2018

INVESTIGATING MIDAIR VIRTUAL KEYBOARD INPUT USING A HEAD MOUNTED DISPLAY

Jiban Krishna Adhikary
Michigan Technological University, jiban@mtu.edu

INVESTIGATING MIDAIR VIRTUAL KEYBOARD INPUT USING A HEAD MOUNTED DISPLAY

By

Jiban Krishna Adhikary

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Computer Science

MICHIGAN TECHNOLOGICAL UNIVERSITY

2018

This thesis has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Computer Science.

Department of Computer Science

Thesis Advisor: *Dr. Keith Vertanen*

Committee Member: *Dr. Scott Kuhl*

Committee Member: *Dr. Elizabeth Veinott*

Department Chair: *Dr. Zhenlin Wang*

Dedication

To my parents

who have always been there and provided support in every ups and downs in my life - without which I would neither be who I am nor would this work be what it is today.

Contents

List of Figures

List of Tables

Acknowledgments

At first I would like to express my gratitude towards my adviser Dr. Keith Vertanen for providing me the opportunity to work as his student in the field of text entry, for providing me funding opportunities in the last few semesters and for being a great mentor. I am lucky to have him as my adviser.

I want to express my gratitude towards Dr. Scott Kuhl for providing me the apparatus used for my study, for giving me the opportunity to use his lab and for all the valuable feedback and suggestions he provided.

I also want to thank Dr. Elizabeth Veinott and Dr. Philart Jeon for being in my committee and always keeping me on my toes. It is too bad that Dr. Jeon had to left Michigan Tech before my defense. I am grateful to both of them for the constructive feedback they provided on my thesis proposal.

Lastly, I want to express my gratitude to my family and friends who have always been the source of encouragement and inspiration towards my graduate studies.

List of Abbreviations

API	Application Programming Interface
VR	Virtual Reality
AR	Augmented Reality
HMD	Head Mounted Display
WPM	Word Per Minute
CER	Character Error Rate
SD	Standard Deviation

Abstract

Until recently text entry in virtual reality has been limited to using hand-held controllers. These techniques of text entry are feasible only for entering short texts like usernames and passwords. But recent improvements in virtual reality devices have paved the way to varied interactions in virtual environment and many of these tasks include annotation, text messaging, etc. These tasks require an effective way of text entry in virtual reality. We present an interactive midair text entry system in virtual reality which allows users to use their one or both hands as the means of entering text. Our system also allows users to enter text on a split keyboard using their two hands. We investigated user performance on these three conditions and found that users were slightly faster when they were using both hands. In this case, the mean entry rate was 16.4 words-per-minute (wpm). While using one hand, the entry rate was 16.1 wpm and using the split keyboard the entry rate was 14.7 wpm. The character error rates (CER) in these conditions were 0.74%, 0.79% and 1.41% respectively.

We also examined the extent to which a user can enter text without having any visual feedback of a keyboard i.e. on an invisible keyboard in the virtual environment. While some found it difficult, results were promising for a subset of 15 participants of the 22 participants. The subset had a mean entry rate of 10.0 wpm and a mean error rate of 2.98%.

Chapter 1

Introduction

Today people frequently use computers or mobile devices to interact with text, for example, reading text, writing new text, editing existing text, etc. At present these kind of interactions are also occurring in virtual reality (VR) head mounted displays (HMDs) or in augmented reality (AR) HMDs. Text entry in VR or AR is an active area of research and there is a lot of room for investigation and development. Existing research in this field focuses on using hand-held controllers, making gestures wearing hand gloves or controllers, or even incorporating a physical keyboard. These techniques suffer from various limitations. Using a hand-controller to enter text is tedious and time consuming. A gesture based technique may require users to memorize gestures and map the gestures to characters. And if we are using a physical keyboard, it is only feasible while seated. To overcome these limitations, we present an interactive system that leverages two hand midair input on a virtual keyboard. We used our system to investigate the performance and usability of text entry in virtual environment. Instead of using hand-controllers or gloves, we used the commercial Leap Motion hand tracker mounted on an HTC Vive. Besides providing easier text input, the goal of our investigation was to find out which mode of text entry, user interacting with one hand, both hands, or splitting the keyboard plane facilitates efficient text interaction in VR.

We also investigated if it was possible for a user to enter text using an invisible keyboard. The keyboard was invisible in the sense that a user could see the keyboard in the beginning and have an idea of the location of each key but was invisible while text was being entered.

This thesis is organized as follows: Chapter 2 discusses existing research in text entry in VR and AR, Chapter 3 describes the study design and various components of the experimental interface, Chapter 4 analyzes the results. Finally, Chapter 5 reflects the system as a whole, discusses its current limitations, and suggests future designs.

Chapter 2

Related Work

Recently there has been significant technological advancements in head-mounted display devices. Oculus Rift, HTC Vive and Samsung's Gear VR are examples of popular devices currently on the market. There are a lot of applications where text entry in VR environment can be beneficial. These applications include training [13] [8] [12], prototyping [4], rehabilitation [17], education [7], and data visualization [2]. Though text entry in physical devices such as, desktop computers and smartphones, is an active research field, relatively little work has been done related to text entry in VR. In this section, we review the existing literature on text entry in VR and AR.

First, we will discuss a few works related to midair text input which do not involve virtual environment. Next, we will discuss some works which investigated one hand versus two hand text input on mobile devices or touch-screen surfaces. Finally, we will discuss existing text entry techniques in virtual environment.

2.1 Midair Text Entry Outside Virtual Environments

2.1.1 Selection Based Techniques

Selection based techniques involve series of movements and activation of different user interface (UI) functions by selecting a key. Markussen et al. [14] proposed a selection based midair text entry on large displays using the OptiTrack[1] motion capture system. They analyzed the design space for midair text entry and proposed three different techniques: (i) H4 midair in which text is entered with the thumb using four buttons of a physical game controller, (ii) MultiTap in which there are 9 keys having multiple character mapped in a single key and a cursor that can be controlled by moving the hand and (iii) Projected QWERTY in which a QWERTY keyboard layout is projected on the display with a dot cursor and the selection of keys can be controlled by moving the hand. The mean entry rates for these techniques were 13.2 wpm in Projected QWERTY, 9.5 wpm in MultiTap and 5.2 wpm in H4 midair.

[1] http://optitrack.com/

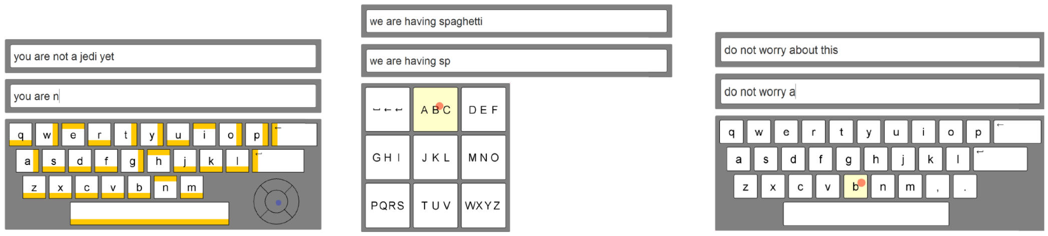

Figure 2.1: The user interface for H4 mid-air (left), MultiTap (middle) and Projected QWERTY (right). Images taken from Markussen et al. [14].

2.1.2 Gesture Based Techniques

Vulture [15] investigated if word gesture keyboards (e.g., SlideIT, Swype, ShapeWriter [25] [11] [26]) can be beneficial to midair text entry. A word gesture keyboard allows users to input word by drawing a pattern formed by the characters in that word instead of typing the characters. Vulture proposed a touch based word-gesture algorithm that works in midair by projecting users' movement in a display and using pinch gestures as word delimiters. In this system, the participants achieved an average entry rate of 20.6 wpm in a 10-session study.

Yi et al. [23] proposed a ten finger freehand typing mechanism in midair based on 3D hand tracking data. First they analyzed users' ten finger midair typing patterns. These included fingertip kinematics during tapping, correlated movement among fingers and 3D distribution of tapping endpoints. Based on the analysis, they proposed a probabilistic tap detection algorithm. They conducted the study with eight participants in four blocks. Experiment results showed that participants could type at

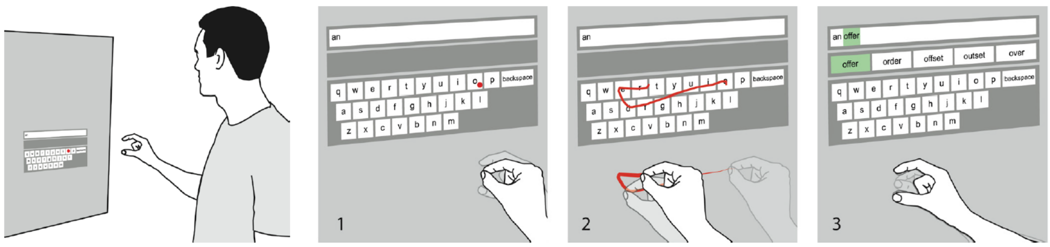

Figure 2.2: Text entry using word gestures in Vulture: by moving the hand, the user places the cursor over the first letter of the word and (1) makes a pinch gesture with thumb and index finger, (2) then traces the word in the air - the trace is shown on the screen. (3) Upon releasing the pinch, the five words that best match the gesture are proposed; the top match is pre-selected. Images taken from [15].

29.2 wpm with a low error rate of 0.4%.

2.2 Bimanual Text Entry

Effects of one hand and two hands text entry have been investigated in touch surfaces and game controllers. However, to the best of our knowledge the effects of one hand and both hands have not been investigated in VR.

Bi et al. [1] proposed a two hand gesture text entry system, based on multi-stroke gesture recognition algorithm on touch tablets using multiple fingers. Their study results showed that 42% participants (15 out of 36) preferred the two hand system in comparison to the one hand system. They also reported that the two hand system raised the comfort level and reduced the physical effort level.

Figure 2.3: The 2-Thumb Gesture keyboard. The user enters text by using both thumbs to perform drag gestures across the letters in a word on their respective half of the keyboard. Image taken from [19].

In a similar work Truong et al. [19] proposed a 2-Thumb Gesture (2TG) system and compared 2TG to Swype (1-finger gesture virtual keyboard) and confirmed that two hand approach reduced the fatigue level. Their 2TG keyboard achieved an entry rate of 24.4 wpm and an error rate of 0.65%.

Sandness et al. [16] investigated a two-finger QWERTY typing technique where text was entered using two hand-held game controllers or joysticks. The keyboard layout and joysticks were positioned such that one could easily assume that joystick positions were the place where index fingers rest while typing on a standard QWERTY keyboard. They achieved a mean entry rate of 6.75 wpm.

2.3 Text Entry in Virtual Environments

A significant amount of works have been done in investigating the interactions with buttons and menus in the virtual environment, but there have been limited works with keyboards in virtual environments. Recent work in this field includes using physical keyboard, touch screen keyboards, virtual keyboard with HMDs, etc.

Grubert et al. [10] studied the performance and user experience of desktop keyboards and touchscreen keyboards for use in VR. They found that novice users were able to retain about 60% of their typing speed on a desktop keyboard and 40-50% of their typing speed on a touchscreen keyboard. In two of their experimental conditions with a desktop keyboard, the mean entry rate were 26.3 wpm and 25.5 wpm, mean error rate were 2.1% and 2.4%. On the other hand, in the two conditions with touchscreen keyboard, the entry rate averaged 11.6 wpm and 8.8 wpm and mean error rate were 2.7% and 3.6%. The advantage in their study was that participants benefited from the tactile and touch feedback from the desktop and touchscreen keyboards.

In another study Grubert et al. [9] investigated several methods for virtually representing a user's hand in VR. They used four hand representations in VR: no hand representation, inverse kinematic hand model, fingertip visualization and augmented virtually (video inlay) representation. For all of the cases, the participants used a physical keyboard to enter text. The mean entry rate in these conditions were 36.1 wpm, 34.4 wpm, 36.4 wpm and 38.7 wpm respectively. They found that fingertip visualization (6.3%) and video inlay (5.1%) resulted in lower error rates in comparison to no hand (15.2%) and inverse kinematic hand (11.5%) representation.

Bowman et al. [3] investigated task performance and usability characteristics in VR with four techniques: pinch key board, a chorded keyboard, a virtual hand-held controlled keyboard and speech. Speech technique was the fastest (13.2 wpm) but none of the techniques showed high levels of performance or satisfaction.

Yu et al. [24] investigated three head based text entry techniques for HMDs: TapType where a user selected a letter by pointing to it using the HMD and tapping a button in a game controller, DwellType where a user selected a letter by pointing to it and dwelling over it for a period of time and GestureType where a user performed a word-level input using a gesture type style. The mean entry rate under these conditions were 10.6 wpm, 15.6 wpm and 19.0 wpm respectively. However, they further achieved a higher entry rate of 24.73 wpm in GestureType by improving the gesture-word recognition algorithm. To do so they used the head movement pattern obtained in the first study and allowed participants a 60 minutes training.

Walker et al. [22] developed a system to assist users type on a physical keyboard while wearing a head mounted device. Their study had two independent variables, whether the virtual keyboard was shown, and whether visual feedback was via an HMD or via

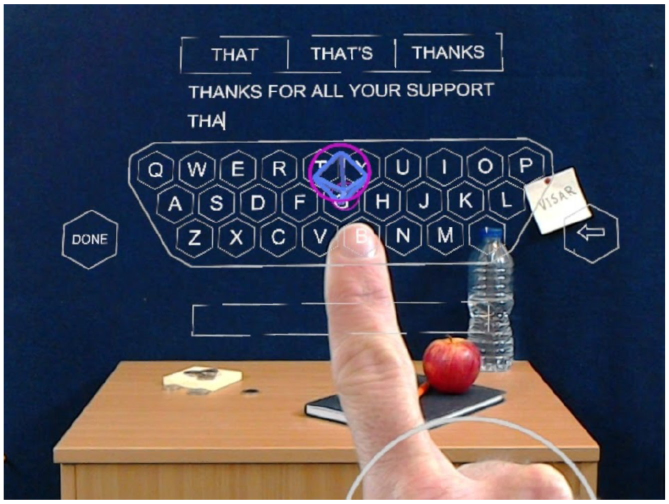

Figure 2.4: Appearance of the VISAR keyboard condition as viewed through the HoloLens in the final experiment of Dudley et al. [5].

a desktop monitor. This resulted in four within-subject conditions. In two desktop display conditions, participants typed on a keyboard that was occluded by a cover either with the virtual keyboard assistant or without the assistant. In two HMD conditions, participants wore an HMD and typed with the virtual keyboard assistant or without the assistant. The entry rates were 44.7 wpm, 44.7 wpm, 41.2 wpm, and 43.7 wpm respectively. The mean error rates were 2.6%, 3.9%, 2.6% and 4.0% respectively.

Dudley et al. [5] designed a keyboard named VISAR (Virtualized Input Surface for Augmented Reality) and performed a series of experiments on a midair virtual keyboard rendered on a see through Microsoft HoloLens HMD (Figure 2.4). They tracked only one hand via the wrist position with the finger location being a fixed offset. They first investigated if allowing users to engage with the keyboard through direct touch is more intuitive than gaze-then-gesture interaction. Due to the fact that the users had to correct errors in the previous step they did not achieve a significant difference between the conditions. To mitigate this problem, in their second experiment, they provided a fall back mechanism for precisely selecting keys. In the following experiment, they minimized the keyboard occlusion by removing the key labels and also the key outlines. In the final experiment, they modified the design from the results of the previous three experiment and provided word suggestions. Their final experiment resulted in an average entry rate of 17.8 wpm with character error rate less than 1% which was a 19.6% increase to the performance relative to the baseline investigated in Experiment 1.

Chapter 3

System Design and Experiment

3.1 Motivation

The goal of our system was to provide an efficient text entry system in VR. As discussed in the previous chapter, most of the previous work related to text entry in virtual environment used gestures, game controllers or gloves to enter text. VISAR [5] supported single-hand typing on a virtual keyboard in AR. But the limitation of this system was that the finger location was a fixed offset from the wrist position and it did not feature two hand typing. Similar to our work, Sridhar et al. [18] and Feit et al. [6] used a Leap Motion controller. However, both required users to learn specific multi-finger gestures. That's why we wanted our system to be feasible for text input

Figure 3.1: A left handed participant entering text on the virtual keyboard in the UNIMANUAL condition. When the participant touches a key it is highlighted yellow. When he takes the finger off the keyboard surface, it is added to the observation and shown in the display.

from both hands such that it does not require to learn a new input technique or require wearable input devices like gloves or controllers. We also wanted to minimize visual occlusion of other virtual environment contents. Besides these, our two hand text entry design principle was motivated by a few other questions.

- **Question 1:** Which text entry condition in VR - user entering text using a single hand (Figure 3.1), user entering text using two hands (Figure 3.2), or splitting a keyboard into two halves (Figure 3.3) enhances typing performance? This question is to investigate which of the three keyboard conditions is better in terms of higher text entry rate and reduced error rate.

  Our primary hypothesis was two hand input and split keyboard would be better than one hand in terms of higher entry rate, less character error rate and higher user satisfaction. Our hypothesis is based on the results of Bi et al. [1] and

Figure 3.2: A participant entering text on a virtual keyboard in the Bimanual condition.

Sandnes et al. [16] where two hand text entry resulted in better performance.

- **Question 2:** Does a split keyboard with two hand input provide ergonomic advantages compared to a single keyboard? In other words, does having two keyboard planes at each hand's proximity improve performance?

  Our hypothesis was split keyboard would have more ergonomic advantages over one hand input keyboard. While there is no prior work to support our hypothesis, we felt that partitioning the keyboard plane would allow a more natural arm position where each hand only needs to be moved small distances to type keys on that hands's side of the keyboard.

- **Question 3:** Provided that a user is familiar with the QWERTY keyboard layout and can locate any key without looking at the keyboard when typing on a physical keyboard, is it possible for the user to enter text in a virtual keyboard with an acceptable entry rate and character error rate?

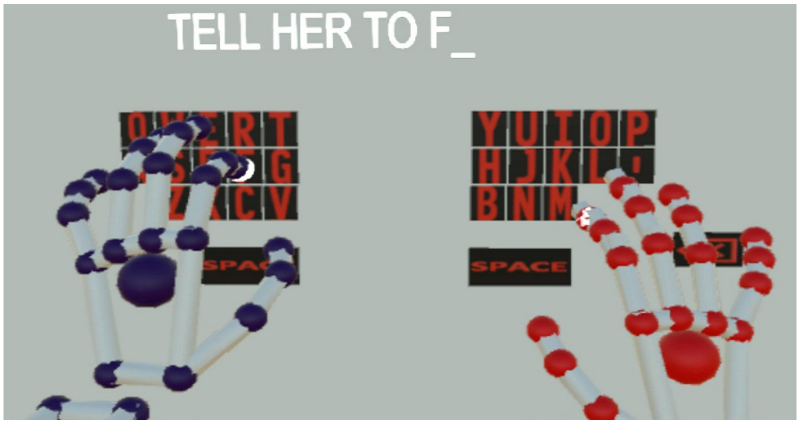

Figure 3.3: A participant entering text on a split keyboard.

Our hypothesis was that, a user could enter text on an invisible keyboard maintaining almost the same level of performance on a visible keyboard. Our hypothesis was inspired by the results from VISAR [5] keyboard where it was shown that certain users achieved comparable entry rates without the system displaying any key outlines or key labels.

3.2 Study Design

We divided our study into two experiments. The first experiment included three conditions whereas the second experiment had a single condition. The goal of the first experiment was to find out the answers to the first two questions in the previous section. The goal of the second experiment was to find out the answer of the third question. The experiments are described below.

Experiment 1: Single Hand, Both Hands or Split Keyboard

We used a 1x3 (keyboard entry) within-subjects design with the following three conditions:

- **UNIMANUAL -** In this condition, participants entered text in VR using the index finger of their dominant hand (Figure 3.1). Capsule hand from the Leap Motion asset library was used to render a participant's hands. In the beginning of this condition, participants tapped the keys of a single QWERTY keyboard projected in VR. The tap event occurred when the finger tip crossed the virtual keyboard plane. The (x,y) position of the tap point was registered and the nearest key to the (x,y) position was lit yellow. Participants could also hear a click sound during a tap event. We had an auto-correction algorithm or decoder working with the system. Participants tapped a series of (x,y) points in the system. These (x,y) points were later sent to the decoder for auto-correction once the spacebar was hit.

  A backspace key allowed the participants to perform two actions. When a participant was in the middle of entering characters of a word, the backspace removed the character just entered. But if a user entered backspace after auto-correction (i.e. the spacebar had just been entered) then backspace removed the entire previous word. Consecutive backspaces could remove other previous words. We provided this functionality so that a participant did not have to type backspace repeatedly to remove all the characters of an incorrect word.

- **BIMANUAL -** Participants entered text using both of their index fingers (Figure 3.2). The backspace key worked the same way as in UNIMANUAL condition.

- **SPLIT -** The participants entered text using both of their index fingers on a split keyboard (Figure 3.3). The motivation behind this design condition was two fold. First, we wanted to use minimum visual user interface (UI) in order to prevent the keyboard from occluding other content that may be in the virtual environment. Second, we wanted to study if their is any advantage of splitting the keyboard into two and if we could incorporate hand information in the existing auto-correction algorithm for better recognition. The left side of the keyboard included the keys - Q, W, E, R, T, Y, A, S, D, F, G, Z, X, C and V. The right side of the keyboard included the keys - Y, U, I, O, P, H, J, K, L, ', B, N and M (Figure 3.3).

Experiment 2: Invisible Keyboard

The participants entered text using an invisible keyboard (INVISIBLE). But before entering text, they were allowed to define the keyboard area by drawing a rectangle using their index finger (Figure 3.4). They could also give a thumbs up to confirm the keyboard size. Additionally, using the UI components, they could redefine the keyboard if necessary. There were a few visual elements in the invisible keyboard: a spacebar, a backspace key, a key to move to the following screen, and a text result area (Figure 3.6). The next key was only visible after any pending input had been

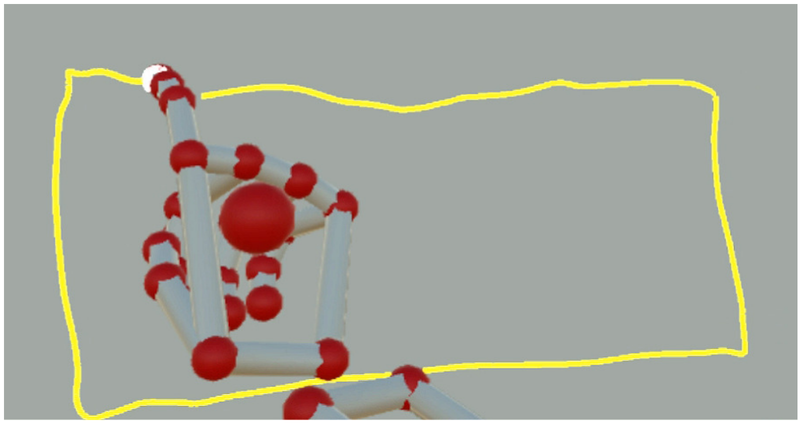

Figure 3.4: In Invisible condition, participants were allowed to define the keyboard rectangle. Above picture shows how a participant drew a rectangle extending the index finger of his dominant hand.

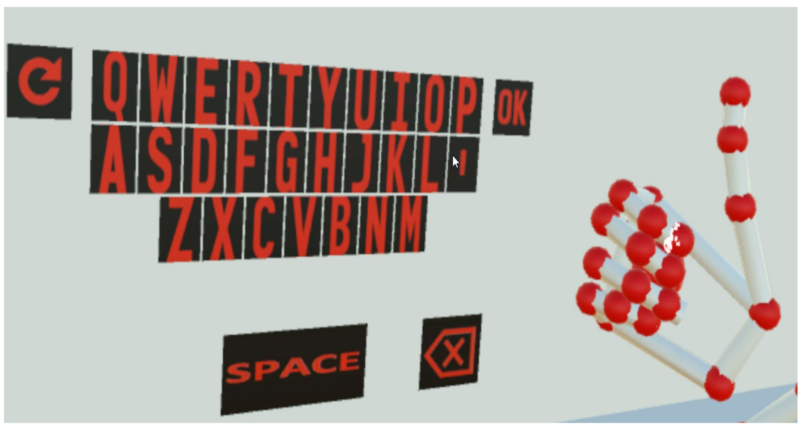

Figure 3.5: In Invisible condition, participants gave a thumbs up to confirm the keyboard area and viewed the keyboard before they proceeded to entering text on the invisible keyboard.

recognized by the user first tapping the spacebar key.

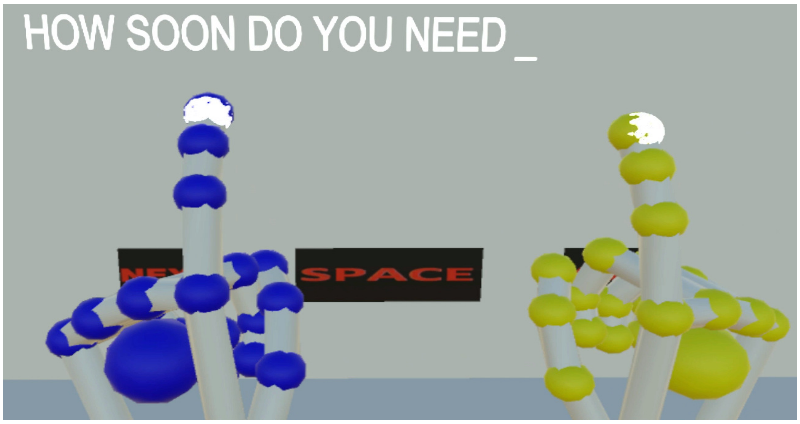

Figure 3.6: A participant entering text on an INVISIBLE keyboard.

3.3 Participants

For Experiment 1, We recruited 24 participants via convenience sampling. No participant had uncorrected vision deficits or motor impairments. Due to technical issues, we had to replace two of the original participants. Participants were aged between 18 and 44 (7 female, 17 male, mean age 26.5, SD = 6.8). 22 participants were right handed and 2 participants were left handed.

All participants were familiar with the QWERTY keyboard layout but 5 participants reported that they could not locate keys without looking at the QWERTY keyboard. 15 participants reported that they used virtual reality headset before while 9 participants reported that they had never used a virtual reality headset.

For Experiment 2, we used the same pool of 24 participants from Experiment 1. But we had to drop two additional participants due to logging issues.

3.4 Apparatus

We used the latest Leap Motion Sensor (Orion Beta, SDK version 3.2) as the primary hand and finger tracking device. We used Unity (Version 5.5.3) as the platform for getting the visual feedback of the hands and fingers. The VR environment was run on an HTC Vive. We developed our software using Monodevelop Unity and Microsoft Visual Studio 2015. To integrate Unity, Leap Motion and HTC Vive, we used Unity Core Assets (version 4.1.5). The programming language for Unity, HTC Vive and Leap Motion related tasks was $C\#$. We also used an auto-correction algorithm named VelociTap [21] which was written in Java. A brief discussion of VelociTap is provided in Section 3.5. We used TCP socket communication to send the tap data from Unity to VelociTap and to send back the recognized text.

3.5 Decoder

We used two tracking devices namely the Leap motion sensor for hand tracking and HTC Vive for projection of virtual keyboard. Both tracker introduced uncertainties into the input data. We used the VelociTap [21] decoder to auto-correct users' noisy

Figure 3.7: A user wearing a HMD with a leap motion sensor attached.

tap sequence into the most likely text.

VelociTap allows users to enter all the characters of a word by touching on the keys of a virtual keyboard. After entry, a series of noisy touch locations is fed to VelociTap as input observations for recognition. VelociTap then searches for the most likely word which is the closest match to the input observations as well as the most probable word according to a language model. To tune the parameters of VelociTap for our study, we ran a pilot and collected some data before the main study. The members of the Future Interaction Lab led by Dr. Keith Vertanen at Michigan Technological University participated in the pilot. Based on the collected data, we fine-tuned VelociTap for our virtual keyboards and used those parameter values in the main study. The parameters were same across all conditions in Experiment 1. However, the size of the keyboard in Experiment 2 was variable and for that reason we needed to change the values of a few parameters.

3.6 Procedure

In the beginning of the experiment we obtained each participant's informed consent. We informed the participants about the purpose of the research, what tasks they had to do, any foreseeable risks of harm, and that the study was voluntary.

We then asked each participant to fill out a paper questionnaire which asked demographics questions, whether the participant was left or right handed and whether the participant had any previous experience with VR text input.

After that, we helped the participant to put on and adjust the HMD. The Leap Motion device was attached to the HMD (Figure 3.7). We gave each participant 1-2 minutes to become familiar with the virtual environment. The participant did 3-4 minutes of practice in each of the three conditions of Experiment 1. The three conditions occurred in the same order the participant would experience them in the evaluation portion of the study.

For each practice condition in Experiment 1, we showed each participant four memorable sentences and for each study conditions, we showed 12 sentences from Enron mobile test set [20]. We chose sentences which had been memorized correctly by at least 6 out of 10 workers in [20]. Table 3.1 shows a list of example sentences from Enron mobile test set. A participant never saw the same sentence twice in any of

(i)	i think that is the right answer
(ii)	keep me posted
(iii)	are you being a baby
(iv)	do you need it today
(v)	neil has been asking around
(vi)	have a great trip
(vii)	she called and wants to come over
(viii)	a gift isn't necessary
(ix)	i'm glad you liked it
(x)	are you feeling better

Table 3.1

A list of example sentences from the Enron mobile test set.

the practice or main study conditions. Each participant went through the three conditions of Experiment 1 in a counterbalanced order. We also counterbalanced and randomized the order of appearance of the sentences to enter. We asked participants to enter the sentences "quickly and accurately" in each of the practice and study conditions.

For each condition, we instructed participants to enter text using one hand, both hands or on a split keyboard. When they entered all the sentences in a condition, they filled out a questionnaire. The questionnaire sought comments and the perception of physical exertion pertaining to that condition. The participants took a break before proceeding to the next condition.

In UNIMANUAL condition, we instructed the participants to enter text using the index finger of their dominant hand.

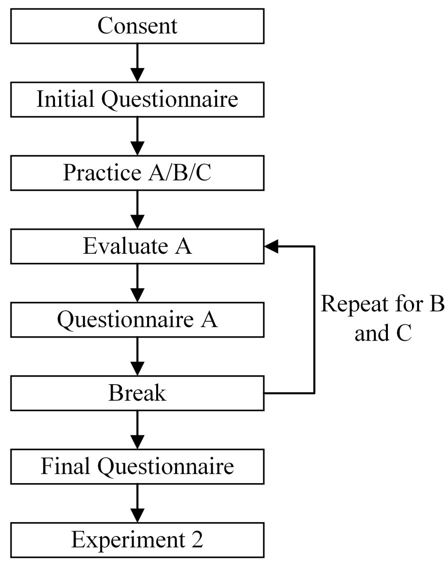

Figure 3.8: Diagram showing different steps of the study.

In BIMANUAL condition, we asked participants to enter text using both of their index fingers.

In SPLIT condition, the keyboard was split into two parts and we asked the participants to enter text using both of their index fingers.

When text entry under all these three conditions was done, the participants filled out a final questionnaire. The questionnaire sought overall comments on those three conditions. Figure 3.8 shows the progression through the study.

For Experiment 2, we asked participants to define the keyboard area first. To do so, a screen with a four second countdown timer appeared before the participants in VR. We instructed each participant to extend the index finger of the dominant hand and position it such that a rectangle could be drawn in the vision space. The count down timer was provided so that a participant could choose a comfortable position. After the countdown timer had stopped, the participant drew a rectangle. A yellow line followed the finger and the trace of the rectangle being drawn was visible. When the participant was done drawing the rectangle, he gave a thumbs up gesture and a keyboard of the same size as the drawn rectangle appeared.

There were two extra buttons on the keyboard - a reload button and a OK button on left and right end of the keyboard respectively. If the participant was happy with the keyboard, the participant could touch OK to confirm the keyboard. If the participant was not happy with the keyboard, the participant could redraw the rectangle and define the keyboard area again by touching the reload button. After the participant touched OK button the keyboard was shown and the participant was allowed to look at the keyboard to visualize the key positions. Next the keyboard disappeared and the participant entered sentences in a keyboard where all the keys except spacebar, backspace and next key were invisible. There was no visual feedback of the keys when the participant tapped a key, but the participant could hear a click sound.

3.7 Metrics

We calculated the following list of metrics to measure the performance and user experience in our study:

- **Words-per-minute:** Entry rate in words-per-minute (wpm) is a standard measure in text entry. Since the length of words can be variable, we considered each word as of five characters including space. Time was measured for each phrase from a user's first key entry until the recognized text was displayed. Then we divided the number of words by the time to calculate entry rate in wpm.

- **CER:** Character Error Rate (CER) compares the entered text phrase by the user to the target text phrase. First we calculated the minimum edit distance between the entered phrase and the target phrase. Then we divided this distance by the number of characters in the target phrase. Finally, it was multiplied by 100 and expressed as a percentage (%).

- **Backspaces-per-character:** While entering text in VR it is pretty normal that a participant will mistype a key. We allowed the participant to remove an incorrect entry by hitting the backspace key. We logged the occurrence of such a behaviour and later used to calculate the backspaces-to-character ratio.

- **Perceived Exertion:** Borg CR10 is a scale that asks users to rate their perception of physical exertion after completing a specified task. This scale is used to measure different kinds of sensations including pain, agitation, taste, smell, loudness, etc. After each of the study conditions in Experiment 1, we asked participants to rate their physical exertion level according to the Borg scale. The form that we used for Borg CR10 rating is included in Appendix A.

3.8 Data

To calculate the metrics, we logged the following data:

- Hand information i.e. which hand did a participant use to enter a character in the experimental interface.
- The 3D coordinate of a point when a finger tapped a key on the virtual keyboard.
- The time a participant started looking at a sentence.
- The time a participant was done looking at a sentence.
- The time a participant touched the first letter of a sentence on the keyboard.
- The time a participant was done entering a complete sentence.
- How frequent a participant was using backspace to remove a character that was entered on the keyboard.

Chapter 4

Results

4.1 Experiment 1: Single Hand, Both Hands or Split Keyboard

Based on the information in the log files, we calculated the entry rate and character error rate for each of the sentences. We found in 9 phrases out of 864, participants left off two or more words in the target sentence or entered a completely different phrase. Inevitably for such cases, the character error rates were so high that they disproportionally increased the overall error rate in each condition. We excluded these input sentences from our analysis. Table 4.1 shows the complete list of such input sentences compared to the reference sentences.

Condition	Reference	User input
Unimanual	take what you can get	hi there
Bimanual	keep me posted	I
	do you need it today	do you
	hope your trip to florida was good	i hope you are doing great
	have a great trip	have
Split	she called and wants to come over	i dong know
	they have capacity now	they had calixty right now
	she called and wants to come over	mckaldidk
	hope your trip to florida was good	fu do

Table 4.1
A complete list of phrases where participants left off two or more words in the target phrase in Experiment 1.

Table 4.2 provides numeric results and statistical tests after excluding input sentences in Table 4.1 and their related information.

4.1.1 Entry Rate

We measured entry rate from a participant's first tap until the recognition was displayed. The average recognition delay including the round trip network delay was 312 milliseconds. Participants spent on an average 3768 milliseconds, 3974 milliseconds and 4372 milliseconds looking at the stimuli sentence in Unimanual, Bimanual and Split respectively.

Participants were slightly faster in Bimanual at 16.4 wpm compared to Unimanual at 16.1 wpm (Figure 4.1 left). They were slower in Split at 14.7 wpm. Differences between Split versus Bimanual were statistically significant (Table 4.2).

Condition	Entry rate (wpm)	Error rate (CER%)	Backspaces-per-character
UNIMANUAL	16.1 ± 2.9 [10.7, 21.9]	0.74 ± 0.87 [0, 3.03]	0.0140 ± 0.0214 [0.0000, 0.0761]
BIMANUAL	16.4 ± 2.3 [10.5, 19.9]	0.79 ± 1.17 [0, 4.37]	0.0169 ± 0.0132 [0.0031, 0.0566]
SPLIT	14.7 ± 2.4 [11.1, 20.5]	1.41 ± 1.51 [0, 5.89]	0.0166 ± 0.0162 [0.0000, 0.0449]
ANOVA	$F_{2,46} = 5.52$, $\eta_p^2 = 0.19$, $p < 0.01$	$F_{2,46} = 2.31$, $\eta_p^2 = 0.09$, $p = 0.11$	$F_{2,46} = 0.36$, $\eta_p^2 = 0.015$, $p = 0.7$
Post-hoc	UNIMANUAL ≈ BIMANUAL, p = 1.00	Not applicable	Not applicable
	UNIMANUAL ≈ SPLIT, p = 0.079		
	SPLIT < BIMANUAL, p < 0.05		

Table 4.2

Results from Experiment 1. Results formatted as: mean ± SD [min, max]. The bottom section of each table shows the repeated measures ANOVA statistical test for each dependent variable. For significant omnibus tests, we show pairwise post-hoc tests (Bonferroni corrected).

Condition	Entry rate (wpm)	Error rate (CER%)	Backspaces-per-character
UNIMANUAL	16.0 ± 2.9 [10.7, 21.9]	1.06 ± 1.62 [0, 07.54]	0.0141 ± 0.0214 [0.0000, 0.0762]
BIMANUAL	16.3 ± 2.3 [10.5, 19.9]	1.98 ± 3.27 [0, 13.16]	0.0168 ± 0.0131 [0.0031, 0.0566]
SPLIT	14.7 ± 2.4 [11.1, 20.5]	2.58 ± 4.27 [0, 19.34]	0.0161 ± 0.0157 [0.0000, 0.0446]
ANOVA	$F_{2,46} = 5.84$, $\eta_p^2 = 0.20$, $p < 0.001$	$F_{2,46} = 1.33$, $\eta_p^2 = 0.05$, $p = 0.13$	$F_{2,46} = 0.27$, $\eta_p^2 = 0.011$, $p = 0.77$
Post-hoc	UNIMANUAL ≈ BIMANUAL, p = 1.00	Not applicable	Not applicable
	UNIMANUAL ≈ SPLIT, p = 0.072		
	SPLIT < BIMANUAL, p < 0.05		

Table 4.3

Results from Experiment 1 before excluding input sentences in Table 4.1 and information related to those sentences. Results formatted as: mean ± SD [min, max].

4.1.2 Error Rate

We measured error rate by computing the CER of the recognition against the stimuli sentences. The error rate was similar and low across all conditions. CER was lowest in UNIMANUAL at 0.74%, followed by BIMANUAL at 0.79% and, finally SPLIT at 1.41% (Figure 4.1 middle). These differences were not statistically significant (Table 4.2).

We plotted entry rate versus character error rate of all the participants in three different conditions in Experiment 1 (Figure 4.2). All participants obtained a low

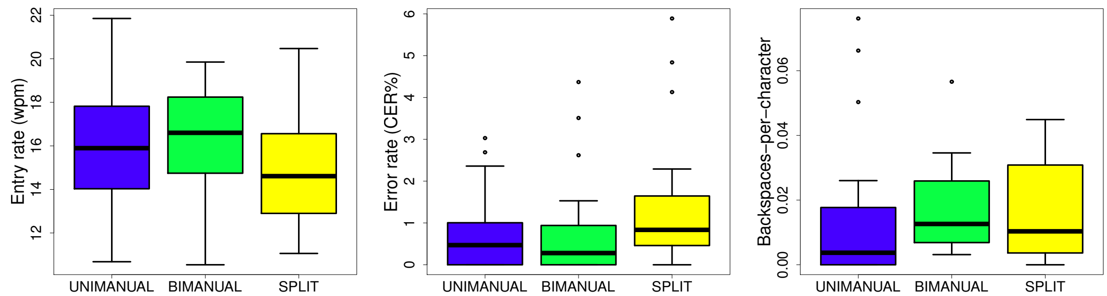

Figure 4.1: Entry rate, character error rate (after recognition) and backspaces-per-character in Experiment 1.

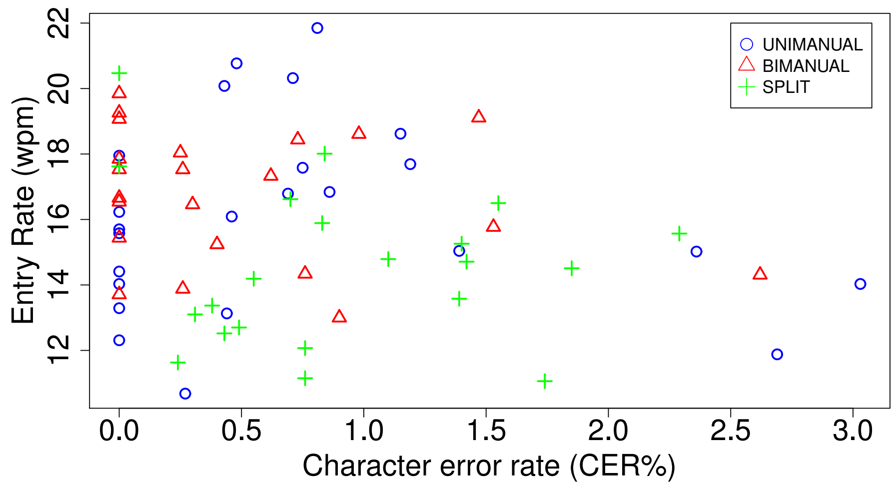

Figure 4.2: Error rate and entry rate of all participants in Experiment 1.

CER of 3% or less with many achieving near perfect accuracy. Error rate was more variable in SPLIT. We conjecture this might be due to hand tracking problems near the sensors periphery, or by participants being less accurate at targeting keys at the periphery.

4.1.3 Backspaces-per-character

Participants could hit the backspace key to remove the last entry. But the participants hit the backspace key only infrequently. Participants' average backspaces-per-character were 0.014 in Unimanual, 0.0169 in Bimanual, and 0.0166 in Split (Figure 4.1 right). These differences were not statistically significant (Table 4.2).

Recall we removed 9 sentences from our analysis. However, we also calculated numeric results including these sentences. Table 4.3 shows these results. We could see that there is not much difference between Table 4.3 and Table 4.2 in terms of entry rate. But there is a decrease of character error rate in Table 4.2 because removing the invalid input sentences from the data lowered the character error rate.

4.1.4 Questionnaire

We asked participants to rate their experience for each of the conditions. Responses were recorded on a Likert scale from 1 (Strongly disagree) to 5 (Strongly agree). For each condition, participants were asked whether they thought they entered text quickly in the experimental interface, whether they thought they entered text accurately, whether they thought the experimental interface provided accurate visual feedback and finally whether they thought the experimental interface detected a key press accurately. The questionnaire is included in Appendix A. Figure 4.3 summarizes the results calculated from the Likert ratings. For these ratings, we tested for significance using Friedman's test.

The mean rating for the statement "I thought I entered text quickly" was 4.17 in Unimanual, 4.08 in Bimanual and 3.75 in Split. These differences were not statistically significant (Table 4.4).

The mean rating for the statement "I entered text accurately" was 3.88 in Unimanual, 3.54 in Bimanual and 3.17 in Split. These differences between Split and Unimanual were statistically significant (Table 4.4).

The mean rating for the statement "I thought the experimental interface provided accurate visual feedback of my hands" was 4.12 in Unimanual, 3.79 in Bimanual and 3.67 in Split. These differences were not statistically significant (Table 4.4).

Finally, the mean rating for the statement "I thought the experimental interface detected a key press accurately when I typed a key" was 3.88 in Unimanual, 3.62 in Bimanual and 3.42 in Split. These differences were not statistically significant (Table 4.4).

At the end of Experiment 1, participants were asked which of the three keyboard conditions they preferred. We also asked the participants to order their preferred

	Condition	Mean (SD)	1	2	3	4	5
Detected key	SPLIT	3.42(1.2)	4.2%	20.8%	25.0%	29.2%	20.8%
	BIMANUAL	3.62(0.97)	0.0%	16.7%	20.8%	45.8%	16.7%
	UNIMANUAL	3.88(1.12)	4.2%	12.5%	4.2%	50.0%	29.2%
Visual Feedback	SPLIT	3.67(1.1)	0.0%	20.8%	16.7%	37.5%	25.0%
	BIMANUAL	3.79(0.88)	0.0%	8.3%	25.0%	45.8%	20.8%
	UNIMANUAL	4.12(0.68)	0.0%	0.0%	16.7%	54.2%	29.2%
Accurately	SPLIT	3.17(1.0)	4.2%	25.0%	29.2%	33.3%	8.3%
	BIMANUAL	3.54(0.83)	0.0%	8.3%	41.7%	37.5%	12.5%
	UNIMANUAL	3.88(0.99)	0.0%	12.5%	16.7%	41.7%	29.2%
Quickly	SPLIT	3.75(0.9)	0.0%	8.3%	29.2%	41.7%	20.8%
	BIMANUAL	4.08(0.72)	0.0%	0.0%	20.8%	50.0%	29.2%
	UNIMANUAL	4.17(0.64)	0.0%	0.0%	12.5%	58.3%	29.2%

Figure 4.3: Subjective ratings on condition Unimanual, Bimanual and Split on a Likert scale from 1 (Strongly disagree) to 5 (Strongly agree).

Condition	Quickly	Accurately	Visual feedback	Detected key
Unimanual	4.17± 0.64 [3.00, 5.00]	3.88 ± 0.99 [2.00, 5.00]	4.12 ± 0.68 [3.00, 5.00]	3.88 ± 1.12 [1.00, 5.00]
Bimanual	4.08± 0.72 [3.00, 5.00]	3.54 ± 0.83 [2.00, 5.00]	3.79 ± 0.88 [2.00, 5.00]	3.62 ± 0.97 [2.00, 5.00]
Split	3.75 ± 0.90 [2.00, 5.00]	3.17 ± 1.05 [1.00, 5.00]	3.67 ± 1.09 [2.00, 5.00]	3.42 ± 1.18 [1.00, 5.00]
Friedman's Test	$\chi^2 = 3.9636, df = 2, p = 0.1378$	$\chi^2 = 10.265, df = 2, p = 0.0059$	$\chi^2 = 5.5088, df = 2, p = 0.0637$	$\chi^2 = 4, df = 2, p = 0.1353$
Post-hoc	Not applicable	Bimanual-Split, obs. diff = 11.5, False Bimanual-Unimanual, obs. diff = 7.0, False Split-Unimanual, obs. diff = 18.5, True Critical difference in all conditions = 16.59	Not applicable	Not applicable

Table 4.4

Subjective results on Experiment 1. Results formatted as: mean ± SD [min, max]. The bottom section of each table shows the Friedman's test.

keyboard in terms of quickness, accuracy and minimal effort.

In terms of quickness, 8 participants preferred Unimanual, 10 participants preferred Bimanual and 6 participants preferred Split. In terms of accuracy, 9 participants preferred Unimanual, 9 participants preferred Bimanual and 6 participants preferred Split. In terms of minimal effort, 11 participants liked Unimanual, 8 participants liked Bimanual and 5 participants liked Split. We also asked them to rate the interfaces overall. 10 participants preferred Bimanual, 7 participants preferred

Rating	Exertion
0	No exertion at all
1	Very light exercise
3	Moderate exercise
5	Heavy exercise
7	Very hard and strenuous
10	Extremely strenuous

Table 4.5
Borg CR10 scale rating and perceived exertion.

UNIMANUAL, and 7 participants preferred SPLIT.

4.1.5 Perceived Exertion

We also asked participants to rate their perceived exertion level while performing in each condition. Table 4.5 describes the Borg CR10 rating and their corresponding meaning. Figure 4.4 shows the results from our study. The average exertion level in each conditions are nearly same. But SPLIT had the lowest average Borg rating with 3.04. BIMANUAL had an average rating of 3.08 and UNIMANUAL had an average of 3.38. All of the average ratings correspond to a rating of "Moderate exercise". We also ran Friedman's test on these ratings but found no statistical significance ($\chi^2 = 3.2698$, $df = 2$, $p = 0.195$).

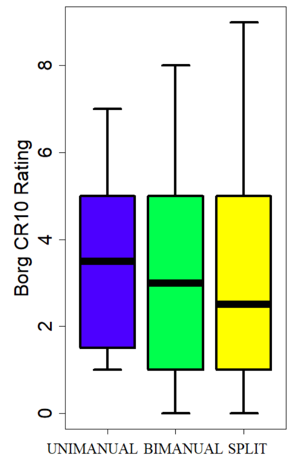

Figure 4.4: Subjective ratings of perceived exertion using Borg CR10.

4.1.6 User input behaviour

We were also interested in how frequent participants used their right index to tap a key on the left side of the keyboard and used their left index finger to tap a key on the right side of the keyboard. We took all the tap points in Bimanual and Split into consideration and made scatter plots to visualize the behaviour. Figure 4.5 shows the scatter plot of all the tap locations by 24 participants in Bimanual. Since 22 of the 24 participants were right handed, we can see that the right hand moving to the left side was fairly frequent while vice versa was rare.

In Bimanual condition, for the given reference phrases, the percentage of letters on the left side of the keyboard was 53.9% and percentage of letters on the right side was 46.1%. In 50.2% cases the participants used a left hand and in 49.8% cases the participants used a right hand to tap a key. So, there is an increased usage of right

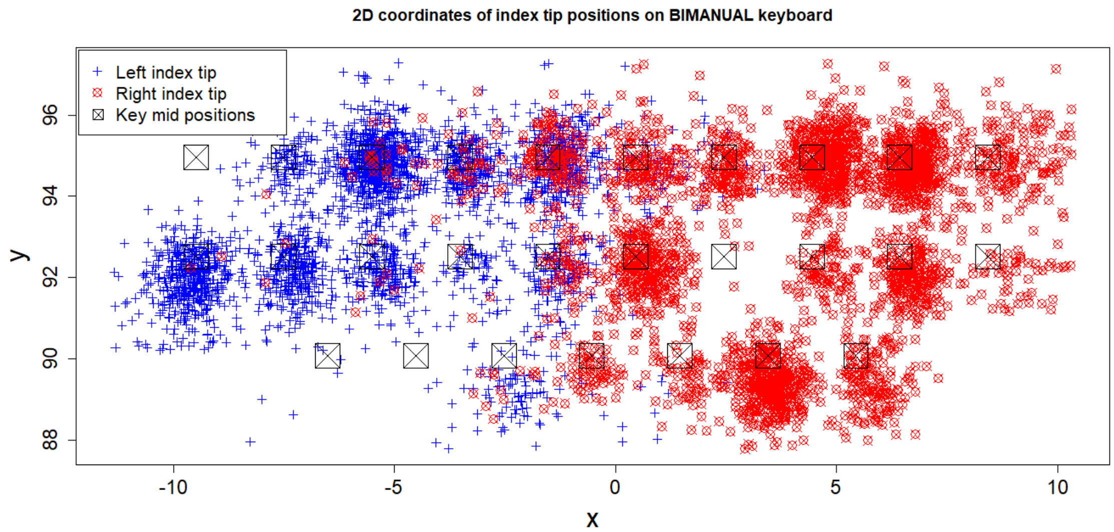

Figure 4.5: Scatter plot of all the index tip positions of the participants when entering a letter on the BIMANUAL keyboard.

hand of about 3%.

Figure 4.6 shows the scatter plot of the tap locations under SPLIT condition. Even if the split parts of the keyboard was a distance away from each other, we can see some movement of the right index finger to the left side of the split keyboard. In case of the left hand typing on the wrong side of the keyboard was very rare and overall there were 4-5 occurrences. This suggests that the participants had a tendency to use their dominant hand once in a while instead of their non-dominant hand.

In SPLIT condition, from the given reference phrases 54.2% letters were on the left split of the keyboard and 45.8% letters were on the right split of the keyboard. But the taps performed by the participants indicate that the usage of left hand was 53.8% and usage of right hand was 46.2%. This corresponds with the observed instances of

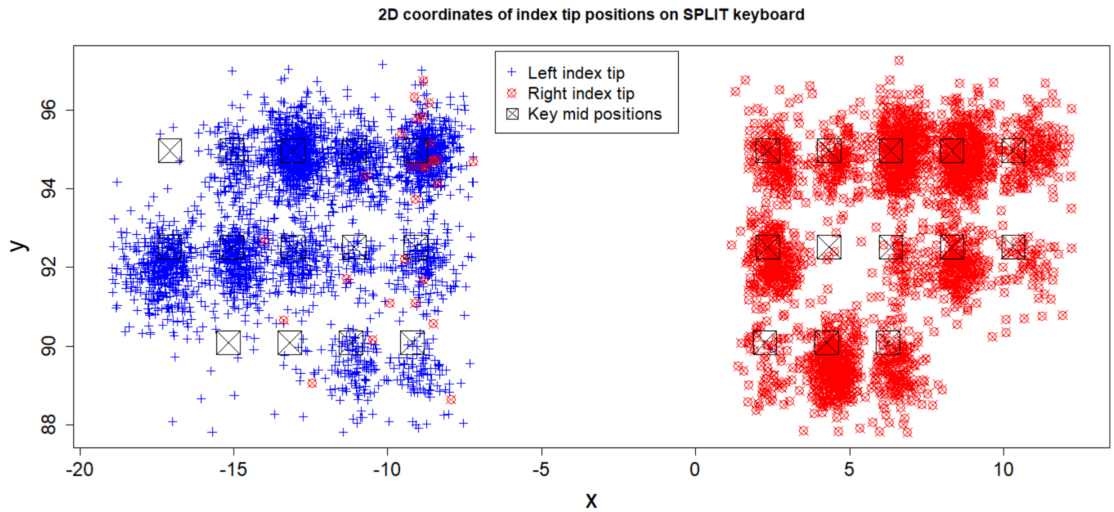

Figure 4.6: Scatter plot of all the index tip positions of the participants when entering a letter on the SPLIT keyboard.

participant typing on the left side with their right hand.

4.1.7 Open comments

We also asked participants to share their comments on each of the conditions in Experiment 1. Most of the negative comments they shared were related to hand tracking. Many of the participants complained that when they used both hands the hand tracker was inaccurate. Sometimes the tracker assumed the right hand was the left hand and vice versa, sometimes the fingers would flip away, sometimes the orientation of the hands were not correct, etc. The reason behind these might be related to one hand occluding the depth camera from seeing the other hand. Also when participants used both hands then the tracker had more things to track

Unimanual
+ "This felt faster and easier"
+ "I became more comfortable using this system. Trust the correction system"
– "It's tiring while using only one index finger. At times, it felt as if the visual feedback was not syncing with my perception"
– "The system confused the orientation of my hand and confused my right hand for my left hand. At time the virtual fingers would move on their own"

Bimanual
+ "This felt easier for me because I type on a computer most of the time. This felt more natural"
+ "I definitely felt more comfortable in this position because of the layout of the keyboard. I was faster and confident in my ability"
– "Seemed to flash when both hands appear on screen at the same time"
– "I felt clogged down by one hand needing to be pulled back before typing the next characters. Hands got in each other's way."

Split
+ "With more practice, it becomes easier to play with the keyboard"
+ "I have never typed in VR before so this has been a learning experience"
– "I had trouble typing accurate with my left hand and felt like I was moving my head a lot to see the different halves of the keyboard
– "Overall, I felt my dominant hand was more accurate"

Table 4.6
Selected positive and negative comments provided by participants about each condition in Experiment 1.

compared to using a single hand. We believe these kind of confusions resulted in less effective performance than we hoped. If the tracker were not prone to anomalous behaviour, we think the entry rate would have been faster and character error rate would have been lower in the Bimanual condition. A list of positive and negative comments on Experiment 1 are given in Table 4.6.

Condition	Entry rate (wpm)	Error rate (CER%)	Backspaces-per-character
INVISIBLE	10.3 ± 3 [5.2, 17.8]	7.96 ± 9.56 [0.52, 40.28]	0.0520 ± 0.0398 [0.0031, 0.1284]

Table 4.7
Results from Experiment 2. Results formatted as: mean ± SD [min, max].

4.2 Experiment 2: Invisible Keyboard

Experiment 2 was a stand alone study to investigate if participants can enter text on an invisible keyboard without any visual occlusion. For this experiment, we allowed participants to define their own keyboard. Table 4.7 summarizes the average entry rate and character error rate in this condition.

Among the 22 participants, the mean entry rate was 10.3 wpm and error rate was 7.96%. In the Figure 4.7, we can see that the worst participant had an error rate of 40%. It was due to the fact that this participant defined a keyboard so small that he was struggling to enter text. The participant self-reported this information on the open comments section. Experiment 2 also revealed that even though some participants self-reported that they could locate a key without looking at the QWERTY layout, we observed many of the participants were not that familiar with the spatial location of the keys. This contributed to their low entry rates and high error rates.

We removed 7 highest error rate participants and calculated mean entry rate and error rate for the remaining 15 participants. This subset of 15 participants had a

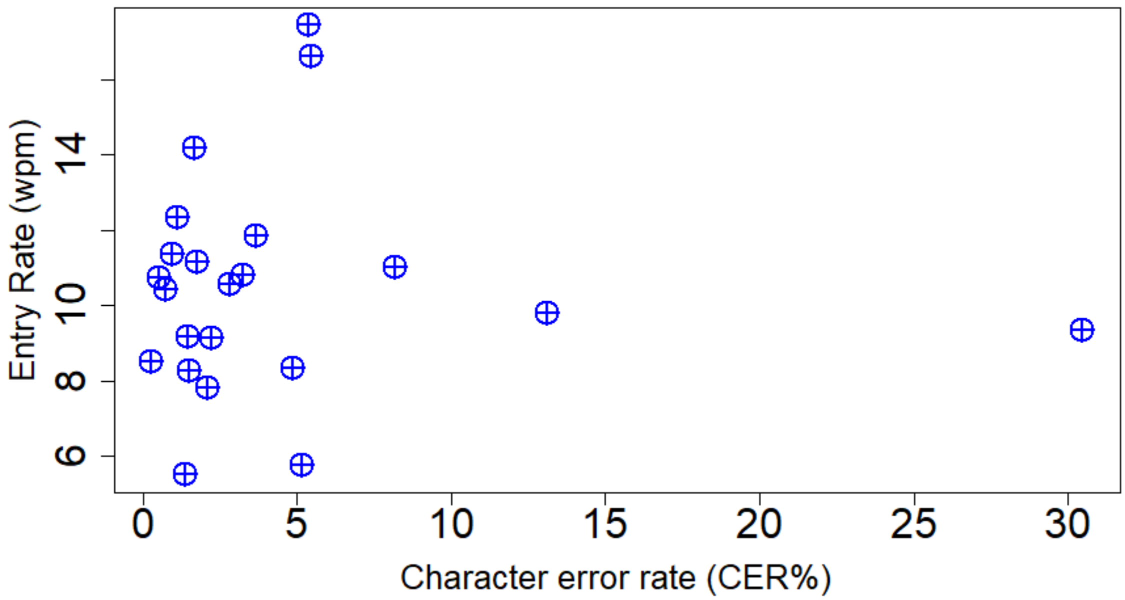

Figure 4.7: Entry rate vs character error rate scatter plot for Experiment 2.

mean entry rate of 10.0 wpm and a mean error rate of 2.98%.

We asked participants to rate their experience in INVISIBLE condition using a Likert scale from 1 (Strongly disagree) to 5 (Strongly agree). The results from the subjective feedback is showed in Figure 4.8. Most of the participants were satisfied with their ability to draw the keyboard as they wanted. They felt that it was difficult to enter text without any visual feedback and felt they could not enter text accurately.

The mean rating for the statement "I successfully obtained the keyboard size" was 4.18. The mean rating for the statement "I found it easy to enter text without any visual feedback" was 2.95. The mean rating for the statement "I was able to easily understand when and what key I typed" was 3.18. The mean rating for the statement

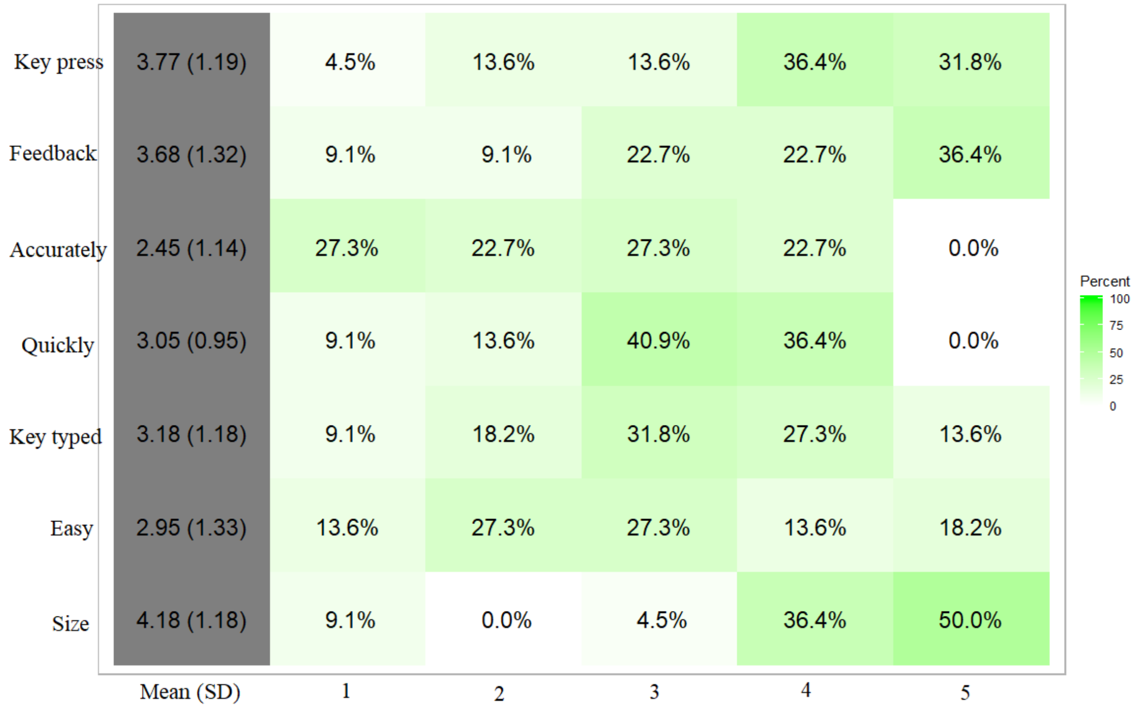

Figure 4.8: Subjective feedback of Experiment 2 on a Likert scale from 1 (Strongly disagree) to 5 (Strongly agree).

"I entered text quickly" was 3.05. The mean rating for the statement "I entered text accurately" was 2.45. The mean rating for the statement "I got accurate feedback of my hands" was 3.68. Finally, the mean rating for the statement "The experimental interface detected a key press accurately when I typed a key" was 3.77.

If we analyze the open comments (Table 4.8), most of the people felt that the invisible keyboard was easier than they thought it would be. The people who struggled mostly had trouble defining the keyboard properly or had trouble remembering the spatial location of the keys.

Invisible	
+	"This part went a lot better than I thought it would. Finding the correct key was more successful than I anticipated but I was still very thankful for auto-correct"
+	"I found watching the key result instead of trying to see the keyboard helped and accepted that close enough was good enough with auto correct"
+	"The adjustable size of the keyboard is good, one can use the one one wants. Relying on the position of key and not the actual location is good"
+	"It got easier with practice - when I understood what my keyboard was"
+	"It was hard at first but I was able to remember the keyboard better with each sentence"
+	"It was fun"
+	"Easier than I thought it would be"
–	"Sometimes, it detected too sensitively so even when I didn't intend to, 'next' button was pressed"
–	"I think the keyboard was actually too small to use for me
–	"There were some errors when I put the buttons ('the finger was folded')"

Table 4.8

Selected positive and negative comments provided by participants about Invisible condition in Experiment 2.

Chapter 5

Discussion

We presented two experiments in this thesis to explore the design of an effective system for text entry in VR. In Experiment 1, we investigated one hand (Unimanual) and two hand (Bimanual) text input. We also investigated text input on a split (Split) keyboard. The results from Experiment 1 differs a little from our initial hypothesis. For example, we thought that we would achieve faster entry rate and lower error rate in the Bimanual condition compared to the Unimanual condition. But we got similar entry rates and error rates for both these conditions (not statistically significant). We think the reason behind getting similar results in Unimanual and Bimanual was tracker challenges with two hands. A slightly higher backspaces-to-character ratio in Bimanual (0.0169) than in Unimanual (0.014) and subjective feedback from both conditions support this conjecture.

We hypothesized that Split would be more ergonomic, but even though Split had the lowest average Borg rating, it was the least preferred condition. We think the poor performance in Split condition was due to three reasons.

First, the tracker issue. We mentioned earlier that the Leap Motion sensor which was detecting the hands sometimes provided inaccurate feedback. That is why the participants faced difficulties to enter text under the Split condition. For example, when a participant was going to hit a key, the hand suddenly flipped or the fingers got crooked. Though this happened for both Split and Bimanual, it was worse in Split. One conjecture is that hands separated by empty space was not well modelled by Leap Motion.

Second, since the keyboard was split into two parts maintaining a significant distance in between the split planes, the participants had to move their head constantly to look at a key on different parts of the keyboard. Rather than concentrating on a single plane the participants had to concentrate on two different planes. Head movement could have lowered the overall performance.

Third, it could be due to moving hands towards the edge of the sensing range.

In Experiment 2, we tried to minimize keyboard occlusion and allow user to choose their own keyboard size. While some participants were natural in entering text in the Invisible condition, some of the participants struggled due to no visual feedback of the keys. Also some participants defined smaller keyboards and some drew keyboards which did not have the proper ratio of height and width like a regular desktop keyboard. Moreover, some participants were not aware of the spatial location of the keys in the QWERTY keyboard layout. Some suggested adding a keyboard boundary for visual clue. But the majority of the participants (15 of 22) entered text with a mean entry rate of 10 wpm and a mean error rate of 2.98%. We think this is a very promising result. We believe that with better tracker and more practice, better performance can be achieved.

5.1 Limitations

Our proposed design eases text entry in virtual reality by allowing users to use their fingers as input an device. But our design is still limited to using only two fingers. Though it is always desirable to use all the fingers of the hands, it is still a difficult task because accurate tracking of multi-fingers is challenging. Our system also suffered from inaccuracies due to tracker field of view. Tracking of hands was not accurate at the edge of tracker's field of view.

Our design also handles only the keys from a-z and apostrophe. For a full featured virtual keyboard, the other keys for example, the punctuation and case sensitivity must be allowed and the performance needs to be evaluated.

Also, prolonged use of the system may result into a higher level of perceived exertion by the user. It is expected because users need to lift their hands and touch a key in midair. Various positions and size of the keyboard can be investigated to find the position with minimal physical effort.

5.2 Future work

From the analyzed results and limitations of our design we plan to perform the following works:

- We plan to incorporate hand information to our existing auto-correction algorithm to investigate if hand information can contribute to better recognition. We will create a training dataset from 8 of the participants' input data and a test dataset from the remaining 16 participants' data. Using these two dataset, we will train and test a model that incorporates which hand the user tapped the keyboard with a goal of reducing character error rate. When entering the characters for a target sentence, each character is assigned a probability based on VelociTap's [21] keyboard model and language model. We will extend this, adding a penalty for characters entered with the wrong hand (i.e. the user entered a key on the left side of the keyboard with their right hand). This penalty will add a new free parameter to VelociTap. We will optimize this parameter with respect to the data in the training set and analyze whether it reduces recognition errors on the test data.

- We will investigate the different sizes and positions of the keyboard to minimize physical effort and maximize performance. We also want to add word suggestions, i.e. words that complete the currently typed partial entry of a word to yield better entry rates.

- We will extend the existing system to work with multi-finger mid air input to investigate if 10 finger text entry in VR can be achieved.

- We will add a keyboard outline as a minimal visual feedback of the keyboard in the virtual environment for the invisible condition.

- We also plan to compare a commodity sensor like Leap Motion against a very accurate Vicon setup.

5.3 Conclusion

In this thesis, we explored four different text entry methods for midair typing in VR. We explored their limitations and their advantages. While in some cases the performance of the system was negatively affected by the tracker, our design paves the way to multi-finger text input research in virtual environment. There are three key contributions of this thesis. First, we designed the system to allow finger inputs instead of using any hand-held controller or gloves. To our knowledge, we are the first to study two handed virtual keyboard input on a midair keyboard. Second, our study compared single hand, two hand and split keyboard text entry techniques and their performance in VR. Third, for users with good knowledge of the QWERTY layout, we found input on an invisible keyboard to be a promising text entry solution in VR.

References

[1] Xiaojun Bi, Ciprian Chelba, Tom Ouyang, Kurt Partridge, and Shumin Zhai. Bimanual gesture keyboard. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, pages 137–146. ACM, 2012.

[2] Wolfgang Birkfellner, Michael Figl, Klaus Huber, Franz Watzinger, Felix Wanschitz, Johann Hummel, Rudolf Hanel, Wolfgang Greimel, Peter Homolka, Rolf Ewers, et al. A head-mounted operating binocular for augmented reality visualization in medicine-design and initial evaluation. *IEEE Transactions on Medical Imaging*, 21(8):991–997, 2002.

[3] Doug A Bowman, Christopher J Rhoton, and Marcio S Pinho. Text input techniques for immersive virtual environments: An empirical comparison. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 46, pages 2154–2158. SAGE Publications, 2002.

[4] Antonino Gomes De Sa and Gabriel Zachmann. Virtual reality as a tool for

verification of assembly and maintenance processes. *Computers & Graphics*, 23(3):389–403, 1999.

[5] John Dudley, Keith Vertanen, and Per Ola Kristensson. The VISAR keyboard: Fast and precise touch-based text entry for head-mounted augmented reality with variable occlusion. *ACM Transactions on Computer-Human Interaction*.

[6] Anna Maria Feit, Srinath Sridhar, Christian Theobalt, and Antti Oulasvirta. Investigating multi-finger gestures for mid-air text entry. In *ACM womENcourage*, 2015.

[7] Laura Freina and Michela Ott. A literature review on immersive virtual reality in education: state of the art and perspectives. In *The International Scientific Conference eLearning and Software for Education*, volume 1, page 133. "Carol I" National Defence University, 2015.

[8] Luciano Gamberini, P Cottone, A Spagnolli, Diego Varotto, and Guiseooe Mantovani. Responding to a fire emergency in a virtual environment: different patterns of action for different situations. *Ergonomics*, 46(8):842–858, 2003.

[9] Jens Grubert, Lukas Witzani, Eyal Ofek, Michel Pahud, Matthias Kranz, and Per Ola Kristensson. Effects of hand representations for typing in virtual reality. *arXiv preprint arXiv:1802.00613*, 2018.

[10] Jens Grubert, Lukas Witzani, Eyal Ofek, Michel Pahud, Matthias Kranz, and

Per Ola Kristensson. Text entry in immersive head-mounted display-based virtual reality using standard keyboards. *arXiv preprint arXiv:1802.00626*, 2018.

[11] Per-Ola Kristensson and Shumin Zhai. SHARK 2: a large vocabulary shorthand writing system for pen-based computers. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 43–52. ACM, 2004.

[12] Ming C Leu, Michael G Hilgers, Sanjeev Agarwal, Richard H Hall, Terry Lambert, Robert Albright, and Kyle Nebel. Training in virtual environments for first responders. *education and training*, 1:10, 2003.

[13] R Bowen Loftin and P Kenney. Training the hubble space telescope flight team. *IEEE Computer Graphics and Applications*, 15(5):31–37, 1995.

[14] Anders Markussen, Mikkel R Jakobsen, and Kasper Hornbæk. Selection-based mid-air text entry on large displays. In *IFIP Conference on Human-Computer Interaction*, pages 401–418. Springer, 2013.

[15] Anders Markussen, Mikkel Rønne Jakobsen, and Kasper Hornbæk. Vulture: a mid-air word-gesture keyboard. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 1073–1082. ACM, 2014.

[16] Frode Eika Sandnes and Andre Aubert. Bimanual text entry using game controllers: Relying on users spatial familiarity with qwerty. *Interacting with Computers*, 19(2):140–150, 2006.

[17] Maria T Schultheis and Albert A Rizzo. The application of virtual reality technology in rehabilitation. *Rehabilitation psychology*, 46(3):296, 2001.

[18] Srinath Sridhar, Anna Maria Feit, Christian Theobalt, and Antti Oulasvirta. Investigating the dexterity of multi-finger input for mid-air text entry. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3643–3652. ACM, 2015.

[19] Khai Truong, Sen Hirano, Gillian R Hayes, and Karyn Moffatt. 2-Thumbs gesture: The design and evaluation of a non-sequential bi-manual gesture based text input for touch tablets. Technical report.

[20] Keith Vertanen and Per Ola Kristensson. A versatile dataset for text entry evaluations based on genuine mobile emails. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, pages 295–298. ACM, 2011.

[21] Keith Vertanen, Haythem Memmi, Justin Emge, Shyam Reyal, and Per Ola Kristensson. VelociTap: Investigating fast mobile text entry using sentence-based decoding of touchscreen keyboard input. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 659–668. ACM, 2015.

[22] James Walker, Bochao Li, Keith Vertanen, and Scott Kuhl. Efficient typing on a

visually occluded physical keyboard. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 5457–5461. ACM, 2017.

[23] Xin Yi, Chun Yu, Mingrui Zhang, Sida Gao, Ke Sun, and Yuanchun Shi. ATK: Enabling ten-finger freehand typing in air based on 3d hand tracking data. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pages 539–548. ACM, 2015.

[24] Chun Yu, Yizheng Gu, Zhican Yang, Xin Yi, Hengliang Luo, and Yuanchun Shi. 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*, pages 4479–4488. ACM, 2017.

[25] Shumin Zhai and Per-Ola Kristensson. Shorthand writing on stylus keyboard. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 97–104. ACM, 2003.

[26] Shumin Zhai and Per Ola Kristensson. The word-gesture keyboard: reimagining keyboard interaction. *Communications of the ACM*, 55(9):91–101, 2012.

Appendix A

Questionnaire

QUESTIONNAIRE

Text Entry in Virtual Environments

Participant number ________ **Order** ____________________

Age (approximate) ________ **Gender** ____________________

I am currently studying at a university:

- ☐ YES
- ☐ NO

If you answered YES, what is your major? ______________________

I am:

- ☐ Left handed
- ☐ Right handed
- ☐ Ambidextrous

I have used a ***virtual reality headset*** (e.g. Oculus Rift, HTC Vive, Samsung Gear VR, Playstation VR) (only check one):

- ☐ Never
- ☐ Occasionally
- ☐ Many times

I most commonly enter text in ***virtual reality*** (only check one):

- ☐ By using a handheld VR controller
- ☐ By typing on a virtual keyboard
- ☐ By typing on a physical keyboard
- ☐ I don't enter text in virtual reality
- ☐ Other: ____

I most commonly enter text on a ***mobile phone*** (only check one):

- ☐ By tapping letters on a virtual keyboard
- ☐ By gesturing words on a virtual keyboard (e.g. Swype, Shapewriter, SwiftKey Flow)
- ☐ By speech recognition
- ☐ I don't enter text on a mobile phone
- ☐ Other: ____

When I enter text on a ***mobile phone***, I most commonly use (only check one):

- ☐ A single finger
- ☐ Two thumbs
- ☐ Two other fingers (but not my thumbs)
- ☐ I don't enter text on a mobile device
- ☐ Other: ____

When I enter text on a ***mobile phone***, I most commonly use (only check one):

- ☐ A standard QWERTY keyboard with tap gestures
- ☐ A standard QWERTY keyboard with swipe gestures
- ☐ Speech recognition
- ☐ I don't enter text on a mobile phone
- ☐ Other: ____

I enter text on a mobile phone ***without looking at the device*** (only check one):

- ☐ Several times per day
- ☐ Once per day
- ☐ A few times per week
- ☐ Once per week or less
- ☐ I always look at the device
- ☐ I don't enter text on a mobile phone
- ☐ Other: ____

I enter text on a mobile phone ***without looking at the device*** most commonly because:

I am familiar with the QWERTY keyboard layout and can locate any key without looking at the keyboard:

- ☐ Yes
- ☐ No

How much do you agree or disagree with the following statements (X a single circle):

I consider myself a fluent speaker of ***English***

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

I frequently enter text on a ***desktop keyboard***

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

I frequently need to look at the keys when I enter text on a ***desktop keyboard***

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

I consider myself a fast typist on a ***desktop keyboard***

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

I frequently enter text on a ***mobile phone***

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

I frequently need to look at the keys when I enter text on a ***mobile phone***

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

I consider myself a fast typist on a ***mobile phone***

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

Condition ________

The questions on this page refer to the experimental condition you just completed.

How much do you agree or disagree with the following statements (X a single circle):

In this part of the study, I thought I entered text ***quickly*** using the experimental interface

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

In this part of the study, I thought I entered text ***accurately*** using the experimental interface

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

In this part of the study, I thought the experimental interface provided ***accurate visual feedback of my hand(s)***

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

In this part of the study, I thought the experimental interface ***detected a key press accurately when I typed a key***

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

In the space below, please share any comments you have about this part of the study:

Perceived Exertion

Condition:

Instruction for Perceived Exertion Using the Borg CR10 Scale

We want you to rate your perception of exertion, that is, how heavy and strenuous the exercise feels to you. The perception of exertion depends mainly on the strain and fatigue in your muscles and on your feeling of breathlessness or aches in the chest.

We want you to use this scale from 0 to 10 and "•", where 0 means "no exertion at all" and 10 means "extremely strong—max P", that is, the maximal exertion you have previously experienced.

1 corresponds to "very light" exercise. For a normal, healthy person it is like walking slowly at his or her own pace for several minutes.

3 on the scale is "moderate" exercise, it is not especially hard, it feels fine, and it is no problem to continue exercising.

5 corresponds to "heavy" exercise; it feels hard and you are tired, but you don't have any great difficulties in going on.

7 is "very hard" and very strenuous. A healthy person can still go on but he or she has to push him- or herself a lot. It feels very heavy and the person is very tired.

10 on the scale is an extremely strenuous exercise level. It is "max P." For most people this is an exercise as strenuous as they have ever experienced before in their lives.

• The dot denotes a perceived exertion that is stronger than 10, "extremely strong." It is your "absolute maximum," for example, 12, 13, or even higher. It is the highest possible level of exertion.

Try to appraise your feeling of exertion as honestly as possible, without thinking about what the actual physical load is. Don't underestimate it, but don't overestimate it either. It's your own feeling of effort and exertion that's important, not how it compares to other people's. What other people think is not important either. Look at the scale and the expressions and then give a number.

What "max exertion"—your "max P"—have you previously experienced in your life? Use that as "10".

Any further questions?

Please rate your perceived exertion level:

Please write the name of conditions in your order of preference from most preferred to least preferred.

UNIMANUAL – using only the index finger of your dominant hand

BIMANUAL – using the index fingers of both hands

SPLIT – using the index fingers of both hands to type on a split keyboard

(a) Preferred condition in terms of quickness:

Most preferred condition (1) (2) (3) Least preferred condition	**Name of the conditions:** UNIMANUAL BIMANUAL SPLIT

(b) Preferred condition in terms of accuracy:

Most preferred condition (1) (2) (3) Least preferred condition	**Name of the conditions:** UNIMANUAL BIMANUAL SPLIT

(c) Preferred condition in terms of minimal physical effort:

Most preferred condition (1) (2) (3) Least preferred condition	**Name of the conditions:** UNIMANUAL BIMANUAL SPLIT

(d) Overall Preferred condition:

Most preferred condition (1) (2) (3) Least preferred condition	**Name of the conditions:** UNIMANUAL BIMANUAL SPLIT

Please share any comments you have about the study in general:

Condition: INVISIBLE

The questions on this page refer to the experimental condition you just completed.

How much do you agree or disagree with the following statements (X a single circle):

In this part of the study, I was able to successfully obtain ***the keyboard size*** I wanted

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

In this part of the study, I found it easy to enter text ***without any visual feedback*** of the letter keys

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

In this part of the study, I was able to ***easily understand when and what key*** I had typed

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

In this part of the study, I thought I entered text ***quickly*** using the experimental interface

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

In this part of the study, I thought I entered text ***accurately*** using the experimental interface

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

In this part of the study, I thought the experimental interface provided ***accurate visual feedback of my hand(s)***

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

In this part of the study, I thought the experimental interface ***detected a key press accurately when I typed a key***

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

In the space below, please share any comments you have about this part of the study: