Development and Evaluation of Mixed Reality Interaction Techniques

Edward Ishak* Hrvoje Benko† Steven Feiner‡

Department of Computer Science, Columbia University
500 W. 120th St., 450 CS Building, New York, NY

ABSTRACT

We present our thoughts on the development and evaluation of novel interaction techniques for mixed reality (MR) systems, particularly those consisting of a heterogeneous mix of displays, devices, and users. Interaction work in MR has predominantly focused on two areas: glove or wand-based virtual reality (VR) interactions and tangible user interfaces. We believe that MR users should not be limited to these approaches, but should rather be able to utilize all available devices and interaction methods in their environment to make use of the most relevant ones for the current task at hand. Furthermore, we will discuss the difficulty in finding appropriate ways to evaluate new techniques, primarily due to the lack of standards of comparison.

CR Categories: H.5.1 **[Information Interfaces and Presentation]**: Multimedia Information Systems—Artificial, augmented, and virtual realities, Evaluation/methodology; H.5.2 **[Information Interfaces and Presentation]**: User Interfaces—Graphical user interfaces (GUI), Interaction styles

Keywords: hybrid user interfaces, mixed reality, augmented reality, tabletop interaction, gesture-based and touch-based interaction, user evaluation.

1. INTRODUCTION

The issues we present here have grown out of our experiences developing VITA (Visual Interaction Tool for Archaeology) [2], a collaborative MR environment that supports multiple users wearing head-worn displays and tracked gloves to visualize and interact with archaeological data. As we built our first prototype, it became clear that users would benefit from being able to regularly "context switch" between 2D and 3D representations of archaeological data. For example, a high-resolution 2D image of a pot does not allow a user to fully grasp its contours and overall 3D structure, which is much better accomplished through a 3D representation. Our first prototype, which relied in part on the traditional glove-based VR/MR user interface of our previous systems [6, 9], did not provide our users with usable and intuitive ways to perform these "cross-dimensional interactions." This led us to develop new cross-dimensional gestures [1], an example of which is shown in Figure 1, which allow users to move objects from 2D images viewed on a projected tabletop to 3D objects experienced on a stereo see-through head-worn display.

Although VITA users generally found these gestures to be useful and enjoyable, we found it difficult to evaluate the usability of these gestures. Since there is no established standard for the task that these gestures accomplish, we began to recollect the alternative sets of gestures that we identified during the early development stages, but were previously discarded because of their lack of intuitiveness or usability. We contemplated performing an evaluation comparing these discarded sets of interactions with our newly developed gestures, but questioned how well such an evaluation would be accepted, since we had designed both the experimental and control sets.

Figure 1: Frames from a cross-dimensional pull gesture implemented in VITA. (left) Selecting a 2D object and starting to form a grabbing gesture. (middle) The 3D object appears from the table. (right) Holding the 3D object. (Table is covered with black paper to provide a darker background for imaging the projected display through the live tracked video see-through display.)

2. LIMITATIONS OF VR INTERACTIONS IN MR ENVIRONMENTS

Users of many current MR systems are limited to conventional glove or wand-based interactions derived from traditional VR systems. Although MR systems inherit a number of characteristics from their VR predecessors, there are inherent features of MR systems that cause these traditional interaction tools to be quite limiting or simply unintuitive. For example, there are interactions that transition data across devices [8] and displays (e.g., [3, 5, 12, 13]), or even across dimensionalities (e.g., 2D and 3D [7, 11]), rather than just manipulating data on a single device or display. These types of interactions can make use of the surrounding environment, including any physical device that is visible to the user. For example, while tracked pads, both active and inactive, have been used in VR [10] and MR [11, 14], a conventional *untracked* touchpad can also be used in MR, since the user can see it even though the system might not know its location.

Because MR intentionally includes the surrounding environment, there is a need for interaction metaphors that facilitate seamless transitions between the virtual portion of the MR world and the surrounding displays and interaction devices of the physical world. For example, VITA uses a Mitsubishi DiamondTouch table [4], a large 2D projected touch-sensitive surface that distinguishes among multiple users and multiple touches of the same user. This device's capabilities make it an excellent candidate for aiding multiple users in performing cross-dimensional interactions. Although our transitional gestures differ in the tasks they perform, they all share several high-level features: (1) a source object, (2) a gesture performing the transition, (3) a destination, and (4) an intuitive way to transition back to the original state. Furthermore, our initial intent was to design the interactions so that the gesture to transition back was symmetric to its counterpart. For example, consider our "cross-dimensional push" (Figure 1) and "cross-dimensional pull" gestures, which allow a user to transfer a 2D object displayed on the DiamondTouch surface into

* e-mail: ishak@cs.columbia.edu
† e-mail: benko@cs.columbia.edu
‡ e-mail: feiner@cs.columbia.edu

its 3D representation in the surrounding 3D world, and vice versa, respectively. Our original implementation allowed recognition of a pull gesture to immediately follow a drag of a 2D object without releasing the touch surface. This was symmetric to the recognition of a push gesture, which could immediately follow a drag of a 3D object by a clenched tracked fist. We often noticed accidentally recognized pull gestures when the 3D-tracked hand (which inherently lives in 3D space) was removed from the currently touched table to perform another task immediately following a 2D drag. In contrast, users rarely touched the 2D surface accidentally immediately following a 3D drag and would intuitively release a 3D-dragged object before interacting with the surface. We, therefore, redesigned the logic to disallow a pull immediately following a 2D drag (without first breaking contact with the 2D surface), making our once symmetric gestures now asymmetric.

Furthermore, we feel that as future MR systems begin to allow users to use more of the surrounding environment, these systems will evolve to recognize a wider variety of interaction devices rather than the traditional VR interaction tools, providing users a wider range of possibilities for completing their tasks.

3. Evaluating Novel MR Techniques for Novel Tasks

While designing and prototyping our cross-dimensional interaction techniques, we struggled to find meaningful ways to evaluate them in comparison with others, because we found no other systems that performed these tasks. Developing interaction techniques for which there are no comparable standards often forces the developer to implement alternate sets of interactions that accomplish the same goals solely for evaluation. We feel that, depending on the nature of this alternate interaction set, this method of evaluation makes the entire comparison less meaningful.

Formulating alternate techniques is challenging because the range of possibilities can potentially be quite vast. For example, consider our "cross-dimensional pull" gesture (Figure 1). It involves explicitly selecting the 2D source object on the surface, making the grab gesture (during which the object transitions to 3D), and dragging the 3D object to its destination, all in one seamless, continuous motion. This was inspired by observing users who we asked to pick up real objects from a table to help us design the interaction to closely mimic a real world gesture. Although speed and accuracy were of concern, we felt the gesture most importantly needed to provide high user satisfaction. However, to adequately evaluate it, we have to consider suitable alternate techniques and compare them to the cross-dimensional pull. One way to do this is to consider alternative "seamless" cross-dimensional gestures that perform the same task. Although such gestures may not be in current use, we can design them by observing how a user might naturally attempt to transform a 2D object into 3D. For example, a user may initially grab the 2D object from a 2D surface using a single hand, and then try to extrude the depth from the object using two hands that start together and slowly separate.

Another way to identify alternative interaction gestures is to construct techniques based on current interaction design heuristics. For example consider the following sequence: (1) using a single finger, the user selects a 2D object on the 2D touch surface; (2) using a context-sensitive menu, the user selects a "transfer to 3D" option; and (3) the user grabs the resulting 3D object and drags it to its destination. Although some might agree that menu selection is a straightforward method of interaction, and therefore, a likely candidate for comparison with new techniques, unfortunately, the evaluation results can vary greatly depending on how the alternate set is implemented. For example, the implementation of a touch-based menu system can range from a single-level one with very few choices, to a multi-level one with numerous choices.

Conclusions

We have presented our thoughts on the need for MR interactions not to be limited to VR-like interaction methods. We feel it is important to understand the nature of MR environments and to design interactions around the specific motivation for creating these environments. We also feel that, in some situations, user interface evaluations are less meaningful than in others, especially when they are potentially biased, unavailable for verification, and rely on the novel and "comparable alternative" techniques all being developed and being evaluated by the same designers. We speculate that that is one reason why many 3DUI interactions have not been formally evaluated. Evaluation is necessary, but it is very difficult to accomplish in many instances. Hopefully, as this relatively young field matures, we will see standards emerge in the coming years.

References

[1] Benko, H., Ishak, E. and Feiner, S., "Cross-Dimensional Gestural Interaction Techniques for Hybrid Immersive Environments," *Proc. IEEE Virtual Reality*, Böhn, Germany, 2005.

[2] Benko, H., Ishak, E. W. and Feiner, S., "Collaborative mixed reality visualization of an archaeological excavation," *Proc. ISMAR 2004 (IEEE and ACM Int. Symp. on Mixed and Augmented Reality)*, Arlington, VA, November 2–5, 2004, pp. 132–140.

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

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

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

[6] 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. International Conference on Multimodal Interfaces*, Vancouver, BC, Canada, November 5–7, 2003, pp. 12–19.

[7] Krueger, M., *VIDEOPLACE and the interface of the future*, in B. Laurel, ed., *The art of human computer interface design*, Addison Wesley, Menlo Park, CA, 1991, pp. 417–422.

[8] Myers, B. A., *Using Handheld Devices and PCs Together*, Commun. ACM, 44 (2001), pp. 34–41.

[9] 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, pp. 300–301.

[10] Poupyrev, I., Tomokazu, N. and Weghorst, S., "Virtual Notepad: Handwriting in Immersive VR," *Proc. VRAIS*, Atlanta, GA, 1998.

[11] Regenbrecht, H. T., Wagner, M. T. and Baratoff, G., *MagicMeeting: A Collaborative Tangible Augmented Reality System*, Virtual Reality - Systems, Development and Applications, 6 (2002), pp. 151–166.

[12] Rekimoto, J., "Pick-and-Drop: A Direct Manipulation Technique for Multiple Computer Environments," *Proc. User Interface Software and Technology*, 1997, pp. 31–39.

[13] Rekimoto, J. and Saitoh, M., "Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments," *Proc. ACM CHI '99*, Pittsburgh, PA, May 15–20, 1999.

[14] Schmalstieg, D., Fuhrmann, A., Hesina, G., Szalavari, Z., Encarnação, L. M., Gervautz, M. and Purgathofer, W., *The Studierstube Augmented Reality Project*, Presence: Teleoperators and Virtual Environments, 11 (2002).