# A Pinch-based Text Entry Method for Head-mounted Displays


Haiyan Jiang[*] [1] Dongdong Weng[✉] [†] [1] Xiaonuo Dongye [1] Yue Liu [1]

[1]Beijing Engineering Research Center of Mixed Reality and Advanced Display, Beijing Institute of Technology


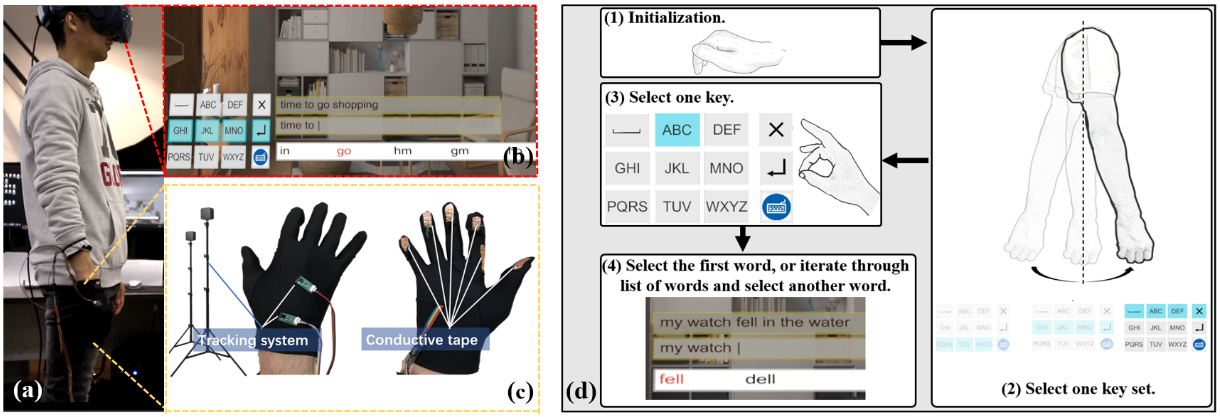


Figure 1: (a) A user is typing with the proposed text entry method. (b) The virtual environment in HMDs when typing. (c) The tracking system and glove used in the prototype. (d) Workflow of the proposed method.


## Abstract

Pinch gestures have been used for text entry in Head-mounted displays (HMDs), enabling a comfortable and eyes-free text entry. However, the number of pinch gestures is limited, making it difficult to input all characters. In addition, the common pinch-based methods with a QWERTY keyboard require accurate control of the hand position and angle, which could be affected by natural tremors and the Heisenberg effect. So, we propose a new text entry method for HMDs, which combines hand positions and pinch gestures with a condensed key-based keyboard, enabling one-handed text entry for HMDs. With this method, users move their hands with a naturally comfortable posture between three large different spaces in the air to choose one key set and then execute one of the pinch gestures to choose one character, where hand jitter does not affect the selection, helping to improve the input speed. The results of a primary study show that the mean input speed of the proposed method is 7.60 words-per-minute (WPM).



**Index Terms:** Human-centered computing—Human computer interaction—Interaction paradigms—Virtual reality; Human-centered computing—Human computer interaction—Interaction techniques—Text input; Human-centered computing—Human computer interaction— Interaction devices—Haptic devices


## 1 Introduction

Text entry is an important task for many applications in HMDs. There are several common text entry methods, including pointing with pressing a button (HTC VIVE[1]) on the controller or with a pinch gesture (Oculus Quest[2]) and clicking on the floating virtual keyboard (HoloLence2[3]) with the index finger for selecting one character. However, they are not efficient or comfortable enough. For example, those involving the whole upper extremity are prone to cause fatigue, such as "Gorilla arm" for long text input, and hand jitter would appear when pressing a button or executing a pinch gesture, causing obvious Heisenberg effect [4] which would lead to bad performance.

A promising text entry method for HMDs is to appropriate the thumb-to-finger-based methods [2, 3, 5]. Except for being always available, this type of method enables a comfortable and eyes-free text entry by leveraging proprioceptive feedback. As the number of pinch gestures is limited, however, it seems impossible to input 26 and more characters directly. So, we propose a new one-handed text entry method, a word-level method in mid-air. By combining hand positions with pinch gestures, this method enables the input of 26 and more characters. Users move their hands to choose one key set and execute one of the pinch gestures to choose a character. Since this method only requires users to move their hands between three large different spaces in the air, hand jitter does not affect the selection of key sets, and pinch gestures are accurate, enabling users to input accurately and quickly. In addition, the virtual keyboard of the proposed method and the hand position are decoupled, enabling users to input when the hands are naturally placed vertically, which helps to decrease fatigue and makes it possible to place the virtual keyboard anywhere and reduce its size.

## 2 A novel Text entry method

It is more comfortable for users to place their hands naturally than to raise their arms in common text entry methods. In addition, according to the arm movement habits that people are more accustomed to swinging their arms back and forth. In our proposed method, the user's hand is naturally placed vertically. The user swings the arm back and forth to change the hand position to select one key set, as shown in Figure 1(d).


*e-mail: jianghybit@163.com
†e-mail: crgj@bit.edu.cn


[1]https://www.vive.com/
[2]https://www.oculus.com/quest/
[3]https://www.microsoft.com/en-us/hololens/hardware

## 2.1 Input mechanism

The proposed method is based on word-level, and the workflow is shown in Figure 1(d).

### 2.1.1 Input space initialization

When the user touches the side area of the index finger (Side-key) over 2 seconds, the position of the hand at this moment is regarded as the initial position wherever the user's hand is placed. The user can initialize the hand position anytime if he/she wants to change the input space, even during the process of input.

### 2.1.2 Key selection

The position of the user's hand is used to decide the chosen key set. The different pinch gestures are used to choose one specific key in the chosen key set.

When the user's hand is in the initial position, the user can choose the middle key set, including the "GHI", "JKL", "MNO" and "Enter" keys. At this time, the thumb-to-index finger, thumb-to-middle finger, thumb-to-ring finger, and thumb-to-pinky pinch gestures are used to choose "GHI", "JKL", "MNO", and "Enter" keys respectively. When the user move his/her hand forward to the position ahead of the initial position, four different pinches are used to choose one of the keys in the upper key set including the "Space", "ABC", "DEF" and "Delete" keys; when the user move their hands backward to the position behind the initial position, four different pinches are used to choose one of the keys in the lower key set including the "PQRS", "TUPV", "WXYZ" and "Change" keys. The "Change" key is used to change the keyboard to other layouts, such as a numeric keypad or symbol keyboard.

### 2.1.3 Word selection

Since each key is associated with multiple characters, it is necessary to adopt word disambiguation for successful typing. We adopt the disambiguation method in [1]. When a key sequence is chosen, the disambiguation word list is displayed below the input box (Figure 1(b)), and at this time, the color of all words is black. After finishing a word sequence input, the user can touch Side-key to select the first candidate word as the preselected word. And users can iterate through the list to choose next or last word as the preselected word. The color of the preselected word turns red as visual feedback. If the user touches Side-key again, the preselected word is selected. In other words, the user can directly touch Side-key twice to enter the first candidate word most of the time. After one word is input, a "space" is added automatically. However, if the word contains only one character, a "space" is not added. Thus, out-of-vocabulary (OOV) words can be input. **Next word:** The next word can be chosen as the preselected word by the thumb-to-index finger pinch gesture. **Last word:** The last word can be chosen as the preselected word by the thumb-to-ring finger pinch gesture. **Delete:** If there is a preselected word, choosing the "Delete" key makes the user can continue to input the key sequence. If there is no preselected word but a key sequence, choosing "Delete" deletes the last key of the key sequence. If there is no key sequence, the last word entered is deleted when choosing "Delete".

## 3 Implementation

We implement our prototype using conductive tape made of copper which are placed on the five fingertips (Figure 1(c)) and on the side of the index finger. When the thumb touches any other fingertip or the side of the index finger, a pinch gesture signal would be activated. A tracking system is used to track the position of the user's hand (Figure 1(c)). The tracking system with 1mm accuracy works the same way as OptiTrack[4]. Others could use OptiTrack or other commercial products to track the position of the hand.

[4]https://www.optitrack.com/

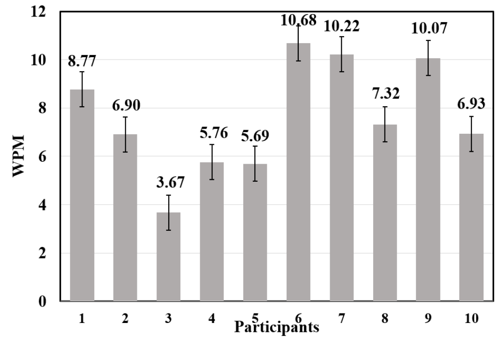


Figure 2: The input speed of 10 participants.

We conducted a primary study to evaluate the input speed of the proposed method. 10 participants (5 females) aged 21-32 (M=24.60) were recruited in this study. The virtual keyboard was placed 5 metres away in the center of the field of view, and the virtual keyboard occupied a field of view of 55°(horizontal)*32°(vertical). Figure 1(a)(b) shows a user is typing with the proposed method and the corresponding virtual environment in HMDs. The results show that the mean speed of the proposed method is 7.60 WPM and Figure 2 shows the input speed of each participant.

## 4 Conclusion and future work

We propose a new text entry method, which combines hand positions in the air and pinch gestures, enabling accurate and quick one-handed text entry for VR, AR, and MR. This method could be used when the hand is naturally placed vertically, benefiting comfortable long text input. In the future, we will conduct user studies to explore the performance and workload of the proposed method with more arm postures and more directions of hand movement and to verify the impact of the virtual keyboard with different sizes and positions.


## Acknowledgments

This work was supported by the Key-Area Research and Development Program of Guangdong Province(No.2019B010149001)and the National Natural Science Foundation of China (No.62072036, No.62007001) and the 111 Project (B18005).


## References


[1] H. Jiang and D. Weng. Hipad: Text entry for head-mounted displays using circular touchpad. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 692–703, 2020.

[2] M. Prätorius, D. Valkov, U. Burgbacher, and K. Hinrichs. Digitap: An eyes-free vr/ar symbolic input device. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, VRST '14, page 9–18, New York, NY, USA, 2014. Association for Computing Machinery.

[3] E. Whitmire, M. Jain, D. Jain, G. Nelson, R. Karkar, S. Patel, and M. Goel. Digitouch: Reconfigurable thumb-to-finger input and text entry on head-mounted displays. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(3), Sept. 2017.

[4] D. Wolf, J. Gugenheimer, M. Combosch, and E. Rukzio. Understanding the heisenberg effect of spatial interaction: A selection induced error for spatially tracked input devices. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–10, New York, NY, USA, 2020. Association for Computing Machinery.

[5] P. C. Wong, K. Zhu, and H. Fu. Fingert9: Leveraging thumb-to-finger interaction for one-handed text entry on smartwatches. In *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications*, SA '17, New York, NY, USA, 2017. Association for Computing Machinery.