Title	Tap Tip Toe : supportive input method based on foot interaction
Sub Title	
Author	Wu, Ang Kunze, Kai
Publisher	慶應義塾大学大学院メディアデザイン研究科
Publication year	2018
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2018年度メディアデザイン学 第672号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002018-0672

Master's Thesis
Academic Year 2018

TapTipToe:Supportive Input Method based on Foot Interaction

Keio University
Graduate School of Media Design

Ang Wu

A Master's Thesis
submitted to Keio University Graduate School of Media Design
in partial fulfillment of the requirements for the degree of
Master of Media Design

Ang Wu

Master's Thesis Advisory Committee:

Associate Professor Kai Kunze	(Main Research Supervisor)
Senior Assistant Professor Roshan Peiris	(Sub Research Supervisor)

Master's Thesis Review Committee:

Associate Professor Kai Kunze	(Chair)
Senior Assistant Professor Roshan Peiris	(Co-Reviewer)
Professor Ichiya Nakamura	(Co-Reviewer)

Abstract of Master's Thesis of Academic Year 2018

TapTipToe:Supportive Input Method based on Foot Interaction

Category: Science / Engineering

Summary

With the development of computer technology, Humans have never stopped exploring human-computer interaction. With the migration of PC to the mobile side, human-computer interaction has also changed, and traditional human-machine interfaces such as keyboards and mice can no longer support the current complex interactive situations. Researchers have done much research on human-machine interaction and put forward such as eye tracking interaction technology, voice recognition interaction technology, etc., although they widen the idea of human-computer interaction, there are some limitations. Foot interaction is a beneficial alternative interaction method when hands are pre-occupied.In the real world, feet are often used with hands for many tasks. However, in a computer environment, this form of interaction is often overlooked. Based on the natural human-computer interaction concept, In an AR/VR environment，feet are ideal as a supportive input method. In this paper, the usability of the wearable input device based on foot interaction is studied. I present a set of foot wearable devices based on proximity sensors, with the aim of exploring how foot interaction can be better adapted to the current computer environment(especially in the AR/VR environment). The possibility of using feet as a supportive input method in different computer environments is discussed. I compared the foot input method with the traditional input method, analyzed its performance and usability, and pointed out its defects.

Keywords:

Foot Interaction, Wearable Device, Input Method, Text Entry

Keio University Graduate School of Media Design

Ang Wu

Contents

Acknowledgements **vii**

1 Introduction **1**

1.1. Introduction 1

1.2. Foot Interaction 2

1.3. Research Contribution 4

1.4. Structure of the Thesis 5

2 Related Works **6**

2.1. Foot Interactions 6

2.2. Foot-Based System 7

2.2.1 Summary 10

3 “TapTipToe” System **11**

3.1. Concept 11

3.1.1 Foot Gestures 12

3.1.2 Multi-Modality 14

3.1.3 Fatigue of foot interaction 14

3.1.4 Design principles 15

3.2. Detection 16

3.3. Prototyping 16

3.4. Application scenario 20

3.4.1 As the main input interface 23

3.4.2 As an additional input method 23

4 Research on Usability of Foot Interaction **25**

4.1. Usability Experiment 25

4.1.1 Goal 27

4.1.2 Participant 27
4.1.3 Average lift height of Toe tapping and Heel tapping 27
4.1.4 Experimental design 27
4.1.5 Result 30
4.2. Text entry based on foot interaction 30
4.2.1 Goal 30
4.2.2 4-key text entry method based on foot interaction 31
4.2.3 Text entry assessment and model 32
4.2.4 Experimental design 33
4.2.5 Result 34

5 Conclusion **39**
5.1. summary 39
5.2. limitation 40
5.3. Future work 41

References **43**

List of Figures

1.1 Common foot interaction system in daily life 3

2.1 Foot Joysticks [1] 9
2.2 Foot text input padel [2] 9
2.3 Foot input system for standing desk [3] 10

3.1 Focus group mind map 11
3.2 Foot gestures summed up by Velloso [4] 12
3.3 Photograph of proximity sensor LBR-127HLD 16
3.4 Photograph of TapTipToe System Slipper module（Sensors） . . 17
3.5 Photograph of TapTipToe System Slipper module 18
3.6 Photograph of TapTipToe System Main Controller 19
3.7 Schematic diagram of the circuit of TapTipToe System(Master Controller) 20
3.8 Schematic diagram of the circuit of TapTipToe System(Slave Controller) 21
3.9 Schematic diagram of TapTipToe system 21
3.10 Photograph of TapTipToe prototype 22
3.11 Photograph of wearing a TapTipToe prototype 24

4.1 Photograph of toe tapping gesture 26
4.2 Photograph of usability experiment 28
4.3 Photograph of 4-key text entry program 31
4.4 Common text entry devices in daily life 33
4.5 Log-log plots of learning rates for TapTipToe for each participant 35
4.6 Learning Rates of TapTipToe and Game controller 35
4.7 Contrast diagram of Word per minute for text input devices . . . 36
4.8 Photograph of text entry test in VR environment 38

List of Tables

3.1 Explanation and function of foot gestures [4] 13

4.1 Date of text entry 36

Acknowledgements

My deepest gratitude goes first and foremost to Associate Professor Kai Kunze, my supervisor, for his constant encouragement and guidance. Without his patience, enthusiasm and enlightening guidance, this thesis could not have reached its present form.

I would also like to thank Project Senior Assistant Professor Roshan Peiris for giving the technical guidance and advice to my prototype. Also, I would like to thank Professor Ichiya Nakamura for his guidance for this thesis.

I would also like to thank Hanlin Zhang, who offered me advice when I was frustrated. I thank Takuro Nakao, who gave me encouragement and support to share his knowledge with me.

Next, I would like to thank all the members of Geist and all the participants in my study for their help with this research.

Finally, a special thanks to my family. Thanks for their understanding, support, encouragement, and sacrifice over the years.

Chapter 1
Introduction

1.1. Introduction

The core of human-computer interaction research is how to transfer information to the computer easily. It can be said that the rapid development and popularization of computer industry benefits from the study of human-computer interaction technology. Since the advent of the first computer, human-computer interaction has experienced three significant stages of evolution: keyboard, mouse, and multi-touch technology. Mouse in the graphical interface of the application of more humane than the manual, it is the ancestor of natural human-computer interaction, now it has become the standard configuration of the computer. The popularity of multi-touch technology has upended traditional methods of human-computer interaction and led to a new gesture-based interactive experience. Until now, the human has never stopped exploring more natural ways of human-computer interaction. motion-sensing technology uses techniques such as computer vision technology to allow people to use body language for input. Before learning grammar and writing, humans have used body language to communicate with each other. The most representative of these is gesture recognition technology, which usability has been proven, and gesture recognition products and solutions have begun to appear in the market. The gradual maturation of big data technology and artificial intelligence technology has brought a brand-new opportunity to voice recognition technology, and the critical factor that is not thoroughly popularized is that the ability of voice recognition technology is to be improved, and the recognition in multiple contexts has yet to be perfected. Eye tracking technology has been around for quite a long time and has been widely used in the field of psychology and neuroscience. With the advent of wearable devices(exceptionally smart eyewear), eye tracking technology is used in the human-computer interaction of wearable devices.

In order to accelerate the popularization of virtual reality, smart home, and the internet of things, it is urgent to improve the usability of man-machine interaction technology. Although there have been many attempts at new methods of interaction, most of the interactions have limitations on the scope of use, low usage, and no real commercial application to popularity. For example, voice recognition technology, environmental noise, and interference will lead to low recognition rate of voice recognition. motion-sensing technology can only be confined to the entertainment field at present. Motion-sensing is only used in some professional fields and entertainment fields. Including eye-tracking technology, researchers have made many attempts, but at present still stay in some professional research institutions or laboratory research. Human-computer interaction technology in the field of entertainment and implementation in the field of life is different levels, in life and work cannot be efficient and accurate, there are many technical needs to be solved. It is difficult to be productive and accurate in life and work, to make them become the reality we still need some enabling technologies.

We use all the senses and limbs to interact with the natural environment. In a computer environment, we try the multimodal input method. There are limitations to a single input method, multiple input methods can increase usability, can provide additional input channels, and the advantages of another input method can remedy the disadvantage of one input method. In a voice recognition environment of a multiplayer scene, the computer vision technology is used for lip recognition to separate the voice instruction and judge the source of the instruction. The multimodal interaction can break the bottleneck of the development of the human-computer interaction technology.

1.2. Foot Interaction

Since humans learned to use tools, we have used our feet to interacting with tools, such as shovels. In the modern living environment, we often use our feet to interact with the device. When driving a car, we tap the throttle and brakes with feet. When riding a bike, we use the foot pedal as a motivator. When playing the piano, we use the foot pedal to switch the sound effects. However, foot interaction is not frequent in a computer environment. In most cases, tasks that can be done through foot interaction can be replaced by hand interaction, but this does not mean that the flexibility of the foot is not as good as the hand, because in daily life we use the foot to complete some important tasks.

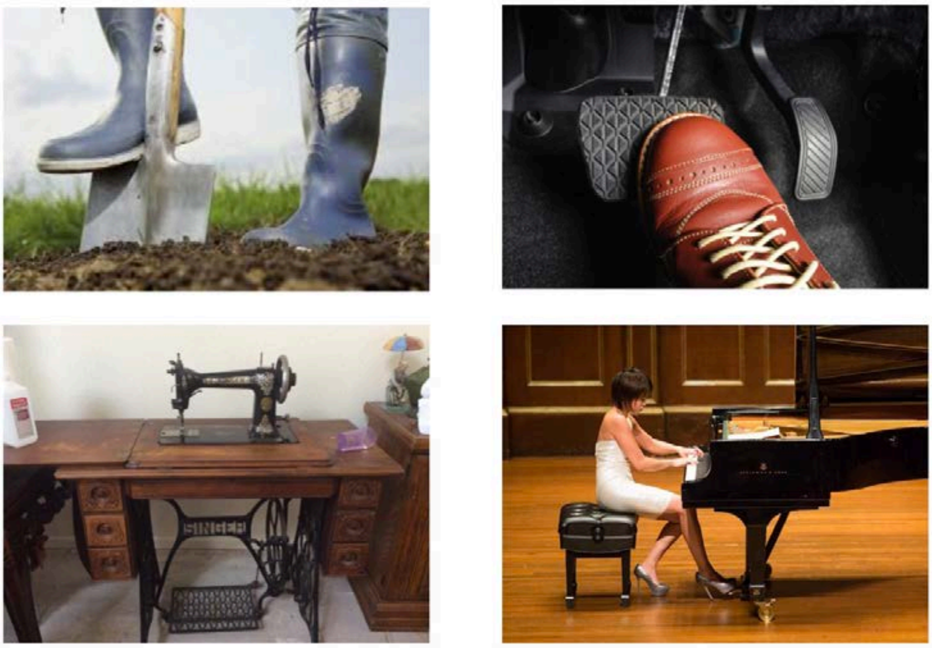

Figure 1.1 Common foot interaction system in daily life

Although there are few foot interaction devices on the market, foot-based interactions are still not universal, and researchers have not given up on the study of foot interactions. Foot interaction can be used as an alternative input mechanism for hands, and to achieve hands-free and eyes-free human-computer interaction. They can be used as the primary input method and can also be used as additional input mechanism in complex tasks. Pakkanen [5] suggest that people choose the input method not only because of its efficiency, the user's subjective satisfaction strongly affects the user's preferred input method. In the study based on Fitts ' law [6], it has been proved that feet can be used as an alternative to hands, but hands are better suited to the accurate task than feet. They believe that foot as an auxiliary input mechanism to handle secondary tasks can significantly reduce disruptions and increase the workflow substantially. We want to explore unconstrained foot interactions that are not limited by physical devices and usage environments (such as foot pedals, foot mice), providing a more natural interaction method. It is not intended to replace the traditional devices, but rather to combine with the conventional method to enhance the experience of existing interaction. We envision this approach to interact in more scenarios and environments, including VR/AR environments, mobile devices, desktop applications, and more.

To achieve this idea, we conducted a series of experiments. First, I formed a focus group of a total of 9 participants to explore foot movements suitable for foot interaction, and which tasks are suitable for the foot interaction. We recorded the foot movements of the nine participants. Based on the feedback from the participants, we envisioned toe tap and heel tap are more suitable for foot interaction. To confirm this view, I created a prototype of a wearable device based on foot interaction and conducted the second experiment to verify the performance of the device. In the third experiment, we compared it with the traditional text entry method. In the fourth experiment, we performed a simulation of walking in VR environment.

1.3. Research Contribution

The contribution points of this study are as follows:

·Provide a set of wearable devices based on foot interaction as an auxiliary input mechanism for the computer environment.

·Provide text input system based on foot interaction wearable device.

·The usability of the foot interactive wearable system as an auxiliary input mechanism in the VR environment is verified, and its performance with the traditional joystick controller in VR is evaluated.

1.4. Structure of the Thesis

Chapter 1 is the introduction to this paper, which introduces the concept of foot interaction and explains the contribution of the study.

Chapter 2 describes the past research in the field of foot interaction, including the method of foot interaction and the system based on foot interaction.

Chapter 3 describes the concept of wearable device ”TapTipToe” based on foot interaction and outlines its system design and the technology used in it.

Chapter 4 describes in detail the two experiments we conducted, analyzed the performance of the system, and assessed its usability.

Chapter 5 presents the summary and conclusion of this thesis.

Chapter 2
Related Works

My inspiration for foot interaction comes from “shaking legs.” Some people have the habit of shaking their legs when they are thinking and sit in a chair, which may be due to concentration. Studies have shown that brain regions that control cognition and exercise overlap with each other, suggesting that people can concentrate when doing repetitive body movement. So can the energy released by this high-frequency motion be used in computer interaction? When a computer worker sits at a table and uses a computer, his feet are resting, while in some similar jobs, feet are used to interact with the device, such as driving, playing the piano, and using a sewing machine. The legs are flexible limbs, and we can also use leg activity while sitting and standing. Some of these modes of motion have been used to interact with the devices, while others are not suitable for foot interaction. In this chapter, we survey realized foot interaction studies and devices based on foot interaction.

2.1. Foot Interactions

Crossan [7] use acceleration sensors to detect the foot taps to interact on mobile devices while the phone remains in the pocket. They choose by tapping the Toe to move a menu item, allowing users to use mobile devices without visual feedback. They point out that this method has high accuracy and is beneficial if the taps are less than five times.

Paelke [8] use the mobile device's camera to detect the movement and the position of feet to interact. They use computer vision technology to analyze the video of kicking action to interact with interactive objects. They said the overall feedback from users is very positive, and this interaction based on computer vision technology is intuitive, and this is more interesting than the traditional button click interaction.

Felberbaum [9]believe that it is necessary to differentiate between the terms foot interaction and foot gesture. They said foot interaction could cover all of the foot-to-device interactions, while foot gestures refer to predefined foot movements that can trigger a particular operation. They defined three gesture sets for three scenarios, each corresponding to standing in front of a large display, sitting in front of a desktop display and standing on a projected surface. At the same time, they suggest a measurement method called Specification score, which can help us understand the specific, preferable and intuitive extent of gestures to action in particular use condition.

Velloso [4] have done a comprehensive survey of the study of foot-based interaction. Their investigation played a crucial role in the study of foot-based interaction. They analyzed the movement of all foot joints and linked them to the corresponding interactions, and the effect of the user on the interaction mode under different postures. At the same time, they investigated the detection and feedback methods of foot-operated devices and classified the interactive sensing methods of the feet. They divide the foot-based interactions into four categories: semaphoric, deictic, manipulative, and implicit. ·Semaphoric-This is one of the most common and essential foot interaction. Such as toe tapping. It is similar to the finger touch with low-effort properties. One variant of it is heel tap. The disadvantage of heel tap is that the weight that needs to carry the legs tends to cause fatigue. ·Deictic-Commonly understood as pointing gestures. ·Manipulative-Performed to change an object's properties. ·Implicit-Its main feature is to track user information and obtain user data. Such as smart footwear. They point out the two purposes of the foot-based interaction in explicit interaction. One is as the primary control, and the other is as a supporting control for other devices.

2.2. Foot-Based System

Also from the survey of Velloso [4], they were divided into three categories according to how the device detects input from the foot: mediated, intrinsic and extrinsic sensing. ·mediated sensing-Instead of directly detecting the foot, it recognizes the device controlled by the foot. ·Intrinsi sensing-They are usually wearable devices that are detected directly from sensors connected to the foot. ·Extrinsic sensing-detect user's feet from a sensor in the environment. Among them, depth camera based on computer vision sensor is the most common. The most widely used foot-operated device is the pedal. The traditional pedal has been used as a force transfer device up to now, but its working principle is different from that of electronic devices. The most common ones are bicycles, sewing machines, and waterwheels. The traditional automobile throttle is connected to the pedal through the throttle cables, and its transmission ratio is 1:1. The electronic throttle it is through the wire to control the opening of the throttle, from the surface, a wire replaces the traditional throttle cable, but in essence, it is not just a simple way to change the connection, but to achieve automatic control of the power output of the entire vehicle.

Göbel [1] proposed a set of multimodal input devices based on gaze and foot interaction to control the scaling of the map using implicit gaze input controlled by the explicit pedal. They proposed three prototypes: a three-pedal device, the foot-joystick based on two-axes foot-based tilting, the foot-rocker based on a two-directional foot pedal. They believe that multi-mode input devices based on gaze and foot interaction to support secondary tasks are very promising. They said foot interaction is ideal for parallel input, which allows precise control of scaling speed.

Dearman [2] have proposed a text input method for mobile devices based on pedals, but the results are not satisfactory. Although this method is faster, its higher error rate reduces the throughput of the text.

Saunders [3] introduced an interactive standing system for controlling traditional desktop applications based on foot interaction. They aim to increase the standing time of computer workers through standing interactive devices. They use depth camera and instrumented shoes based on several foot interactions for indirect and discrete inputs. Demonstrates the desktop application task and evaluates the usability of this method. They believe that the efficiency of foot input cannot be compared with the mouse and keyboard, However, as an auxiliary input channel, you can enhance the mouse and keyboard, perform secondary tasks without interrupting the primary task.

Figure 2.1 Foot Joysticks [1]

Figure 2.2 Foot text input padel [2]

Figure 2.3 Foot input system for standing desk [3]

2.2.1 Summary

Although foot interaction is not common in human-computer interaction, the researchers have studied it in depth. Velloso [4] made a systematic analysis of the related research on foot interaction. They define the classification of foot interaction and foot-operated devices, which provides great help to the study of follow-up researchers. In recent years, the number of research based on foot interaction shows an increasing trend. Some of the research results are not positive for foot interaction, but most of these results indicate that foot interaction is feasible as a secondary input channel to assist traditional interfaces. This provides an important theoretical basis for my research.

Chapter 3
“ TapTipToe ” System

In this chapter, I will introduce how we build a foot interaction system. First, we conducted a focus group discussion on how to use foot actions in human-computer interaction. Through experiments, the principles of foot action in human-computer interaction are summarized. The method of the foot input signals and mappings is explored. Based on these principles, we used the proximity sensor to built the prototype.

3.1. Concept

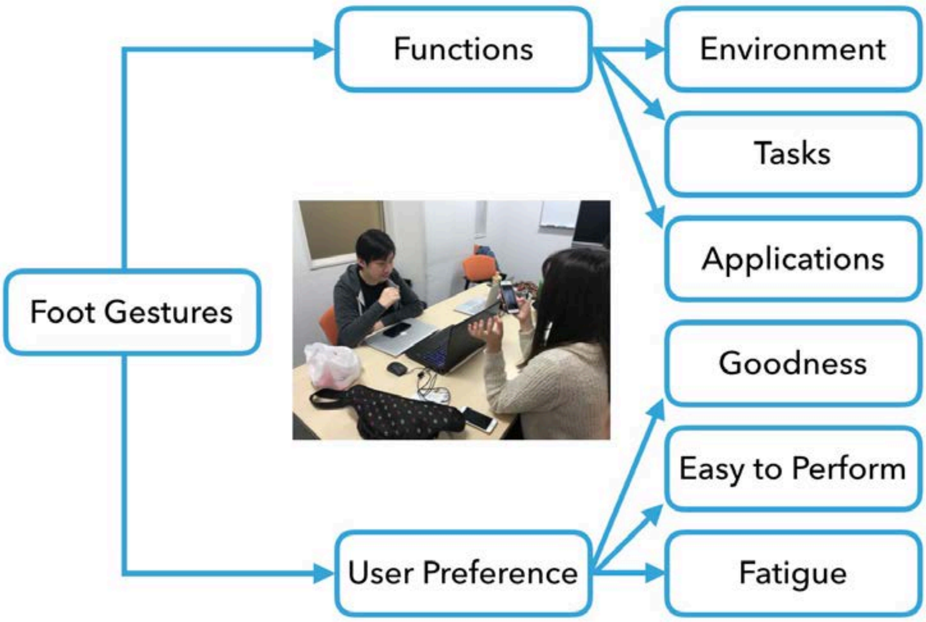

Figure 3.1 Focus group mind map

3.1.1 Foot Gestures

In the survey by velloso et al., the foot interactions were divided into four types. In this thesis, we mainly discuss Semaphoric actions. The advantage of this type of gesture is that in addition to discrete information, and these gestures also can carry continuous information. Such as simple switches or speed control. We discussed the 11-foot gestures they summarized. We showed participants 11-foot gestures and let them match the function of the gesture to understand what the appropriate mapping between gestures and functions is. Participants make a foot gesture and determined its execution command, and fills in the questionnaire accordingly. We conducted a short interview with participants based on the questionnaire and got feedback from participants to understand the foot gestures and mappings. In this experiment, we mainly study three factors. The goodness of the match, ease of performing, and fatigue. Based on these three criteria, the most appropriate foot gestures and mappings are selected. Besides, we also get feedback on the proper tasks and application scenarios from the questionnaire.

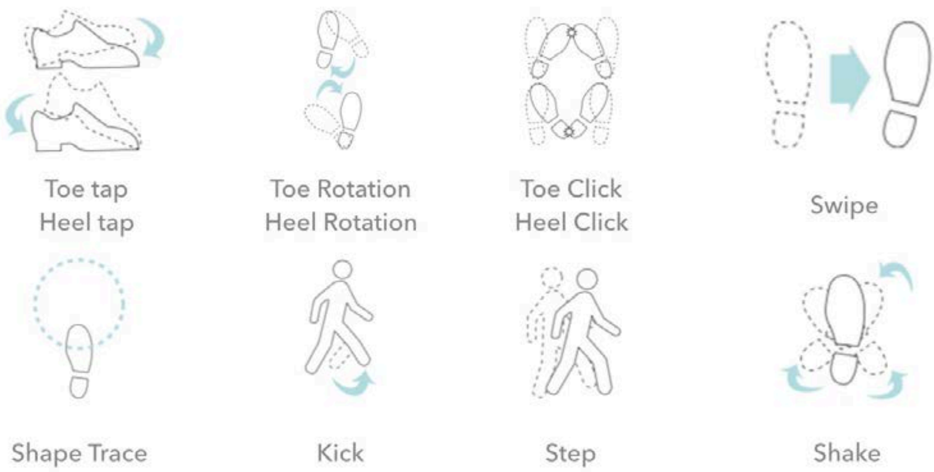

Figure 3.2 Foot gestures summed up by Velloso [4]

Table 3.1 Explanation and function of foot gestures [4]

Gestures	Description	Function
Toe tap	User raises and lowers the toes	touch click
Heel tap	User raises and lowers the heel	touch click
Toe rotation	User pivots the foot around the toes	drag pointing
Heel rotation	User pivots the foot around the heel	drag pointing
Toe click	User touches both toes together	click select
Heel click	User touches both heel together	click select
Swipe	User slides the foot in a certain direction	drag
Shake	User moves the foot with short, quick, irregular vibratory movements	select
Shape Trace	User draws the outline of a shape with the toes	drag pointing
Kick	Vigorous movement of the foot in a certain direction	touch click
Step	User puts one foot in front of the other as if walking pointing	click

3.1.2 Multi-Modality

The feet are good at performing simple tasks that are as important as the tasks performed by the hand, such as car braking. It is also able to perform complex tasks, such as playing musical instruments. However, the foot interaction is not taken seriously. In the study based on Fitts's law, it can be known that the flexibility of the foot is not as good as the hand. Researchers have compared their hands and feet with much research. As a result, the performance of the hand is always better than the foot, but this does not mean that the foot is not qualified for the job. Most users rarely use their feet to perform tasks, which can explain the large performance gap between hands and feet due to lack of training. Furthermore, there is not much research on the use of hands and feet at the same time, but the researchers believe that the foot can be used as a supplement to the hand, providing additional input method. In past studies, the Multiple gestures of foot interaction were not common. Most of the existing foot interaction systems use a single foot gesture. The foot interaction can be used as the primary interface control application and as a secondary interface, as a supplement to the hand. As the primary interface, Multiple gestures can provide more mapping. As a secondary interface, too many foot gestures will distract attention.

3.1.3 Fatigue of foot interaction

There is not much research on the fatigue of foot interaction, but from the experience of daily life, we can know that foot interaction can be used for long-term tasks, such as driving a car and riding a bicycle. From the perspective of human body structure, the weight of the lower limbs is much higher than that of the upper arms, and it is undoubted that the interaction fatigue of the feet is higher than that of the hands. Similar to hand-based interactions, due to the weight of the lower limbs, the ankle-driven actions are less fatigued than the hip joint and knee-actuated actions in the foot interaction. For the heel gesture, the user is required to bear the weight of the entire leg. In a pedal-based interface, the user places the foot on the pedal to minimize fatigue without activation, but this can easily lead to accidental activation. From the perspective of the use posture, the user usually adopts a sitting posture when interacting with the desktop system. In the study by Saunders et al., they used the standing foot interaction to enhance the user's body movements to achieve the effect of exercise.

3.1.4 Design principles

Based on the study of foot gestures, we summarize three design principles.

(1) Ease of performing - First, for most current computing environments, users typically take a sitting position. In this case, the foot movement will be constrained, and the desktop will block the user's field of view, and no visual feedback will be obtained. The movement of the foot is limited and may be disturbed by the wire under the table. Secondly, long -time foot movements increase fatigue and are not conducive to long-term operation.

(2) ease of detecting - For the detection of foot gestures, it should be wearable. Based on the concept of natural interaction, users should not be distracted by the location of the device. In the VR/AR environment, the user's location is continuously updated, and the detection device needs to allow the user to move freely.

(3) ease of learning - Foot interaction is a rare interaction method for users, and it is difficult to achieve the desired accuracy in the absence of practice. Use simple foot gestures to reduce learning costs with intuitive mapping, allowing users to learn how to use the interface quickly.

Based on these three design concepts, our overall design goal is to allow users to move freely and minimize the cost of learning. We chose Toe tapping and heel tapping as basic foot gestures for the prototype. The advantage is that it is similar to a finger touch, enabling the most basic and most commonly used interactive functions. On the other hand, its power consumption is relatively low in the foot gestures, which is more suitable for high frequency and long-term use. Concerning culture, Toe tapping and heel tapping are interesting foot movements. For music, we use Toe tapping to stabilize the beat. For dance, it is one of the most basic moves of tap dancing. I believe that the rhythm of the tap can enhance the interactive experience.

3.2. Detection

There are many methods to detect foot gestures. Among them, the most widely used and highest accuracies are the foot pedal and the depth camera, but they are not in line with the design principle of this experiment. Initially, we considered using a pressure sensor as a detection method. Initially, we thought using a pressure sensor as a detection method. After the trial, we found that the weight of the lower limbs and the daily activity increased the "noise" of the detection, which reduced the accuracy and negatively affected the user experience. After consideration, we decided to explore a new detection method - using proximity sensors for detection. The distance sensor can effectively and accurately sense the distance between the sole and the ground. It has a small, low-cost feature that fits our requirements for wearable devices.

Figure 3.3 Photograph of proximity sensor LBR-127HLD

3.3. Prototyping

Due to the above design principles, the prototype was designed as a modular wearable device. The distance sensor is placed in a pair of slippers that are cut to fit the shape of the user's sole. For ease of movement and easy adjustment of the sensor position, the slipper is divided into two parts, the toe, and the heel, while the two distance sensors are placed in the heel and toe modules respectively. The slipper is fastened to the 25 ± 5CM size sole with a strap. The sensors are

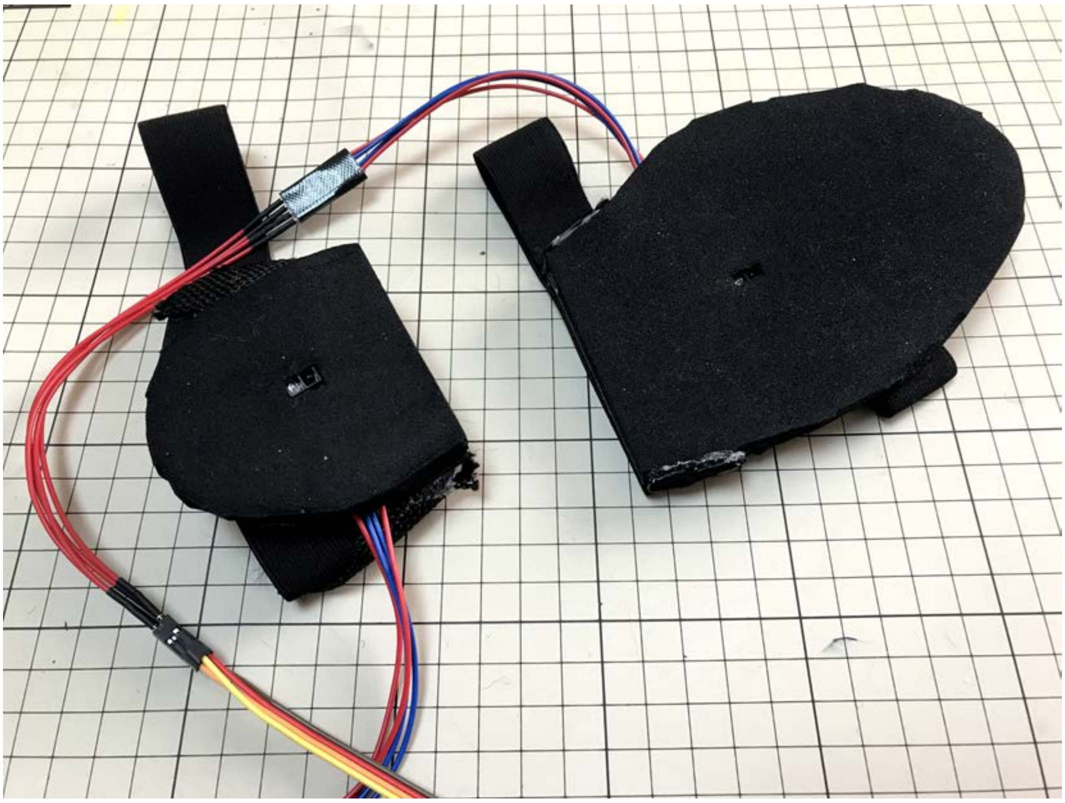

Figure 3.4 Photograph of TapTipToe System Slipper module（Sensors）

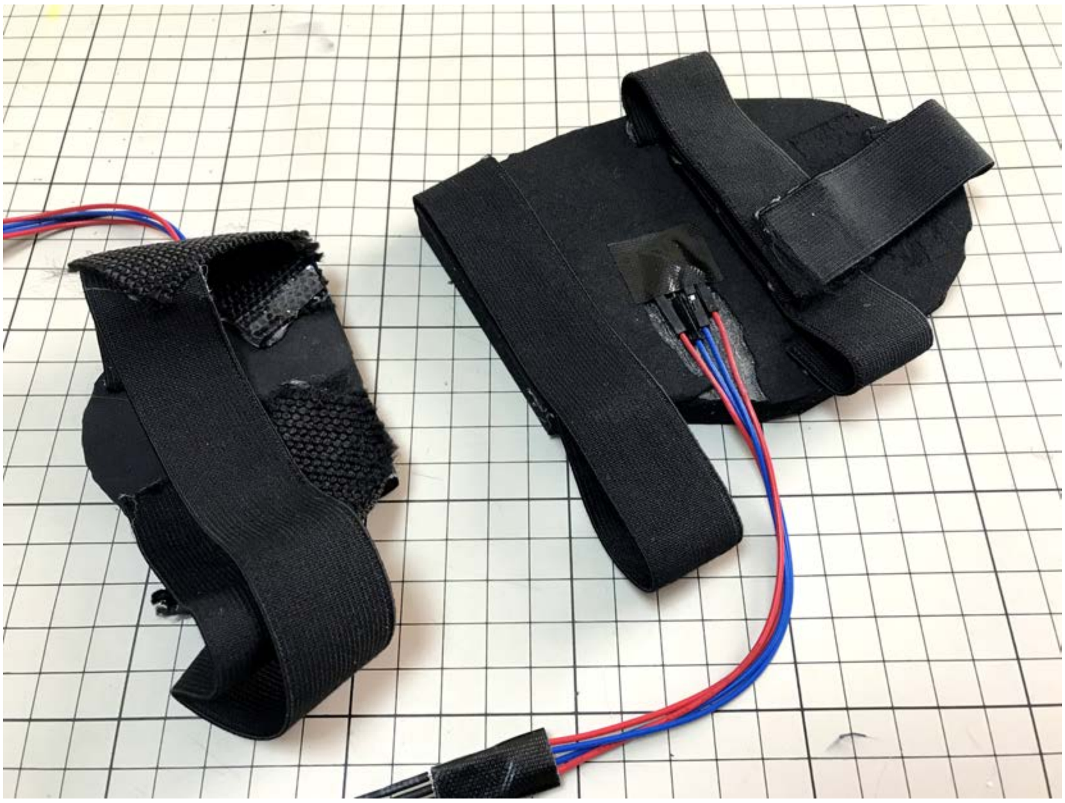

Figure 3.5 Photograph of TapTipToe System Slipper module

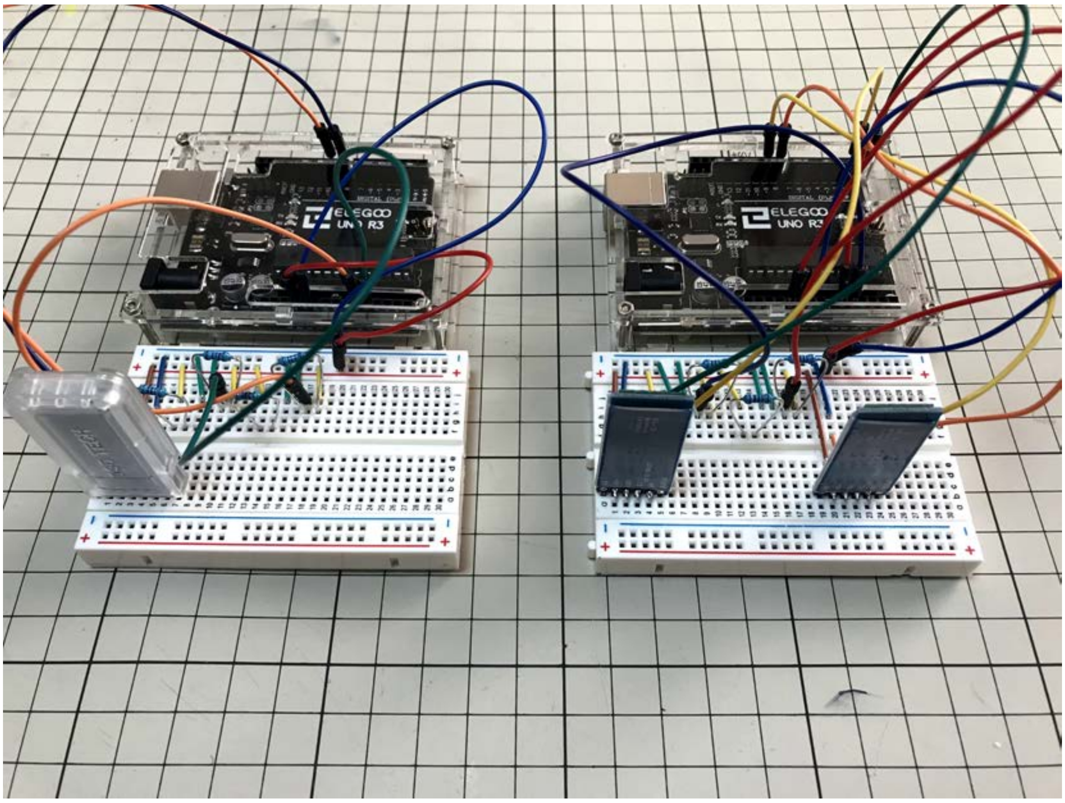

Figure 3.6 Photograph of TapTipToe System Main Controller

connected to 2 Arduino control modules with wires, and the analog signal is transmitted to the computer by Bluetooth module. The four distance sensors located at the sole can respectively sense four sets of analog signals, which can form the most basic four buttons.

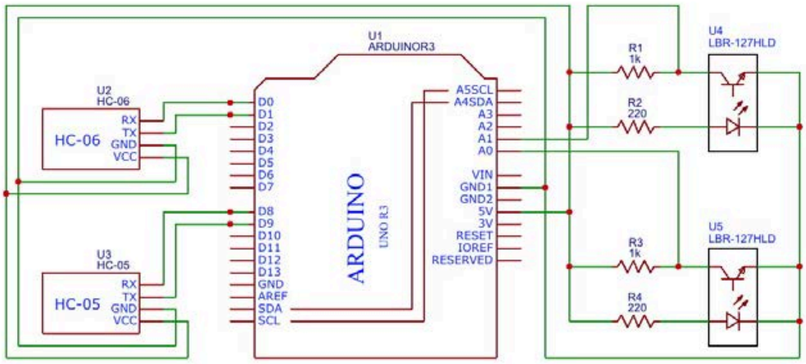

Figure 3.7 Schematic diagram of the circuit of TapTipToe System(Master Controller)

3.4. Application scenario

According to the characteristics of the distance sensor, the TapTipToe system supports the detection of 4 sets of 5-40mm analog signals, which brings us great possibilities. Based on the characteristics of the foot gesture, it is possible to arrange multiple combinations of keys by discrete signals. Through continuous analog signals, we can simulate driving and walking. In addition, it not only provides an optional input method when both hands are occupied but also serves as an accessible input method. Based on these characteristics, we have proposed some possible application scenarios.

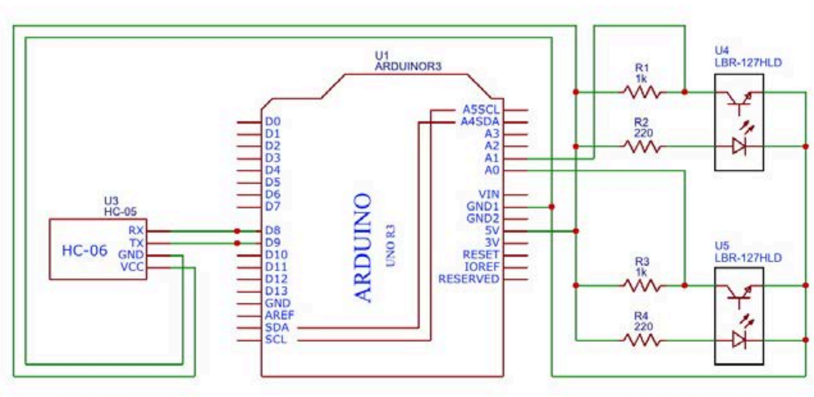

Figure 3.8 Schematic diagram of the circuit of TapTipToe System(Slave Controller)

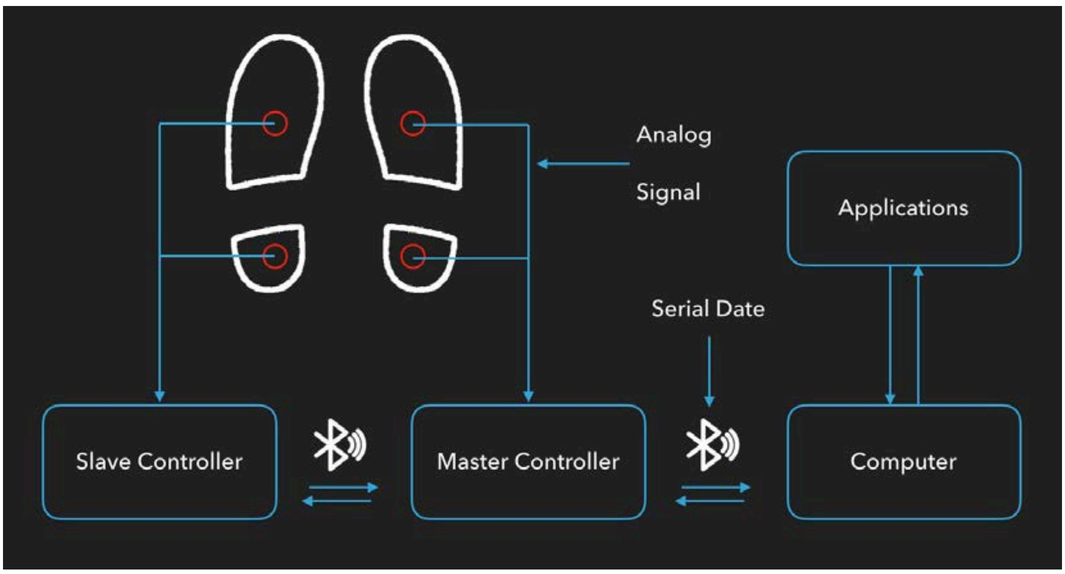

Figure 3.9 Schematic diagram of TapTipToe system

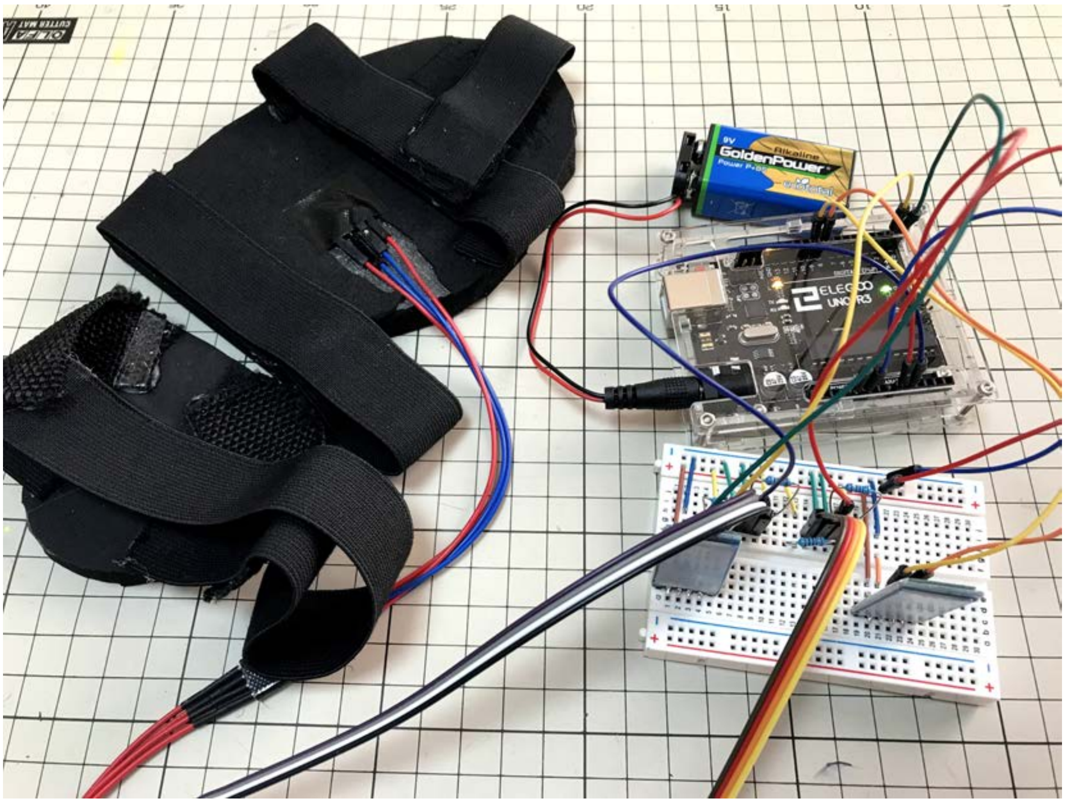

Figure 3.10 Photograph of TapTipToe prototype

3.4.1 As the main input interface

In our daily lives, in some individual cases, our hands cannot interact with the computer. Taptiptoe provides an optional input method when the device cannot be reached by hand and is pre-empted or soiled. For example, when performing maintenance work, both hands are occupied by the tool. We can interrogate documents through the foot and perform tasks such as scrolling and zooming. When painting, when the brush holds both hands and stained by the ink, the foot gesture is used to interact with the mobile device to input text to achieve the operation of replying message and mail. Also, TapTipToe can be used as an accessible solution. For physically challenged people foot interactions can be an effective alternative to hand interaction.

3.4.2 As an additional input method

When we use computers on a daily basis, we often need to deal with some secondary tasks which occupancy our main input method, interrupting our main tasks and causing inefficiency. Alternatively, some complex tasks that cannot be handled even with both hands. We can provide an additional input method through foot interaction to improve the efficiency of the main task execution. For example, scrolling documents while typing text, or offer extra buttons to game operations. In a VR/AR environment, we typically use the gamepad as the primary interface to interact with the virtual world. However, the function of the game controller is limited, and it is impossible to interact naturally. We even need to use ”hands” to walk in the VR environment, which has damage to the experience of the VR environment. In this case, we can use the foot interaction to simulate the two feet, to achieve walking or other foot movements, also, to provide more mapping, increase user productivity and enhance the user experience.

Figure 3.11 Photograph of wearing a TapTipToe prototype

Chapter 4
Research on Usability of Foot Interaction

There are three main kinds of research I have conducted. The first is a usability experiment on the Taptiptoe system based on foot interaction. The second is the experiment of assisting the traditional gamepad when the foot interaction is used as a secondary input method. The third is an experiment of text entry when used as the primary input interface. These experiments have two main purposes: to examine the usability of foot gestures as an interactive method and to analyze the performance of foot interactions in various environments.

4.1. Usability Experiment

As a new interaction method, the foot interface is very different from the traditional interface. Based on Fitts's law predictions, the efficiency of foot interaction is low compared to conventional hand interaction. Especially for our prototype, from the working principle of the button, it is entirely different from the keyboard button. The keystroke distance of the traditional keyboard is 1/2 of the full keystroke. For TapTipToe, the process of lifting and lowering the proximity sensor from the ground is a complete button flow, which is determined as an entire input flow at the moment of contact with the ground. Due to individual differences in humans, the angle of lift and reaction time is much different when doing Toe tapping and heel tapping. Before making an availability assessment for this device, we need to define the lift height. On this basis, we will test the average button time of Toe tapping and compare it with the KLM of the traditional keyboard to evaluate the usability of the TapTipToe system based on the foot interaction.

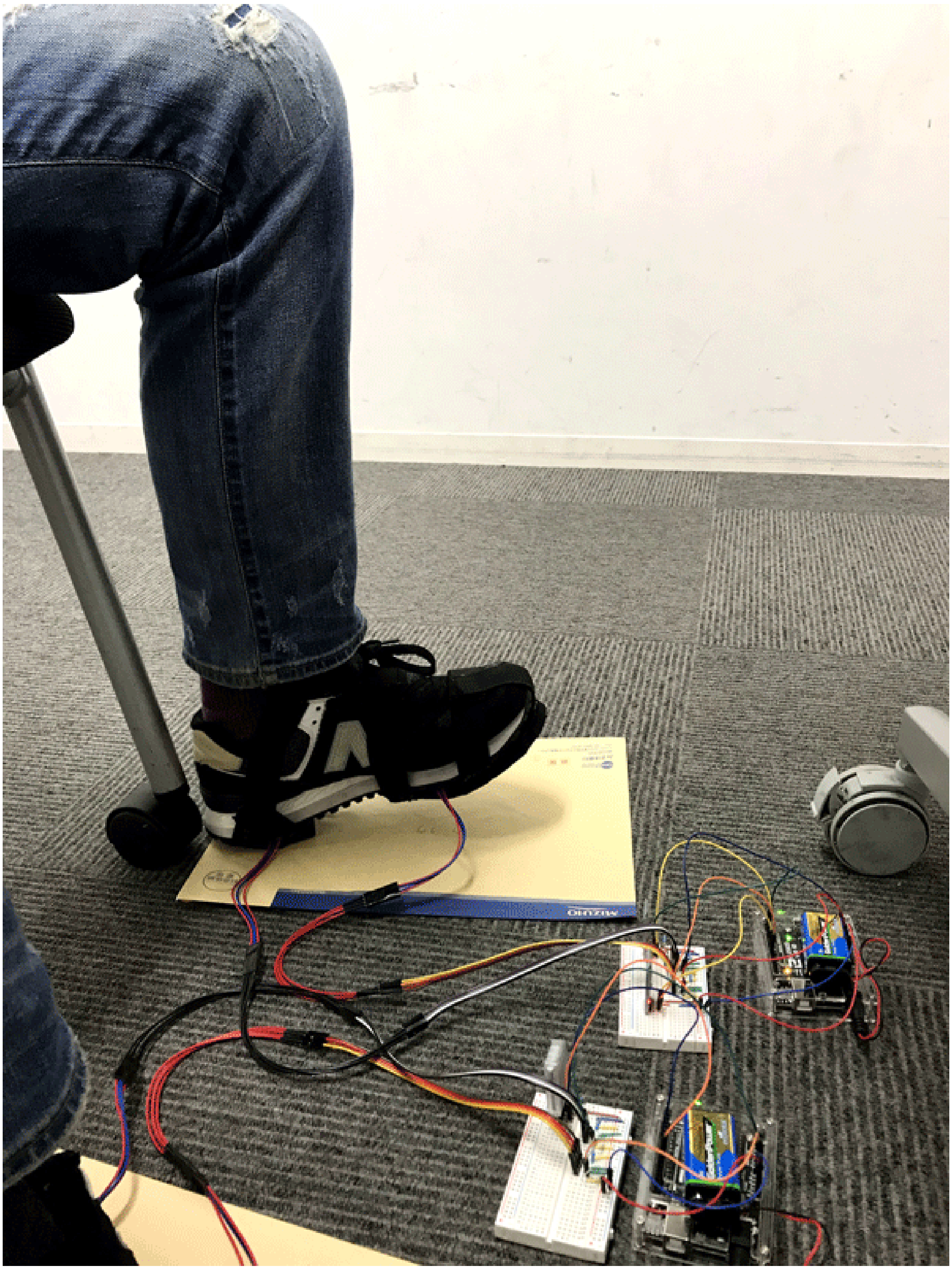

Figure 4.1 Photograph of toe tapping gesture

4.1.1 Goal

The purpose of this experiment was to investigate the top tapping and heel tapping keystrokes in the foot interaction, determine the lift height of the foot gesture, and have a reasonable time and error rate to assess its usability. After determining the standard value, we will use the KLM (keystroke-level model) to evaluate the performance of the system. Moreover, compare the predicted button time of the system with the expected button time of the ordinary keyboard. This knowledge will be applied to the second and third experiments.

4.1.2 Participant

A total of 9 participants (3 female) participated in the experiment, ranging in age from 23 to 30 years old. The dominant foot of 7 of them is the right foot. Shoe sizes range from 22 cm to 27 cm. The position of the distance sensor is adjusted before testing to determine its effectiveness.

4.1.3 Average lift height of Toe tapping and Heel tapping

The effective working distance of the proximity sensor Lbr-127hld used in the prototype is 5-40mm. The return value in Arduino is 900-1018. At the same angle, the size of the shoe is proportional to the lift distance. Under the principle of being lifted and comfortable as much as possible, each participant performs 50 complete processes for each button, taking the minimum average as the user-specified distance. Through experiments, we obtained a minimum average of 953. For dominant feet, the difference is not apparent. The difference between Toe tapping and Heel tapping is close to the predicted result. The minimum average of Heel tapping is lower than the minimum average of Toe tapping. To reduce the probability of accidental activation, the final standard value is set to 960.

4.1.4 Experimental design

Based on the above results, we tested the KLM (keystroke-level model) model [10] of the system. KLM is one of the models for evaluating human-computer interaction. This model attempts to calculate the time of an individual button, including

Figure 4.2 Photograph of usability experiment

the reflection or reflection time before each key press. It is mainly used to test the interaction effects of users who are skilled in the input device. It predicts the execution time of a task through design and task settings. This model decomposes the entire process of button execution into independent execution units according to the detailed execution steps of the user. The time of this execution process can be calculated by experimenting with the time of the unit. The experimental method is as follows:

(1) Obtain the operation instructions: Set the four buttons of the left and right feet in the order of the left foot toe: 1. left foot heel: 2. right foot toe: 3. right foot heel: 4.

(2) Final goal: complete the specified key sequence

(3) Specify the experimental procedure: execute in the order of

$$12344321132442311432 3241$$

(4)M(Metal preparation)$= 1.2S$

(5) The actual execution process is predicted as

$$M12344321M13244231M1432M3241$$

(6) Execution of task flow

(7) Statistics of the standard execution time of each participant

(8) Calculate the average button time

In this experiment, we set the predicted button time to K and Metal preparation to M. The time for executing multiple buttons is set to T_n, and the waiting system response time is set to $W_t = 0$

4.1.5 Result

We measured the standard execution time of 7 participants. After calculation, the average process time is $13.22S$ According to the formula

$$K = (S - 4 * M)/24 \tag{4.1}$$

Determine the button time of the system $K = 0.35S$. For traditional keyboard buttons, the button time for most people is about $0.23S$. It can be seen from the comparison that the button time of the wearable system based on the foot interaction is higher than that of the traditional keyboard, but the availability of the system is proved. It is envisaged that in a computer environment, the system can serve as an additional input method, providing more button mappings.

4.2. Text entry based on foot interaction

Text entry is a critical part of the computer environment. No matter whether it is work or entertainment, you can't leave the text input aside. For most computer environments, their text input is not much different from the typewriters of more than 100 years ago. There is no doubt that the traditional keyboard is the most efficient text input solution. With the development of mobile devices, especially for wearable devices, keyboards have been unable to match their lightweight requirements. For mobile terminals, multi-touch is undoubtedly the most efficient and natural alternative. However, for wearable devices, or VR/AR environments, there is no very efficient solution. Text entry is undoubtedly an essential topic of innovative interaction. In this study, we will use foot interaction as an alternative to text input when both hands are occupied.

4.2.1 Goal

Analyze the text input performance of mainstream input devices, and evaluate the advantages and disadvantages of different input devices by testing input efficiency and other indicators. Analyze the input method of TapTipToe for reference for further foot interaction. The experiment was mainly to test the text input of computer keyboard, mobile phone, game controller, and TapTipToe. Compare the input efficiency and performance of these devices based on the input efficiency of the PC keyboard. Using the TapTipToe system based on foot interaction as the primary or secondary input method, a text entry scheme based on foot interaction is provided. Analyze the performance of the foot interaction for text entry to verify the performance of the foot interaction as a mobile device is a wearable device effective input method in a VR/AR environment. Provides a solution for effective text entry when both hands are occupied.

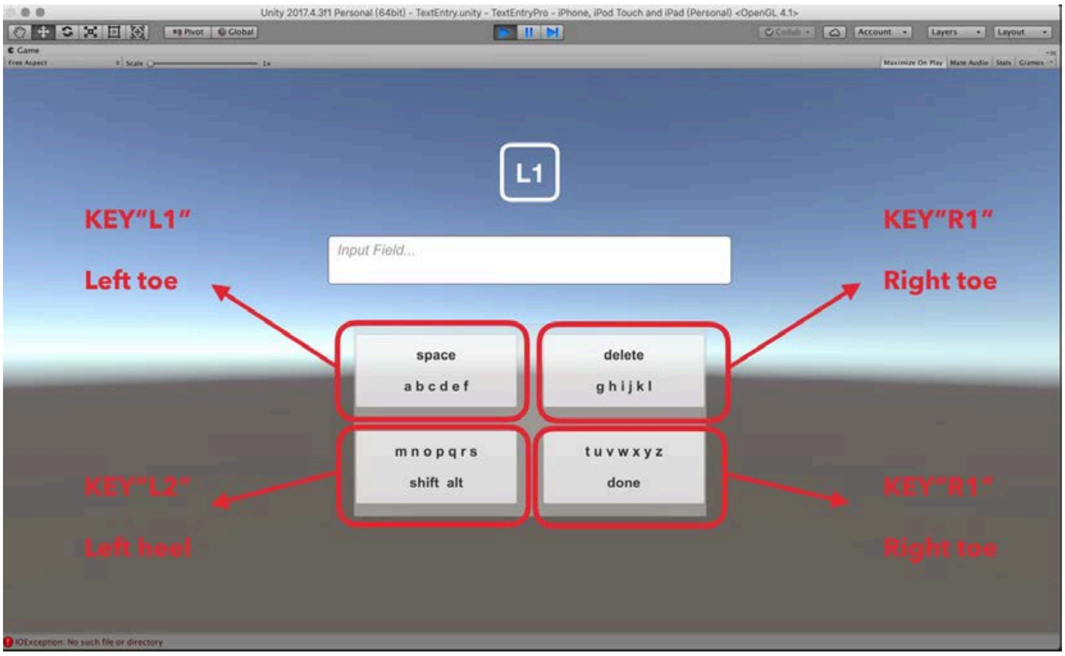

Figure 4.3 Photograph of 4-key text entry program

4.2.2 4-key text entry method based on foot interaction

We designed a 4-button text entry method for foot interaction. Map the four buttons in the TapTipToe system to a four buttons text entry system.
The toes and heels of the left foot are named

$$L1, L2$$

The toes and heels of the right foot are named

$$R1, R2$$

The letter is selected by activating the four buttons. The text entry method system is divided into three levels. The first and second levels are preview layers of letters, numbers, symbols, and functions. The third level is the input layer for letters, numbers, and symbols. Entering each character requires 2-3 taps. For example, if you enter letter A, you need to tap three buttons： L1, R1, and L1. The advantage of this input method is that it provides a Hand-free input method. Practice can be used for eyes-free interaction.

4.2.3 Text entry assessment and model

When evaluating the text entry method, sample examples are usually provided to the participants as input references. While the participant enters the statement, the system records the keystroke and the corresponding time. By dividing the total number of input characters by the time [2] used, you can get the speed of the input. The formula is expressed as

$$CPS = C_n/T_c \tag{4.2}$$

Here C_n is the number of characters to be entered, and T_c is the time used for input. In addition to "character per second" as a statistical method of input speed, there are also "words per minute." At present, WPM [11] is generally used for English statistics. The conversion formula between cps and wpm is

$$WPM = (CPS * 60)/W_c \tag{4.3}$$

Wc is the average number of characters in the word. The words here are not only letters but also punctuation. Usually, W_c takes a value of 5. Unlike the speed of text entry, there is no uniform standard for accuracy testing of text input. The diversity of input errors has made it impossible to determine the accuracy of the data. A method used to measure the efficiency of text input process(including correcting errors), is KSPC [12]. The formula is

$$KSPC = K_n/C_n \tag{4.4}$$

Here Kn is the number of keystrokes in the input process, and C_n is the number of characters entered. In general, the KSPC for text input is higher than 1.0 because most input interfaces use composite keys. For this study, KSPC can be used as a perfect criterion for input efficiency, because all characters in this input method use the composite key input.

Figure 4.4 Common text entry devices in daily life

4.2.4 Experimental design

We selected an example and invited 7 participants to conduct a text entry experiment. Each participant used four input devices (Taptiptoe, game Controller, smartphone, keyboard) for the experiments. Of the 7 participants, only 3 had experience using game controllers for text entry. Seven participants, all with keyboard and mobile text entry experience in daily life, and can be considered that all participants are proficient in keyboard and mobile phone text entry. To improve the fairness of the experiment and make the experimental data more accurate, we set up a learning experiment for the text input of Taptiptoe and the game controller. All participants used each of the two text input devices for a 15-minute practice experiment. For learning experiments, we recorded the average character input speed per minute for two input devices to compare the learning efficiency of the two input devices. For text input experiments, we select a 217-character example. To improve the fairness of the experiment and make the experimental data more accurate, we set up a learning experiment for the text entry of Taptiptoe and the game controller. All participants used each of the two text input devices for a 15-minute practice experiment. For learning experiments, we recorded the average character input speed per minute for two input devices to compare the learning efficiency of the two input devices. For text entry experiments, we select a 217-character sample example. 4 Complete text input processes in the order of keyboards, mobile phones, game controllers, Taptiptoe, respectively. We recorded the input time and the number of keystrokes for each input device. Based on the experimental data, we compared and analyzed three kinds of text entry efficiency models (CPS, KSPC, WPM) for four types of input devices.

4.2.5 Result

Because all participants are experienced in the text entry of the keyboard and mobile phone, the learning efficiency of these two input methods is not analyzed.

Figure 4.5 is the Taptiptoe learning curve for all participants. The x-axis represents 15 minutes of study time. The y-axis represents the input speed (CPM characters per minute) in the test. Apparently, after 15 minutes of practice, most participants can reach speeds above 20CPM. According to the feedback of the participants, Taptiptoe's text entry system is cumbersome, but after practice can remember part of the text entry order, which is very helpful to improve the input speed. It can be predicted that after a long period of training, the system can achieve eye-free text entry.

Figure 4.6 is a comparison of the learning rate of the Taptiptoeand the game controller ' s text input method, which can be seen from the graph, the game controller input method based on the soft keyboard is easy to understand and learn, the early speed is much higher than the Taptiptoe system, after 3 minutes of study tends to flatten. Based on this trend, it can be predicted that while the

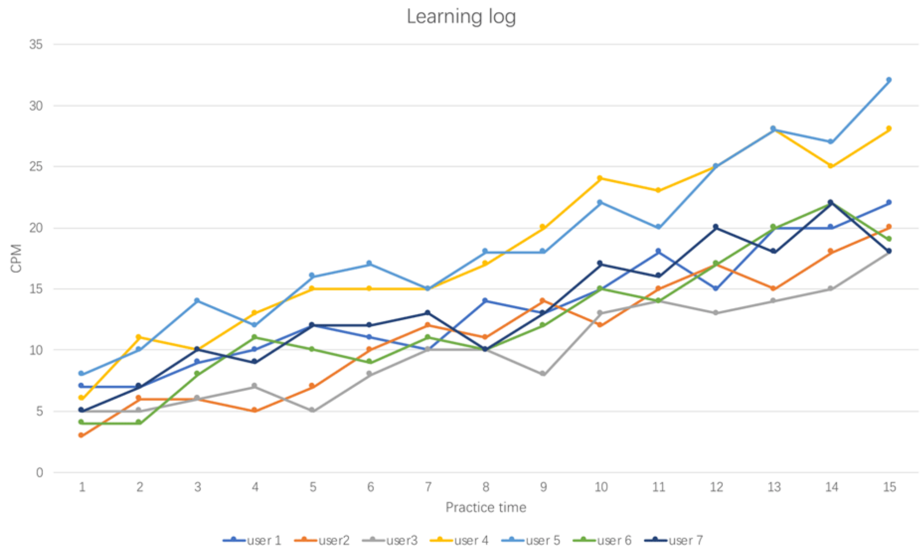

Figure 4.5 Log-log plots of learning rates for TapTipToe for each participant

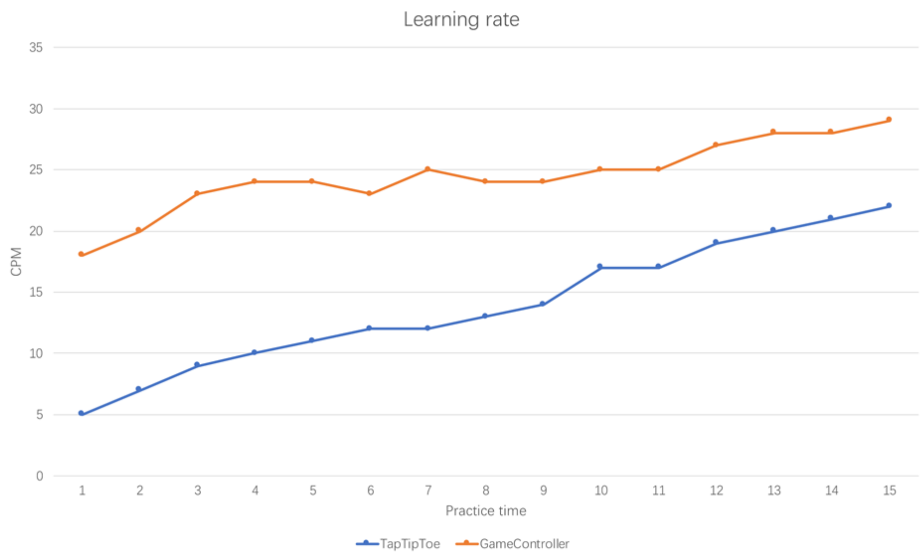

Figure 4.6 Learning Rates of TapTipToe and Game controller

input method of the game controller is profoundly learning efficient, the upper limit of the input speed is not high because of its operating principle, and it is necessary to maintain most of the attention in the soft keyboard area. Taptiptoe text entry Learning speed is not high, the input process is cumbersome, according to the principle of its input text, can predict the speed limit of this input method is low, but in the aspect of Eyes-free interaction has the potential to develop, it takes much time to remember and practice.

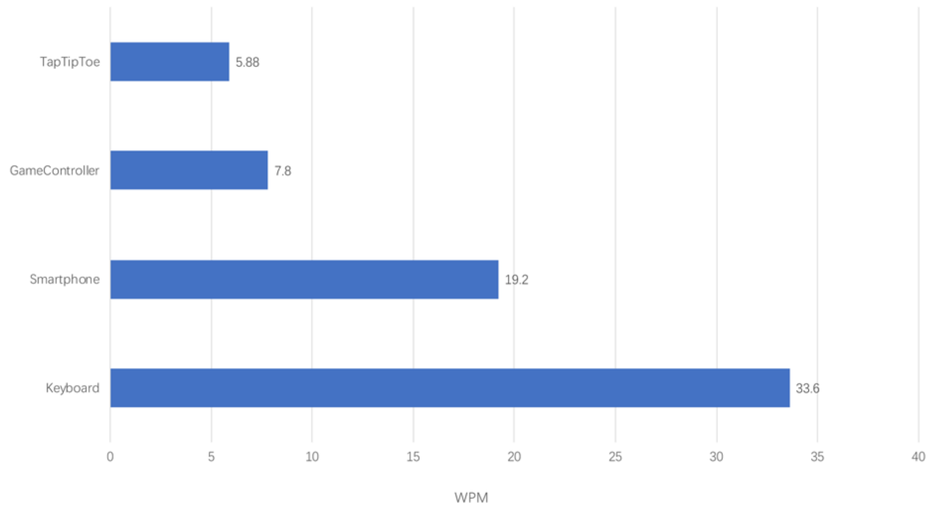

Figure 4.7 Contrast diagram of Word per minute for text input devices

Figure 4.7 is a comparison of the average input speed of 4 text input methods in the experiment. Taptiptoe is inefficient as a text entry interface and is not suitable as a primary input interface for use as an efficiency tool.

Table 4.1 Date of text entry

	Average Time	Average Keystrokes	CPS	KSPC
TapTipToe	440s	723	0.49	3.33
GameController	335s	923	0.65	4.25
Smartphone	135s	281	1.60	1.29
Keyboard	77s	238	2.80	1.10

Table 4.1 is a statistic of the average speed of experiments for all participants. We use CPS and KSPC two indicators as a criterion for determining the efficiency of text input methods. We use the standard PC keyboard as the judgment benchmark. We can find that Taptiptoe text entry is not efficient, the speed is about 17.5% of the standard PC keyboard. To use it to enter 217 characters, participants need to press the key 723 times on average. This data contains the number of backspace times. Depending on how it works, the 1-character input requires an average of 3 keystrokes, which can be confirmed in the KSPC model. As an input method of non-corresponding mapping, the KSPC value of game controller is higher than Taptiptoe system. However, its CPS value is higher than Taptiptoe system, which does not coincide with the law we get. This result can be explained by the KLM (keystroke level models) model, which analyzes the efficiency of human-computer interaction. When using the game controller for text entry, most of the time by repeatedly moving the cursor to switch between letters. This process consumes metal preparation lower than the Taptiptoe system. Based on our analysis of the learning efficiency of the Taptiptoe system, this problem can be improved through high-intensity exercises although training can significantly enhance the text entry speed of Taptiptoe systems. Judging by the standard PC keyboard, it is difficult to be called an efficient text entry method. On the other hand, the text input of foot interaction is not efficient, but it cannot be denied that it does work. As a text entry method for Hands-free interaction, we believe that the Taptiptoe system has a high potential to develop. It is also expected to perform in eyes-free interactions after high-intensity exercises.

Figure 4.8 Photograph of text entry test in VR environment

Chapter 5
Conclusion

5.1. summary

From the finding in this thesis, Users tend to interact with the computer using several foot interaction gestures driven by the ankle joint. The main reasons for this are three points.

easy to understand. Toe tapping and Heel tapping similar to finger touch, suitable for performing clicks, tap, select and other functions.

Easy to execution. Foot gestures driven by ankle joints have excellent performance. When performing this type of action, it is less limited by the physical environment and can be easily performed in a variety of postures and environments. Also, after the implementation of the effect is significant, can be detected in a variety of ways.

Strong anti-fatigue ability. The ankle joint is the most flexible joint of the lower limb of the human body and is located at the end of the leg, consuming less energy when performing foot gestures driven by the ankle joint. Gestures such as Toe tapping take on less limb weight than other foot gestures.

In this study, we designed, developed and evaluated a set of foot interaction systems using Toe tapping and hell tapping foot gestures. The aim is to further study the application of foot interaction in computer interaction through this system. We have evaluated the usability of this system, which has a fast key response and is suitable for use in a variety of computer systems. To evaluate the efficiency of this system, we have designed a set of text entry methods for this system. After testing, we confirmed that the system could effectively carry out text input. After comparing with the efficiency of other text entry methods, the conclusion that the efficiency of this text input method is not ideal is obtained. From the results of the experiment, we have identified several characteristics of the system.

(1) The number of mappings is small, used as the primary input interface, the efficiency of dealing with complex tasks is not ideal.

(2) Able to achieve hands-free interaction. When the hands are occupied, they can be used as an effective alternative input method. As a secondary input method, it can be used in conjunction with the primary input method. Can also be used as an accessibility input method.

(3) Ability to achieve eye-free interaction. It is simple to operate and easy to learn. Users have a strong sense of foot interaction, after a long period of training and memory, can effectively carry out eye-free input.

Through this study, we further understand the application of foot interaction in human-computer interaction. Foot interaction is not comparable to the efficiency of hand interaction. Foot intersection is not as flexible as fingers. The characteristics of eye-free interaction and hands-free interaction can be performed based on foot interaction. Find the right task to interact with the foot can get maximize benefits.

5.2. limitation

For the Taptiptoe system designed in this study, the function of Toe tapping and Heel tapping foot gestures is single, and many essential functions cannot be performed. There are limitations to detection methods. This system uses a single detection method, in some special physical environment, the activation accuracy will be reduced, or even can not work correctly. When used on the ground with low reflectivity, due to low reflectivity, error activation occurs， leakage activation, inability to activate. When used on the ground with low reflectivity, due to low reflectivity, error activation occurs ， leakage activation, failure to activate. For the text entry method based on foot interaction designed in this study, it can realize the most basic text input function. After our test, it proved that it could function properly. However, as an efficiency tool, it is inefficient. To implement an application in a normal working environment, in addition to optimizing the input method itself, we need to improve its fundamental problem of foot interaction gestures.

5.3. Future work

After the study of foot interaction in this thesis, we understand the characteristics of some foot interaction. Based on these characteristics, we will focus on the following studies in the future.

(1) For foot interaction gestures. We need to explore more foot interaction gestures. There are limitations to a single gesture. With a single foot interaction gesture, we can perform tasks more accurately, while learning more quickly and mastering how they are used. However, concerning functionality, it is not possible to achieve many of the effects we want to achieve. In the future, we will deepen our research on foot interaction gestures and try to add more foot gestures to achieve functionality.

(2) Detection of foot interaction. At present, our understanding of foot interaction is not thorough enough, and the research on the detection of foot interaction appears to be too immature. In the future, we need to improve the depth of our foot interaction research, explore more sensors, and try to use a variety of sensors to detect foot movements at the same time. To better study human-computer interaction.

(3) Application of foot interaction. In this study, we understand the characteristics and design principles of some foot interactions. We will strengthen the advantages of foot interaction, cooperate with more human-computer interaction

interfaces, expand the breadth of foot interaction applications, and explore the possibility of foot interaction.

References

[1] Fabian Göbel, Konstantin Klamka, Andreas Siegel, Stefan Vogt, Sophie Stellmach, and Raimund Dachselt. Gaze-supported foot interaction in zoomable information spaces. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 3059–3062. ACM, 2013.

[2] David Dearman, Amy Karlson, Brian Meyers, and Ben Bederson. Multi-modal text entry and selection on a mobile device. In *Proceedings of Graphics Interface 2010*, pages 19–26. Canadian Information Processing Society, 2010.

[3] William Saunders and Daniel Vogel. Tap-kick-click: Foot interaction for a standing desk. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, pages 323–333. ACM, 2016.

[4] Eduardo Velloso, Dominik Schmidt, Jason Alexander, Hans Gellersen, and Andreas Bulling. The feet in human–computer interaction: A survey of foot-based interaction. *ACM Computing Surveys (CSUR)*, 48(2):21, 2015.

[5] Toni Pakkanen and Roope Raisamo. Appropriateness of foot interaction for non-accurate spatial tasks. In *CHI'04 extended abstracts on Human factors in computing systems*, pages 1123–1126. ACM, 2004.

[6] I Scott MacKenzie. Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction*, 7(1):91–139, 1992.

[7] Andrew Crossan, Stephen Brewster, and Alexander Ng. Foot tapping for mobile interaction. In *Proceedings of the 24th BCS Interaction Specialist Group Conference*, pages 418–422. British Computer Society, 2010.

[8] Volker Paelke, Christian Reimann, and Dirk Stichling. Foot-based mobile interaction with games. In *Proceedings of the 2004 ACM SIGCHI Interna-*

tional Conference on Advances in computer entertainment technology, pages 321–324. ACM, 2004.

[9] Yasmin Felberbaum and Joel Lanir. Better understanding of foot gestures: An elicitation study. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 334. ACM, 2018.

[10] Laura Giordano, Valentina Gliozzi, and Gian Luca Pozzato. Klmlean 2.0: A theorem prover for klm logics of nonmonotonic reasoning. In *International Conference on Automated Reasoning with Analytic Tableaux and Related Methods*, pages 238–244. Springer, 2007.

[11] I Scott MacKenzie, Shawn X Zhang, and R William Soukoreff. Text entry using soft keyboards. *Behaviour & information technology*, 18(4):235–244, 1999.

[12] I Scott MacKenzie. Kspc (keystrokes per character) as a characteristic of text entry techniques. In *International Conference on Mobile Human-Computer Interaction*, pages 195–210. Springer, 2002.