A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment

Steven Feiner, Blair MacIntyre, Tobias Höllerer

Department of Computer Science
Columbia University
New York, NY 10027
{feiner,bm,htobias}@cs.columbia.edu
http://www.cs.columbia.edu/graphics/

Anthony Webster
Graduate School of Architecture, Planning and Preservation
Columbia University
New York, NY 10027
acw18@columbia.edu
http://www.cc.columbia.edu/~archpub/BT/

Abstract

We describe a prototype system that combines together the overlaid 3D graphics of augmented reality with the untethered freedom of mobile computing. The goal is to explore how these two technologies might together make possible wearable computer systems that can support users in their everyday interactions with the world. We introduce an application that presents information about our university's campus, using a head-tracked, see-through, head-worn, 3D display, and an untracked, opaque, handheld, 2D display with stylus and trackpad. We provide an illustrated explanation of how our prototype is used, and describe our rationale behind designing its software infrastructure and selecting the hardware on which it runs.

Keywords: Augmented Reality, Virtual Environments, Mobile Computing, Wearable Computing, GPS.

1. Introduction

Recent years have seen significant advances in two promising fields of user interface research: virtual environments, in which 3D displays and interaction devices immerse the user in a synthesized world, and mobile computing, in which increasingly small and inexpensive computers and wireless networking allow users to roam the real world without being tethered to stationary machines. We are interested in how virtual environments can be combined with mobile computing, with the ultimate goal of supporting ordinary users in their interactions with the world.

To experiment with these ideas, we have been building the system described in this paper. The kind of virtual environment technology with which we have been working is augmented reality. Unlike most virtual environments, in which a virtual world replaces the real world, in augmented reality a virtual world supplements the real world with additional information. This concept was pioneered by Ivan Sutherland [27], and is accomplished through the use of tracked “see-through” displays that enrich the user’s view of the world by overlaying visual, auditory, and even haptic, material on what she experiences.

The application that we are addressing is that of providing users with information about their surroundings, creating a personal “touring machine.” There are several themes that we have stressed in this work:

- Presenting information about a real environment that is integrated into the 3D space of that environment.
- Supporting outdoor users as they move about a relatively large space on foot.
- Combining multiple display and interaction technologies to take advantage of their complementary capabilities.

Our prototype assists users who are interested in our university’s campus, overlaying information about items of interest in their vicinity. As a user moves about, she is tracked through a combination of satellite-based, differential GPS (Global Positioning System) position tracking and magnetometer/inclinometer orientation tracking. Information is presented and manipulated on a combination of a head-tracked, see-through, head-worn, 3D display, and an untracked, opaque, handheld, 2D display with stylus and trackpad.

Our emphasis in this project has been on developing experimental user interface software, not on designing hardware. Therefore, we have used commercially available hardware throughout. As we describe later, this has necessitated a number of compromises, especially in the accuracy with which the user’s 3D position and orientation is tracked. These have in turn affected the design of our user interface, which relies on approaches that require only approximate, rather than precise, registration of virtual and real objects.

In Section 2 we present related work. Section 3 describes a scenario in our application domain, including pictures generated by a running testbed implementation. In Section 4, we describe both our high-level approach in designing our system and the specific hardware and software used. Finally, Section 5 presents our conclusions and
the directions that we will be taking as we continue to
develop the system.

2. Related Work

Previous research in augmented reality has addressed a
variety of application areas including aircraft cockpit con-
trol [12], assistance in surgery [26], viewing hidden build-
ing infrastructure [10], maintenance and repair [9], and
parts assembly [5, 29]. In contrast to these systems, which
use see-through headworn displays, Rekimoto [23] has
used handheld displays to overlay information on color-
coded objects. Much effort has also been directed towards
developing techniques for precise tracking using tethered
trackers (e.g., [16, 2, 28, 25]).

Work in mobile user interfaces has included several
projects that allow users to explore large spaces. Loomis
and his colleagues have developed an application that
makes it possible for blind users to navigate a university
campus by tracking their position with differential GPS
and orientation with a magnetometer to present spatialized
sonic location cues [18]. Petrie et al. have field-tested a
GPS-based navigation aid for blind users that uses a speech
synthesizer to describe city routes [22]. The CMU Wear-
able Computer Project has developed several generations
of mobile user interfaces using a single handheld or
untracked headworn display with GPS, including a campus
tour [24]. Long et al. have explored the use of infrared
tracking in conjunction with handheld displays [17]. Mann
[20] has developed a family of wearable systems with
headworn displays, the most recent of which uses optical
flow to overlay textual information on automatically recog-
nized objects.

Our work emphasizes the combination of these two
streams of research: augmented reality and mobile user
interfaces. We describe a prototype application that uses
tracked see-through displays and 3D graphics without
assuming precise registration, and explore how a combina-
tion of displays and interaction devices can be used
together to take advantage of their individual strengths.

Prior to the development of VRML, several researchers
experimented with integrating hypertext and virtual envi-
ronments [7, 8, 1]. All investigated the advantages of pre-
senting hypertext on the same 3D display as all other
material, be it headworn or desktop. In contrast, our current
work exploits the different capabilities of our displays by
presenting hypertext documents on the relatively high-res-
olution 2D handheld display, which is itself embedded
within the 3D space viewed through the lower-resolution
headworn display.

Figure 1. Prototype campus information system. The
user wears a backpack and headworn display, and
holds a handheld display and its stylus.

3. Application Scenario

Consider the following scenario, whose figures were
created using our system. The user is standing in the mid-
dle of our campus, wearing our prototype system, as shown
in Figure 1. His tracked see-through headworn display is
driven by a computer contained in his backpack. He is
holding a handheld computer and stylus.

As the user looks around the campus, his see-through
headworn display overlays textual labels on campus build-
ings, as shown in Figures 2 and 3. (These image were shot
through the headworn display, as described in Section 4.3,
and are somewhat difficult to read because of the low
brightness of the display and limitations of the recording
technology.) Because we label buildings, and not specific
building features, the relative inaccuracy of the trackers we
are using is not a significant problem for this application.

At the top of the display is a menu of choices: “Colum-
bia?”, “Where am I?”, “Depts?”, and “Buildings?”. When
selected, each of these choices sends a URL to a web
browser running on the handheld computer. The browser
then presents information about the campus, the user’s cur-
rent location, a list of departments, and a list of buildings,
respectively. The URL points to a custom HTTP server on
the handheld computer that generates a page on the fly con-
taining the relevant information. The generated pages con-
tain links back to the server itself and to regular web pages.
elsewhere. (An additional menu item, “Blank”, allows the headworn display to be blanked when the user wants to view the unaugmented campus.) Menu entries are selected using a touchpad mounted on the back of the handheld computer. The touchpad’s x coordinates are inverted to preserve intuitive control of the menus.

Labels seen through the headworn display are grey, increasing in intensity as they approach the center of the display. The one label closest to the center is highlighted yellow. If it remains highlighted for more than a second, it changes to green, indicating that it has been selected, and a second menu bar is added below the first, containing entries for that building. A selected building remains selected until the user’s head orientation dwells on another building for more than a second as indicated by the color change. This approximation of gaze-directed selection can be disabled or enabled through a menu item.

When a building is selected, a conical green compass pointer appears at the bottom of the headworn display, oriented in the building’s direction. The pointer turns red if the building is more than 90 degrees away from the user’s head orientation. This allows the user to find the building more easily if they turn away from it. The pointer is especially useful for finding buildings selected from the handheld computer. This is made possible by our custom HTTP server, which can tell the backpack computer to select a building on the headworn display.

The building’s menu bar contains the name of the building, plus additional items: “Architecture”, “Departments”, and “Miscellaneous”. Selecting the name of the building from the menu using the trackpad sends a relevant URL to the handheld computer’s browser. Selecting any of the remaining menu entries also sends a URL to the browser and creates a collection of items that are positioned near the building on the headworn display.

4. System Design

While we wanted our system to be as lightweight and comfortable as possible, we also decided to use only off-the-shelf hardware to avoid the expense, effort, and time involved in building our own. Consequently we often settled for items that were far bulkier than we would like them to be, in return for the increased flexibility that they offered. The combined weight of the system is just under 40 pounds.
Figure 4. After the "Departments" menu item is selected, the department list for the Philosophy Building is added to the world, arrayed about the building. The three figures show the label animation sequence: (a) a fraction of a second after selection, (b) approximately half a second later, and (c) after the animation has finished.

Figure 5. (a) Selecting the "Departments" menu item causes an automatically-generated URL to be sent to the web browser on the handheld computer, containing the department list for the Philosophy Building. (b) Actual home page for the English and Comparative Literature department, as selected from either the generated browser page or the department list of Figure 4.

The following subsections describe some of the hardware and software choices that we made in designing our system, whose hardware design is diagrammed in Figure 6.

4.1. Hardware

Backpack computer. It was important to us that our main computer not only be portable, but also capable of working with readily available peripherals, including high-performance 3D graphics cards. We chose a Fieldworks 7600, which includes a 133MHz Pentium, 64Mbyte memory, 512K cache, 2GB disk, and a card cage that can hold 3 ISA and 3 PCI cards. While this system is our biggest compromise in terms of weight and size, it has significantly simplified our development effort.
many applications, we feel that augmented reality systems will become commonplace only when they truly add to reality, rather than subtract from it. In our work we have selected the relatively lightweight Virtual I/O i-glasses head-worn display. This is a 60,000-tri color display. We are also experimenting with a Virtual I/O 640x480 resolution greyscale display.

Orientation tracker. We use the built-in tracking provided with our head-worn display. This includes a magnetometer, which senses the earth’s magnetic field to determine head yaw, and a two-axis inclinometer that uses gravity to detect head pitch and roll.

Position tracking. We use a Trimble DSM GPS receiver to obtain position information for its antenna, which is located on the backpack above the user’s head. While normal GPS generates readings that are accurate only within about 100 meters, it can be routinely coupled with correction information broadcast from another receiver at a known location that contains information about how far it is off. We subscribe to a differential correction service provided by Differential Corrections Inc., which allows us to achieve about one-meter accuracy.

Network. To provide communication with the rest of our infrastructure we use NCR WaveLan spread-spectrum 2Mbit/sec radio modems in both the backpack and handheld PCs, which operate with a network of base stations on campus.

Power. With the exception of the computers, each of the other hardware components has relatively modest power requirements of under 10 watts each. We run them all using an NRG Power-MAX NiCad rechargeable battery belt. It has the added advantage of allowing a fully charged replacement powerpack to be plugged in prior to unplugging the depleted powerpack, without interrupting power.

4.2. Software

Infrastructure. We use COTERIE [19], a system that provides language-level support for distributed virtual environments. COTERIE is based on the distributed data-object paradigm for distributed shared memory. Any data object in COTERIE can be declared to be a shared object that either exists in one process, and is accessed via remote-method invocation, or is replicated fully in any process that is interested in it. The replicated shared objects support asynchronous data propagation with atomic serializable updates, and asynchronous notification of updates. COTERIE runs on Windows NT/95, Solaris, and IRIX, and includes the standard services needed for building virtual environment applications, including support for assorted trackers, etc. This software is built on top of Modula-3 [14] and Obliq [4].

Graphics package. We use a version of Obliq-3D [21], a display-list based 3D graphics package, which we have modified both to provide additional features needed for vir-
Application software. The prototype comprises two applications, one running on each machine, implemented in approximately 3600 lines of commented Obliq code. Figure 7 shows the overall software structure.

The tour application running on the backpack PC is responsible for generating the graphics and presenting it on the headworn display. The application running on the handheld PC is a custom HTTP server in charge of generating web pages on the fly and also accessing and caching external web pages by means of a proxy component.

One of the main reasons that we run our own HTTP server on the handheld display is that it gives us the opportunity to freely react to user input from the web browser. For example, when a URL is selected on the handheld display, the HTTP server can call a network object method that selects corresponding graphical items on the headworn display. Thus data selection works in both directions: from the backpack PC to the handheld PC (by launching relevant URLs from the headworn display’s menus) and vice versa (selecting buildings, departments, etc. on the headworn display from a link on the handheld’s browser).

As shown in Figure 7, the HTTP server has two components: the campus information server, responsible for the dynamic generation of HTML pages, and a caching proxy server. The purpose of the proxy server is to cache the data returned by external HTTP requests to mitigate the slowness of the radio network link. In addition, commonly accessed pages, such as department home pages and building descriptions, can be pre-cached without relying on the browser’s own caching mechanisms.

The HTTP server is initialized by the tour application running on the backpack PC. Each piece of information (buildings, departments, their whereabouts, and assorted URLs) in the tour data on the backpack PC is sent to the handheld PC’s HTTP server with an accompanying procedure closure. The closure executes a procedure on the backpack PC when the corresponding link is selected on the web browser. This makes it possible for the handheld display to control the headworn display, as described in Section 3.

The web browser on the handheld PC is a totally separate process. It can be pointed at URLs from within the campus information server, which we currently accomplish by forking off a separate URL pusher process. The web browser then issues a request back to the HTTP server to obtain either a locally generated, cached external, or uncached external HTML document.

The tour application continuously receives input from the GPS position tracker and the orientation tracker. It also takes user input from the trackpad that is physically attached to the back of the handheld PC. Based on this input and a database of information about campus buildings, it generates the graphics that are overlaid on the real world by the headworn display.
4.3. Figures

Figures 2–4 were created using a dummy head whose right eyeocket contains a Toshiba 1K-M41A 410,000 pixel miniature color CCD camera, shown in Figure 8. The Virtual I/O display was worn on the head, which was carried by one of the experimenters. The video was recorded in Hi-8 format, which was later framegrabbed to create the images.

5. Conclusions and Future Work

We have described a prototype mobile, augmented-reality application that explores approaches to outdoor navigation and information-seeking on our campus. Thus far, our project has been used only experimentally by the authors as a research prototype with which to explore issues in software design for future user interfaces. Although we feel that it provides a good testbed environment, there are many technical issues that will need to be addressed for commercial versions of such systems to become practical:

Quality of displays. The low brightness of the headworn display's LCD necessitates the use of neutral density filters. Low brightness of the handheld display makes reading outside difficult in sunlight. Headworn display resolution is currently quite low, with color VGA resolution systems only beginning to become affordable.

Quality of tracking. Although we believe that approximate tracking can be extremely useful, there are many applications that require precise tracking. We are in the process of replacing the magnetometer/inclinometer contained in the headworn display with a higher-quality unit, and are considering obtaining a gyroscopic system for hybrid tracking. We will also be exploring 3D tracking of the handheld computer [11] and the user's stylus. Better outdoors position tracking can be addressed through real-time kinematic GPS systems, which can achieve centimeter-level accuracy. These are largely temporary solutions, given the inherent problems of electromagnetic and inclinometer-based approaches, and the line-of-sight restrictions of GPS mentioned below. However, we believe that camera-based approaches [28, 20] are a promising way to address the problem.

Loss of tracking. While GPS doesn't present any practical range restrictions for our work, it does not work if an insufficient number of satellites are directly visible. GPS satellite signals are weak and are blocked by intervening buildings and even foliage. While our system works on a large portion of our campus, there are far too many areas in which it does not, including outdoor sites shaded by trees and nearby buildings, and most indoor sites.

Currently, we indicate loss of tracking, but do not attempt to compensate for it. For example, we could point the user back to where they were last tracked, based on their orientation. Since tracked sites can be predicted based on satellite ephemeris information broadcast to the GPS receiver, combined with known campus topology, we could also direct the user toward other reliably tracked sites either on the headworn display or on the 2D absolute space of a map viewed on the handheld computer. GPS can also be used with inertial systems that temporarily extrapolate position when tracking is lost. Eventually GPS techniques may be used with spread-spectrum radio transmitters to support precise tracking in large indoor spaces [3].

We are working on several extensions to our work:

Overlaying virtual objects on the real world can potentially create a good deal of confusion if they interfere with the user's view of the real world and of each other. For example, even the relatively sparse overlaid graphics of Figures 2–4 evidence problems caused by self-occlusion. We are currently incorporating the Snap-Together Math constraint-based toolkit [13] into our system to explore how automated satisfaction of geometric constraints among objects could help maintain display layout quality as the user moves about.

We are extending our application domain to include 3D models of underground campus infrastructure, in the spirit of our earlier indoor work on using augmented reality to present hidden architectural infrastructure [10]. In another direction, we are beginning to work with our colleagues in the Graduate School of Journalism to explore the potential for presenting additional multimedia information in the spatial context of the campus.

6. Acknowledgments

We would like to thank Xinshi Sha for assistance in developing COTERIE, and Ruigang Yang for Windows 95 utilities. Christina Vernon and Alex Kleviskky helped create the campus databases. Sean Eno, Damijan Saccio and Scott
Sindorf developed VRML building models that we are incorporating. Jim Foley of MERL, a Mitsubishi Electric Research Laboratory, generously provided a Mitsubishi Amity and Reunen Kobloch of Mitsubishi Electric ITA Horizon Systems Laboratory assisted us with the Amity; Jim Spohrer of Apple generously provided us with several Apple Newton handhelds, currently being integrated into the project, and Steve Weyer provided software advice. The folks at Digital Equipment Corporation’s System Research Center and at Critical Mass, Inc., helped during development of our infrastructure, especially Marc Najork and Bill Kalsow.

This work was supported in part by the Office of Naval Research under Contracts N00014-94-1-0564 and N00014-97-1-0838, the Columbia Center for Telecommunications Research under NSF Grant ECD-88-11111, a Columbia University Provost’s Strategic Initiative Fund Award, and a gift from Microsoft.

7. References


