A Virtual Environment for Studying Immersion with Low-Cost Interaction Devices

Connie Lu[1]
University of California San Diego
Human-Computer Interaction
(Cognitive Science)

Craig Disselkoen[2]
Dordt College
Computer Engineering
Mathematics

Lanya Butler[3]
Colby College
Interdisciplinary Computing:
Photography

Meisha Rosenberg[4]
Iowa State University
Mechanical Engineering
Human-Computer Interaction

Judy Vance, Ph. D.[5]
Iowa State University
Mechanical Engineering
Human-Computer Interaction

ABSTRACT

Our research explores various low-cost virtual reality interface technologies, such as the Microsoft Kinect, Nintendo Wiimote, and Razer Hydra, to determine the unique features and limitations of each device with regards to usability and ease of interaction. We set these devices up in three different configurations—one with a Wiimote used for head tracking and two Hydra wands used for hand tracking, another with the Hydra used for both head and hand tracking, and the last one simply with the Kinect. Based on our tests of these three configurations, we hope that we can provide meaningful insights into the usability of low-cost VR technologies.

KEYWORDS

Virtual reality, low cost, 3D input devices, tracking

1. INTRODUCTION

The keyboard and mouse, which have traditionally dominated as input devices for computer interfaces, fail to provide smooth and intuitive ways to interact with a 3-dimensional environment. Virtual reality devices allow users to interact with these environments and the objects in them fluidly, just as they would in the real world. However, existing systems are too expensive to be used on a wide scale, costing thousands of dollars or more [1][6]. The continuous growth of the gaming industry and its investment in the development of gaming technology has facilitated the emergence of new interaction devices including the Nintendo Wiimote, Microsoft Kinect, and the Razer Hydra. Although these have traditionally been used in entertainment systems, they can also be repurposed to function in virtual reality (VR) applications. We tested three different VR systems based on these devices to evaluate which of them provides the most immersive VR experience and most fluid interface. All of these systems can be assembled at a fraction of the cost of other commercially available systems. Our systems could make virtual reality more accessible, ultimately leading to more wide-scale adoption of these interfaces.

2. RELATED WORK

The need for low-cost VR solutions is not new. As early as 1991, Pausch designed a system using a PC, a Polhemus Isotrak, two Reflection Technology Private Eye displays, and a Mattel Power Glove, in an effort to make immersive VR available to a broader range of researchers. Although the limited technology at his disposal led to issues with performance and usability, his approach inspired others to think about inexpensive systems that could be implemented for “five dollars a day” [2]. Technological advancements have since dramatically improved the capabilities of computing hardware, and the increased availability of mass-produced tracking devices has spurred a renewed interest in inexpensive VR systems.

[1]col003@ucsd.edu [2]crgdsslk@dordt.edu
[3]llbutler@colby.edu [4]meishar@iastate.edu
[5]jmvance@iastate.edu

One of the new systems that has potential for low-cost VR is the Microsoft Kinect. While some researchers have worked with the Kinect's tracking capabilities by proposing new techniques for rapidly and accurately predicting 3D positions of the human body [3], others have focused on the possible use of the Kinect in interaction applications. For instance, in 2011 Blanchard *et al* took advantage of the nature of the Kinect to develop Kinoogle, a natural user interface for navigating Google Earth using hand and bodily gestures [4]. This system is an excellent example of the use of the Kinect to achieve low-cost VR. Our study compares interfaces like these for the Kinect with interfaces made possible by other devices.

Low-cost tracking devices are not limited to the Kinect. The Wii, which incorporates tracking through its controller, the "Wiimote", is one of the fastest selling gaming consoles in history, and has gained popularity not only as an entertainment device but also as a platform for exploring interaction techniques [5]. Researchers such as [5] and [8] have utilized the Wiimote in its traditional gaming format - with the remote held in hand and stationary IR emitters placed in the environment. Other researchers have employed "Wiimote hacking," reverse-engineering the input device to use the underlying technology in alternative applications. For instance, rather than using the Wii in its traditional format, Calderwood *et al.* incorporated IR emitters into transparent glasses for the user to wear, while statically mounting the Wiimote in the environment. In this way he achieved head tracking using the Wiimote, without overly encumbering the user [14]. Similarly, Lee created a multi-touch interactive whiteboard surface by mapping the Wiimote's camera coordinate system to that of the whiteboard display's coordinates. Unfortunately, his approach was limited by the tracking resolution and the tracking quality of the Wiimote due to sensitivities in positioning and occlusions, issues inherent in using the Wiimote for tracking [7].

The Razer Hydra is a recently released system that is just beginning to be explored by researchers. As one example, Altenhoff *et al.* used the Hydra as an input device with an eye toward evaluating its possible use to control robotic surgical systems [13]. This study, however, focused on the effect of stereoscopic viewing on task performance, rather than the effect of the use of the Hydra as an input device. In a different study, Basu *et al* used the Hydra as a second tracking source alongside the iPod Touch 4G sensor system, and used the resulting tracking information to allow the user to interact with a virtual environment rendered in a head mounted display (HMD) [1]. Separately, Kuntz and Ciger explored using the Hydra for hand tracking and button-based input along with a HMD for low-cost immersive VR [10]. These two systems, however, are still costly enough to be a barrier for some, costing approximately €1000, mostly due to their use of HMDs.

Our research builds on past studies involving low-cost VR technologies to explore the unique features and limitations of three specific interaction devices—the Wiimote, the Kinect, and the Hydra. In doing this, we hope our research will allow others to gain a better understanding of these technologies, as well as demonstrate the flexibility and potential of modern low-cost tracking devices.

3. METHODOLOGY

To explore the differences between the various interaction technologies, we created a conceptual product design application that serves to showcase the advantages as well as the limitations of these technologies. We built the application using the VR JuggLua framework developed by Ryan Pavlik [11] and the Virtual Reality Private Network (VRPN) [12] to use with three different configurations of input technologies. In the application, the user can build complex models using 3D primitives in a 3D environment. After selecting a shape and color via a menu, the primitives can be drawn onto the environment using a series of extrusions. Each primitive drawn can also be manipulated by moving, rotating, and scaling; there is also the capability to stretch a primitive along one of its axes to create more unusual looking primitives. Primitives can also be duplicated or deleted. Primitives can be placed together to create a more complicated shape, and with the multiple shape selection can be moved as easily as a single primitive. In addition, the camera can be freely manipulated in the space through translation, rotation, and zooming.

This type of application requires an involved interface, which makes it a useful tool for evaluating different interaction devices. It requires heavy use of 3D visualization, manipulation of objects in 3D space, navigation in 3D space, and pointing in a 3D environment, functions that are awkward and overly complicated when implemented in 2D, as with a keyboard and mouse. The complexity of the interaction also sharply brings out the differences between interfaces. Therefore, we used this application as a case study for evaluating different interaction devices.

We formed our chosen input devices into three configurations for testing, each of which was a complete interface. We chose the configurations for their low cost, immersiveness, wide-scale availability, and intuitiveness. We believed they would provide acceptable performances at a fraction of the cost of high-end devices. The component devices are all easily accessible and simple to set up. In addition, the existence of open source SDK's for the Microsoft Kinect and the Razer Hydra make them easy to reverse engineer and manipulate.

The Razer Hydra is an affordable ($80 per unit) solution for magnetic tracking. It consists of two hand-held controllers wired to a base. Each controller reports full 3DOF position as well as 3DOF orientation at 250 Hz [1]. In addition, each controller is outfitted with seven digital buttons, one analog trigger, and one analog stick with 2-dimensional motion. The hand-held controllers ("wands") are also lightweight, weighing only 800g each. The Hydra is entirely powered via a USB connection to its base.

The Wiimote, at $30 per unit, is also a useful device for tracking. It contains an IR camera and is capable of calculating and reporting its position relative to two IR LEDs via Bluetooth. When used in video gaming, it is traditionally held in the hand while a pair of IR emitters are positioned in a static location. In our system, we placed the Wiimote in a static location in the environment, and then mounted IR LEDs on the user's head so as to achieve head tracking without overly encumbering the user, similar to the system in [14].

Figure 1: The Nintendo Wiimote and Razer Hydra configuration.

In the first configuration, we combined the Hydra and the Wiimote - the two wands of the Hydra were used for hand tracking, and the Wiimote was used for head tracking. We appropriated all eight buttons and the analog stick of the right-hand Hydra remote, each featuring a different function that could be used in our demonstration application.

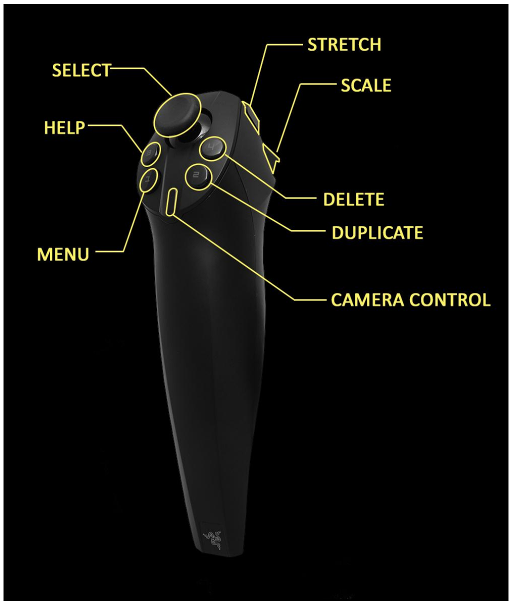

Figure 2: Diagram of the Razer Hydra controls.

We chose not to use the buttons on the left-hand remote, opting instead to use it only as a tracking device to allow for greater flexibility in manipulating 3D objects by taking advantage of a second hand. This had the additional benefit of providing continuity with the Hydra-only configuration detailed next. Finally, since we are merely using the Wiimote as a camera for head tracking, the buttons on that device were left unmapped.

Figure 3: The Hydra-only configuration.

For the second configuration, we used only the Hydra device. We removed the magnetic tracking sensor out of the body of one Hydra wand and attached it to the temple of a pair of glasses, using a 3D-printed plastic housing. The other wand was used as a hand-held input device, as it was in the first configuration, including the use of all of the buttons on the one remaining hand. By doing so, we were able to perform all of the tracking with the Hydra alone. This reduced the cost of the system while still providing accurate hand and head tracking; the main difference was the availability of only one hand to interact with the application.

Figure 4: The Microsoft Kinect.

The third configuration consisted of one Microsoft Kinect unit responsible for all motion tracking and user input. The Kinect utilizes an RGB camera with 1280x960 resolution, as well as an IR emitter and IR sensor to obtain depth information. It has a 43° vertical by 57° horizontal field of view and can report video data at 30 FPS. It also contains a multi-array microphone with noise canceling. At $250 per unit, it is capable of providing full-body motion tracking, gesture recognition, and voice recognition. We implemented a library of gestures and voice controls to be used in our application.

4. RESULTS

In our preliminary tests, we found that the application was much easier to use with the configurations involving the Razer Hydra than with the Kinect. One reason for this is that the Hydra's digital buttons provide a fluid way to issue commands while moving the wand to manipulate the object. For instance, it was easy to quickly create many duplicates of a primitive by simply moving it into each position using the wand and then pressing the duplicate button. Another reason is that digital controls such as buttons have a much faster response time and are much easier to activate than voice-recognition or gesture-recognition controls. Buttons respond almost instantly, while processing creates a delay for voice or gesture commands. The tracking was also smoother when using the Hydra than when using the Kinect.

Gesture recognition with the Kinect was, in our experience, not robust enough to support a reliable complex interface. Developers and users alike still face challenges when it comes to defining a clear gesture library for the Kinect that is appropriate and intuitive [9]. We substituted voice commands for some functions because of this difficulty. Head tracking felt the most natural with the Kinect, as opposed to in the other configurations, because no glasses or other wearable devices were required. It worked smoothly when the correct operational range was observed. The hand tracking, however, was insufficient to create a fluid interaction experience due to noise. Additionally, the voice recognition was not natural since the inherent latency was too great.

We found that adding tracking for the second hand, as we did in the Hydra and Wiimote configuration, did not provide significant advantages with regards to usability. All of the interactions necessary for our application could be performed using only one hand; we struggled to write an interface for two-handed tracking that seemed to have any advantage over a one-handed system. For head tracking, the magnetic tracking of the Hydra was smoother and offered greater range than that of the Wiimote. The Hydra-only system is also cheaper to implement because it does not require a Wiimote.

In short, we found that the configuration involving the Hydra for both hand and head tracking provided the most advantages with regards to usability. The Kinect configuration is promising in concept, but the current implementation has fundamental deficiencies in usability, at least for our application. The Hydra provided the most fluid interface for user input.

5. FUTURE WORK

Our research developed an application to explore the differences between virtual reality input devices. Further research could expand our informal testing into a full user study to compare the immersiveness, learning speed, and ease of use of these technologies. This user study could also incorporate timing tasks to compare each of the configurations.

Although our application provides interactions representative of a wide variety of tasks typically required in VR input scenarios, further research could be done to verify or expand on our findings for specific applications currently in use in the field. Our research does not address any specific use cases for VR input, and the optimal input technology may vary by use case.

We struggled with implementing a robust gesture library for the Kinect with regards to our application. Further work could be done to build a more complete gesture library and compare it to other forms of VR input. Future hardware developments may also make optical tracking technologies, such as those used by the Kinect, more comparable in tracking performance to magnetic trackers such as the Razer Hydra. If the tracking were improved, this technology has the potential to provide a more natural and immersive user experience.

6. CONCLUSION

Our work suggests that for many applications, 3D motion tracking should be supplemented with digital controls such as buttons to create a more fluid interface for the user. The recently released Razer Hydra is a desirable input method combining digital buttons with 3D motion tracking. Our work also reaffirms the feasibility of low-cost equipment for implementing VR experiences, which we hope will lead to greater adoption of these interfaces in the future.

ACKNOWLEDGMENTS

We would like to thank Patrick Carlson, Timothy Morgan, Diana Jarrell, and Ryan Pavlik from the Virtual Reality Applications Center Haptics Lab for their contributions and support.

This research was performed at Iowa State University as part of the SPIRE-EIT research internship sponsored by the university’s Human Computer Interaction Program and the Virtual Reality Applications Center during the summer of 2013. This material is supported in part by the National Science Foundation Grant CNS-1156841. All findings and conclusions presented are those of the authors and do not necessarily reflect the ideas, beliefs, and opinions of the National Science Foundation.

REFERENCES

[1] Basu, A.; Saupe, C.; Refour, E.; Raij, A.; Johnsen, K., "Immersive 3DUI on one dollar a day," *3D User Interfaces (3DUI), 2012 IEEE Symposium on* , vol., no., pp.97,100, 4-5 March 2012.

[2] R. Pausch, “Virtual reality on five dollars a day,” *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology, CHI ’91*, pp. 265-270, 1991.

[3] Shotton, J.; Fitzgibbon, A.; Cook, M.; Sharp, T.; Finocchio, M.; Moore, R.; Kipman, A.; Blake, A., "Real-time human pose recognition in parts from single depth images," *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on* , vol., no., pp.1297,1304, 20-25 June 2011.

[4] Blanchard, B., Kamel Boulos, M., Gutierrez-Osuna, R., Montero, J., Tripathy, A., Walker, C. (2011). Web GIS in Practice X: A Microsoft Kinect Natural User Interface for Google Earth

Navigation.

[5] Thorpe, A.; Minhua Ma; Oikonomou, A., "History and alternative game input methods," *Computer Games (CGAMES), 2011 16th International Conference on* , vol., no., pp.76,93, 27-30 July 2011.

[6] Suma, E.A.; Lange, B.; Rizzo, A.S.; Krum, D.M.; Bolas, M., "FAAST: The Flexible Action and Articulated Skeleton Toolkit," *Virtual Reality Conference (VR), 2011 IEEE* , vol., no., pp.247,248, 19-23 March 2011.

[7] Lee, J.C., "Hacking the Nintendo Wii Remote," *Pervasive Computing, IEEE* , vol.7, no.3, pp.39,45, July-Sept. 2008.

[8] Ardito, C.; Buono, P.; Costabile, M.F.; Lanzilotti, R.; Simeone, A.L., "Comparing low cost input devices for interacting with 3D Virtual Environments," *Human System Interactions, 2009. HSI '09. 2nd Conference on* , vol., no., pp.292,297, 21-23 May 2009.

[9] Cruz, L.; Lucio, D.; Velho, L., "Kinect and RGBD Images: Challenges and Applications," *Graphics, Patterns and Images Tutorials (SIBGRAPI-T), 2012 25th SIBGRAPI Conference on* , vol., no., pp.36,49, 22-25 Aug. 2012.

[10] Ciger, J., Kuntz, S., "Low-Cost and Home-Made Immersive Systems," *International Journal of Virtual Reality*, pp. 1-9, 2012.

[11] Pavlik, R.A.; Vance, J.M., "VR JuggLua: A framework for VR applications combining Lua, OpenSceneGraph, and VR Juggler," *Software Engineering and Architectures for Realtime Interactive Systems (SEARIS), 2012 5th Workshop on* , vol., no., pp.29,35, 5-5 March 2012.

[12] Taylor, R. M., II, Hudson, T. C., Seeger, A., Weber, H., Juliano, J., & Helser, A. T., "VRPN: A device- independent, network-transparent VR peripheral system," Proceedings of the ACM Symposium on Virtual Reality Software & Technology, VRST, 2001.

[13] Altenhoff, B. M., Sewall, A. S., Shannon, B., & Duchowski, A. T., "Effects of Haptic Feedback and Stereoscopy on a Surgical Training Task in Virtual Reality," 2012.

[14] Calderwood, M. D., Kelly, J., McNamara, T. P., & Bodenheimer, B., "Adding head tracking to desktop virtual reality with the Wii remote as an aid to spatial cognition," Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization - APGV '09, p. 125, 2009.