An Annotated Situation-Awareness Aid for Augmented Reality

Blaine Bell   Tobias Höllerer   Steven Feiner

Department of Computer Science, Columbia University

500 W. 120th St., 450 CS Building

New York, NY 10027

+1-212-939-7000

{blaine, htobias, feiner}@cs.columbia.edu

ABSTRACT

We present a situation-awareness aid for augmented reality systems based on an annotated "world in miniature." Our aid is designed to provide users with an overview of their environment that allows them to select and inquire about the objects that it contains. Two key capabilities are discussed that are intended to address the needs of mobile users. The aid's position, scale, and orientation are controlled by a novel approach that allows the user to inspect the aid without the need for manual interaction. As the user alternates their attention between the physical world and virtual aid, popup annotations associated with selected objects can move freely between the objects' representations in the two models.

Keywords

World in miniature, augmented reality, mobile computing, user interaction.

INTRODUCTION

In augmented reality systems, in which synthesized graphics annotate the surrounding world, user interfaces contain both virtual and physical elements. We are interested in how an augmented reality user interface can assist a user in understanding their environment. This is an especially important task if the user is mobile and the environment unfamiliar: the virtual overlays need to enrich and explain, rather than clutter and confuse, the user's physical surroundings.

We introduce a situation-awareness aid for augmented reality that is intended to provide the user with an overview of the surrounding environment and the ability to discover, select, and inquire about objects that may not be directly visible to the user within that environment. Our aid is based on a *world in miniature* (WIM) [13] or exocentric "god's eye view" [4]: a miniature overview model of the surrounding environment that is embedded within the user's view of that environment. The work described here addresses two key issues. First, we introduce an approach for controlling the position, scale, and orientation of our aid as a function of head orientation alone. This is designed to make it easy for the user in a mobile, (mostly) hands-free application to

determine how much attention they wish to devote to the aid. Second, we describe how these changes in user focus produce complementary modifications in the ways in which overlaid annotations are presented in the full-scale physical environment and in the miniature virtual aid.

In the rest of this paper, we begin by describing related work and how our approach builds on it. Then, we introduce our method for controlling the user's view of the aid, and describe how annotations are managed in the surrounding environment and the aid. Finally, we present our experience with the system and describe ongoing and future work.

RELATED WORK

The concept of a miniature "god's eye view" model of the user's environment, embedded within the full-scale environment viewed on a head-worn display, is sketched by Furness in a description of the user interface for the Air Force Visually Coupled Airborne Systems Simulator [4]. Similarly, Darken and Sibert [2] place a scaled map of the user's environment at a fixed position at the bottom of the user's view in a head-tracked environment intended to assist in search tasks, and experiment with controlling the map's yaw (*y*-axis orientation) so that it is always aligned with the environment.

Stoakley, Conway, and Pausch [13] expand these ideas to create a *world in miniature* (WIM), a miniature model of the user's environment that is affixed to a 6DOF tracked clipboard held in the user's non-dominant hand and manipulated with a 6DOF device held in the dominant hand. Their WIMs are viewed in context of a full-scale virtual environment and used to support selection and manipulation of objects, including a representation of the user that makes possible rapid locomotion. To avoid the visual confusion that could result when moving the user's representation in a WIM, Pausch and colleagues [10] defer the viewpoint change in the full-scale environment until the manipulation is complete, and then interpolate the user's current viewpoint into the new viewpoint within the WIM itself. Fukatsu and colleagues [3] take an alternative approach in which a hand-held 6DOF tracker is used to manipulate a bird's eye view inset into the user's first person view of a virtual environment. Koller and colleagues [7] use head orientation alone to orbit the user around an object, such as a WIM, and head or hand position to move the camera closer to or farther from the center of rotation. In another approach to hands-free interaction, LaViola and colleagues [8] control a world-stabilized WIM by walking, toe and heel clicking, and leaning.
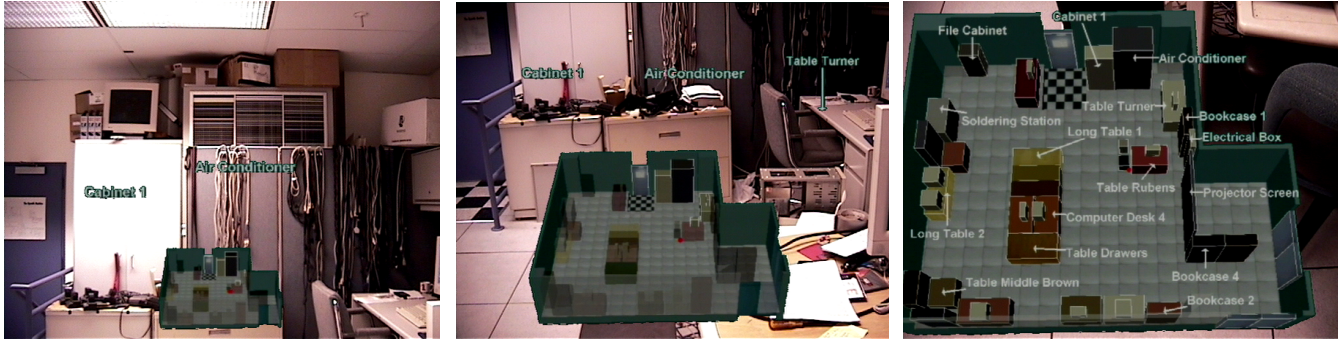
Figure 1. Annotated situation-awareness aid. (a) Looking slightly upward. (b) Looking slightly downward. (c) Looking nearly straight down.

We adopt the head-tracked yaw control implemented by Darken and Sibert, and some of the selection and manipulation capabilities of Stoakley, Conway, and Pausch (although without the self-locomotion capabilities, which would not be appropriate for our augmented reality applications). However, instead of assigning our aid a fixed position in the view frustum or requiring that the user exert full 6DOF control over it (as in previous work), we map user head pitch alone to the aid's pitch, position and scale. Unlike Koller and colleagues, we do not need to use head or hand position to control the perceived size of the aid and we generalize the mapping between head orientation and the way the aid is displayed. Our mapping is separately piecewise linear for each parameter, as described in the next section, and thus has some of the flavor of nonlinear manipulation techniques, such as the *Go-go* technique [12], in which the user's virtual hand can stretch nonlinearly to manipulate distant objects. An analogy can also be drawn to *pre-screen manipulation* [5], in which small changes in head position and orientation relative to a desktop monitor produce disproportionately large changes in the user's view of the entire environment. The ease of controlling our interface through head orientation is also closely related to gaze control [6], in which hands-free interfaces react to the user's eye movement.

Pierce and colleagues [11] describe how one of a set of storage spaces ("toolspaces") positioned around a user's virtual body and outside of the view frustum can be quickly accessed with a touchpad gesture (a manually invoked "glance") to deposit or retrieve objects, potentially including WIMs. Glancing into a toolspace is modal. The original view is restored when the finger is lifted off the pad to terminate the gesture, making it easy to exit the mode, much as in conventional click-and-drag mouse manipulation. Our system supports a similar approach to modality, in that the aid is accessed by looking down, and retreats into the background as the user looks up again.

Objects are annotated in our aid using an interactive annotation placement algorithm based on that of Bell, Feiner, and Höllerer [1]. Unlike this previous work, which treated only a single view of the environment, we address how annotations can be shared and handed off between multiple views, as described in the section on Shared Annotations.

HEAD-MOTION INTERACTION
The user of our system is immersed within a full-scale physical environment, most of which is usually viewed directly through a stereo, optical see-through, head-worn display. Overlaid graphics can be enabled to annotate the surrounding world; for example, to label objects or to provide detailed information about them. The

aid, which can also be selectively enabled, displays the user's environment as a similarly annotated, perspective-projected WIM, containing schematic virtual representations of the physical objects. The WIM is located significantly ahead of the user to make stereo fusion easier (currently 4m, determined pragmatically for our application). A red dot represents the user's current position, and objects in the WIM are colored differently depending upon whether they are (partly) visible or fully invisible in the real world from the user's actual viewpoint, using an analytic visible surface algorithm.

We are interested in mobile applications that might require the user's hands to be free for real world tasks. Furthermore, accurate position tracking might not always be available. Therefore, we decided that the aid should be controlled by head orientation only. The aid's yaw is fixed to that of the surrounding world, and thus changes directly with the user's head yaw [2]. In contrast, the position, scale, and pitch of the aid, as well as the decision to annotate its contents, are controlled by head pitch.

Figure 1 shows our laboratory, viewed through our system, with the aid enabled. In (a), the user looks slightly upward and the unannotated aid stays close to the bottom of the viewport, pitched towards the view plane by a default angle (22.5˚) about the user's position in the aid. In (b), the user looks slightly downward and the aid is scaled to be slightly larger, angled slightly closer to parallel to the view plane, and moves up slightly higher in the viewport. In (c), the user looks nearly straight down, and in response receives a zoomed-in, annotated view, that is nearly parallel to the view plane, with the user's position in the aid located at the center of the viewport.

The roll component of the aid's orientation is kept at zero, so that its ground plane is always parallel to the bottom of the viewport, but not necessarily parallel to the ground plane in the physical world. Note that the only effect of the user's position on the aid is to control the "you are here" point about which the aid orients (in yaw and pitch). This point is positioned at the center of the viewport's width and at a head-pitch–determined location relative to the viewport's height.

Figure 2 shows how the user's head pitch determines the aid's pitch, scale, viewport *y*-position, and labeling status. The parameters shown were used to make the images in Figures 1 and 3. The viewport height is *yres,* and the *max_s* scale factor is computed automatically for a given environment so that the aid fills the viewport when centered and scaled fully. The other parameters were determined pragmatically through informal tests with visitors to our lab. A threshold value determines when annotation is triggered (annotations are turned on in the aid when
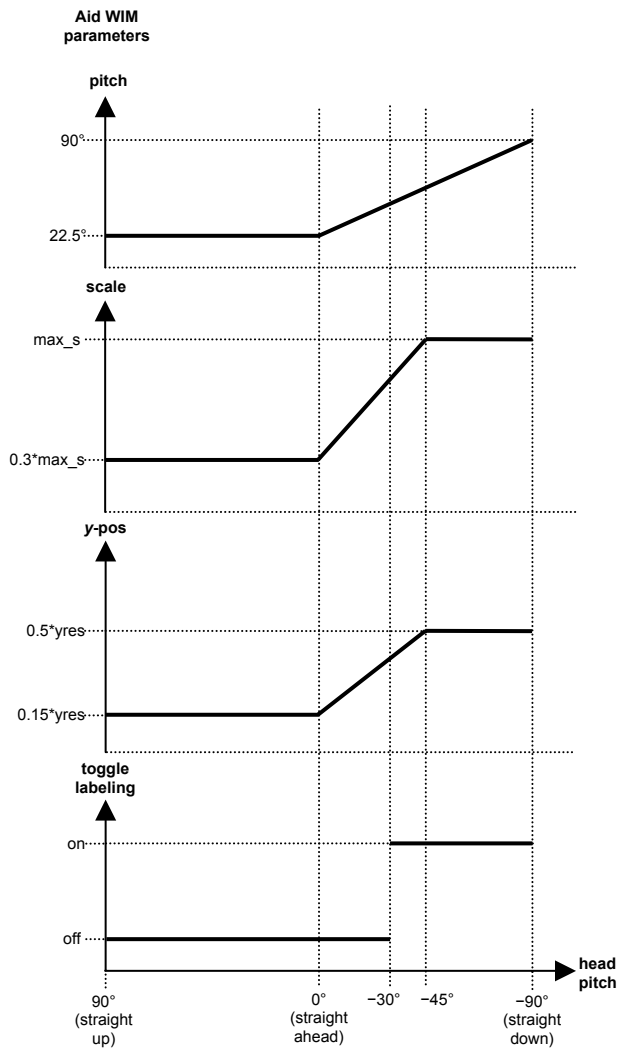
Figure 2. Mapping user head pitch to the aid's placement and labeling parameters.

they are turned off in the real world). For each of the other three variables (pitch, scale, and y position in the viewport), there is a set of four parameters: min and max values and two head pitch trigger values. We linearly interpolate the dependent variables from min to max for head pitch values in between these triggers. Note that these mappings are independent of each other, even though we noticed during our tests that it makes sense to correlate some of the trigger values for WIM animations, as shown in Figure 2.

No head pitch trigger value is set above zero degree head pitch. That means that as long as the user looks straight ahead or up, the aid occupies the smallest possible amount of screen space allowed by the parameters and stays near the bottom of the viewport. Both the scale and the vertical position of the WIM are set to reach their maximum values when the user looks down at a 45° angle.

SHARED ANNOTATIONS

Specific objects can be selected either in the physical world or the aid, to bring up additional information about the object. For example, in Figure 3 (a), the user has selected a label annotating the monitor of a computer named "rembrandt," using a wireless trackball. This interaction causes live information about the computer's resources to pop up in an annotation. The popup information is placed with an arrow that points to the computer's monitor. As the user looks down at the WIM, the annotated physical monitor is no longer visible, but the popup remains displayed. As the aid scales up, its labels appear, and the popup's arrow now points to the monitor's virtual representation in the aid, shown in Figure 3 (b). As the user looks up, the popup is dissociated from the aid and recaptured by the physical monitor in the full-scale physical environment. Thus, popup annotations can move freely between the physical world and its scaled virtual representation.

EXPERIENCE AND CONCLUSIONS

Our system is implemented in Java 1.3 and Java 3D 1.2.1.03 and runs on a variety of platforms, including the dual 1.6GHz Athalon with 512 MB of memory and *n*VIDIA GeForce3 graphics card used for the images in this paper and the accompanying videotape. The system generates between 18–60 fps, limited by the annotation placement algorithm. In the indoor examples shown here, head position and orientation tracking is performed with an InterSense IS 600 Mark 2 Plus hybrid tracker, and graphics are displayed on a Sony LDI D100B stereo optical see-through head-worn display. To avoid problems in creating sufficiently bright videos and stills when imaging through a display that significantly attenuates light, the images in this paper and in the accompanying videotape were created by tracking a video camera with the abovementioned tracker and compositing overlaid graphics in real time by chromakeying with a Videonics MXPro digital video mixer.

Our annotated aid has been tried, as part of an experimental augmented reality system, by hundreds of attendees at live demonstrations indoors at *ACM SIGGRAPH 2001* and approximately thirty attendees at live demonstrations indoors and outdoors at *IEEE and ACM ISAR 2001*. These demonstrations used annotated models of the actual demonstration venues that we prepared to allow users to find out information about their environment. Based on informal verbal user feedback obtained during the demonstrations, we found that the vast majority of these first-time users were able to use head-pitch–based interaction once we explained it to them. (A small number of users initially stooped down to look at the aid, even though they were informed that head position has no effect on the aid's scale or orientation. This is not particularly surprising, since head position does influence the user's view of the real world annotations that were demonstrated prior to showing the aid.)

To increase demonstration throughput, we typically operated the system's menus and selection facilities ourselves, rather than teaching users how to use them. However, once an object was selected for detailed annotation, we also found that the majority of users were able to use head orientation to switch between viewing that object and its annotation in the surrounding environment and in the scaled aid.

We are building on the work described here in a variety of ways. Over the coming year, we expect to design and run a formal user study that will examine the effectiveness of our aid and several other user interface techniques when performing tasks indoors and outdoors with our augmented reality system. Like previous research [13, 10], we have not addressed how to determine
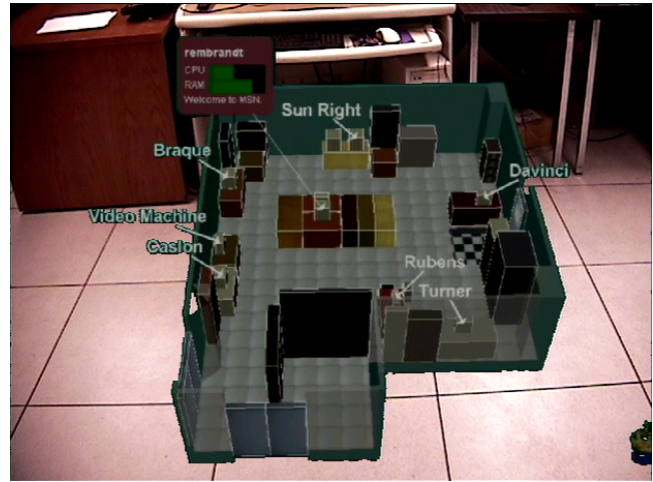
Figure 3. Transfer of annotations between the physical world and virtual aid. (a) Popup annotation provides information about the computer monitor to which its arrow points at the center. (b) As user looks down, annotations are enabled in the aid, and the popup's arrow now points at the monitor's representation in the aid.

appropriate subsets of a large environment to display in the aid; however, we have begun to apply rule-based techniques to make these decisions based on the tasks that the user is performing. Since the aid obscures objects behind it, we are using rules to constrain the aid's position and scale to avoid blocking objects that are deemed to be more important than it. We also intend to explore how multiple WIM aids can be automatically situated within the full-scale environment to annotate and assist the user in selecting environmental details that are below the resolution of our tracking systems and displays; an approach of this sort could rely on an analysis of registration errors, such as that of MacIntyre and Coelho [9].

REFERENCES
[1] B. Bell, S. Feiner, and T. Höllerer. View management for virtual and augmented reality. *Proc. ACM UIST 2001*, Orlando, FL, November 11-14 2001. 101-110.

[2] R. Darken and J. Sibert. A toolset for navigation in virtual environments. *Proc. ACM UIST '93*, New York, NY, USA, Nov. 1993. ACM Press. 157-166.

[3] S. Fukatsu, Y. Kitamura, T. Masaki and F. Kishino. Intuitive control of "bird's eye" overview images for navigation in an enormous virtual environment. *Proc. ACM VRST '98*, 67-76.

[4] T. Furness. The super cockpit and its human factors challenges. *Proc. Human Factors Society 30th Annual Meeting*. Santa Monica, CA, 1986, 48-52.

[5] D. Hix, J. Templeman, and R. Jacob. Pre-screen projection: From concept to testing of a new interaction technique. *Proc.*

CHI '95 (ACM Conf. On Human Factors in Comp. Sys.), May 1995, 226-233.

[6] A. Hyrskykari. Gaze control as an input device. Advanced Course on HCI, *Proc. ACHCI '97*, University of Tampere, Dept. of CS, Report B-1997-4, 22-27.

[7] D. Koller, M. Mine, and S. Hudson. Head-tracked orbital viewing: An interaction technique for immersive virtual environments. *Proc. ACM UIST '96,* Nov. 1996. 81-82.

[8] J. LaViola Jr., D. Feliz, D. Keefe, and R. Zeleznik. Hands-free nulti-scale navigation in virtual environments. *Proc. Int. Symp. on Interactive 3D Graphics,* March 2001. 9-15.

[9] B. MacIntyre and E. M. Coelho. Adapting to dynamic registration errors using level of error (LOE) filtering. *Proc. IEEE and ACM ISAR '00 (Int. Symp. on Augmented Reality)*, Munich, Germany, October 5-6 2000. 85-88.

[10] R. Pausch, T. Burnette, D. Brockway, and M. Weiblen. Navigation and locomotion in virtual worlds via flight into hand-held miniatures. *Proc. ACM SIGGRAPH '95*, 399-401.

[11] J. Pierce, M. Conway, M. van Dantzich, and G. Robertson. Tool spaces and glances: Storing. accessing, and retrieving objects in 3D desktop applications. *Proc. 1999 ACM Symp. on Interactive 3D Graphics*, New York, Apr. 26-28 1999. ACM Press. 163-168.

[12] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The go-go interaction technique: Non-linear mapping for direct manipulation in VR. *Proc. UIST '96*, 1996. 79-80.

[13] R. Stoakley, M. Conway, and R. Pausch. Virtual reality on a WIM: Interactive worlds in miniature. *Proc. CHI '95 (ACM Conf. on Human Factors in Comp. Sys.)*, May 7-11 1995. 265-272.