International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2017)
R. Lindeman, G. Bruder, and D. Iwai (Editors)

Asymmetric Bimanual Interaction for Mobile Virtual Reality

Huidong Bai†1, Alaeddin Nassani‡1, Barrett Ens§2, Mark Billinghurst¶2

[1]HIT Lab NZ, University of Canterbury, New Zealand
[2]School of Information Technology and Mathematical Science, University of South Australia, Australia

Abstract

In this paper, we explore asymmetric bimanual interaction with mobile Virtual Reality (VR). We have developed a novel two handed interface for mobile VR which uses a 6 degree of freedom (DoF) controller input for the dominant hand and full-hand gesture input for the non-dominant hand. We evaluated our method in a pilot study by comparing it to three other asymmetric bimanual interfaces (1) 3D controller and 2D touch-pad, (2) 3D gesture and 2D controller, and (3) 3D gesture and 2D touchpad in a VR translation and rotation task. We observed that using our position aware handheld controller with gesture input provided an easy and natural experience.

CCS Concepts

•***Human-centered computing*** → ***Virtual reality; Interface design prototyping;***

1. Introduction

Our research is concerned with how to enable intuitive input in mobile Virtual Reality (VR) interfaces. Mobile VR provides an inexpensive way for people to experience VR, but with limited hardware options, current mobile VR systems provide only a few input methods. For example, with the Google Cardboard or the first generation of Samsung Gear VR, the user can only click a button or touch a navigation pad on the side of the VR headset. This simple input is often combined with head pointing to select virtual objects or menus items in the VR application, but there is no way to support direct 6 degree of freedom (DoF) manipulation.

Addressing this problem, the later generation of the Gear VR and the Google Daydream mobile VR headsets have a new 3DoF handheld controller which provides orientation tracking and touchpad input. Using this controller a user can select a virtual object by pointing at it, and then move it by waving their wrist around. However, this type of 3DoF controller cannot provide full 6DoF input and is not intuitive for many VR tasks. For example, to move a selected virtual object away, the user has to lift up the front end of the controller instead of moving it forward. Thus, there is a need for new types of input devices and methods for mobile VR experiences.

We are interested in using different combination of input methods to provide support for bimanual input in mobile VR. This because with current mobile VR devices, only one hand is occupied with input while the other hand is not used. Research in desktop VR has shown that two handed input methods can significantly improve the usability of VR experiences [MM95]. There has also been significant research work in asymmetric cooperative bimanual interfaces [HPGK94] where users use different tools in each hand. However, it has been challenging to provide similar two handed input in a mobile VR setting. In this paper, we describe a novel asymmetric bimanual interface that combines input from a handheld controller in the dominant hand with natural gesture input from the non-dominant hand. The use of two different input methods allows us to support more natural and efficient user interaction.

The main contributions of our work are using a handheld controller for 6DoF input based on visual tracking, and developing a bimanual interface combining the handheld controller and the natural gesture. We have also conducted a preliminary pilot study of our method compared with three other asymmetric bimanual input methods.

2. Related Work

Our research builds on earlier work with wearable interfaces, mobile VR, two handed VR input and asymmetric input methods. In this section, we provide a summary of some related research from each of these fields. Two handed input has been explored for a long time in VR interfaces. For example, the Chordgloves interface allowed people to manipulate virtual objects using special gloves on each hand [MM95], emulating a pair of 3D mice and a keyboard for creating CAD models. The VLEGO interface used two handheld controllers to support modelling in VR, and researchers found

† huidong.bai@canterbury.ac.nz
‡ alaeddin.nassani@pg.canterbury.ac.nz
§ barrett.ens@outlook.com
¶ mark.billinghurst@unisa.edu.au

DOI: 10.2312/egve.20171343

that performance was 20% faster than a one-handed interface and produced less errors [KTK*98]. Similarly, Schultheis found that a two-handed interface was over four times faster for 3D virtual object manipulation than a standard mouse interface [SJT*12], and up to two and a half times as fast as a standard one-handed wand interface like that of the Google Daydream.

Many two-handed VR input systems have the same controllers for both hands and focus on symmetric input. For example, the iSith interface uses a virtual ray from each handheld controller to select and interact with objects in a VR environment [WBB06]. However, Hinckley et al. demonstrated that bimanual asymmetric input can also be valuable where there are two different controllers [HPGK94]. Poupyrev demonstrated pen input in VR where the user held a tablet in one hand and stylus in the other [PTW98], and others have shown similar VR interfaces with two different physical controllers [SEaS99]. In these asymmetric interfaces the dominant and non-dominant hands often have different input devices and roles.

Current mobile VR interfaces typically use single button input (Google Cardboard) or a handheld controller (Google Daydream). More recently, gesture input can also be supported through adding the Leap Motion hand tracking, or through using the phone camera [YKD*16] or acoustic sensing [ZTF16]. However, while there has been significant research on asymmetric two handed interfaces for VR, none of this has been implemented in a mobile VR setting. In the next section, we describe our prototype which combines natural handheld gesture and controller input in mobile VR. This is one of the first examples of a mobile VR interface that combines gesture input with a handheld controller.

3. Prototype System Design

We have developed a prototype system that supports asymmetric bimanual input in mobile VR, combining three main components, as shown in Figure 1: (1) A standard mobile VR HMD (Samsung Gear VR headset + Samsung Galaxy S7) with its default 3DoF controller; (2) Intel RealSense camera connected to the Intel Compute Stick, (3) Hand gesture tracking sensor (Leap Motion) connected to the Galaxy S7. For the pilot study comparison, a wireless touchpad (Logitech T650) was also included in our setup (4).

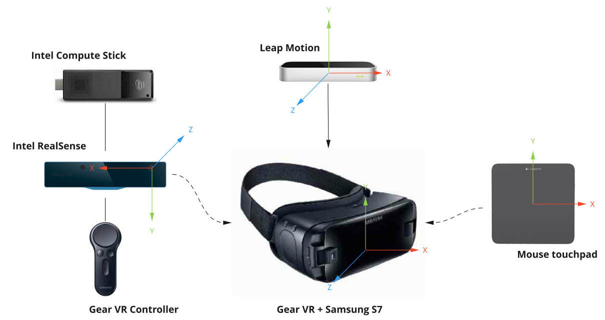

Figure 1: *System components.*

The Gear VR HMD with Galaxy S7 Android phone tracks the user's head orientation and displays the VR scene in stereoscopic camera view. It serves as the main unit of the system which other sensors connect to. We then extended the Gear VR controller's capability by adding an Intel RealSense (RS) camera, which is connected to an Intel Compute Stick (ICS) powered by a portable power bank. The RS and ICS provides SLAM-based tracking which calculates the RS's position and rotation relative to the static environment, and wirelessly transfers the real-time tracking result to the main unit. The SLAM tracking does not provide absolute position tracking, but position and orientation sensing relative to the start point. This combination provides 6DoF input in 3D space and enables a wider range of interaction than the standard mobile VR handheld controller.

Meanwhile, we support natural gesture input using the Leap Motion sensor which tracks real-time free-hand gestures in 3D space. The Leap Motion is able to track hand position and joint motion at an accuracy of less than 1mm up to 200 fps. The sensor is connected directly to the Galaxy S7 via a USB cable. The position and rotation of the user's hand joints are represented by a virtual hand in the VR scene. The 2D touchpad tracks the user's single finger touch input and sends the input data to the Gear VR phone via wireless USB receiver connected to the Galaxy S7 USB port. The touchpad is mounted on the user's pocket for easy access.

This combination of hardware provides 3DoF head tracking, along with 6DoF input from the handheld controller held in the user's dominant and natural gesture input from the non-dominant hand (see Figure 2). The system components were integrated together using a Unity3D application running on the main mobile unit. Based on this, we developed asymmetric bimanual input methods for mobile VR.

To demonstrate the system, we developed a simple asymmetric manipulation technique using raycasting and gesture input. The user can use the new vision-based controller to point at a target virtual object, highlight it with a raycast ray, and then pull the trigger of the controller to activate the selection. The object is selected while the trigger is held down. The top left corner of Figure 2 shows the RS tracking in a visual result. Once an object is selected, the user can move the controller to translate it, or make a pinch gesture with their hand to manipulate the object, rotating it in response to the user's real hand motions. The bottom of Figure 2 shows the screenshot of the example application with our bimanual input. Using our hardware platform, a range of different asymmetric methods could be implemented.

4. Preliminary Pilot Study

To evaluate the prototype system, we conduct a preliminary pilot study by comparing our method with three other interfaces. Seven participants (4 female) were recruited for the experiment, ranging in age from 27 years to 40 years with an average of 31.9 years. These users all had experience with mobile VR and its traditional input methods. The experiment used a simple VR scene consisting of three cubes (with different colored faces) scattered around the user at a certain distance away from them and a highlighted target square area (Figure 3). The task was to move all the cubes into the target area and rotate them so that they all had the same color

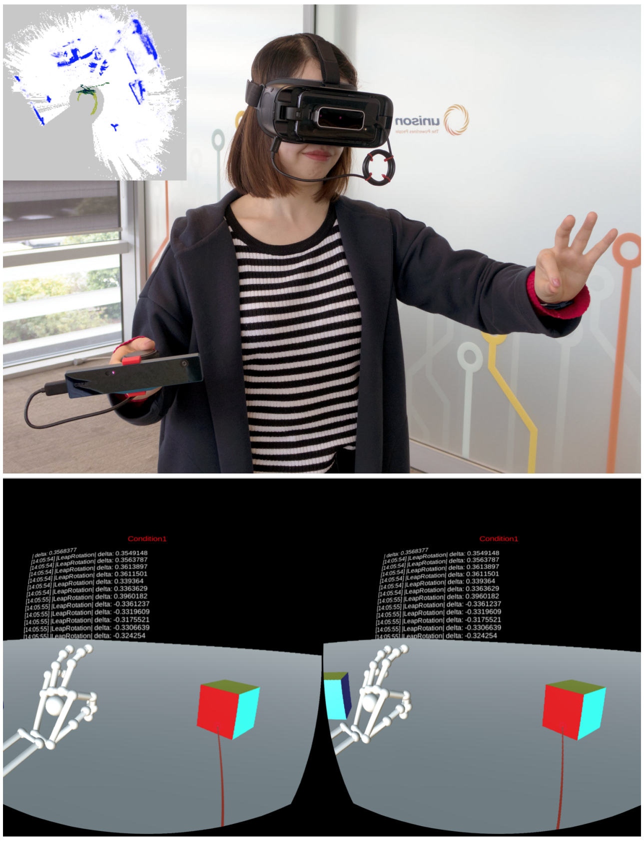

Figure 2: *Prototype system and screenshot.*

facing the user. Participants were told to perform the task as fast and as accurately as possible.

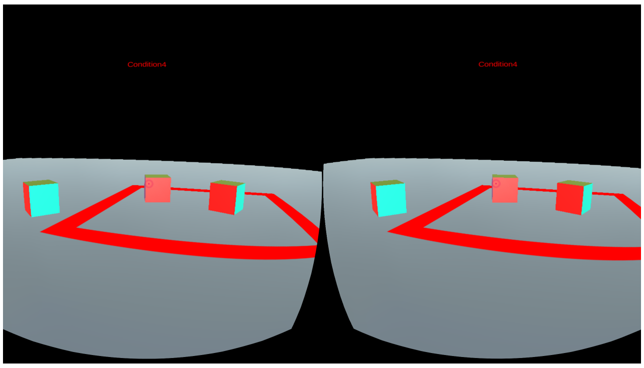

Figure 3: *Task example.*

The participants wore the Gear VR on their head and used two hands to perform the task. To simplify the task and focus on asymmetric bimanual interaction, the dominant hand was used for selecting and moving cubes in 3D virtual space, and the non-dominant hand was used to rotate the virtual cube around its y-axis. We wanted to compare our prototype input method (C1) with other techniques, so the experiment consisted of four conditions (Table 1), the order of which was randomized. Each of these conditions had a different combination of gesture and controller input, but they all used asymmetric bimanual input.

Table 1: *Pilot study conditions*

	Dominant Hand for 3D translation	**Non-Dominant Hand for one-axis rotation**
C1	Intel RealSense Controller	Leap Motion
C2	Intel RealSense Controller	Trackpad
C3	Leap Motion	Intel RealSense Controller
C4	Leap Motion	Trackpad

Figure 4 demonstrates how each interface was used by participants. Our input method C1 is described in Section 3. In condition C2, the user uses the handheld controller to select and move the target object, but instead of using a pitch gesture to rotate the object, the user can scroll on the touchpad with a single finger from the non-dominate hand to complete the rotation. For conditions C3 and C4, the user selects objects by using the head gaze and making a pitch gesture while looking at the object. Once the object is selected, it can be translated by the movement of the dominant hand, and rotated with the controller input (C3) or by scrolling on the touchpad (C4). Note that with all of these asymmetric bimanual interaction conditions the user can translate and rotate the object at the same time.

Figure 4: *Pilot study conditions.*

Users were asked to complete the sample task and then answered a number of questions in terms of how easy the interface was to use, if it was mentally or physically stressful, and if it was quick to learn. These were answered on a Likert-scale of 1 to 5, (1: strongly disagree, 5: strongly agree). Users were also asked to rank the four conditions in terms of how easy it would be to complete similar tasks using them, and we also interviewed them to capture their impression.

We used a Friedman test to analyze the Likert-scale results, but due to the small number of subjects there was no significant difference between conditions. However, from the ranking results, our bimanual input method (C1) was preferred (3 votes) with C4 getting 2 votes, while C2 and C3 only got 1 vote.

In the subject questioning afterwards, and by observing user behavior it was clear that the users felt that each bimanual interface that we evaluated had its own advantages and disadvantages, in the rest of this section we discuss this.

In terms of 3D translation, both the handheld controller and using hand gesture can provide direct 3D input in mid-air. The users mentioned that using raycasting with the controller "made the task easier to do", especially when the target object was located far from the user. The raycasting method is considered straightforward for selecting items with less physical movement required for translation. Once objects were selected it was easy to rotate them with the natural hand gesture.

In contrast, using the pitch gesture with the head gaze for selection was more difficult. The user has to move their head around to locate the target and then precisely target it with their head when it is in view. The changing VR view can distract the user between these two interfaces, and may confuse users on how to coordinate gaze and gesture input.

Using the controller and the touchpad provides accurate and efficient task input because of the device-centric interface, but participants claimed that "it is not as natural as the way we used in the daily life". As we expected, the 2D touchpad is not good for controlling 3D virtual content but it could be suitable for choosing the rotation direction during tasks and rotating about one axis.

Users also identified some challenges with our system. The vision-based controller and natural gesture input methods are performed in the mid-air without any physical anchor, which gives a much larger interactive volume compared with the small touchpad. However, this can also produce fatigue sooner. We also observed how the gesture input can also be easily moved out of sensor working range and lose tracking, badly interrupting the user experience.

Some participants expected the system to show the controller in VR environment in the correct position in the real world. However, in order to do this we would need to track the controller relative to the mobile VR headset which currently has not been implements. Some users also suggested that "it would be good to combine all conditions into one adaptive interface, using gesture to select the object, the controller to move it and the touchpad for rotation." This could be something we can explore in the future.

5. Conclusion

This research present an asymmetric bimanual interface using a custom 6DoF controller and natural hand gesture in a mobile VR setting. Asymmetric bimanual interfaces are common in desktop VR, but this is one of the first examples for mobile VR. We compared our interface with three other bimanual interaction methods in a preliminary pilot study. The number of users was too small to get significant results, but our technique was ranked first among the four methods. From observation and interview questions we found that this was because using our position aware handheld controller with gesture input provided an easy and natural experience. Users could easily select objects with the controller and at the same time rotate them with their hand gestures.

This is just a work in progress so there are many ways this research could be extended in the future. For example, we would like to conduct a formal user study comparing our asymmetric bimanual interface to more traditional mobile VR input with just one device. The prototype interaction methods were relatively simple, so we would like to explore more complicated 3D user interface techniques. Finally, we could also further explore other combinations of different interaction modalities, such as adding speech input.

References

[HPGK94] HINCKLEY K., PAUSCH R., GOBLE J. C., KASSELL N. F.: Passive real-world interface props for neurosurgical visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, April 1994), CHI '94, ACM, pp. 452–458. doi:10.1145/191666.191821. 1, 2

[KTK*98] KIYOKAWA K., TAKEMURA H., KATAYAMA Y., IWASA H., YOKOYA N.: Vlego: A simple two-handed 3d modeler in a virtual environment. *Electronics and Communications in Japan (Part III: Fundamental Electronic Science) 81*, 11 (1998), 18–28. doi:10.1002/(SICI)1520-6440(199811)81:11<18::AID-ECJC3>3.0.CO;2-N. 2

[MM95] MAPES D. P., MOSHELL J. M.: A two-handed interface for object manipulation in virtual environments. *Presence: Teleoperators and Virtual Environments 4*, 4 (Jan. 1995), 403–416. doi:10.1162/pres.1995.4.4.403. 1

[PTW98] POUPYREV I., TOMOKAZU N., WEGHORST S.: Virtual notepad: handwriting in immersive vr. In *Proceedings of IEEE Virtual Reality Annual International Symposium* (Washington, DC, USA, Month 1998), VR '98, IEEE Computer Society, pp. 126–132. doi:10.1109/VRAIS.1998.658467. 2

[SEaS99] SCHMALSTIEG D., ENCARNAÇÃO L. M., SZALAVÁRI Z.: Using transparent props for interaction with the virtual table. In *Proceedings of the Symposium on Interactive 3D Graphics* (New York, NY, USA, April 1999), I3D '99, ACM, pp. 147–153. doi:10.1145/300523.300542. 2

[SJT*12] SCHULTHEIS U., JERALD J., TOLEDO F., YOGANANDAN A., MLYNIEC P.: Comparison of a two-handed interface to a wand interface and a mouse interface for fundamental 3d tasks. In *Proceedings of IEEE Symposium on 3D User Interfaces* (Washington, DC, USA, March 2012), 3DUI '12, IEEE Computer Society, pp. 117–124. doi:10.1109/3DUI.2012.6184195. 2

[WBB06] WYSS H. P., BLACH R., BUES M.: isith - intersection-based spatial interaction for two hands. In *Proceedings of IEEE Symposium on 3D User Interfaces* (Washington, DC, USA, March 2006), 3DUI '06, IEEE Computer Society, pp. 59–61. doi:10.1109/VR.2006.93. 2

[YKD*16] YOUSEFI S., KIDANE M., DELGADO Y., CHANA J., RESKI N.: 3d gesture-based interaction for immersive experience in mobile vr. In *Proceedings of the 23rd International Conference on Pattern Recognition* (Dec 2016), ICPR '16, pp. 2121–2126. doi:10.1109/ICPR.2016.7899949. 2

[ZTF16] ZAYER M. A., TREGILLUS S., FOLMER E.: Pawdio: Hand input for mobile vr using acoustic sensing. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (New York, NY, USA, Oct 2016), CHI PLAY '16, ACM, pp. 154–158. doi:10.1145/2967934.2968079. 2