Cross-Dimensional Gestural Interaction Techniques for Hybrid Immersive Environments

Hrvoje Benko, Edward W. Ishak, Steven Feiner
Department of Computer Science, Columbia University
500 W. 120th St. 450 CS Building, New York, NY
{benko | ishak | feiner}@cs.columbia.edu
www.cs.columbia.edu/graphics

ABSTRACT
We present a set of interaction techniques for a hybrid user interface that integrates existing 2D and 3D visualization and interaction devices. Our approach is built around one- and two-handed gestures that support the seamless transition of data between co-located 2D and 3D contexts. Our testbed environment combines a 2D multi-user, multi-touch, projection surface with 3D head-tracked, see-through, head-worn displays and 3D tracked gloves to form a multi-display augmented reality. We also address some of the ways in which we can interact with private data in a collaborative, heterogeneous workspace.

KEYWORDS: hybrid user interfaces, augmented reality, tabletop interaction, gesture-based and touch-based interaction.

INTRODUCTION
Augmented Reality (AR) makes it possible to visualize computer-generated graphics overlaid onto the surrounding environment. Many AR systems focus on the use of personal displays (such as see-through head-worn displays), and their user interfaces often consist of purely virtual elements, such as 3D widgets or 3D interaction metaphors. However, since the user is not completely immersed in the virtual world, AR also allows for the simultaneous use of all existing user interfaces that surround the user. These surrounding user interfaces are often better suited for some tasks than the personal AR user interface, but the lack of simple ways to move data to and from them, discourages the AR user from interacting with them.

In this paper, we present an experimental system that combines multiple see-though head-worn displays with a 2D user interface projected on a shared touch-sensitive surface. We refer to a heterogeneous combination of displays and interaction technologies as a *hybrid* user interface [4], in this case forming a hybrid immersive environment.

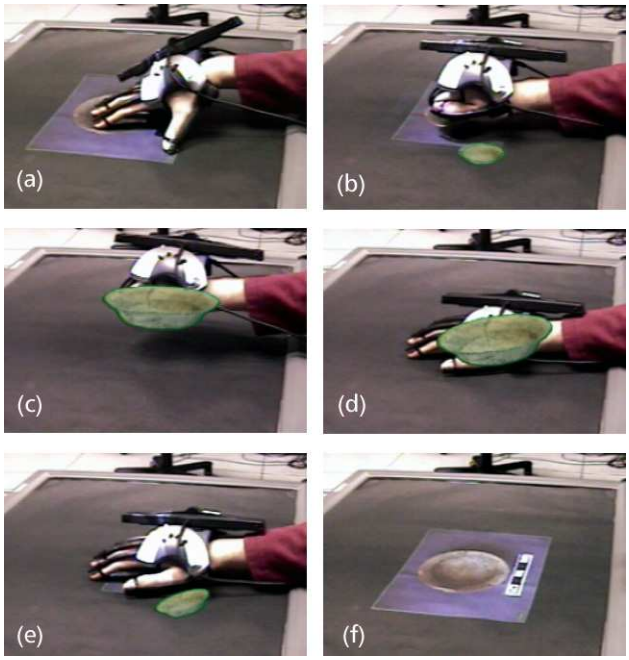

Figure 1: Frames from a sequence of cross-dimensional pull and push gestures. (a) Selecting the 2D object and beginning to form a grabbing gesture. (b) The 2D object disappears, while the 3D object appears from the table. (c) Holding the 3D object. (d) Push gesture begins by pressing on the table through the 3D object. (e) The 3D object disappears and the 2D object appears. (f) The 2D object, projected on the table. (Table surface is covered with black paper to provide a darker background for imaging the projected display through the live tracked video see-through display used to create all images in this paper.)

Some researchers have used a single universal tool (much like a common remote control for multiple components in a home entertainment system) [2, 15] to interact in a hybrid environment. In contrast, we are exploring ways to allow users to maintain the familiar interaction techniques that are native to each of the user interfaces being combined. Given the large body of existing work on gestural interaction in 2D (e.g. [22]) and 3D (e.g. [13]), what are missing are hybrid gestural interactions that create a unified interaction space.

Our research focuses primarily on the border between 2D and 3D user interfaces. To bridge this divide, we have designed cross-dimensional gestural interactions that require the use of both 2D and 3D modalities, as shown in Figure 1. For example, to move an object from the 2D display to the 3D AR world, the user starts by interacting with the object in 2D, and then moves into the 3D environment, bringing along the object. These context-aware cross-dimensional gestures build on and integrate the ways in which we already interact in each dimension.

RELATED WORK

The inspiration for our work has come from two different themes: hybrid user interfaces and gesture- and touch-based interactions.

Our first research on hybrid user interfaces embedded a small, stationary, flat-panel "detail" display within the field of view of a monocular, head-tracked, see-through, head-

worn "context" display [4]. Low et al. reversed this approach, using a stereo, head-tracked, opaque, head-worn display to provide detail within the context of a larger monocular, projected display [12]. Neither of these systems addressed gestural interaction, although the former supported seamless mouse-based interaction across the interface between the displays.

EMMIE [2] demonstrated a hybrid user interface that integrated a variety of tracked handheld and stationary displays, head-worn displays, and physical input devices. See-through head-worn displays were used to overlay virtual information on the physical environment, including the displays within it. EMMIE's primary cross-dimensional user interaction technique accomplished *drag-and-drop* across displays using a 3D mouse. Privacy was supported by manipulating 3D "privacy lamps" that caused the virtual objects they illuminated to become invisible to others.

The StudierStube collaborative AR environment [19] provides the Personal Interaction Panel (PIP) as a two-handed interaction tool for manipulating 2D and 3D data. The PIP offers passive tactile feedback and allows for 2D and 3D interactions; however, it does not use any gestural interaction and the visualization remains completely within the 3D context. Rekimoto's *pick-and-drop* [16] and *hyperdragging* [18] techniques were developed as part of an exploration of multi-computer direct manipulation interfaces, in which he examined the hybrid interactions between multiple portable computers, table and wall displays, and other physical objects. However, his interaction set addresses only pointer-based interaction with 2D data.

In attempting to cross the 2D–3D divide, we have also been inspired by the gestures that Ishii et al. have developed for manipulating tangible user interfaces [8], as our work similarly addresses "seamless workspaces" that stretch across different dimensionalities, displays and users [7].

Regenbrecht et al. created MagicMeeting [15], a collaborative tangible AR environment that mixes 2D and 3D displays and employs tangible tools for seamless interaction between 2D and the 3D data. In contrast to our work, all of their interaction techniques rely on the use of tangible tracked tools. Brown et al., in their SCAPE collaborative augmented reality environment [1], use head-worn projective displays that project onto many retroreflective surfaces in a room, including the walls, a tabletop, and several hand-held widgets. Since all material is projected and visible only along a projector's line of sight, all data is potentially private. However, data cannot be overlaid on conventional displays in the room because of the need for a retroreflective surface to display it. Although Brown et al. employ hand gestures for navigation, all interactions with objects in the world are accomplished through their hand-held widgets.

Much research has been performed on gesture-based and touch-based interaction. Krueger's VIDEOPLACE [10] is an early example of multi-hand and multi-finger interaction that recognizes gestures through video silhouettes. The i-LAND [20] and Connectables [21] projects both explore collaboration and interaction with displays built into furniture. Hinckley [6] uses synchronous gestures for collaborating and interacting with several independent mobile displays.

Recent work by Wu and Balakrishnan [22] presented multi-finger and whole-hand gestural interaction techniques for multi-touch, multi-user displays. While their system was entirely 2D, and used a single display, it highlighted the touch- and gesture-sensing capabilities of the DiamondTouch table [3]. SmartSkin technology [17] also addresses tabletop interaction and offers more powerful multi-touch capabilities than DiamondTouch; however, it cannot differentiate input from different users.

ARCHEXPLORER

We have developed a testbed application, called *ArchExplorer*, which allows users to visualize 2D and 3D representations of object data in a hybrid environment. Motivation for this application comes from our ongoing research on developing visualization tools for archaeologists who analyze and interpret information about archaeological finds post-excavation. ArchExplorer uses a sample database of archaeological finds containing images and 3D models of objects excavated on site. Visualization of this information on a traditional computer screen is often sufficient for 2D data, such as images, drawings, and text; however it is often inadequate for 3D data, such as 3D models, for which AR or virtual reality would be more suitable.

Figure 2: Two users, wearing tracked, see-through, head-worn displays and tracked 3D gloves, interact with linked 2D and 3D objects in *ArchExplorer*.

In ArchExplorer, users view both 2D and 3D objects in a shared hybrid collaborative space, which is composed of the multi-user, multi-touch, projection surface, and the entire surrounding 3D space tracked by a 6DOF tracking system, viewed through head-tracked, see-through, head-worn displays (Figure 2). We have built in part on our previous work in multimodal interaction for augmented and virtual immersive environments [9,14], in which our 3D gesture recognition technology was originally developed.

To interact in our environment, users employ a variety of 2D and 3D interaction techniques, although we focus on cross-dimensional techniques that integrate the 2D and 3D contexts, and allow objects to pass between them. While our techniques were prototyped in this sample application, there is no reason why they could not be applied more generally to any hybrid system.

HYBRID INTERACTION TECHNIQUES

The hybrid interaction techniques that we have developed can be categorized into two main classes: one-handed techniques and two-handed techniques. One-handed techniques are generally simpler and more intuitive, thus forming a set of basic interactions for a hybrid environment, while two-handed techniques form a more complex set that addresses higher-level concepts, such as privacy. Furthermore, having a relatively rich set of techniques that each require only a single hand opens up many interesting opportunities for using two one-handed techniques simultaneously.

In the remainder of this paper, we use "2D object" to refer to a pictorial representation of a real object presented on the table surface and "3D object" to refer to a 3D model of a real object presented to the user on the head-worn display.

Building Blocks: Existing 2D and 3D Gestures

We begin by describing the gestures that our system recognizes independently in 2D and 3D. In our 2D environment, we support one-finger touch, tap (typically with one or more knuckles), flat hand, and vertical hand, and in 3D, we support point, grab, and idle state, as shown in Figure 3. The DiamondTouch information that we use for discriminating 2D gestures is the width and height of the axis-aligned bounding box of user's contact points. In addition, we extend these gestures by considering motion, such as vertical hand "sweep" gesture [22] or a "swoop" gesture (the transition between the idle flat hand and a closed fist).

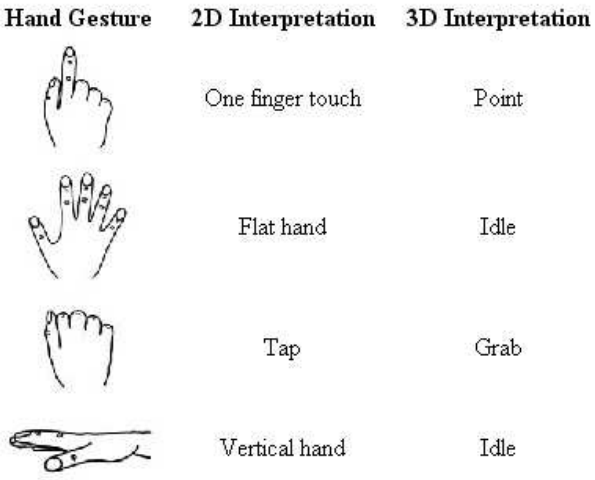

Figure 3: A set of basic hand gestures and their interpretations in 2D and 3D environments.

Our primary considerations for selecting these basic gestures were simplicity and relative lack of ambiguity, rather than completeness. While there are many other useful gestures possible in 2D and 3D, we feel that this set forms a good basis for creating interactions that span the 2D–3D divide in a wide range of potential applications.

While these gestures are recognized independently, both in 2D and in 3D, in a hybrid environment it is important to look at them in combination, rather than in isolation. This is

especially true for hand gestures that have different meanings, depending on whether they are interpreted in 2D or in 3D. In the following sections, we define a set of hybrid interactions (each of which is prefixed by the modifier "Cross-Dimensional") by combining interpretations across the 2D–3D divide.

Single-handed Interactions

Selection and Movement

Selection is probably the most basic interaction in any environment. Because we believe that selection is inherently environment-dependent, we have chosen to implement it differently in 2D and 3D. 2D selection is implemented by a tap gesture on top of a 2D object, while 3D selection is implemented by a grab gesture near a 3D object. These gestures are extended by factoring in hand motion, yielding "touch and drag" and "grab and move," both of which produce movement of the selected object. While these techniques are not "hybrid," selection is a critical component of the interaction set from which we form our hybrid gestures.

Cross-Dimensional Pull

To make an object transition from a 2D representation to a 3D representation, the user employs a *cross-dimensional pull* gesture. Wanting this gesture to mimic a real-world "grab" from a flat surface, we asked several users to perform exaggerated grabs of actual physical objects and observed their hand's location, movement, and finger contractions and expansions as they grabbed items, such as a CD, tennis ball, and a pair of scissors, on a table. Based on these observations, we designed the gesture so that the user places the hand flat on the table on top of the object's image representation, moves the fingers towards the center of the object, and then forms a fist (Figure 1a–c). As this occurs, the 2D image on the table scales down and disappears, while a 3D model appears and scales up, attached to the user's grabbing hand, representing the transformation from 2D to 3D. This gesture exemplifies a theme common to the entire set: preservation of the interaction context. If the intention is to transfer something from 2D to 3D, the user starts to gesture on the 2D surface, and finishes in the 3D world.

Cross-Dimensional Push

As a complement to the *cross-dimensional pull* gesture, an object can transition from 3D to 2D through use of the *cross-dimensional push* gesture. Once the 3D object is placed above the table, a user can position a flat hand above the object and "push" downwards until her hand touches the table (Figure 1d–e), mimicking an attempt to push a real physical object into the table. As this occurs, the 3D model scales down and disappears, while the 2D image appears on the table and scales up. (Alternatives to our uniform-scaling visualization include one-dimensional scaling along the axis perpendicular to the table to "flatten" a 3D object or extrude a 2D one, or clipping the object against a virtual plane coplanar with the table. Visual differences among many of the 2D–3D pairs have led us to select our current approach.)

Cross-Dimensional Connect

Sometimes users want to see both the 2D and 3D object representations at the same time. To establish this relationship between the two environments, the user performs a *cross-dimensional connect* gesture, first requiring a selection (grab) of the 3D object, pulling it

towards the table, and then briefly tapping the table anywhere while continuing to hold the object (Figure 4a). This results in a 2D object appearing on the table surface while the 3D object remains visible and attached to the hand. Once the two representations are connected, a leader line appears between them, highlighting the connection. This leader line visualization is similar to EMMIE's use of leader lines that extend from a search control panel on a hand-held display to 3D virtual objects that meet the search criteria [2]. We designed this gesture to resemble pulling a virtual string between one object and the other.

We also experimented with the idea of having a connected 2D object appear as a "virtual shadow" on the table following the 3D object in the world similar to the ideas presented by Herndon et al. [5]. If the user moves the 3D object away from the table, the "shadow" maintains the position on the table that is closest to the 3D object's position in the world. While the "shadow" representation felt like a natural way to showcase the connection between the two objects, the leader lines alone allow for more freedom, since the position of each object is independent of the other.

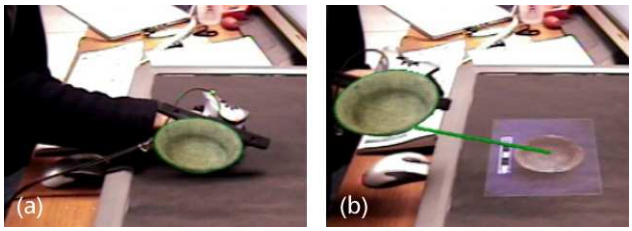

Figure 4: Cross-dimensional connect gesture. (a) Tapping the table while holding a 3D object establishes connection with the table. (b) Leader line connects two representations of the same object.

The same tapping gesture used to connect the two representations can also be used to break the connection between them to perform a *cross-dimensional disconnect*. All that is needed is to tap the 2D representation on the table with the hand that is holding the 3D object. The 2D object disappears, leaving the 3D object in the user's hand. This cross-dimensional interaction thus serves as a simple, but powerful toggle action.

Cross-Dimensional Pin

Another way that the user can connect the 3D and the 2D environments is by "pinning" the 3D object to the table. In contrast to the cross-dimensional push, which transforms a 3D object into a 2D object, a *cross-dimensional pin* attaches the 3D object to the 2D surface. Pinning an object is accomplished by placing the 3D object close to the table, making a pointing gesture with the hand, and then pushing through the object until touching the table (Figure 5a). Three colored rings appear around the object to signal that the object is pinned to the table (Figure 5b). To *cross-dimensionally unpin* the 3D object, the user needs to select it (i.e., grab it) and move it away from the table surface. As soon as the object is grabbed in 3D, it is no longer pinned, which can be seen by the disappearance of the three colored rings around it.

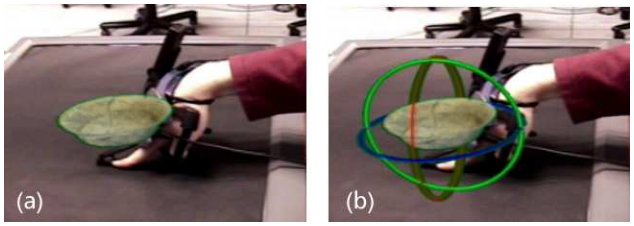

Figure 5: (a) Cross-dimensional pin gesture associates the object with the table. (b) Rings appear once the object is pinned to provide visual feedback to the user.

Although the differences between the pin and the connect techniques might appear subtle, we feel that they are distinctly different. The connect technique allows for simultaneous viewing and manipulation of both 2D and 3D representations. By pinning the object to the table, there is no 2D representation shown below the 3D object and the focus is on allowing the use of the rich set of existing 2D touch-based interaction for manipulating the 3D model.

3D Drag

When a 3D object is pinned to the table, the user can drag it around in exactly the same way that the 2D object can be moved. Simple tap and drag will translate the pinned object anywhere on the table.

Context-Sensitive Menus

We extend the gestures developed by Wu and Balakrishnan [22] for context-sensitive menus to manipulate 3D objects. This makes it possible to invoke a context-sensitive menu for a pinned 3D object in exactly the same way as for the regular 2D object: by tapping and holding on the table at the location of the object. A context-sensitive menu will then appear to the left of the hand. (The placement of the menu is currently determined by the fact that all our tracked gloves are right handed, and therefore minimal occlusion occurs when the menu is shown on the left side. Alternatively, the system could position the menu based on which hand is currently touching the table.)

Parameter Adjustment

Oftentimes, a user may wish to adjust the parameter values of a pinned or connected 3D object; for example, to perform transformations on the object, such as a rotation, scale, or hue-change. However, these can be cumbersome to perform in 3D space using a tracked, instrumented glove without haptic feedback. Instead, by taking advantage of the multi-touch table, a user can benefit from 2D touch-based interactions with passive haptic feedback. Figure 6 shows a user rotating a pinned vessel object.

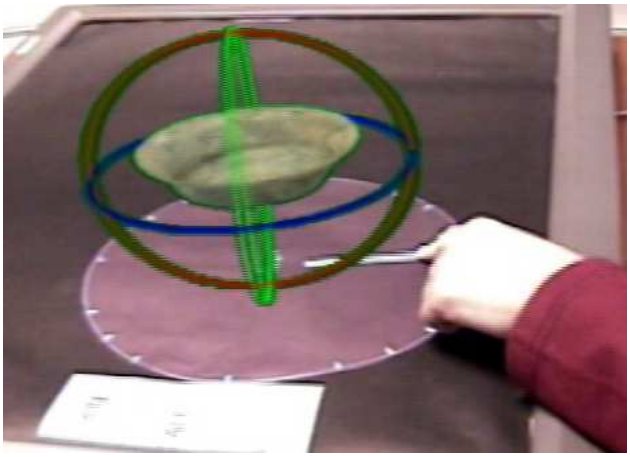

Figure 6: The context sensitive menu is invoked on the 3D object and the 2D rotation tool is applied.

Two-Handed Interactions

Interaction Combo

To create more powerful gestural interactions, we allow the user to use both hands simultaneously to adjust two different independent transformation parameters of an object. For example, a user may wish to perform transformations on the object, such as changing its scale or color while holding the object in her hand and inspecting it from all sides. Users tend to find it cumbersome to adjust these parameters in 3D space using a tracked, instrumented glove without haptic feedback. We take advantage of the multi-touch table to allow users to perform these adjustments on a 2D surface with their non-tracked hand, while their tracked hand can simultaneously manipulate the 3D object representation directly. For example, with remote parameter adjustment, the user can grab and move a 3D object in 3D space to change its location, while simultaneously using the other hand on the multi-touch surface to adjust its scale, as shown in Figure 7.

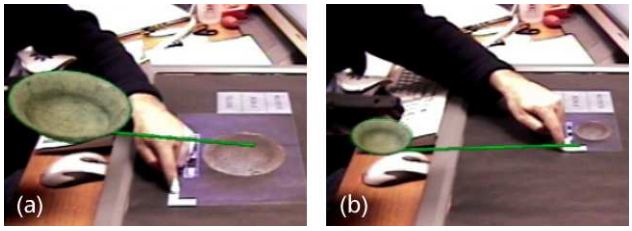

Figure 7: Interaction combo: combining scaling with 3D manipulation. (a) The 3D object is held by the right hand, while the left hand adjusts the scale. (b) The right hand rotates the object to get a better view, while the object is simultaneously being scaled.

Cross-Dimensional Privacy Pull

In a hybrid collaborative environment, the users share a common workspace, both on the 2D surface and in the surrounding 3D volume. This allows each user to clearly observe any other user's actions and object information. However, in certain situations, users may wish to privately view sensitive data and not allow other users to visualize this informa-

tion. Although privacy gestures on a 2D surface have been explored before [22], these gestures do not facilitate absolute privacy if the hand is not correctly positioned or angled on the multi-touch table relative to other users' viewpoints.

To address this, we introduce a two-handed hybrid privacy gesture, called *cross-dimensional privacy pull*, which allows a user to have complete and absolute privacy when visualizing 3D data. Placing a vertical hand on the multi-touch surface, while performing a single-handed cross-dimensional pull with the other hand, allows the user to privately visualize the object information in 3D through his head-worn display (Figure 8a–c). After executing the privacy pull, the 3D object is only visible in the head-worn display of the user who performed the interaction, thus ensuring that the information is private.

Cross-Dimensional Privacy Push

The complementary *cross-dimensional privacy push* allows a user to view 2D object data privately. As mentioned earlier, attempting to visualize private information projected on the multi-touch surface using a single or two-handed gesture may not completely obstruct the information from other users. However, when a user places a vertical hand on the surface, while performing a single-handed cross-dimensional push with the other hand, this allows a private view of 2D information in the head-worn display, rendered geometrically as if it were projected onto the 2D multi-touch surface (Figure 8d–f). This private 2D object behaves just like a normal 2D object and can be manipulated using the standard 2D interaction on the surface, but is completely protected from the other user's view. Although we recognize that current head-worn displays are not the most suitable means to view high-resolution 2D data, this allows for complete user privacy, unlike projection onto the multi-user 2D surface.

SYSTEM ARCHITECTURE

Our ArchExplorer environment consists of two types of modules: a 2D module and a 3D module. Each module powers one display unit; therefore, with two users sharing our interaction table, we require one 2D module and two 3D modules (one for each AR user). The 2D visualization module uses a Mitsubishi Electric Research Laboratory DiamondTouch table [3] for our multi-user multi-touch surface (26"x20") onto which we project with a Proxima Ultralight x350 DLP projector (1024x768 resolution). A black paper covering that we have added to the table surface increases the apparent contrast and improves the quality of images shot through the video see-through display used to create the figures in this paper. Our 3D augmented reality modules consists of a modified Essential Reality right-handed P5 glove to sense hand gestures, a Sony LDI-D100B optical see-through head-worn display to present the AR overlay and two InterSense IS900 six-degree-of-freedom trackers to monitor head and hand position and orientation.

Since each user requires two trackers, one on the head and one on the right hand, we support only two users because our tracking system can track only four trackers at once. However, the techniques described in this paper easily extend to a larger number of users.

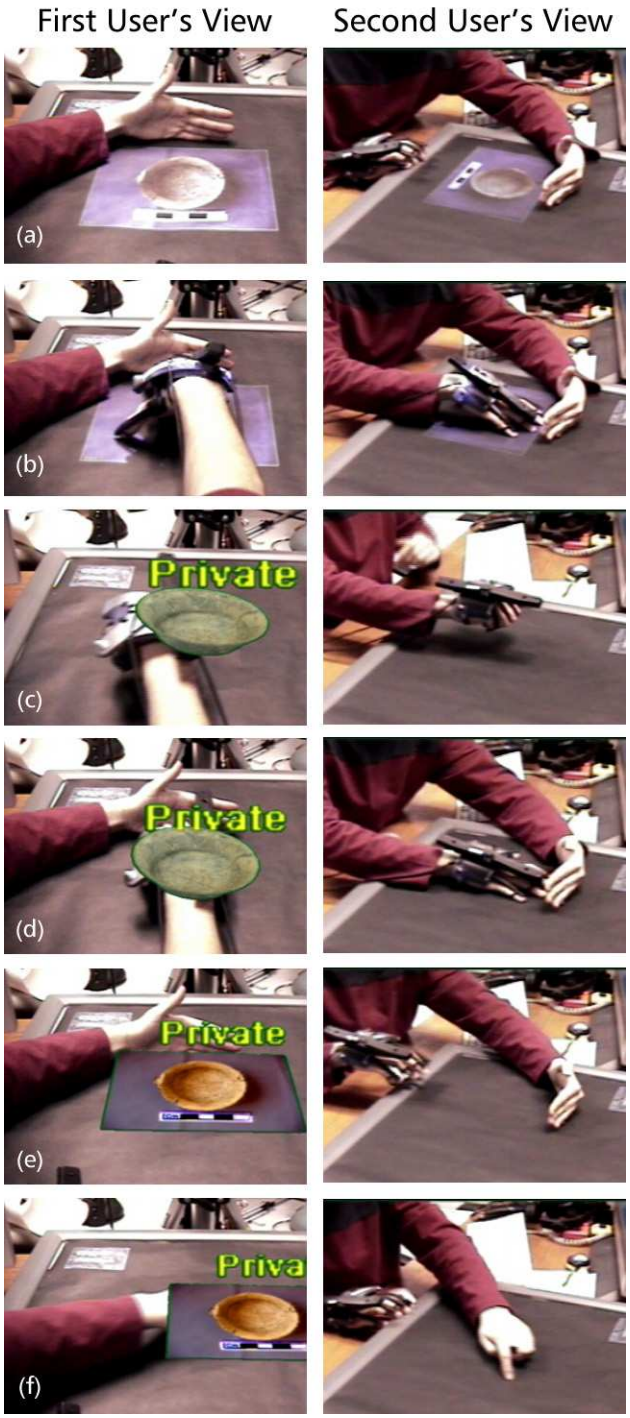

Figure 8: Side-by-side comparison of two users' views when using two-handed privacy gestures. (a) User places a vertical hand on the table to indicate a private gesture. (b) Cross-dimensional grab gesture is performed with the vertical hand above it, yielding a cross-dimensional privacy pull. (c) 3D object is private and invisible to the other user. (d) Performing cross-dimensional push with a vertical hand above it, yielding a cross-dimensional privacy push. (e) 2D object appears as if on the table, but is invisible to the other users. (f) 2D private object behaves like all other 2D objects and is dragged around.

One PC workstation (dual Athlon MP 2.0, 1GB RAM, Nvidia Quadro4 750 XGL) powers each module and all communication is implemented via message passing through a publish-and-subscribe message board system. We use the Adaptive Agent Architecture (AAA) [11] to facilitate easy connection, discovery and communication management of our modules.

Synchronization of Gesture Recognition

All the hybrid gesture techniques presented here require that the 2D module and 3D modules communicate extensively and synchronize their gesture recognitions. Figure 9 illustrates the message passing and synchronization between the two modules during the recognition of a *cross-dimensional pull* gesture. In this example, all messages are assumed to pertain to the same user; however in our multi-user system, each message is tagged with the "user name" tag to allow for disambiguation among users.

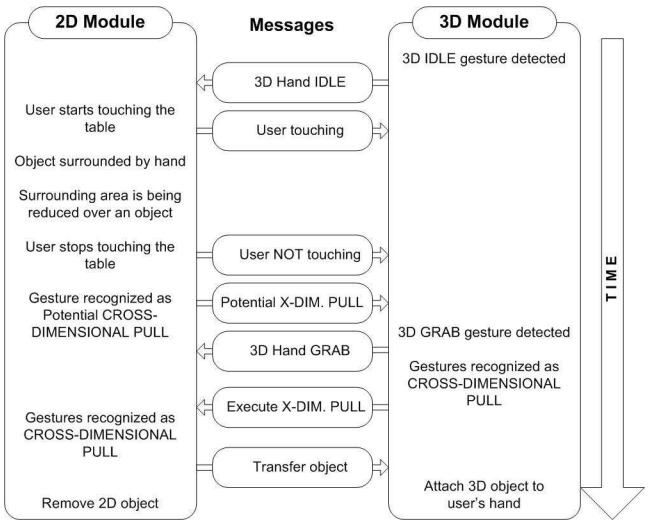

Figure 9: Sample communication between 2D and 3D modules during the recognition of *cross-dimensional pull* (abbreviated as X-DIM. PULL).

DISCUSSION AND FUTURE WORK

Our early test users have pointed out that the visual feedback is essential to making our hybrid interactions feel like seamless transitions (particularly for cross-dimensional push and pull). Our initial implementation of those techniques did not employ animation, and many users commented on the visual discontinuity of the object disappearing in one environment and appearing in the other. Through several iterations, we have added simple animations to the interactions that have improved continuity and visual feedback. Audio feedback could also be incorporated to further enhance cross-dimensional transitions.

We have considered the idea of making the rotation a two-handed or two-fingered technique (similar to Wu and Balakrishnan's freeform rotation technique [22]), but in our experience with the system, we have noticed that rotation is clearly specified by fixing the center of rotation to the center of the object and using one additional point to specify the angle. Hence we have decided that it is most simply performed as a one-handed gesture.

Furthermore this frees up the other hand to make additional simultaneous object transformation.

Our gestures are designed to require only one 3D tracked hand. This is sufficient for a variety of gestures presented here, however obviously more gestures would be possible if both hands were tracked. However, only right-handed P5 gloves are available from Essential Reality. Since the DiamondTouch table requires that users be electrostatically coupled, the use of many other glove-based devices was not possible because they cover the palm and the fingertips of the hand, affecting the table's ability to track the user. While we currently use the tethered P5 glove to determine 3D gestures, we plan to experiment with video-based gesture tracking in the future, which would eliminate the need for a glove and allow for much more convenient setup.

During implementation and refinement of our gestural techniques, several coworkers in our lab participated in the preliminary informal tests. Their initial feedback suggests that learning to perform our hybrid interactions does not require much time, due to their similarity to physical world interactions. However, we plan to perform a more thorough, formal user study in the near future to evaluate effectiveness and ease of use of these interaction techniques.

CONCLUSIONS

This paper presents a set of interaction techniques designed for hybrid user interfaces that use 2D touch-sensitive displays and 3D AR user interfaces. Our system uses synchronized 2D and 3D gestures to facilitate seamless transition of data and interactions across dimensionalities. In addition, our techniques are designed for multi-user collaborative environments and thus address some of the privacy needs in these shared spaces. Initial informal user feedback suggests that learning the techniques does not require much time, due to their close similarity to physical world interactions. While our techniques are presented within a simple application scenario, we believe that they easily extend to a broader range of hybrid environments.

ACKNOWLEDGMENTS

We wish to thank Sajid Sadi and Avanindra Utukuri for their support with the Essential Reality P5 glove, and Ian Morris and the Stanford Archaeology Center for images and data from the Monte Polizzo excavation. Mitsubishi Electric Research Labs generously donated the DiamondTouch table. This research was funded in part by Office of Naval Research Contract N00014-04-1-0005, NSF Grants IIS-01-21239 and IIS-00-82961, and gifts from Microsoft Research and Alias Systems.

VIDEO

All techniques presented in this paper can be previewed in a digital video available for download from www.cs.columbia.edu/graphics/projects/HybridInteraction.

REFERENCES

1. Brown, L., Hua, H., and Gao, C. "A Widget Framework for Augmented Interaction in SCAPE." *Proc. Symp. on User Interface Software and Technology (ACM UIST 2003)*. 2003, 1–10.

2. Butz, A., Höllerer, T., Feiner, S., MacIntyre, B., and Beshers, C. "Enveloping Users and Computers in a Collaborative 3D Augmented Reality." *Proc. International Workshop on Augmented Reality (IWAR '99)*, San Francisco, CA, October 20–21, 1999, 35–44.

3. Dietz, P. and Lehigh, D. "DiamondTouch: a Multi-User Touch Technology." *Proc. Symp. on User Interface Software and Technology (ACM UIST '01)*, 2001, 219–226.

4. Feiner, S. and Shamash, A. "Hybrid User Interfaces: Breeding Virtually Bigger Interfaces for Physically Smaller Computers." *Proc. Symp. on User Interface Software and Technology (ACM UIST '91)*. 1991.

5. Herndon, K., Zeleznik, R., Robbins, D., Conner, D.B., Snibbe, S., van Dam, A. "Interactive Shaddows." *Proc. Symp. On User Interface Software and Technology (ACM UIST 1992)*. ACM Press, November 15–18, 1992, 1–6.

6. Hinckley, K. "Synchronous Gestures for Multiple Persons and Computers." *Proc. Symp. on User Interface Software and Technology (ACM UIST 2003)*. 2003, 149–158.

7. Ishii, H., Kobayashi, M. and Arita, K. "Iterative Design of Seamless Collaboration Media." *Communications of the ACM*. August 1994. Vol. 37, No. 8.

8. Ishii, H. and Ullmer, B. "Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms." *Proc. ACM SIGCHI '97*, Atlanta, GA. March 22–27, 1997.

9. Kaiser, E., Olwal, A., McGee, D., Benko, H., Corradini, A., Li, X., Cohen, P., and Feiner, S. "Mutual disambiguation of 3D multimodal interaction in augmented and virtual reality." *Proc. ICMI 2003* (*Fifth Int. Conf. on Multimodal Interfaces*), Vancouver, BC, November 5–7, 2003, 12–19.

10. Krueger, M. "VIDEOPLACE and the interface of the future." In *The art of human computer interface design,* B. Laurel, Editor. 1991, Addison Wesley: Menlo Park, CA. 417–422.

11. Kumar, S., Cohen, P., and Levesque, H. "The Adaptive Agent Architecture: Achieving Fault-Tolerance Using Persistent Broker Teams." *Proc. Int. Conf. on Multi-Agent Sys.*, July 2000, 159–166.

12. Low, K-L., Ilie, A., Welch, G., and Lastra, A. "Combining head-mounted and projector-based displays for surgical training." *Proc. IEEE VR 2003,* March 22–26, 2003, 110–117.

13. Mine M, Brooks F, Squin C. “Moving objects in space: Exploiting proprioception in virtual-environment interaction.” *Proc. of ACM SIGGRAPH '97*. 1997, 19-26.

14. Olwal, A., Benko, H., and Feiner, S. “SenseShapes: Using Statistical Geometry for Object Selection in a Multimodal Augmented Reality System.” *Proc. IEEE and ACM Int. Symp. on Mixed and Augmented Reality (ISMAR 2003)*. Tokyo, Japan. October 7–10, 2003, 300–301.

15. Regenbrecht, H.T., Wagner, M.T., and Baratoff, G. “MagicMeeting: A Collaborative Tangible Augmented Reality System.” *Virtual Reality—Systems, Development and Applications*, 6(3), Springer-Verlag, 2002, 151–166.

16. Rekimoto, J. “Pick-and-Drop: A direct manipulation technique for multiple computer environments.” *Proc. Symp. on User Interface Software and Technology (ACM UIST ’97)*, 1997, 31–39.

17. Rekimoto, J. “SmartSkin: an Infrastructure for Freehand Manipulation on Interactive Surfaces.” *Proc. ACM SIGCHI ’02*. 2002, 113–120.

18. Rekimoto, J. and Saitoh, M. “Augmented Surfaces: a Spatially Continuous Work Space for Hybrid Computing Environments.” *Proc. ACM SIGCHI ‘99*. 1999.

19. Schmalstieg, D., Fuhrmann, A., Hesina, G., Szalavari, Zs., Encarnação, L. M., Gervautz, M., and Purgathofer, W. “The Studierstube Augmented Reality Project.” *PRESENCE: Teleoperators and Virtual Environments*, 11(1), MIT Press. 2002.

20. Streitz, N., Geißler, J., Holmer, T., Konomi, S., Müller-Tomfelde, C., Reischl, W., Rexroth, P., Seitz, P., and Steinmetz, R. “i-LAND: an interactive landscape for creativity and innovation.” *Proc. ACM SIGCHI*. 1999, 120–127.

21. Tandler, P., Prante, T., Müller-Tomfelde, C., Streitz, N., and Steinmetz, R. “Connectables: dynamic coupling of displays for the flexible creation of shared workspaces.” *Proc. Symp. on User Interface Software and Technology (ACM UIST ’01)*. 2001, 11–20.

22. Wu, M. and Balakrishnan, R. “Multi-Finger and Whole Hand Gestural Interaction Techniques for Multi-User Tabletop Displays.“ *Proc. Symp. on User Interface Software and Technology (ACM UIST 2003)*. 2003, 193–202.