Visualizing the Keyboard in Virtual Reality for Enhancing Immersive Experience

Jia-Wei Lin
National Taiwan University

Ping-Hsuan Han
National Taiwan University

Jiun-Yu Lee
National Taiwan University

Yang-Sheng Chen
National Taiwan University

Ting-Wei Chang
National Taiwan University

Kuan-Wen Chen
National Chiao Tung University

Yi-Ping Hung
National Taiwan University

Figure 1: Typing Interfaces: (a) Reality, (b) VR No Keyboard, (c) VR Keyboard Model, (d) MR Real Hands, (e) MR Full Blending.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**;

KEYWORDS

Virtual Reality, Mixed Reality, Keyboard

ACM Reference format:
Jia-Wei Lin, Ping-Hsuan Han, Jiun-Yu Lee, Yang-Sheng Chen, Ting-Wei Chang, Kuan-Wen Chen, and Yi-Ping Hung. 2017. Visualizing the Keyboard in Virtual Reality for Enhancing Immersive Experience. In *Proceedings of SIGGRAPH '17 Posters, Los Angeles, CA, USA, July 30 - August 03, 2017,* 2 pages.
https://doi.org/10.1145/3102163.3102187

1 INTRODUCTION AND MOTIVATION

Recently, virtual reality (VR) becomes more and more popular and provides users an immersive experience with a head-mounted display (HMD). However, in some applications, users have to interact with physical objects while immersed in VR. With a non-see-through HMD, it is difficult to perceive visual information from the real world. Users must recall the spatial layout of the real surroundings and grope around to find the physical objects. After locating the objects, it is still inconvenient to use them without any visual feedback, which would detract the immersive experience.

Among all physical objects we may interact with in VR, the keyboard is one of the primary input modalities for computers. Although there are many text entry methods for VR, such as voice input, wide adoption of novel modalities may take a long time [McGill et al. 2015]. The Keyboard is a fast, accurate and ubiquitous input device, and it could be used without privacy concerns. Users are familiar with the standard PC keyboard and proficient in its usage. However, it is difficult to find and use a keyboard in VR without any visual information. To solve this problem, we use video see-through HMD to build a mixed reality (MR) environment, which integrates the images of the real world into the virtual world, as shown in Figure 1. In this paper, we visualized the keyboard and users' hands in VR in different ways and conducted a user study to evaluate their helpfulness and users' preference.

SIGGRAPH '17 Posters, July 30 - August 03, 2017, Los Angeles, CA, USA

ACM ISBN 978-1-4503-5015-0/17/07.
https://doi.org/10.1145/3102163.3102187

2 APPROACH

We use Unity 5.4 to build the virtual environment. For the video see-through HMD, we choose Intel RealSense F200 RGB-D Camera and attach it to HTC VIVE HMD. To bring the keyboard into the virtual world, we use Vicon Motion Capture System to track the position of an Apple keyboard with numeric keypad.

We design three interfaces that visualize the keyboard in VR: First, VR Keyboard Model (Figure 1(c)) is an interface showing a 3D model of the keyboard. The model provides visual feedback that the keys would turn gray while being pressed. Also, its position is corresponding to the position of the keyboard in the real world. Hence we have a "cloned keyboard" in the virtual world. Second, MR Real Hands (Figure 1(d)) not only shows an aforementioned 3D model of the keyboard but also draws the user's hands. We render a point cloud mesh of user's hands by using the depth images acquired by Intel RealSense F200 and segmenting the hands with YCbCr skin detection algorithm [Basilio et al. 2011]. However, Lai et al. [Lai et al. 2017] have shown that users may lose the sense of distance when using video see-through HMD without any process on the view due to the distance between the camera and user's eyes. Thus, in our implementation, we adjust the position of the mesh manually to make the point cloud mesh of the hands well-aligned with the keyboard. Third, MR Full Blending (Figure 1(e)) renders

Table 1: The Statistics of Typing Performance Green denotes p < 0.05).

Typing Metric	(1) Reality Baseline	(2) VR No Keyboard	(3) VR Keyboard Model	(4) MR Real Hands	(5) MR Full Blending	RM-ANOVA	Tukey Post-Hoc
WPM (words per minute)	39.8 (13.3)	28.1 (18.9)	24.3 (16.2)	27.4 (16.8)	27.8 (17.6)	F=1.91, p=0.118	N/A
Total Error Rate	0.07 (0.04)	0.28 (0.16)	0.2 (0.1)	0.2 (0.1)	0.2 (0.11)	F=6.93, p<0.001	1-2, 1-3, 1-4, 1-5 2-3, 2-4, 2-5
Corrected Error Rate	0.07 (0.04)	0.28 (0.16)	0.2 (0.1)	0.19 (0.1)	0.2 (0.11)	F=6.81, p<0.001	1-2, 1-3, 1-4, 1-5 2-3, 2-4, 2-5
Not Corrected Error Rate	0.001 (0.001)	0.003 (0.01)	0.001 (0.001)	0.006 (0.013)	0.001 (0.001)	F=1.33, p=0.265	N/A
Duration To First Key	1.5 (0.3)	4.3 (1.9)	3.0 (1.3)	3.2 (1.6)	3.6 (1.1)	F=9.10, p<0.001	1-2, 1-3, 1-4, 1-5 2-3, 2-4
First Key Accuracy	0.95 (0.06)	0.57(0.24)	0.66 (0.23)	0.65 (0.23)	0.75 (0.21)	F=7.64, p<0.001	1-2, 1-3, 1-4, 1-5 2-3, 2-4

the view of reality at the position of the keyboard. Users could see their hands and the keyboard when they face toward there.

3 USER STUDY

We conducted a text entry study with the following interfaces (as shown in Figure 1): (1) Reality: Baseline typing performance in reality; (2) VR No Keyboard: Wearing an HMD and no keyboard view; (3) VR Keyboard Model: Providing a 3D model of the keyboard; (4) MR Real Hands: Providing a 3D model of the keyboard and the mesh of user's hands; (5) MR Full Blending: Blending the full view of the keyboard and user's hands.

A total of 16 participants (10 males and 6 females) were recruited into this study. Their ages range from 20 to 24 years old (mean = 22.1, SD = 1.02). For the baseline performance, participants were asked to finish the typing task without HMD. They have an intermediate typing speed. The other four modes were in random order. Each task was to enter fifteen phrases randomly chosen from MacKenzie 500 phrase set [MacKenzie and Soukoreff 2003]. For each phrase, when participants pressing the button of the controller right beside their legs, the system would display the phrase and start recording typing time. After finishing a phrase, participants were required to place their hands back on their legs, then we would arbitrarily move the keyboard. There was a two-minute break between each task. Meanwhile, participants were asked to answer a questionnaire.

4 RESULTS AND DISCUSSION

The result is shown in Table 1. The significant difference of Total Error Rate shows that both a virtual keyboard and an image of the real keyboard could reduce the error rate when users are typing in VR. Additionally, the improvement of Duration To First Key indicates that providing a virtual keyboard in VR could help users to find the physical keyboard. However, such improvement is relatively small in MR Full Blending mode. Because the FOV of the video see-through camera is not wide enough, users need to bow their heads to see the keyboard. Such action would cause some additional delay. Nevertheless, MR Full Blending mode is most helpful to First Key Accuracy. Although MR Real Hands mode also provides the information about the keyboard and the hands, the mesh of the hands is created from the depth image, and the fingertips are often occluded by the fingers. Therefore, it does not help users to press the first key correctly.

In the questionnaire result, participants mostly prefer VR Keyboard Model mode. In a further interview, they showed a disinclination for the mesh of the hands in MR Real Hands mode due to the same issue of FOV. When typing in reality, users can check the relative position of their hands and the keyboard with their peripheral vision. While in our VR system, users need to bow their heads when they want to acquire such visual information. Therefore, VR Keyboard Model mode and MR Real Hands mode have similar performance on the quantitative results. To understand the influence of hands information on the typing performance in VR, the system needs to be improved for the further study.

5 CONCLUSIONS AND FUTURE WORK

From the results and discussion, we have discovered that the FOV of the camera for video see-through will influence the performance and user preference. Reconstructing the hands by other methods or using another camera that has larger FOV might yield better results. We will try to improve our method through these approaches and conduct a further study.

ACKNOWLEDGEMENTS

This work was financially supported by the Ministry of Science and Technology of Taiwan under Grants MOST 105-2622-8-002-002, and sponsored by MediaTek Inc., Hsin-chu, Taiwan.

REFERENCES

Jorge Alberto Marcial Basilio, Gualberto Aguilar Torres, Gabriel Sánchez Pérez, L. Karina Toscano Medina, and Héctor M. Pérez Meana. 2011. Explicit Image Detection Using YCbCr Space Color Model As Skin Detection. In *Proceedings of the 2011 American Conference on Applied Mathematics and the 5th WSEAS International Conference on Computer Engineering and Applications (AMERICAN-MATH'11/CEA'11)*. World Scientific and Engineering Academy and Society (WSEAS), Stevens Point, Wisconsin, USA, 123–128. http://dl.acm.org/citation.cfm?id=1959666.1959689

Chun-Jui Lai, Ping-Hsuan Han, Han-Lei Wang, and Yi-Ping Hung. 2017. *Exploring Manipulation Behavior on Video See-Through Head-Mounted Display with View Interpolation.* Springer International Publishing, Cham, 258–270. https://doi.org/10.1007/978-3-319-54526-4_20

I. Scott MacKenzie and R. William Soukoreff. 2003. Phrase Sets for Evaluating Text Entry Techniques. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems (CHI EA '03)*. ACM, New York, NY, USA, 754–755. https://doi.org/10.1145/765891.765971

Mark McGill, Daniel Boland, Roderick Murray-Smith, and Stephen Brewster. 2015. A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2143–2152. https://doi.org/10.1145/2702123.2702382