ThumbsUp: Integrated Command and Pointer Interactions for Mobile Outdoor Augmented Reality Systems

Wayne Piekarski and Bruce H. Thomas

Wearable Computer Laboratory
School of Computer and Information Science
University of South Australia
Mawson Lakes, SA, 5095. Australia
{wayne, thomas}@cs.unisa.edu.au

Abstract

This paper presents a new user interface technology known as *ThumbsUp*, which we have designed and developed for use with mobile outdoor augmented reality systems. Using a simple pair of vision tracked pinch gloves and a new menuing system, a user is able to control an augmented reality system in outdoor environments under poor tracking conditions with a high level of accuracy. Highly interactive 3D augmented reality applications can now be operated outdoors with our new easy to use interface technology.

1 Introduction

ThumbsUp is our new user interface technology for use with mobile outdoor augmented reality (AR) systems. User interfaces to date for outdoor AR systems have been quite simple, but our investigations into modelling 3D geometry outdoors (Piekarski and Thomas 2001) have required complex user interfaces on par with what is currently available on desktop workstations. The ThumbsUp user interface technology utilises a tracked set of pinch gloves that combine command entry and 3D manipulation into one user interface device. ThumbsUp enables the user to enter commands via a hierarchical menu with mapped pinch gestures and perform 3D manipulations through the tracking of the user's thumbs relative to their head position, as shown in Figure 1 and Figure 2.

Operating user interfaces for mobile computers outdoors is inherently difficult due to the large and dynamic nature of outdoor environments. The computing technology must be mobile to allow the user to roam freely in this environment. Restrictions on size, performance, electric power consumption, weight,

Figure 1 – Example showing a mobile AR user performing an interactive rotation operation on a virtual table at a distance outdoors, wearing the Tinmith-Endeavour backpack
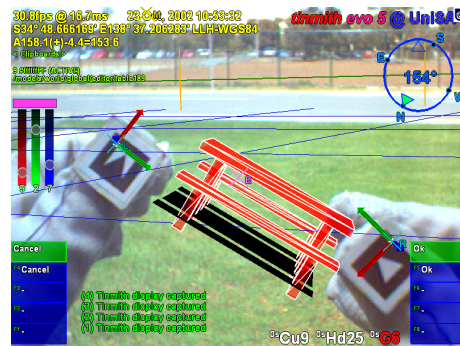

Figure 2 – Immersive AR view of figure 1, showing the virtual table at a distance being rotated with the hands, implemented using new AR extensions to existing image plane techniques

and magnetic interference limit the options of devices for use outdoors. Although technology improves from year to year, we are designing user interfaces based on vision tracked gloves that takes advantage of technology available today. Other recent input devices for mobile user interfaces are implemented using a variety of hardware, such as with ultrasonics (Foxlin and Harrington 2000) or accelerometers (Cheok, Kuman and Prince 2002). Mobile computers can now perform realistic rendering for augmented reality; therefore good user interfaces are now required to support powerful new applications, with particular attention to the limitations of the technology outdoors.

One application domain we are currently investigating is outdoor augmented reality 3D modelling, where a user can capture the models of existing large structures (such as buildings), or prototype plans for new objects that may be constructed in the future. We see this form of application improving design and planning in areas such as landscape design, building construction, and surveying. In order to control a complex modelling application with many features, we developed the ThumbsUp user interface and evaluated it on a number of users to iteratively refine it. Other application areas are explored in systems such as the Touring machine (Feiner, MacIntyre and Hollerer 1997) and Studierstube system (Reitmayr and Schmalstieg 2001).

Our user interface is made up of three components: a 3D tracked pointer using gloves on the user's hands; a command entry system where the user's fingers interact with a menu for performing actions; and an augmented reality display that presents the results back to the user. These components are used to interact with a virtual environment, in this case implemented as outdoor augmented reality. The hands free nature of ThumbsUp does not require interaction props, allowing the user to freely move about the real world without restrictions. Our investigation into ThumbsUp has leveraged current research into different 3D interaction techniques, and complements rather than replaces existing techniques. Interaction techniques for outdoor augmented reality (OAR) are a subset of the augmented reality (AR) and virtual reality (VR) domains, which are a further subset of virtual environments (VE) and 3D interfaces.

Our interaction techniques use natural head and hand movements to specify operations during the construction of 3D graphical models outside. 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 when dealing with the physical world. By using techniques such as the pointing and grabbing of objects in positions relative to the body, user interfaces can leverage the user's inbuilt knowledge (known as Proprioception) as to what their body is doing. Mine et al (Mine, Brooks and Sequin 1997) demonstrated that designing user interfaces to take advantage of these human proprioceptive capabilities produced improved results. We were also inspired by the elegant combination of commands and pointing by the *Put-That-There* system (Bolt 1980). The user interface we have developed uses similar techniques, with the focus being the user's region of interest framed by their current field of view. Commands and pointing both operate within this view, and building on this is the use of the user's physical presence (location and orientation) to aid with their interactions. For example, in a scenario where a user wants to create an outdoor scene of a garden, the order of operations would be as follows: first, specify the prefabricated object to create, such as a table; second, use AR image plane techniques to slide the table into position from different angles; third, scale and rotate the table (such as shown in Figure 1 and Figure 2); and finally, walk away from the table to preview at a distance its placement.

2 Current applications

As previously mentioned, one application domain we believe augmented reality will be used in the future for is modelling of 3D geometry, allowing the preview of non-existent objects, and the capture of existing geometry that can then be modified to view what the proposed changes will be. The Tinmith-Metro application (Piekarski and Thomas 2001) implemented simple building construction using the *infinite planes technique* (by placing down large planes and combining them with constructive solid

geometry operations), and the placement and manipulation of *street furniture* objects, both with the user interface described in this paper. These techniques allow users to capture the geometry of outdoor objects without having to actually stand next to or on top of them. The user can model the objects from a distance, with partial occlusion of the real world as if the objects were physically present. This is an advantage over existing techniques, such as: 1) photo and laser based scanning, requiring a full view of the object; 2) using GPS waypoints, not working well near large buildings; and 3) standing on top of the building, possibly being not possible or too dangerous.

3 Interface overview

The user interface can be described as two separate components, a tracked 3D cursor for selection and manipulation, and a special menu for controlling the system and entering commands. The menu is fixed to the user's display and presents up to ten possible commands that are available at that moment. Eight of these commands are mapped to the fingers as shown in Figure 3, and the user activates a command by pressing the appropriate finger against the thumb. At this point, the menu refreshes to reflect the selection made, and the next series of commands are then made available to the user. *Ok* and *cancel* operations are indicated by pressing the fingers into the palm of the appropriate hand, depending on which is selected by the user as being their dominant hand, and these are indicated in the topmost boxes in the menu. The 3D cursor is implemented using vision tracking techniques (Kato and Billinghurst 1999) and fiducial markers placed on the tips of the thumbs. Using this tracking, and combining this with the previously described command system, the user interface can perform selection, manipulation, and creation operations by pointing into the virtual environment.

The design of the menu is based around users executing commands through direct finger mappings, without requiring them to lift up their hands to interact with the menu. This allows users to perform cursor operations with their hands without having to move them to execute commands. Traditional VR systems require the user to select a mode from a menu, and then interact with an object. With our design, operations may be performed without having to take the hands away from the task at hand.
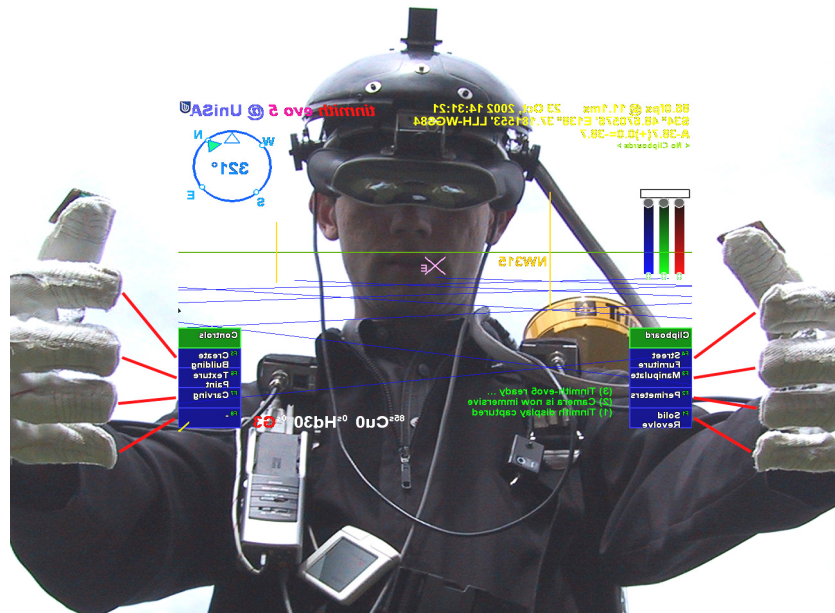

Figure 3 – Each finger maps to a menu option, the user selects one by pressing the appropriate finger against the thumb, and does not rely on the position of the hands when navigating through the menu hierarchy

4 Cursor operations

Tinmith-Metro is our original AR outdoor 3D modelling application (Piekarski and Thomas 2001) that performed the placement of outdoor street furniture and the capture of simple building shapes. The Tinmith-Metro application extends previous image plane techniques (Pierce, Forsberg, Conway, Hong, Zeleznik and Mine 1997) to support object manipulation (translate, rotate, scale) and object selection in mobile augmented reality. Figure 1 and Figure 2 is an example showing a virtual table that has been placed down on the ground and is being manipulated into the correct position using these techniques.

This section discusses the features of the 3D cursor that forms an integrated part of the user interface, performing direct manipulation operations such as selection and object transformation. Interacting with 3D graphical objects in an outdoor environment is implemented using vision based hand tracking. To create new objects and then edit them (scale, rotate, translate, carving), we provide a number of interactions, using a combination of zero, one, and two handed input techniques, depending on what is most appropriate for the task. We implement each transformation technique as a separate command. This is on the assumption the user will wish to work with certain degrees of freedom without affecting the others. This constraining of degrees of freedom is useful to compensate for most users inability to maintain the exact position and orientation of their hands simultaneously (Hinckley, Pausch, Goble and Kassell 1994), and in environments with poor vision tracking.

Using a single hand, an object can be translated in the scene. To perform translation, it must be selected first, and the hand is brought into view so the cursor can be placed on top of the object. Using extended AR image plane techniques, a user is able to move the object against the view plane fixed to their head. By rotating the head and keeping the hands at the same point in the image plane, the object can be dragged around the user's body since our techniques maintain the same distance.

In order to provide more natural manipulation techniques for operations like scaling and rotation, it has been shown that using two hands for interaction can improve performance and accuracy. The two handed interaction ideas used for these transformations were initially pioneered in a study in 2D environments by (Buxton and Myers 1986). Although our work is different in that we are working at a distance in absolute coordinates (rather than directly on the object), the previous work is very useful in showing possible approaches, and how these tasks can be improved with two hands. We make use of the two hands by having the angle between the dominant and non-dominant hand control the rotation. This technique is also implemented using our AR extensions to image planes, and can be configured for either left or right hand dominance. Figure 1 and Figure 2 show a rotation operation being performed on a virtual object at a distance from the user. Since the tracking system used produces high quality position and low quality rotation values, the use of two hands allows rotations to be specified through only two position values, maximising the accuracy of the operation.

The user interface has very powerful command menus to perform a number of manipulation and creation operations without requiring the hands to be visible. The nudge commands allow the user to use precise manipulations based on fixed increments in order to accurately work with objects, and are most useful for altering the distance of an object (which is not possible using image planes since the distance is fixed) or for when very precise fixed movements are required. The eye cursor is used to create objects relative to the front of the user's body.

5 Conclusion

In this paper, we have presented a set of new user interface technologies we have developed for use in mobile outdoor augmented reality systems. Manipulation of 3D virtual artefacts in an outdoor setting requires different user interface technologies to traditional indoor AR and VR systems, due to the difference in tracking hardware and input devices. ThumbsUp is a new user interface technology that integrates a 3D cursor for selection and manipulation with a special menu for entering commands to

support mobile outdoor augmented reality systems. The 3D cursor is controlled by a vision tracking system with fiducial markers placed on the tips of the thumb. The user interface can perform selection, manipulation, and creation operations by pointing into the virtual environment. A number of interaction modes (zero, one, or two handed input techniques) are provided to manipulate objects, such as translation, scaling, and rotation. The menu system is screen relative and presents up to ten possible commands that are available at that time. Each of these commands is mapped directly to the user's fingers, and the user activates a command by pressing the appropriate finger against the thumb.

6 Acknowledgements

7 References

Bolt, R. A. (1980): "Put-That-There" : Voice and Gesture at the Graphics Interface. In *ACM SIGGRAPH 1980,* pp 262-270, Seattle, Wa, Jul 1980.

Buxton, W. and Myers, B. A. (1986): A Study In Two-Handed Input. In *CHI - Human Factors in Computing Systems,* pp 321-326, Boston, Ma, 1986.

Cheok, A. D., Kuman, K. G., and Prince, S. (2002): Micro-Accelerometer Based Hardware Interfaces for Wearable Computer Mixed Reality Applications. In *6th Int'l Symposium on Wearable Computers,* pp 223-230, Seattle, Wa, Oct 2002.

Feiner, S., MacIntyre, B., and Hollerer, T. (1997): A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment. In *1st Int'l Symposium on Wearable Computers,* pp 74-81, Cambridge, Ma, Oct 1997.

Foxlin, E. and Harrington, M. (2000): WearTrack: A Self-Referenced Head and Hand Tracker for Wearable Computers and Portable VR. In *4th Int'l Symposium on Wearable Computers,* pp 155-162, Atlanta, Ga, Oct 2000.

Hinckley, K., Pausch, R., Goble, J. C., and Kassell, N. F. (1994): A Survey of Design Issues in Spatial Input. In *7th Int'l Symposium on User Interface Software Technology,* pp 213-222, Marina del Rey, Ca, Nov 1994.

Kato, H. and Billinghurst, M. (1999): Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System. In *2nd Int'l Workshop on Augmented Reality,* pp 85-94, San Francisco, Ca, Oct 1999.

Mine, M., Brooks, F. P., and Sequin, C. H. (1997): Moving Objects In Space: Exploiting Proprioception In Virtual-Environment Interaction. In *ACM SIGGRAPH 1997,* pp 19-26, Los Angeles, Ca, Aug 1997.

Piekarski, W. and Thomas, B. H. (2001): Tinmith-Metro: New Outdoor Techniques for Creating City Models with an Augmented Reality Wearable Computer. In *5th Int'l Symposium on Wearable Computers,* pp 31-38, Zurich, Switzerland, Oct 2001.

Pierce, J. S., Forsberg, A., Conway, M. J., Hong, S., Zeleznik, R., and Mine, M. R. (1997): Image Plane Interaction Techniques in 3D Immersive Environments. In *1997 Symposium on Interactive 3D Graphics,* pp 39-43, Providence, RI, Apr 1997.

Reitmayr, G. and Schmalstieg, D. (2001): Mobile Collaborative Augmented Reality. In *Int'l Symposium on Augmented Reality,* pp 114-123, New York, NY, Oct 2001.