MIND-WARPING: Towards Creating a Compelling Collaborative Augmented Reality Game

Thad Starner, Bastian Leibe, Brad Singletary, and Jarrell Pair
Contextual Computing Group, College of Computing
Georgia Institute of Technology
Atlanta, GA 30332-0280
(404) 894-3152
{thad,leibe,bas,jarrell}@cc.gatech.edu

ABSTRACT

Computer gaming offers a unique test-bed and market for advanced concepts in computer science, such as Human Computer Interaction (HCI), computer-supported collaborative work (CSCW), intelligent agents, graphics, and sensing technology. In addition, computer gaming is especially well-suited for explorations in the relatively young fields of wearable computing and augmented reality (AR). This paper presents a developing multi-player augmented reality game, patterned as a cross between a martial arts fighting game and an agent controller, as implemented using the Wearable Augmented Reality for Personal, Intelligent, and Networked Gaming (WARPING) system. Through interactions based on gesture, voice, and head movement input and audio and graphical output, the WARPING system demonstrates how computer vision techniques can be exploited for advanced, intelligent interfaces.

Keywords

Augmented reality, wearable computing, computer vision

1. INTRODUCTION: WHY GAMES?

Computer gaming provides a unique prototyping arena for human-computer interactions. Due to the entertaining nature of the interactions, users are willing to explore innovative metaphors, modalities, and hardware even when they are not as apparent or fluid as the designer might have hoped. In addition, there is a certain universality of a sense of play that entices users who would not be interested in testing prototype systems normally.

Another advantage is that game play can be designed so as to hide limitations in the current implementation of a system, allowing researchers to rapidly implement and field test new techniques. Such limited but practical experience may demonstrate which concepts are worth pursuing and which are not. In fact, games tend to draw out possibilities that might not have been considered before. Also, such exposure to real users often demonstrates whether or not a particular technique degrades gracefully in adverse or unexpected conditions.

Computer gaming provides an excellent application for working in wearable computing and augmented reality [5,10]. In the constructs of a game, CSCW, HCI, graphics, autonomous agents, mobile sensing and pattern recognition, wireless networking, and distributed databases can be explored in a community of wearable computer users, without the need to convince the community of the usefulness of the equipment. As infrastructure needs are determined and the infrastructure improved, new metaphors and uses can be introduced. With time, these improvements can be directed towards development of longer term, everyday-use wearable computing - for example, through real-time, real-space, role-playing games. Game design becomes more complex and, correspondingly, the weaknesses of the system are revealed, encouraging continuing iterations of improvements and rapid deployment. The Wearable Augmented Reality for Personal, Intelligent, and Networked Gaming (WARPING) system presented below is a first attempt at developing an extensible infrastructure along these lines.

2. AUGMENTED REALITY AND WEARABLE COMPUTING IN GAMING

In augmented reality, the virtual world is composited with the physical [8], resulting in obvious restrictions to the simulation, but not without significant benefits. First, the game designer already has a pre-made environment for his game, that of the user's physical surroundings. Thus, the designer can concentrate on creating just the virtual artifacts that deviate from the physical world in the game. For example, an augmented reality game may be as simple as a scavenger hunt of virtual flags hidden in the physical environment. More compelling effects may emulate the abilities of the protagonists in myths and modern stories in science fiction and fantasy. For example, the augmented reality hardware may seem to augment the senses of the player, giving him the ability to see through physical walls to virtual objects in the game, hear a whisper at great distances, see other players' “good” or “evil” auras to avoid deception, follow lingering “scents'”, or detect “dangerous” hidden traps. In addition, the system may associate graphics with the player's physical body, allowing him to shoot fireballs from his hands, envelop himself in a force field, or leave a trail of glowing footprints.

While such effects can be striking, some of the biggest benefits result from the input possibilities from the mobility and proximity to the user of a wearable computer. By mounting cameras near the user's eyes and microphones near the user's ears, the wearable computer can “see” as the user sees and “hear” as the user hears. By recognizing the user's current context, the wearable can adapt both its behavior and its input and output modalities as appropriate to the situation. This technique grounds user interface design in perception and may encourage more seamless or intelligent-seeming interfaces as the computer begins to understand its user from his everyday and, almost by definition, mobile behavior.

Wearable computers also allow the observation of the player himself. Body posture, hand gestures, lip motion, voice, and biosignals might be used as input to the game. For example, a player taking on the role of a wizard may use hand gestures and recite particular incantations which, when recognized by the computer, causes a spell to be cast. In order to create a particular mood, the system may measure a player's heart beat and respiration and play amplified heart beat and respiration sounds in time to, or a little faster than, the player's own.

3. WARPING

The WARPING system combines several advanced perception and interaction techniques into a foundation for creating augmented reality games. Currently, WARPING has two available platforms: an augmented desktop and a body-centered augmented reality. Both platforms make extensive use of computer vision in their interfaces. A more complete description of the Perceptive Workbench can be found in [3].

3.1 The Perceptive Workbench

WARPING's augmented desktop extends the utility of the Responsive Workbench [2] which is based on the Fakespace Immersive Workbench, a semi-immersive 3D display device. The Perceptive Workbench consists of a large wooden desk, a large frosted glass surface for display, an equally large mirror, and a projector (see Figure 1). Images are displayed on the workbench by an Electrohome Marquee 8500 projection system by reflecting the inverted image onto a mirror mounted diagonally under the desk, thus correctly displaying the projection to a viewer opposite the side of projection and above the frosted glass surface. The unit may display stereoscopic images through the use of shutter glasses. In the past, the workbench employed wired pinch gloves and electromagnetic trackers for user input. WARPING attempts to replace that functionality through computer vision, eliminating the need for encumbering wires and the restriction of the to the number of participants number of input devices.

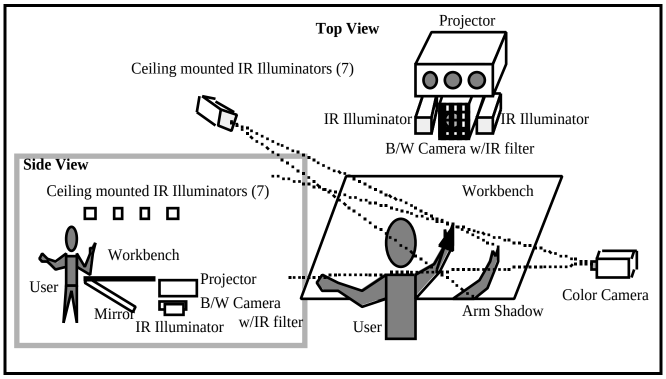

Figure 1: Apparatus for the Perceptive Workbench

Extending a technique initially designed by the first author for use in Ullmer and Ishii's MetaDesk [11], we have adapted the workbench to use near-infrared (IR) light and computer vision to identify and track objects placed on the desk's surface. The camera used for our vision algorithms is a low-cost black and white CCD and is mounted behind the desk, close to the projector. Two near-infrared light sources are mounted adjacent to the projection and camera systems behind the desk Illumination provided by the IR spot-lights is invisible to the human eye and does not interfere with desktop projection in any way. Meanwhile, most black and white cameras can capture this near-IR light very well. An IR-pass filter is placed in front of the lens of the camera so that visible light is blocked. This allows the IR vision system to be used regardless of what is displayed on the workbench. From the reflections of IR light off the bottom surface of the object in contact with the table, a contour-based computer vision system can recognize and track a wide variety of objects as they are placed on the surface of the desk by the user. Currently tracking latency is 0.25-0.33 seconds but seems adequate for the context of the game.

This vision system has been extended for three-dimensional reconstruction of the objects placed on the desk. This ability is desired for several reasons. First, not all objects reflect well in the infrared. While the bottom camera may see a change in intensity when an object is placed on the surface of the desk, the resulting blob may be darker than the normal, uncovered desk surface instead of lighter. Thus, another method is necessary for identification. Secondly, recovering a three-dimensional shape allows distant, networked workbench users to introduce and share novel objects freely. Given the stereoscopic capabilities of the workbench, such 3D object recovery may be a convenient and natural way to include physical objects in a virtual collaboration.

A ring of 7 infrared light sources is mounted to the ceiling, each one of which can be switched independently by computer control. The camera under the desk records the sequence of shadows an object on the table casts when illuminated by the different lights. By approximating each shadow as a polygon (not necessarily convex)[6], we create a set of polyhedral "view cones", extending from the light source to the polygons. Intersecting these cones creates a polyhedron that roughly contains the object [4].

The reconstruction process is triggered by the discovery of a new object placed in the center of the desk. When such an object is discovered, a virtual button is rendered on the desk's surface. If the user "presses" the button, the workbench automatically switches the light sources in rapid succession and reconstructs the object. The reconstruction process requires 3-4 seconds to complete. Figure 2 shows the reconstruction of a watering can placed on the desk. For test cones and pyramids, there is a mean error of 4.7mm (1.2%) for the reconstructed vertices using the Hausdorff distance metric (the two-sided distance between the point and reference surface).

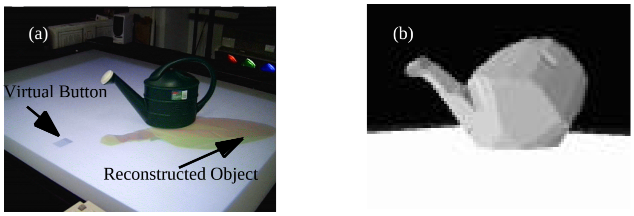

Figure 2: (a) 3D reconstruction of an object placed on the workbench display; (b) resulting polygonal object.

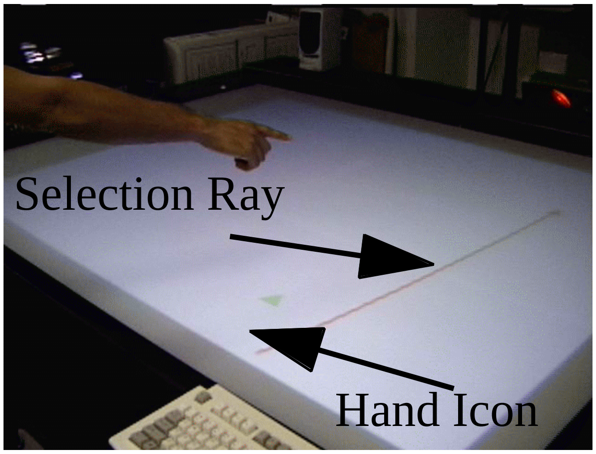

Figure 3: Recovering 3D pointing gestures

Through the addition of a color camera mounted at the side of the desk (see Figure 1), the Perceptive Workbench can recovery 3D pointing gestures made by the user above the desk's surface (Figure 3). Such gestures are often useful to refer to objects rendered out of the user's reach on the surface of the desk. Note that the side camera is not necessary for slow gestures since the infrared spotlights could also be used to recover the user's arm position in 3D. However, the switching speed of the lights used would limit the frame rate.

3.2 The Mobile Interface

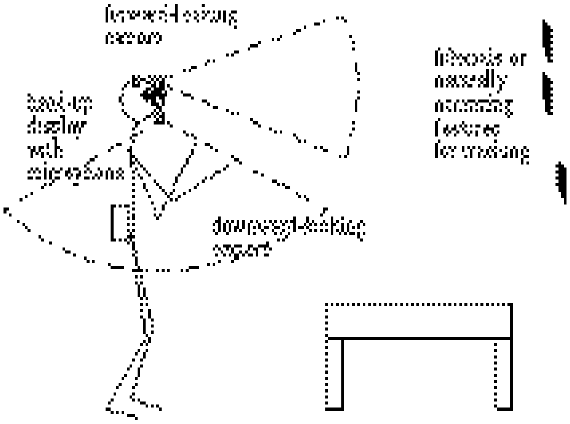

Figure 4: Mobile user apparatus

In order to create a powerful mobile augmented reality, WARPING currently uses a 800x600 pixel "see-through" Sony Glasstron equipped with two cameras (see Figure 4). The cameras are mounted so as to be directed toward what the player sees in the physical world and towards the player's hands. In this manner, the player's head position can be tracked and his hand gestures watched. A microphone is also mounted on the display to provide the user with voice control. In order to maintain mobility and the illusion of a self-contained wearable computer, we are experimenting with wireless video transmission as in [8].

Currently, the player's head rotation is recovered by visually tracking fiducials mounted on the walls of his room. In current WARPING implementations, the player is assumed to stay in relatively the same place, allowing the use of the head rotation information for rendering graphics that seem registered to the physical world. In the future, it is hoped that registration will be adequate to clip virtual creatures to physical objects in the room such as the coffee table in Figure 4.

We are adapting a computer vision gesture toolkit developed by the first author to recognize player gestures. Based on hidden Markov models, this system has been shown to recognize 40 American Sign Language signs in real-time with over 97% accuracy [7]. However, in the implementation below, a simple template matching method was sufficient. Gestures are recognized through tracking the hands which are segmented through color.

The 3D graphical and sound component of the current WARPING system was created using the Simple Virtual Environment (SVE) toolkit. SVE is a programming library developed by the Georgia Institute of Technology Virtual Environments Group. Built upon OpenGL, the library supports the rapid implementation of interactive 3D worlds. SVE allows applications to selectively alter, enhance, or replace components such as user interactions, animations, rendering, and input device polling [1]. SVE applications can be compiled to run on both PC and SGI platforms.

Silicon Graphics O2's are used as the hardware platform for the WARPING system. Depending on the desired functionality, 1-3 machines are used for the Perceptive Workbench and 1-2 machines for the mobile user.

4. EXAMPLE IMPLEMENTATION

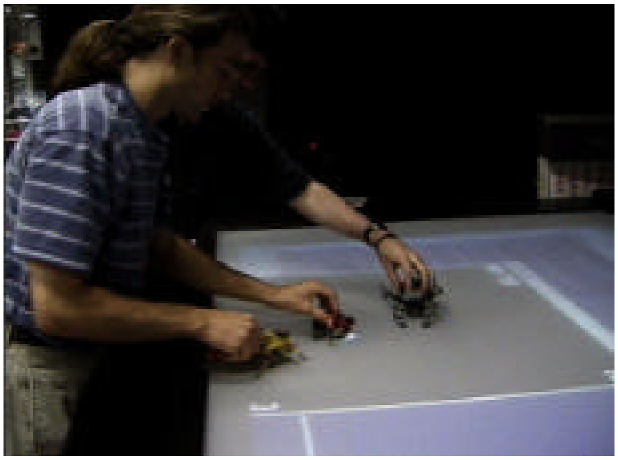

Figure 5: "Magicians" controlling monsters on the Perceptive Workbench.

Multiple-Invading Nefarious Demons (MIND) is our first game implementation in the WARPING system. The game pits a "Kung Fu" fighter against an evil magician who directs monsters through his magical desk. The fighter performs magical gestures and shouts to fend off the attacking demons. When the mobile player puts on the head-up display, a status screen is visible which allows the user to prepare for the magician (workbench user) to start the next fight. The Perceptive Workbench provides the magician with an overhead view of the playfield while the fighter sees the graphics registered on the physical world from a first-person perspective. When the magician places one or more opponents (in this case, plastic bugs – see Figure 5) on to the virtual gaming table, the player sees the opponents appear to be on one of 3 planes of attack: high, middle, and low. For MIND-WARPING, the plane of attack is indigenous to the type of attacking monster, for example: ants on the floor, spiders on the ceiling, flies in the middle etc. To effect attacks on the virtual opponents, the fighter performs one of 3 hand gestures in combination with an appropriate Kung Fu yell ("heee-YAH") to designate the form of the attack. Figure 6 demonstrates the gestures used in MIND WARPING. The yell is a functional part of the recognition system, indicating the start and end of the gesture. Audio thresholding is sufficient to recognize the yell, though speech recognition technology may be employed for future games. Once the correct gesture is performed, the corresponding monster is destroyed, making its death sound and falling from the playfield. However, if the magician is successful in navigating a monster so that it touches the player, the player loses health points and the monster is destroyed. Figure 7 shows the view from the player's perspective.

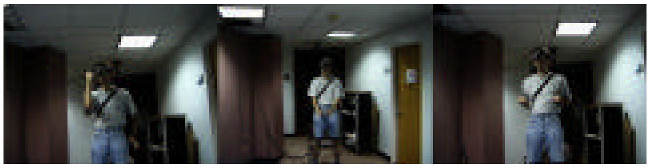

Figure 6: High, low, and middle level attack gestures.

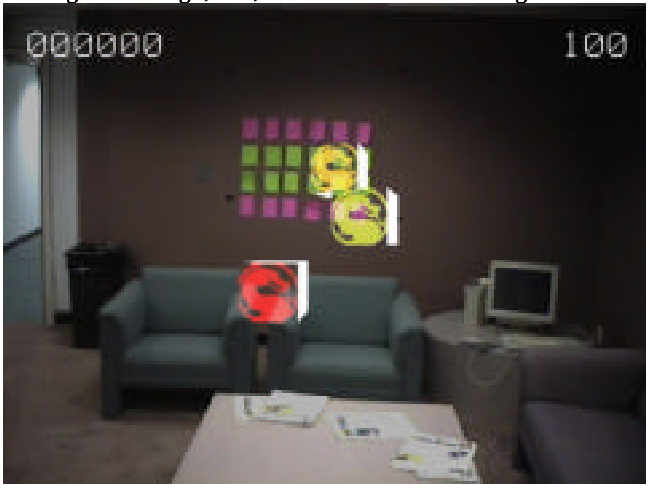

Figure 7: View from the mobile player's perspective.

5. GAME PLAY

While the system is used mostly by the authors and graduate students in the same laboratory, we have found that the experience is surprisingly compelling and balanced. The magician who controls the monsters has difficulty keeping pace with the player, which is opposite of what was expected. In fact, two workbench users sometimes team to attack the mobile player.

The combination of modalities is engaging for the mobile player. He or she sees and hears the monsters overlaid on the physical room and responds with sharp gestures and yells for defense.

6. FUTURE WORK

Currently, MIND-WARPING only uses the object tracking abilities of the Perceptive Workbench. Future implementations may allow the desk player to insert novel objects into the game and manipulate them through pointing gestures. Future testing will explore the speed and effectiveness of using gestural and voice input for the mobile and desk based players

7. CONCLUSIONS

This paper demonstrates wearable computing and augmented reality interfaces, based mostly on computer vision techniques, in the context of computer gaming. The resulting implementation suggests the potential of such techniques for future intelligent interfaces.

8. ACKNOWLEDGMENTS

The authors would like to thank Larry Hodges, Bill Ribarsky, and the Virtual Environments group for the use of SVE and the workbench; Paul Rosin and Geoff Vest for their line fitting code; and Martti Mantyla for his code on boolean set operations on polyhedra solids.

9. REFERENCES

[1] Kessler, D., Kooper, R., and Hodges, L. The Simple Virtual Environment Library User's Guide V2.0. GVU Center, Georgia Tech, 1997.

[2] Kruger, W., Bohn, C., Frohlich, B., Schuth, H., Strauss, W., and Wesch G. The Responsive Workbench: A virtual work environment. IEEE Computer 28(7), July 1995, 42-48.

[3] Leibe, B., Starner, T., Ribarsky, W., Wartell, Z., Krum, D., Singletary, B., Hodges, L. The Perceptive Workbench. Submitted to VR2000.

[4] Mantyla, M. An Introduction to Solid Modeling Computer Science Press, 1988

[5] Ohshima, T. AR2 Hockey. Proc. of SIGGRAPH'98.

[6] Rosin, P., and West, G. Nonparametric Segmentation of Curves into Various Representations. IEEE PAMI 17, 12.

[7] Starner, T., Weaver, J., and Pentland, A. Real-Time American Sign Language Recognition Using Desk and Wearable Computer-Based Video. IEEE PAMI 20, 12.

[8] Starner, T., Mann, S., Rhodes, B., Levine, J., Healey, J., Kirsch, D., Picard, R., and Pentland, A. Augmented Reality Through Wearable Computing. Presence 6, 4.

[9] Sullivan S., and Ponce, J. Automatic Model Construction, Pose Estimation, and Object Recognition from Photographs Using Triangular Splines. To appear IEEE PAMI.

[10] Szalavari, Z., Eckstein, E., and Gervautz, M. Collaborative Gaming in Augmented Reality. Proc. Of VRST'98 (Taipei, Taiwan 1998), 195-204.

[11] Ullmer B., and Ishii, H. The metaDESK: Proc. of UIST, 1997.