave behind all their fixed equipment and independently perform 3D modelling in real-time. This dissertation explores user interface issues for AR, and makes a number of contributions toward making AR systems able to operate independently in the future, particularly for use in outdoor environments.

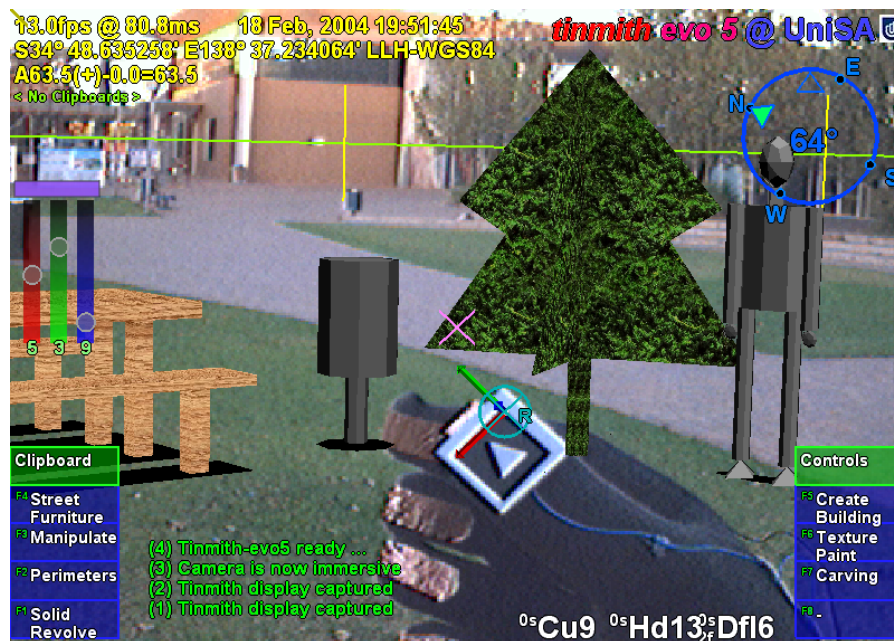


Figure 1-2 Example of outdoor augmented reality with computer-generated furniture

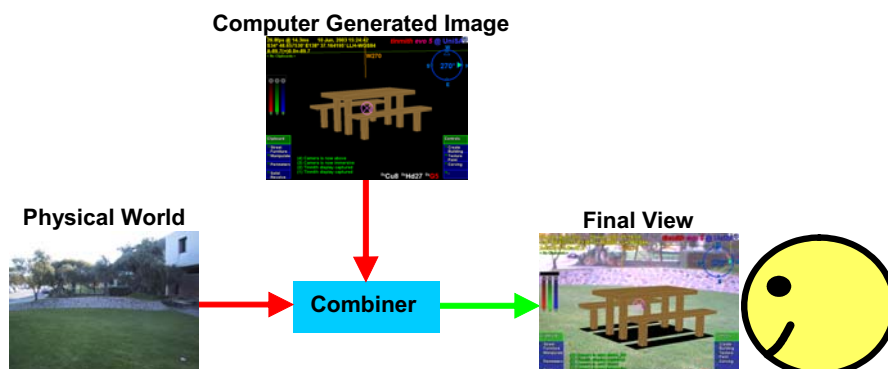


Figure 1-3 Schematic of augmented reality implementation using a HMD

1.1 Problem statement

The development of user interfaces for mobile outdoor AR systems is currently an area with many unsolved problems. When technology is available in the future that solves existing registration and tracking problems, having powerful applications that can take advantage of this technology will be important. Azuma et al. state that “we need a better understanding of how to display data to a user and how the user should interact with the data” [AZUM01]. In his discussions of virtual reality technology, Brooks stated that input devices and techniques that substitute for real interactions were still an unsolved problem and important for interfacing with users [BROO97]. While VR is a different environment in that the user is fully immersed and usually restricted in motion, the 3D interaction problems Brooks discusses are similar and still relevant to the AR domain. Working in an outdoor environment also imposes more restrictions due to its mobile nature, and increases the number of problems to overcome.

On desktop computers, the ubiquitous WIMP interface (windows, icons, menus, and pointer - as pioneered by systems such as the Xerox Star [JOHN89]) is the de-facto standard user interface that has been refined over many years. Since mobile outdoor AR is a unique operating environment, many existing input methodologies developed for AR/VR and desktop user interfaces are unavailable or unsuitable for use. In an early paper about 3D modelling on a desktop, Liang and Green stated that mouse-based interactions are bottlenecks to designing in 3D because users are forced to decompose 3D tasks into separate 1D or 2D components [LIAN93]. Another problem is that most desktop 2D input devices require surfaces to operate on, and these are unavailable when walking outdoors. Rather than trying to leverage WIMP-based user interfaces, new 3D interfaces should be designed that take full advantage of the environment and devices available to the user. An advantage of AR and VR is that the user’s body can be used to control the view point very intuitively, although no de-facto standard has emerged for other controls in these applications. While research has been performed in the VR area to address this, many techniques developed are intended for use in immersive environments (which do not have physical world overlay requirements) and with fixed and limiting infrastructure (preventing portability).

With AR systems today (and with virtual environments in general), a current problem is the supply of 3D models for the computer to render and overlay on the physical world. Brooks mentions that 3D database construction and modelling is one of four technologies that are crucial for VR to become mainstream, and is still an unsolved problem [BROO97]. He mentions that there is promising work in the area of image-based reconstruction, but currently modelling is performed using CAD by-products or hard work. An interesting problem area to

Chapter 1 - Introduction

explore is what types of models can be created or captured directly while moving around outdoors. By integrating the modelling process and user interface, the user is able to control the modelling process directly and take advantage of their extensive knowledge of the environment.

1.1.1 Research questions

For this dissertation I have investigated a number of different unsolved problems in the field of AR, and then combined the solutions developed to produce real world applications as demonstrations. I have formulated these different problems into research questions that will be addressed in this dissertation:

- How can a user intuitively control and interact with a mobile AR system outdoors, without hampering their mobility or encumbering their hands?
- How can a user perform manipulation tasks (such as translate, rotate, and scale) with an outdoor mobile AR system of existing 3D geometry, in many cases out of arm's reach and at scales larger than the user's body?
- How can a user capture 3D geometry representing objects that exist in the physical world, or create new 3D geometry of objects that the user can preview alongside physical world objects, out of arm's reach and at scales larger than the user's body?
- What is an appropriate software architecture to develop this system with, to operate using a wide variety of hardware and software components, and to simplify application development?
- What hardware must be developed and integrated into a wearable platform so that the user can perform AR in the physical world outdoors?
- What application domains can take advantage of the novel ideas presented in this dissertation, and for what real world uses can they be applied to?

1.1.2 Research goals

The main goal of this dissertation is to answer the questions discussed previously and contribute solutions towards the problems facing augmented reality today, predominantly in the area of user interfaces for mobile outdoor systems. The specific research goals of this dissertation are as follows:

- *Mobile user interface* - The user interface should not unnecessarily restrict the user's mobility, and should use intuitive and natural controls that are simple to learn and use.

Chapter 1 - Introduction

Requiring the user to carry and manipulate physical props as controls may interfere with the user's ability to perform a required task, such as holding a tool.

- *Real-time modelling enhanced with proprioception* - The user should be able to interactively create and capture the geometry of buildings using the presence of their body. Using solid modelling operations and the current position and orientation of the head and hands will make this process intuitive for the user.
- *Mobile augmented reality* - Outdoor augmented reality requires a mobile computer, a head mounted display, and tracking of the body. These technologies currently suffer from a number of limitations and applications must be designed with this in mind to be realisable. By using current technology, new ideas can be tested immediately rather than waiting for future technology that may not appear in the short term.

1.2 Thesis statement

Using parts of the body such as the head and hands to perform gestures is a natural way of interacting with 3D environments, as humans are used to performing these actions when explaining operations to others and dealing with the physical world. By using techniques such as pointing and grabbing with objects in positions relative to the body, user interfaces can leverage proprioception, the user's inbuilt knowledge as to what their body is doing. Mine et al. [MINE97a] demonstrated that designing user interfaces to take advantage of a human's proprioceptive capabilities produced improved results. Using an input device such as a mouse introduces extra levels of abstraction for the direct manipulation metaphor (as discussed by Johnson et al. [JOHN89]), and so using the head and hands allows more intuitive controls for view point specification and object manipulation.

Trying to leverage existing 2D input devices for use in a naturally 3D environment is the wrong approach to this problem, and designing proper 3D user interfaces that directly map the user to the problem (as well as taking advantage of existing interactive 3D research) will yield improved results. The use of 3D input devices has been demonstrated to improve design and modelling performance compared to 2D desktop systems that force the user to break down 3D problems into separate and unnatural 2D operations [CLAR76] [SACH91] [BUTT92] [LIAN93].

In VR environments the user is able to use combinations of physical movement and virtual flying operations to move about. In contrast, in AR environments the user is required to always move with their physical body otherwise registration between the physical and artificial worlds will be broken. Using direct manipulation techniques, interacting with objects

Chapter 1 - Introduction

that are too large or too far away is not possible. Desktop-based CAD systems rely on 2D inputs but can perform 3D operations through the use of a concept named working planes. By projecting a 2D input device cursor onto a working plane the full 3D coordinates of it can be calculated unambiguously. By extending the concept of working planes to augmented reality, both the creation of geometry and interaction with objects at a distance can be achieved. These working planes can be created using the physical presence of the body or made relative to other objects in the environment. Accurate estimation of the depth of objects at large distances away has been shown to be difficult for humans [CUTT95], and AR working planes provides accurate specification of depth. This functionality is achieved using a slightly increased number of interaction steps and reduction in available degrees of freedom.

By combining AR working planes with various primitive 3D objects (such as planes and cubes) and traditional constructive solid geometry techniques (such as carving and joining objects), powerful modelling operations can be realised, which I have termed construction at a distance. These operations give users the capability to capture 3D models of existing outdoor structures (supplementing existing surveying techniques), create new models for preview that do not currently exist, and perform editing operations to see what effect changes have on the environment. By taking advantage of a fully tracked AR system outdoors, and leveraging the presence of the user's body, interactive modelling can be supported in an intuitive fashion, streamlining the process for many types of real world applications.

1.3 Contributions

This dissertation makes a number of research contributions to the current state of the art in augmented reality and user interfaces. Some of the initial contributions in this dissertation also require a number of supporting hardware and software artefacts to be designed and developed, each with their own separate contributions. The full list of contributions is:

- The analysis of current techniques for distance estimation and action at a distance, and the formulation of a technique named *augmented reality working planes*. This new technique can create objects accurately at large distances through the use of line of sight techniques and the projection of 2D cursors against planes. This technique is usable in any kind of virtual environment and is not limited to augmented reality. [PIEK03c]
- The design and implementation of a series of techniques I term *construction at a distance*, which allow users to capture the 3D geometry of existing outdoor structures, as well as create 3D geometry for non-existent structures. This technique is based on

Chapter 1 - Introduction

AR and uses the physical presence of the user to control the modelling. Objects can be modelled that are at scales much larger than the user, and out of arm's reach. [PIEK03c]

- The iterative design and development of an augmented reality user interface for pointing and command entry, allowing a user wearing gloves to navigate through and select menu options using finger presses, without requiring high fidelity tracking that is unavailable outdoors. This user interface can operate without tracking, but when tracking is available it allows interaction at a distance with 3D environments. [PIEK03d]
- The development of a vision-based hand tracking system using custom designed pinch gloves and existing fiducial marker tracking software that can work reliably under outdoor wearable conditions. [PIEK02f]
- The development of applications that allow users to model buildings with the techniques described in this dissertation, and in some cases being able to model objects previously not possible with or faster than existing surveying techniques. [PIEK01b] [PIEK03c]
- The design and implementation of a software architecture that is capable of supporting the research for this dissertation, culminating in the latest design of the Tinmith software. Current software architectures are still immature and do not support all the requirements for this dissertation, and so this architecture was designed to support these requirements and implements many novel solutions to various problems encountered during the design. [PIEK01c] [PIEK02a] [PIEK03f]
- The iterative design and development of hardware required to demonstrate the applications running outdoors. Wearable computing is currently in its infancy and so the devices that are required to be used outdoors are not necessarily designed for this, and numerous problems have been encountered for which novel solutions to these are implemented. [PIEK02h]

1.4 Dissertation structure

After this introduction chapter, chapter two contains an overall background discussion introducing the concepts and technology that form a core of this dissertation. This overview provides a general discussion of related research, and each chapter then includes other more specific background information when relevant. Chapter three discusses the problems associated with a user interacting with objects at large distances – while most VR systems take advantage of a human's ability to work well within arm's reach, outdoor AR work tends

Chapter 1 - Introduction

to be performed away from the body where depth perception attenuates very rapidly. The novel idea of using the concept of CAD working planes for AR is introduced, along with various techniques that can be employed to perform interaction at a wide range of scales and distances beyond the reach of the user. Chapter four takes the concepts previously developed in the dissertation as well as constructive solid geometry and develops a new series of techniques named construction at a distance that allow users to perform modelling of outdoor objects using a mobile AR system. These techniques are then demonstrated using a series of examples of outdoor objects to show their usefulness. Chapter five explains how the previously developed techniques are interfaced to the user through pointing with the hands, command entry using the fingers, and the display of data to the HMD. The user interface is required to develop applications that are useable tools - an important part of this dissertation is the ability to test out the techniques and improve them iteratively. Chapter six introduces the software architecture used to facilitate the development of virtual environment applications. The software architecture contains a number of novel features that simplify the programming of these applications, unifying all the components with a consistent design based on object-oriented programming, data flow, and a Unix file system-based object repository. By tightly integrating components normally implemented separately, such as the scene graph, user interface, and internal data handling, capabilities that are normally difficult to program can be handled with relative ease. To complete the research, chapter seven describes the hardware components that are an important part of the overall implementation since the software relies on these to execute. The research and development of the mobile backpack, user input glove, and vision tracking system are explained to show some of the important new innovations that have been developed in these areas for use outdoors. After concluding with a discussion of the numerous contributions made and future work, this dissertation contains an appendix with a history of my previous hardware and software implementations, as well as links to where further information can be found about the project.