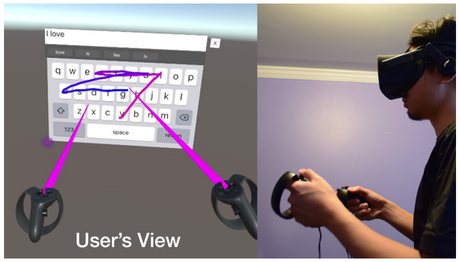

Figure 1: Illustration of our implemented word-gesture input method using 6-DOF VR Controllers. The left image shows user's view in the head-mounted display.

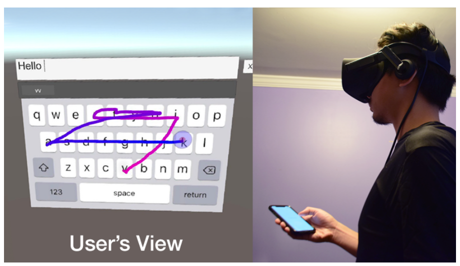

Figure 2: Illustration of our implemented word-gesture input method using pressure-sensitive touchscreens. The left image shows user's view in the head-mounted display.

Exploring Word-gesture Text Entry Techniques in Virtual Reality

Sibo Chen
Junce Wang
University of California, Berkeley
Berkeley, CA, USA
The Chinese University of Hong Kong, Shenzhen
Shenzhen, China
sibo_chen@berkeley.edu
junce@berkeley.edu

Santiago Guerra
Neha Mittal
Soravis Prakkamakul
University of California, Berkeley
Berkeley, CA, USA
s.guerra@berkeley.edu
neha01mittal@berkeley.edu
soravis@berkeley.edu

ABSTRACT

Efficient text entry is essential to any computing system. However, text entry methods in virtual reality (VR) currently lack the predictive aid and physical feedback that allows users to type efficiently. The state of the art methods such as using physical keyboards with tracked hand avatars require a complex setup which might not be accessible to the majority of VR users. In this paper, we propose two novel ways to enter text in VR: 1) Word-gesture typing using six degrees of freedom (6DOF) VR controllers; and 2) word-gesture typing using pressure-sensitive touchscreen devices. Our early stage pilot experiment shows that users were able to type at 16.4 WPM and 9.6 WPM on the two techniques respectively without any training, while an expert's typing speeds reached up to 34.2 WPM and 22.4 WPM. Users subjectively preferred the VR controller method over the touchscreen one in terms of usability and task load. We conclude that both techniques are practical and deserve further study.

CHI'19 Extended Abstracts, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5971-9/19/05.
https://doi.org/10.1145/3290607.3312762

KEYWORDS

Virtual Reality; Head-mounted displays; Text entry; Word-gesture keyboard; Touchscreen

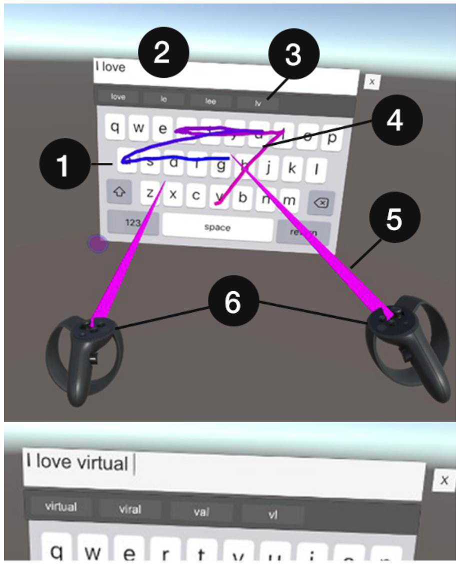

Figure 3: (Top) The user interface presented in the HMD for *Controller WGK* which consists of (1) a virtual keyboard with QWERTY layout (2) an output text box (3) the list of the predicted words for the most recent completed stroke (4) the stroke path (5) the rays cast from the controllers (6) visualizations of the user's controllers. The stroke path currently shows an attempt to enter the word 'virtual' by passing through the characters 'v', 'i', 'r', 't', 'u', and 'a'.

(Bottom) A close-up view of the output text and the suggested words after the path 'v-i-r-t-u-a-l' was completed and the trigger was released.

1. INTRODUCTION

Text entry in VR environments is becoming increasingly important as a wider variety of applications, such as messaging, are appearing on VR platforms. However, VR text entry still faces numerous challenges. The head-mounted displays which are essential for creating an immersive experience block users from viewing their physical environment. Therefore, users can no longer see their hands or traditional input devices such as keyboard and mouse. Common VR input devices, such as motion-tracked hand-held controllers, on the other hand, do not provide the same level of tactile and haptic feedback as keyboards. Users feel less confident when they type into thin air.

Text entry with word-gesture keyboards (WGK) has been shown to be useful in mid-air situations [7]. It was also introduced to VR HMDs by using a head pointing method [10]. However, there are no studies that explore different ways to apply word-gesture text entry in VR. Word-gesture with hand-held controllers could have advantages over head pointing because of its higher accuracy and lower physical demand [8]. Text entry with the help of external touchscreen devices , as explored by Kim et al. [4], could also be promising as touchscreens provide a feedback surface for the fingers and users are familiar with word-gesture typing on smartphones.

We propose two techniques for text entry in VR based on word-gesture text entry method using commodity hardware such as VR controllers and pressure-sensitive touchscreen devices. Our implementation includes subtle improvements such as custom-built WGK predictions, auto-punctuation, ability to add custom words to the existing dictionary and visual highlighting that improve the user experience and typing speed. We have conducted a pilot experiment to compare these techniques. The result shows that word-gesture text entry with controllers outperforms that with touchscreens in terms of typing speed, accuracy and usability. Our current investigation has two important limitations: First, our performance metrics are collected in a first use study with participants for whom English is a secondary language. The performance achievable by trained native typists is likely higher. Second, we do not yet directly compare performance of our techniques to other baseline techniques. We report comparison data from prior studies, and leave a more complete experiment to future work.

2. RELATED WORK

2.1 Text Entry in Virtual Reality

The most prevalent approach for text entry in VR is the *Point and Select* method where users use a spatial controller to aim and select single character keys on a virtual keyboard. Speicher et al. [8] found the typing speed of *Controller Point and Select* to be at around 15.44 WPM. Knierim et al [5] explored the use of physical keyboards in virtual reality with (39.8WPM) and without hand avatars (31.8WPM). However, showing the avatar requires advanced hand motion tracking and the full geometry of the

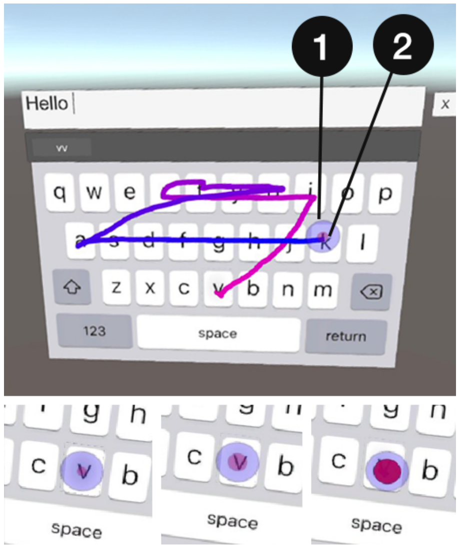

Figure 4: (Top) The user interface presented in the HMD for the *Touchscreen WGK* which consists of the virtual keyboard, the output text box and the list of suggested words just like in that of *Controller WGK*. The position of the blue circle (1) visualizes the user's touch position on the touchscreen while size of the red circle (2) visualizes the amount of pressure applied.

The images on the bottom row visualizes the different states of the red circle when a finger touches the screen without any pressure applied (Bottom-left), when a pressure smaller than the threshold was applied (Bottom-middle), when a pressure >= the threshold was applied (Bottom-right).

keyboard which might not be available to all VR users. The word-gesture technique has also been applied to VR by Yu et al. [10] which allows users to type at 24.7WPM after 60 minutes of training.

2.2 Word-gesture Text Entry

Word-gesture typing or Word-gesture Keyboard (WGK) refers to a text entry method that lets users draw the shape of a word on the input surface and was first explored in, SHARK [11]. The method was shown to perform well and has been widely adopted in mobile user interfaces, for example, in smartphone applications such as Swype. The use of word-gesture typing in mid-air interfaces was also explored by Vulture [7] (typing at 20.6 WPM). Hand-controlled word-gesture text entry has been shown to increase typing speed significantly while taking nearly no effort to learn. We hypothesize that the same technique could be effectively applied to VR.

2.3 Touchscreens as Text Input Device for VR

Smartphones have become a prevalent commodity hardware owned by many and are often paired with VR headsets such as the Oculus Go. We propose that touchscreen devices are a good platform for leveraging the word-gesture technique in VR as they have been the most widely used devices for word-gesture typing and that their surfaces could provide haptic feedback that makes users feel more confident while typing. The use of smartphones as the text input device for VR has also been explored in R. Kim and J. Kim's HoVR-Type [4]. Their proposed method using the hovering function available in some models outperformed the traditional Point and Select method and could achieve the typing speed of up to 9.2 WPM. Yeo et al. [9] also mentioned that tilt-based gesture keyboard on smartphones (typing at 32 WPM max.) could be a good candidate for VR text entry.

3. PROPOSED METHODS

3.1 Word-gesture Text Entry with 6-DOF Controllers (Controller WGK). – In this method illustrated in Fig. 1 and 3, the user uses hand-held 6-DOF VR controllers to aim a ray cast from the controller onto the virtual keyboard plane. To enter a word, the user draws the shape of the word on the plane while holding the trigger button located at the index finger. Once the trigger is released, a word is predicted and inserted to the input field. The user can also select a word from the suggested words list (shown as label 3 in Fig. 3) if the predicted word is not the right one. As a fallback, the user can push the trigger button while aiming at a character key to enter single characters.

3.2 Word-gesture Text Entry with a Pressure-sensitive Touchscreen Device (Touchscreen WGK). – As illustrated in Fig. 2 and 4, the user holds a touchscreen device (e.g. a phone) that acts as a proxy for the virtual keyboard plane. The virtual cursor provides two kinds of visual feedback to address the lack of visibility of the touchscreen device once the user is wearing a headset. In Fig. 4, the blue cursor

Figure 5: The experimental setup

The setup consisted of a Windows laptop connected to an Oculus Rift. The participants hold the default Oculus Rift's 6-DOF controllers in the Controller WGK technique. They hold an Apple iPhone X in the Touchscreen WGK one. The typing interface is shown in the HMD; laptop graphics are shown here to clarify what the user sees inside the HMD.

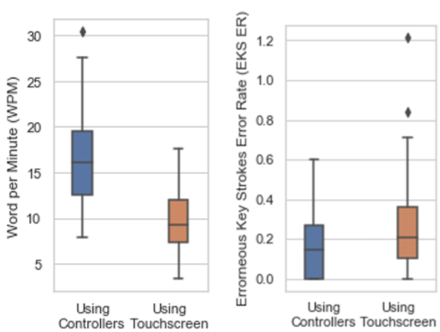

Figure 6: The typing speed in WPM (Left) and the error rate in Erroneous Keystrokes Error Rate (EKS ER) (Right) for each method.

on the keyboard visualizes the position of the finger on the touchscreen. The red dot in the center visualizes the normalized pressure applied to the touch. The user could either type a single character by tapping on the touchscreen or gesture-type a word by (1) first applying pressure that exceeds a predefined threshold of 50% of the maximum pressure to activate gesture-typing, (2) followed by drawing the path while holding the touch (as done with usual swipe typing), (3) finally lifting off the finger to end the path. The pressure change is visually conveyed to the user continuously through a concentric red circle within a blue one. The appropriate threshold of 50% was manually selected to create a clear distinction between two pressure levels (no pressure vs maximum pressure).

4. IMPLEMENTATION

Our prototype system was implemented in Unity3D. To implement the Touchscreen WGK technique, we created an iOS app that captures touch events and sends them to Unity over local network using the Open Sound Control (OSC) protocol. Our implementation of WGK broadly follows that of SHARK [11]. In order to identify words typed through strokes (gestures), our algorithm captures the stroke drawn by the cursor on the virtual keyboard plane, finds the condensed path of character traversal, performs matching on a dictionary, and calculates a match score from the combination of characters matched, dwell time, angle, and word frequency. The system would then present words with the highest score to the user. While pressure is used to make the state transition between hover and entry modes in the touchscreen condition, dwell time is only used as a feature in the stroke interpretation algorithm.

5. EVALUATION

We conducted a within-subject lab experiment with one independent variable (Input Method) and multiple dependent variables related to text entry speed, error rate, usability, and task load. Each participant was to perform a text transcription task of 10 phrases as quickly and as accurately as possible using each technique in Latin-square order. The phrases were selected from McKenzie's standard phrases [6] and each participant encountered the same set of phrases while using the two techniques. The pilot study was conducted in Bangkok, Thailand and the subjects were recruited through a signup form distributed over social media. There were 10 participants in our pilot study (1 female, 1 left-handed, 7 new to VR HMDs) with the age range of 28-30 years (M = 28.93). Our participants had some experience in typing in English but were not native speakers. This limitation may have led to lower performance.

The procedure started with the calibration of interpupillary distance (IPD) followed by a 5-minute Oculus Rift tutorial. The user watched a video tutorial and was given a 1-minute warm-up time to type a sentence before beginning the task. After each task, they were asked to take off the headset to complete a survey. The user had the discretion to spend time on correcting the incorrectly typed words but there was no formal setup to force them to correct errors before moving on. We used the

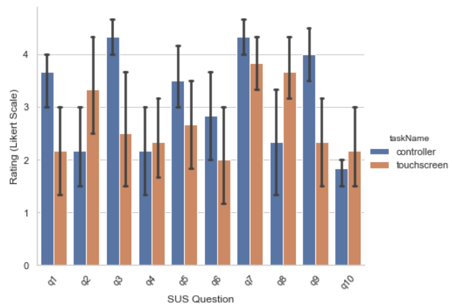

Figure 7: The System Usability Scale (SUS) Questionnaire Response

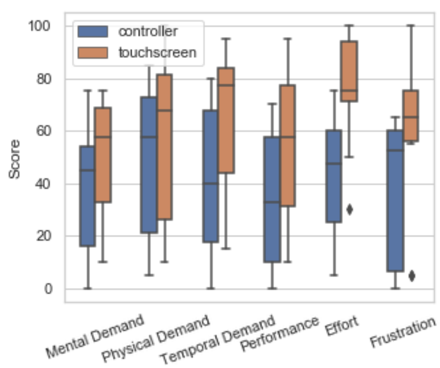

Figure 8: NASA Talk Load Index Rating

System Usability Scale (SUS) [2] and the NASA Task Load Index (TLX) [3] questionnaires to measure usability and task load. Each experimental session took 60 minutes on average.

6. RESULTS

6.1 Text Entry Speed. : Fig. 6 (left) shows the comparison of Word-per-Minute (WPM) typing speed for each technique. A dependent t-test suggests that the performance of Controller WGK (M=16.43WPM, SD=4.69) was significantly higher than that of Touchscreen WGK (M=9.62WPM, SD=3.04) (t=10.23, p=2.51×10^{-15}). These metrics are likely low because of the first-use scenario. In informal testing, an expert user with extensive previous training who was not among the ten subjects could achieve, with the same experimental conditions, 35.44WPM and 22.41WPM on Controller WGK and Touchscreen WGK technique respectively. On average, our controller-based solution shows only slight improvement in typing speed from the *Controller Pointing* technique [5] while our touchscreen version performed as well as the former work, HoVR-Type [4].

6.2 Error Rate. : To measure error rate, a metric derived from Erroneous Keystroke Error Rate (EKS ER) [1] was used, but, instead of checking erroneous key directly, we use the number of times the backspace key was hit as the indicator of the number of characters that had to be backtracked. Fig. 6 (right) shows the comparison of EKS ER for each technique. A dependent t-test shows that the error rate of Controller WGK (EKS ER=0.1564) is lower than that of Touchscreen WGK (EKS ER=0.2390) (t=-2.9088, p=0.0049).

6.3 Subjective Usability and Task Load. : Fig. 7 shows that the users rated the two techniques the same on the SUS when asked about the technical support (q4) and the amount of learning (q10) required to use the system. However, Controller WGK has a higher rating when it comes to ease of use (q3), being well-integrated (q7), and user's confidence (q9). Fig. 8 shows the task load index rating for the two methods. The two methods are rated similarly for Physical Demand, but Touchscreen WGK was rated higher for other demands (Mental, Temporal, Performance, Effort, and Frustration). Quantitative feedback shows that all participants prefer Controller WGK for precision and intuitiveness. However, some pointed out that if they were to use the system without gesture typing, they might have preferred touchscreen for stability and for its similarity with typing techniques used for texting on smartphones.

7. CONCLUSION AND FUTURE WORK

In this paper, we proposed two input techniques that make use of word-gesture text entry. Although the controller method outperforms the touchscreen one in terms of speed, accuracy, and usability, we argue that both have potential for further study. For example, the touchscreen technique could be useful on headsets that do not work with spatial controllers. It would be interesting to conduct a broader experiment comparing the proposed techniques with existing gesture-based technique applied to VR such as Head-based [10], and other traditional methods such as Controller Point and Select. This study also did not explore learning effects for each technique. We found that one ten-phrase session is not enough to determine whether users get better over time. The performance gap between expert and non-expert users also suggests a possible learning effect. Future experiments should address situations where users have ample time to practice the technique.

ACKNOWLEDGMENTS

Our special thanks go to Bala Throravi Kumaravel and our advisor, Björn Hartmann.

REFERENCES

[1] A. S. Arif and W. Stuerzlinger. 2009. Analysis of text entry performance metrics. In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*. 100–105. https://doi.org/10.1109/TIC-STH.2009.5444533

[2] John Brooke. 1995. SUS: A quick and dirty usability scale. *Usability Eval. Ind.* 189 (11 1995).

[3] Sandra G. Hart and L E. Stavenland. 1988. *Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research*. Vol. 52. 139–. https://doi.org/10.1016/S0166-4115(08)62386-9

[4] Youngwon R. Kim and Gerard J. Kim. 2016. HoVR-type: Smartphone As a Typing Interface in VR Using Hovering. In *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology (VRST '16)*. ACM, New York, NY, USA, 333–334. https://doi.org/10.1145/2993369.2996330

[5] Pascal Knierim, Valentin Schwind, Anna Maria Feit, Florian Nieuwenhuizen, and Niels Henze. 2018. Physical Keyboards in Virtual Reality: Analysis of Typing Performance and Effects of Avatar Hands. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 345, 9 pages. https://doi.org/10.1145/3173574.3173919

[6] 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

[7] Anders Markussen, Mikkel Rønne Jakobsen, and Kasper Hornbæk. 2014. Vulture: A Mid-air Word-gesture Keyboard. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1073–1082. https://doi.org/10.1145/2556288.2556964

[8] Marco Speicher, Anna Maria Feit, Pascal Ziegler, and Antonio Krüger. 2018. Selection-based Text Entry in Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 647, 13 pages. https://doi.org/10.1145/3173574.3174221

[9] Hui-Shyong Yeo, Xiao-Shen Phang, Steven J. Castellucci, Per Ola Kristensson, and Aaron Quigley. 2017. Investigating Tilt-based Gesture Keyboard Entry for Single-Handed Text Entry on Large Devices. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4194–4202. https://doi.org/10.1145/3025453.3025520

[10] Chun Yu, Yizheng Gu, Zhican Yang, Xin Yi, Hengliang Luo, and Yuanchun Shi. 2017. Tap, Dwell or Gesture?: Exploring Head-Based Text Entry Techniques for HMDs. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4479–4488. https://doi.org/10.1145/3025453.3025964

[11] Shumin Zhai and Per-Ola Kristensson. 2003. Shorthand Writing on Stylus Keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 97–104. https://doi.org/10.1145/642611.642630