FootUI: Designing and Detecting Foot Gestures to Assist People with Upper Body Motor Impairments to Use Smartphones on the Bed

Xiaozhu Hu
Division of Integrative Systems and Design, Hong Kong University of Science and Technology
Hong Kong, China
Department of Computer Science and Technology, Tsinghua University
Beijing, China
xhubk@connect.ust.hk

Jiting Wang
Department of Computer Science and Technology, Tsinghua University
Beijing, China
while1incrp@hotmail.com

Weiwei Gao
Department of Computer Science and Technology, Tsinghua University
Beijing, China
gww20@mails.tsinghua.edu.cn

Yongquan Hu
School of Computer Science and Engineering, University of New South Wales
Sydney, Australia
yongquan.hu@unsw.edu.au

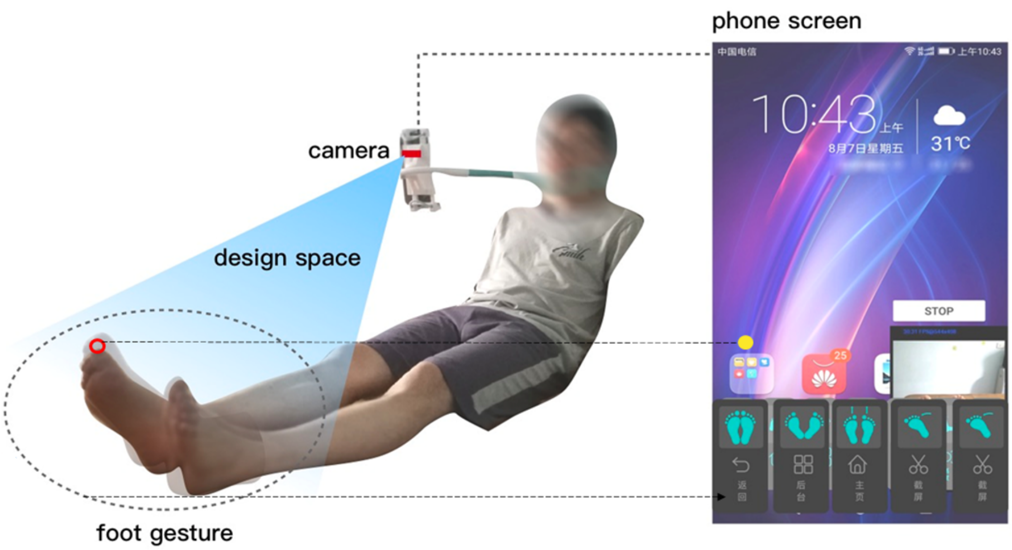

Figure 1: Usage scenario of FootUI

ABSTRACT

Some people with upper body motor impairments but sound lower limbs usually use feet to interact with smartphones. However, touching the touchscreen with big toes is tiring, inefficient and easy to mistouch. Targeting at this pain point, we propose FootUI, which leverages the phone camera to track users' feet and translates the foot gestures to smartphone operations. This technique enables users to interact with smartphones while reclining on the bed and improves the comfort of users. In this paper, we present the design and evaluation of the foot gestures through user studies as well as the development and evaluation of FootUI. Results show that toes-based foot gestures are not only less perceivable but also uncomfortable. FootUI avoid the use of toes-based gestures and is proved to be an easy, efficient and interesting input technique for people with upper body motor impairments but sound lower limbs.

ASSETS '22, October 23–26, 2022, Athens, Greece

ACM ISBN 978-1-4503-9258-7/22/10…$15.00
https://doi.org/10.1145/3517428.3563285

CCS CONCEPTS

• **Human-centered computing → Accessibility technologies.**

KEYWORDS

accessibility, foot-based interaction, upper body motor impairments, smartphone

ACM Reference Format:

Xiaozhu Hu, Jiting Wang, Weiwei Gao, and Yongquan Hu. 2022. FootUI: Designing and Detecting Foot Gestures to Assist People with Upper Body Motor Impairments to Use Smartphones on the Bed. In *The 24th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '22), October 23–26, 2022, Athens, Greece.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3517428.3563285

1 INTRODUCTION

On social media, we can see that some people with upper body motor impairments but sound lower limbs use their feet as compensation of hands to do many daily works, like writing, drawing and picking objects. Although this group of people is small in number, it is worthy of attention. We focus on their smartphone interaction and try to provide them with an interaction technique that is suitable for foot input.

There have been some studies on touch-based [9–12] and hands free [3] interaction techniques for people with upper body motor impairments but there is still a lack of interaction techniques which specifically designed for those with sound lower limbs. Although head-based or voice-based interaction techniques are usable for some of them, those with sever upper body paralysis or those with dysphonia are unable to use these techniques. However, all of them are able to use feet. Additionally, despite there are lots of research about foot based interaction, these people who are awfully dependent on their feet have been overlooked for a long time. Some of them choose to interact with smartphones with their feet by touching the phone screen [2], which is fatiguing and inefficient.

In our prior work, we explored the needs and pain points of people with upper body motor impairments, defined the usage scenario, and proposed a vision-based interaction technique to assist them to interact with smartphones through foot gestures when reclining on the bed. In this work we present the design and evaluation of the foot gestures as well as more details of the final evaluation of FootUI. Results show that toes-based foot gestures are not only less perceivable for the computer vision algorithm but also uncomfortable for users. FootUI avoid the use of toes-based gestures and is proved to be an easy, efficient and interesting input technique which facilitates people with upper body motor impairments but sound lower limbs to interact with smartphones when reclining on the bed.

The contributions of this work are: 1) We propose a foot gesture set based on both technical and user-center perspectives and show the weakness of toes-based gestures. 2) We present the development and performance of FootUI with details of participants' subjective feedback.

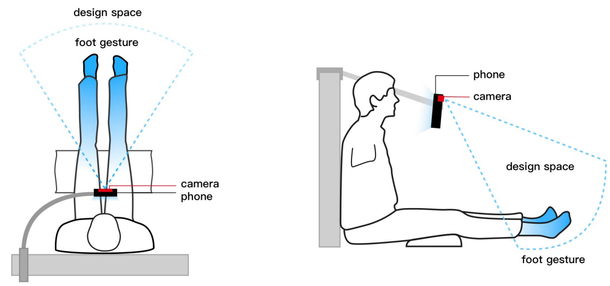

Figure 2: Posture and scenario of using FootUI

2 RELATED WORKS

There are lots of research proposing interaction techniques to assist people with upper body motor impairments to use their motor-impaired upper limbs to interact with information devices. Cursor based input techniques like PointAsist [14], Steady Clicks [16] and Click Control [6] have been developed to reduce the pointing difficulties. Wobbrock et al. [18–20] experimented the use of physical edges to assist motor-impaired users and proposed Barrier Pointing [4]. To improve touch accuracy, researchers explored the users' interaction behavior and developed gesture recognition algorithms such as Shared User Modeling Framework [9], Session Specific Models [10], Smart Touch [11] and Cluster Touch [12]. However, these techniques are unavailable to amputees and people with severe upper body paralysis.

For people with limited motor ability of the upper limbs but sound lower limbs, assisting them to use lower limbs may be a better alternative. Research on foot-based interaction has a long history. Velloso et al. [17] surveyed foot-based interaction from three aspects: characteristics of users, foot-based systems and foot-based interactions. Katsumi et al. [8] compared gaze input, head input and foot input and concluded that foot input is an option for hands-free interaction when seated. As for researches on designing and sensing foot gestures for mobile device controlling, Scott et al. studied the ergonomic characteristics and design space of four gestures: dorsiflexion, plantar flexion, toe rotation and heel rotation for eyes-free interaction with mobile devices in the pocket [15]. Alexander et al. [1] created a user-defined gesture set for mobile device commands and validated that rate-based techniques are significantly faster, more accurate and result in far fewer target crossings compared to displacement-based techniques. However, accessible foot input research focusing on leveraging foot gestures as the main form of input technique to support fully hands-free interaction is lacking. Additionally, there's little known about foot gestures which are suitable for camera detection and practical use when the user reclining on the bed.

3 DESIGN OF FOOTUI

Based on our prior findings [5], we summarized the characteristics of upper body motor-impaired smartphone users and defined the usage scenario. For flexible foot motion, we define that the user reclines on the bed with a smartphone fixed on the smartphone holder. The user's feet are within the view of the phone camera and the camera is always on to track the user's foot gestures. Users can

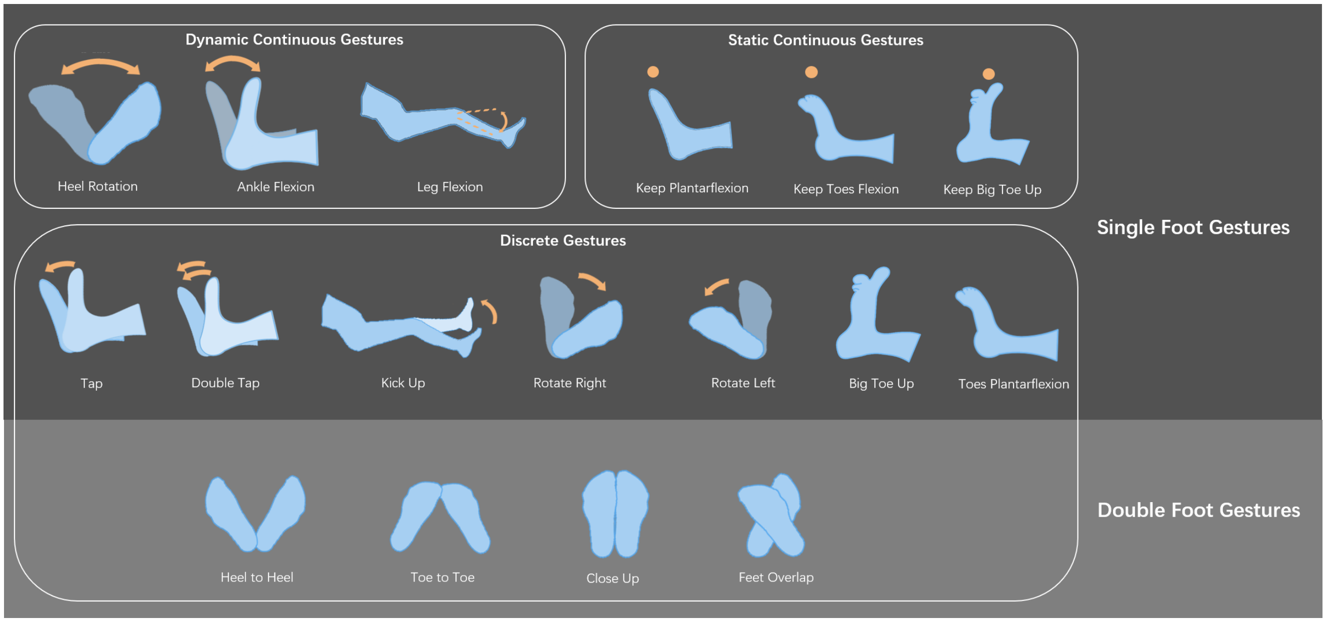

Figure 3: Initial foot gesture vocabulary

put pillows or cushions under legs to facilitate leg lifting (See Fig 2). Based on the usage scenario, we created an initial gesture vocabulary, evaluated the gesture vocabulary through a user study and developed the gesture tracking and recognition algorithm. According to the result of gesture evaluation and recognition accuracy, we selected 11 gestures, created the final foot gesture set and mapped the foot gestures to smartphone operations.

3.1 Initial Foot Gesture Vocabulary

The foot gestures are classified into two categories as continuous foot gestures and discrete foot gestures [1]. We divided the continuous foot gestures into dynamic continuous foot gestures and static continuous foot gestures. Dynamic continuous foot gesture means a state of continuous foot movement, such as "Heel Rotation" and "Ankle Flexion", which is similar to finger drag and can be tracked continuously with realtime dynamic feedback. Static continuous foot gesture means to keep the foot in a specific posture for some time, such as "Keep Plantarflexion" and "Keep Big Toe Up", which is alike to finger long-touch and can be tracked continuously and trigger some discrete feedback. Discrete gestures are also described as semaphoric gestures [13] which can be used as iconic triggers for commands but also carries continuous information [17]. Based on the kinematic analysis and semaphoric gesture summary in [17], we designed 17 kinds of foot gestures (See Fig 3). The foot gestures were divided into three groups: dynamic continuous gestures, static continuous gestures and discrete gestures. Thirteen of them were single foot gestures and four of them were double foot gestures.

3.2 User Study 1: Evaluating Initial Foot Gesture Vocabulary

This study aims at evaluating the foot gestures in the initial foot gesture vocabulary from the perspective of ease-of-use, fatigue, speed and user preference. The result of this study is the basis for selecting foot gestures and creating the final foot gesture set. We also collected video clips of foot gestures in this study for building the foot gesture dataset.

3.2.1 Participants. Since the inconvenience of people with upper body motor impairments and the similar foot motor ability between our target users and able-bodied people, we recruited ten able-bodied participants (5 female, 5 male) and two upper body motor-impaired participants (all female) through online contact. They were 28.50 years old (SD=10.55) on average and half of them were Android smartphone users. We provided the participants with monetary rewards for their participation.

3.2.2 Apparatus. We conducted this study in the participant's living environment. There was a phone holder pre-installed on the bed. An app that displayed foot gestures was installed on an Android smartphone (Huawei Honor 8) and we placed the phone on the phone holder in advance.

3.2.3 Procedure. The participant reclined on the bed with a cushion under the thigh (see Fig 2). Then we started the foot gesture display app. The app displayed the foot gestures one by one. For the 13 single foot gestures in the initial gesture vocabulary, the app showed the pictures of the foot gestures performed by the left foot and right foot respectively. Therefore, there were 30 (13×2+4) foot gestures displayed in the app. The app showed these 30 foot gestures in random order twice for each participant. In the first round of the display, we illustrated each foot gesture to the participant. We required the participant to perform it so that the participant got intuitive experience about the foot gesture. In the second round, after the display of each gesture, the app offered the participant a 5-point Likert scale in terms of ease-of-use (1=the most difficult to perform, 5=the easiest to perform), fatigue (1=the most fatiguing,

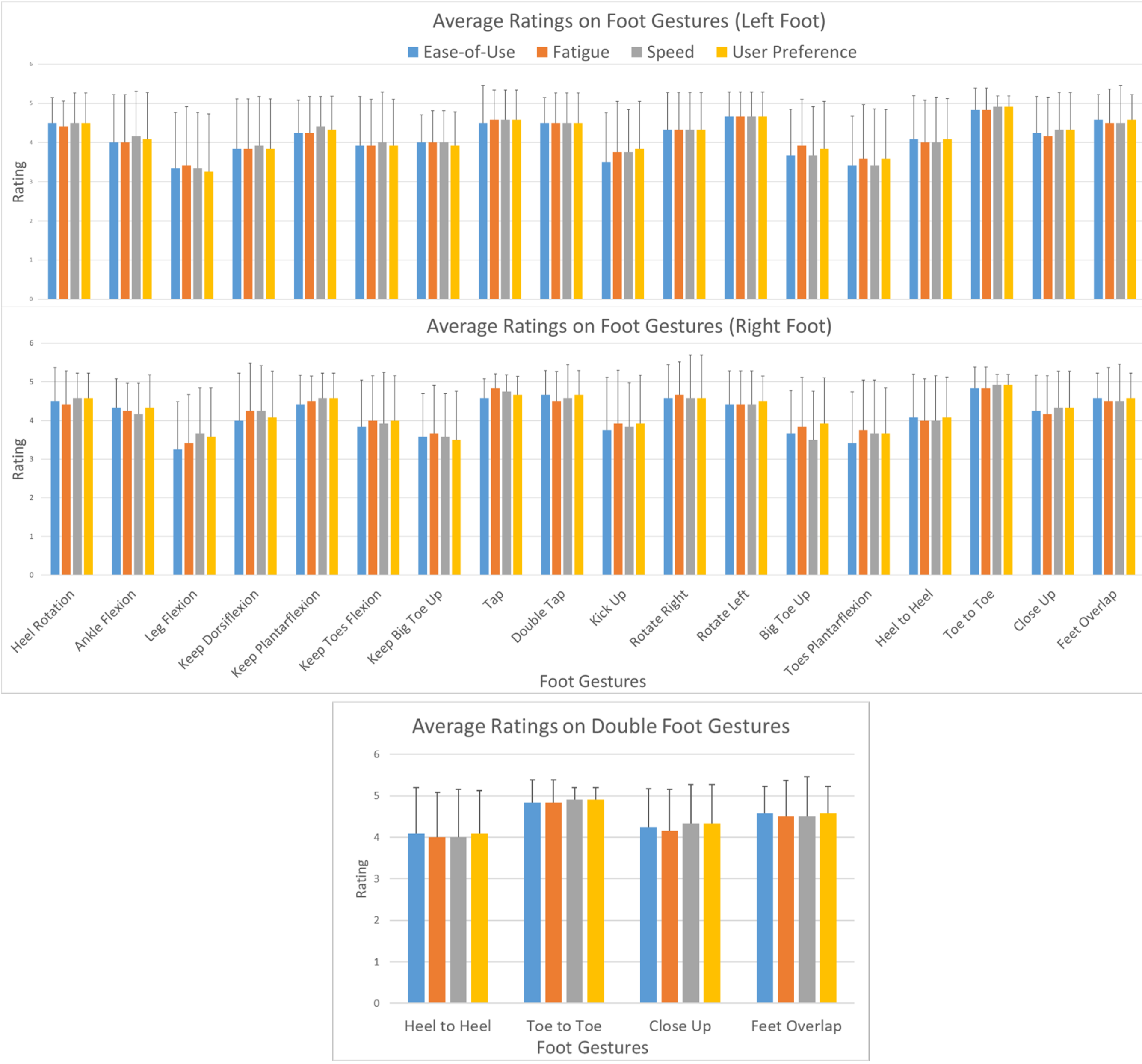

Figure 4: Average ratings on foot gestures

5=the least fatiguing), speed (1=the slowest to finish, 5=the most quick to finish) and user preference (1= the lowest user preference, 5= the highest user preference) of the foot gesture so that the participant can evaluate if conveniently. During the study, the back camera of the phone was always on and participants were required to perform each of the foot gestures in each round. The foot gesture data was recorded for algorithm development. We also conducted a small post-interview for some explanation about the extreme ratings and suggestions for gesture design.

3.2.4 Results. Fig 4 shows the average ratings of the foot gestures in terms of ease-of-use, fatigue, speed and user preference. Wilcoxon signed-rank test revealed that there are no significant difference between the left and right foot in performing all single foot gestures on all indicators. We calculated the average of each foot gesture across the four indicators as the overall score, on which no significant difference between left and right foot appeared. In this case, there are 24 ratings of overall score for each single foot gesture (12 of left foot and 12 of right foot) and 12 ratings of overall score for double foot gestures. Fig 5 shows the average overall score of the foot gestures.

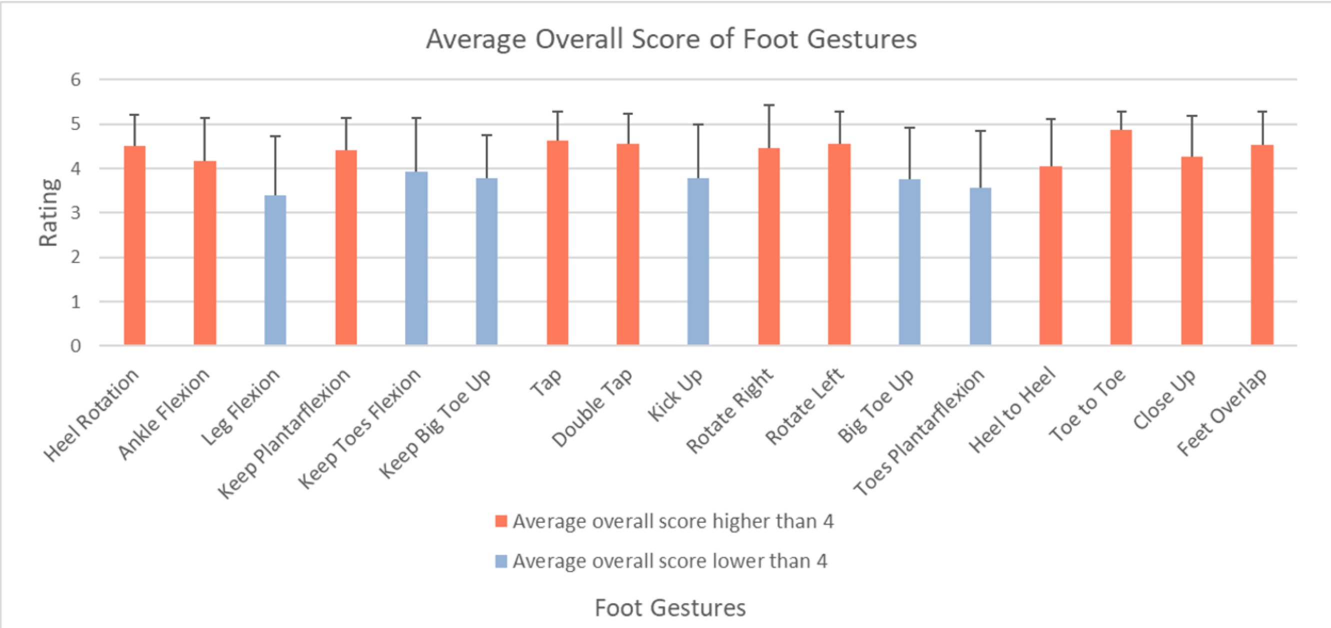

Figure 5: Average overall score of Foot Gestures

Considering the different kinematic characteristics among dynamic continuous gestures, static continuous gestures , single foot discrete gestures, and double foot discrete gestures, we compared the overall score of the gestures inside each group. Friedman tests showed a significant effect on overall score in each group (dynamic continuous gesture χ^2=29.83, p<0.001; static continuous gesture χ^2=16.26, p<0.001; single foot discrete gesture χ^2=70.01, p<0.001; double foot discrete gestures χ^2=12.44, p<0.01).

For dynamic continuous foot gestures, a Bonferroni correction post-hoc analysis showed that the overall scores of "Heel Rotation" (M=4.50, SD=0.70) and "Ankle Flexion" (M=4.17, SD=0.98) are both significantly higher than "Leg Flexion" (M=3.41, SD=1.32) with p<0.01 and the score of "Heel Rotation" (M=4.50, SD=0.70) is also significantly higher than that of "Ankle Flexion" (M=4.17, SD=0.98) with p<0.05. For static continuous foot gestures, "Keep Plantarflexion" (M=4.41, SD=0.73) got a significantly higher overall score than "Keep Big Toe Up" (M=3.78, SD=0.97) and "Keep Toes Flexion" (M=3.94, SD=1.20) with p<0.01. Among the single foot discrete gestures, post-hoc analysis showed that the overall scores of "Tap" (M=4.64, SD=0.64), "Double Tap" (M=4.55, SD=0.69), "Rotate Right"(M=4,47, SD=0.97) and "Rotate Left" (M=4.55, SD=0.73) are significantly higher than "Big Toe Up" (M=3.75, SD=1.17) and "Toes Plantarflexion" (M=3.56, SD=1.28) with p<0.01. "Kick Up" also got a significantly lower overall score than "Tap" (M=4.64, SD=0.64), "Double Tap" (M=4.55, SD=0.69) with p<0.01, and "Rotate Right"(M=4,47, SD=0.97) and "Rotate Left" (M=4.55, SD=0.73) with p<0.05. For double foot discrete gestures, a post-hoc analysis showed that "Toe to Toe" (M=4.88, SD=0.41) got a significantly higher overall score than other foot gestures with p<0.05.

We got two findings in the small interview. 1) Participants don't prefer toe-based gestures like "Big Toe up" and "Toes Plantarflexion", which is consistent with the data analysis results. p7 expressed that she felt difficult to perform toe-based gestures and p8 express that she felt sore calve when performing "Toes Plantarflexion". 2) As for heel rotation related foot gestures such as "Rotate Left" and "Rotate Right", participants usually prefer to rotate in the same direction with their foot. P4, p7 and p9 all prefer "Rotate Right" to "Rotate Left" regarding to the right foot but get opposite regarding to the left foot.

3.3 Gesture Tracking and Recognition Algorithm

3.3.1 Data Processing. During the user study 1, we collected 720 (12 participants x 2 rounds x 30 raw gestures) videos from the back camera as our dataset. We manually segmented the videos and removed the videos and frames when the participants were not performing the required foot gesture. Video segments have lengths varying from 4 to 96 frames, with a resolution of 960x1280 pixels. The final dataset contains 543 videos, 11920 frames in total.

3.3.2 Foot Segmentation and Feature Points Extraction. We firstly segmented the foot from the background by color in HSV (Hue, Saturation and Value) space. To improve the accuracy of segmentation, the participants were asked to wear solid color socks that can be clearly distinguished from the background. Before recording the video, they were asked to put their feet on a fixed position as a calibration. The center color of two feet were recorded, and the pixels that had similar Hues with them were segmented as the feet region.

Then, We detected the contours of the feet region. For each foot in every frame, an ellipse was fit into its contour, determining its size and direction. The fingertip of the big toe was retrieved as the

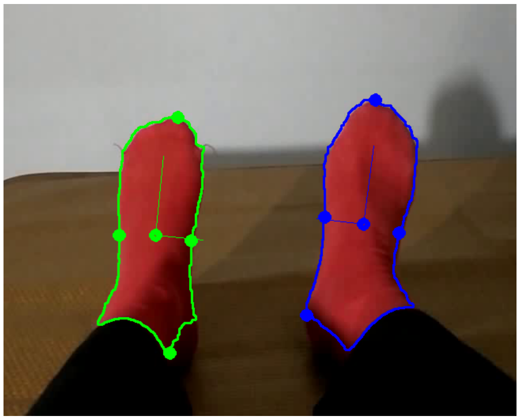

Figure 6: Feature point extracted from foot contours

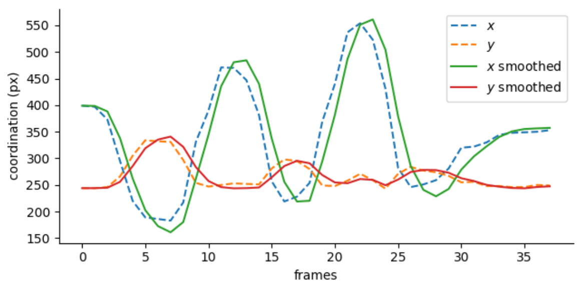

Figure 7: Smoothing tracking result with Kalman filter

furthest point on the contour found in the direction of its main axis. Similarly, we extracted feature points, or the furthest point in different directions, as in Fig 6. Considering that the coordinate of fingertips was going to be used to control the cursor, a Kalman filter was applied to provide a smoother tracking result, as shown in Fig 7.

3.3.3 Gesture Classification. We noticed that different participants performed the same gesture with significantly different speed, some of them may hesitate when performing a tapping or swiping. To condense length-varying feature point sequences into low-dimension features, we broke down the continuous fingertip movement into discrete strokes. A stroke was defined as the cumulative fingertip movement within a short period of time, until the fingertip moved backwards or stops moving, e.g., a "Tap" is composed of a downward stroke and an upward stroke. The most complex gesture is "Double Tap" which is composed of four strokes. Therefore, four latest strokes of each foot were kept as input feature. A random forest classifier was trained on the normalized, 32-dimension feature (4 strokes $\times$ 2 dimensional coordinates $\times$ 2 points, fingertip & heel $\times$ 2 feet) to classify the gestures.

3.4 Gesture Selection and Final Foot Gesture Set

Our tracking algorithm can smoothly track the fingertips, which means users can use their preferred dynamic continuous foot gestures for pointing. For discrete gestures, We trained our classifier on the dataset, and performed leave one person out cross validation, showing big difference among gestures, as in Table 1, Phase 1. "Feet Overlap" is less distinguishable with "Toe to Toe" and "Close Up" so that we removed it from our gesture set. Also, "Keep Plantarflexion" is frequently confused with a single "Tap", as their only difference is the interval between foot strokes, which depends on the individual behaviour of each participant. Combined with the user evaluation result in Section 3.2.4, we included "Keep Plantarflexion" in "Tap".

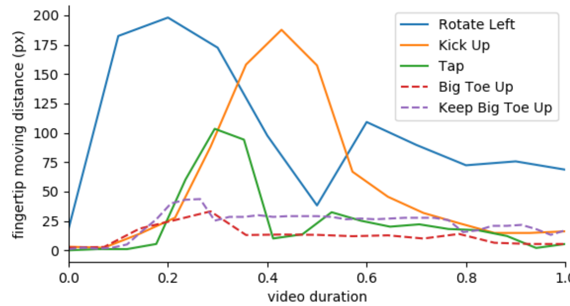

Figure 8: Toes-based gestures are less perceivable, compared with other gestures

When selecting the final foot gestures, we noticed that toes-based gestures added significant difficulty to recognition. Compared with other foot gestures (Fig 8 as an example), the movements can hardly be noticeable. Coincidentally, the user evaluation result in Section 3.2.4 also showed that users were not satisfied with toes-based gestures, which means toes-based gestures are both undetectable for CV techniques and uncomfortable for users. As a result, we removed toes-based gestures from our gesture set.

After removing gestures that might be ambiguous and hard to use, we chose eight discrete foot gestures to our final foot gesture set. Leave-one-out cross validation showed an accuracy of 90.53% (SD=6.62). Table 1, Phase 2 shows the confusion matrix between these gestures. The result showed that simple gestures including "Rotate Left", "Rotate Right", "Kick up" and "Tap" can be detected with high accuracy. Complex gesture recognition are less accurate, thus the user may have to learn to perform these gestures in a more standard way.

Fig 9 shows the final foot gesture set, which contains 3 dynamic continuous foot gestures and 8 discrete foot gestures. Users can mainly use "Heel Rotation" and "Ankle Flexion" for pointing and sometimes use "Leg Flexion" for pointing the top area of a phone screen. Discrete foot gestures can be used as iconic trigger for smartphone operations.

3.5 Mapping from Foot Gestures to Smartphone Operations

We mapped the foot gestures to smartphone operations following the rule of mapping the most user-preferred foot gestures to the most necessary smartphone operations. The smartphone operations were divided into basic operations and shortcuts. Fig 10 shows the mapping from foot gestures to smartphone operations. "Right foot heel rotation" moves the cursor horizontally, the combination of "right foot ankle flexion" (for fine movement) and "right leg flexion"

Table 1: Recognition accuracy and confusion matrix in classification. In Phase 1, gestures with low recognition accuracy or high user effort were removed. The final gestures remained were classified in Phase 2.

Raw Gesture	Phase 1		Phase 2: Confusion Matrix							
	Precision	Recall	CU	RL	RR	KU	DTap	T2T	H2H	Tap
Close up (CU)	0.737	0.933	**0.800**	0.067	0.067	0	0	0	0.067	0
Rotate left (RL)	0.931	0.964	0	**0.929**	0.036	0	0	0	0	0.036
Rotate right (RR)	0.824	0.966	0	0	**0.931**	0	0	0	0	0.069
Kick up (KU)	0.967	1.000	0	0	0	**0.931**	0.034	0	0	0.034
Double tap (DTap)	0.828	0.800	0	0	0	0	**0.900**	0	0	0.100
Toe to Toe (T2T)	0.533	0.533	0.067	0	0	0	0	**0.933**	0	0
Heel to Heel (H2H)	1.0	0.600	0.067	0.067	0.067	0	0	0	**0.800**	0
Tap	0.543	0.655	0	0	0.034	0	0.034	0	0	**0.931**
Keep Plantarflexion	0.636	0.519	-	-	-	-	-	-	-	-
Feet Overlap	0.444	0.286	-	-	-	-	-	-	-	-

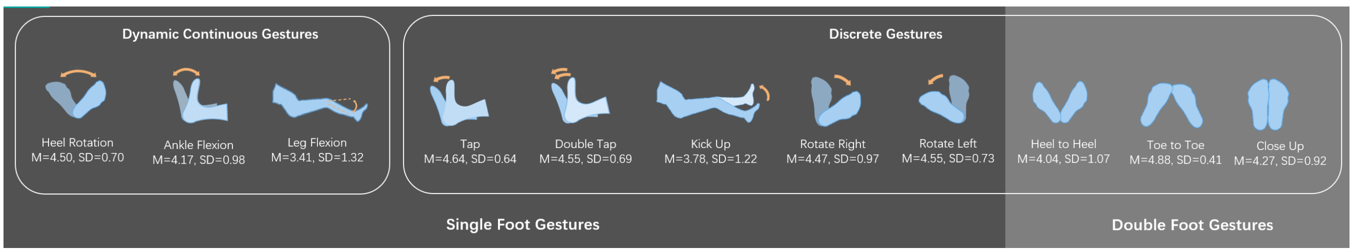

Figure 9: Final foot festure set with average overall score of each gesture

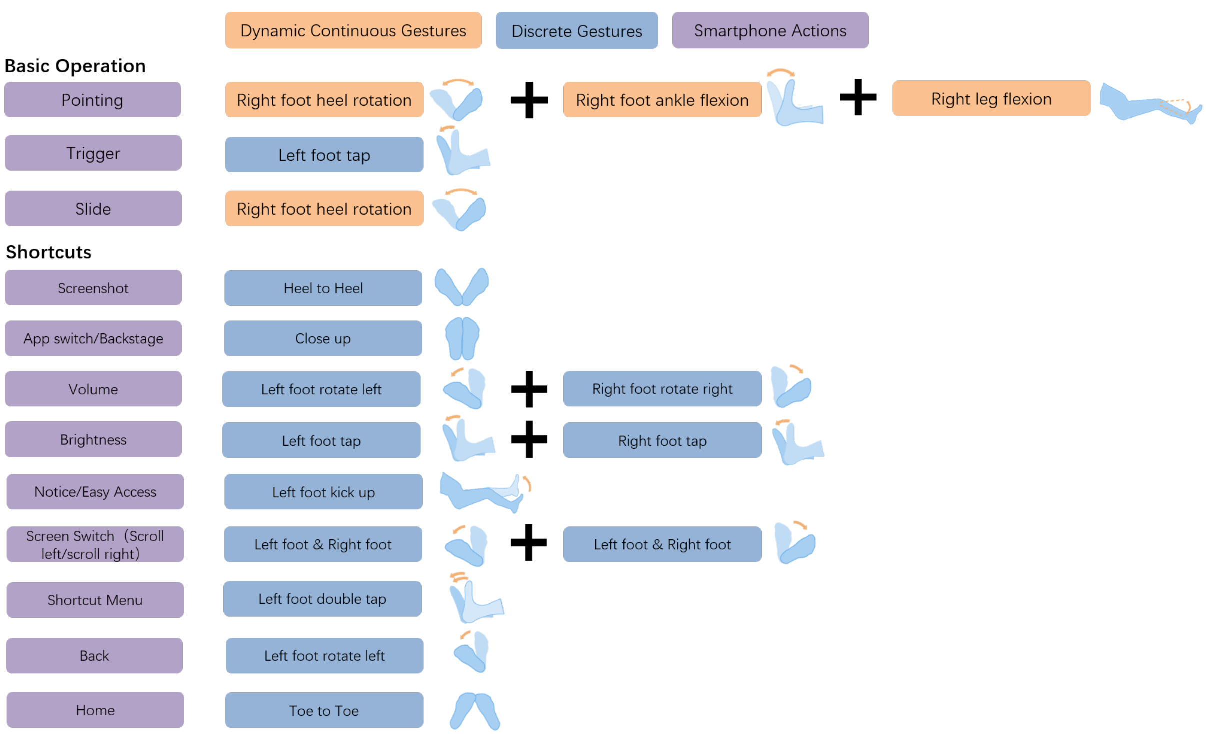

Figure 10: Mapping from foot gestures to smartphone operations

(for rough movement) moves the cursor vertically, and "left foot tap" triggers. When there is a slider on the interface, "right foot heel rotation" slides the slider. For shortcuts, "heel to heel" is mapped to "screenshot", "close up" is mapped to "app switch", "left foot kick up" is mapped to notice, "left foot rotate left" is mapped to "back", "toe to toe" is mapped to "home" and "left foot double tap" evokes shortcut menu. We also mapped the combination of single foot gestures to some shortcuts e.g. both feet tapping evokes brightness slider. The screen scrolls left when the user rotates both feet to the left, and the screen scrolls right, when the user rotates both feet to the right. When the user rotates right foot to the right and left foot to the left simultaneously, the volume slider appears.

4 USER STUDY 2: USABILITY EVALUATION OF FOOTUI

4.1 Participants

We recruited 14 able-bodied participants (7 female, 7 male) who were college students and two upper body motor-impaired participants (all female) through online contact because of the inconvenience of upper body motor-impaired people and similar foot motor ability between able-bodied people and the upper body motor-impaired people. They were 25.94 years old (SD=7.52) on average and ten of them used Android smartphone. Six of the able-bodied participants and two upper body motor-impaired participants (both-arm amputees) participated in the user study 1. We provided them with some monetary rewards for their participation.

4.2 Apparatus

For the both-arm amputees, We conducted this study in their living environment where there was a phone holder installed on the bed. For the able-bodied participants, we conducted the experiment in a room, where there was a convertible bed with a phone holder. We placed a smartphone (Huawei Honor 8) which was installed the FootUI and an E-Reading app, and would be used in the study on the phone holder.

4.3 Procedure

4.3.1 Interactive Task. To simulate the real interaction process between users and smartphones, we designed an interactive task, reading and listening to an e-book with an e-reading app. This task not only covered the foot gestures that we designed, but also simulated the user's interaction behavior in the wild. We divided the task into ten sub-tasks (See Table 2) and allowed the participants to determine the order to complete them by themselves.

4.3.2 Learning FootUI. At the beginning of this study, we explained the foot gestures to the participants and helped them get familiar with FootUI by guiding them to leverage various foot gestures to complete the interactive task. The following are teaching steps. This phase takes about 10 minutes for each participant.

Step I: Helping participants calibrate the feet on the device.

Step II: Teaching foot gestures and corresponding smartphone operations. We divided the foot gestures into 2 groups based on their corresponding smartphone operations (See Fig 10) and taught them how to perform the foot gestures as well as their mapped smartphone operations. 1) Pointing based gestures with which users can move the cursor on the screen and trigger the pointed object. 2) Shortcut gestures, with which users can evoke some smartphone shortcuts.

Step III: Teaching the interactive task. We introduced the sub-tasks of the interactive task (See Table 2) to them and guided them to leverage foot gestures to complete the sub-tasks one by one. When errors occurred due to the incorrect foot gestures, we corrected their foot gestures and taught them how to handle the errors.

Step IV: Extra practice. After one round of teaching with practice, participants were allowed to practice the tasks and review the foot gestures. We answered all of the questions that they asked when they were practice.

4.3.3 User Test. When participants finished practice and were ready for the test, the host would give them a start instruction, start timing and observe their actions. To ensure the continuity of the interaction process, the host can remind the participants of their tasks or foot gestures when they asked for help, but the query for foot gestures was recorded as a memory error.

The upper body motor-impaired participants only need to complete one round of tasks with FootUI while the able-bodied participants need to complete the tasks with finger touch and FootUI each. The order of using the finger touch and FootUI was counter balanced and we recorded the whole task completion time of these two interaction methods. Timing for the interaction process with finger touch and FootUI was a "closed-loop" [7], which means if participants recovered from the errors occurred in the interaction process, the time for handling errors would be counted as a part of task completion time. The whole interaction process was video recorded for subsequent evaluation.

4.3.4 Post-Task Questionnaires and Interviews. After completing the above tasks, we asked the participants to fill out a questionnaire about the user satisfaction (1=the lowest user satisfaction, 5=the highest user satisfaction), fatigue (1=the most fatiguing, 5=the least fatiguing), ease-of-use (1=the most difficult to use, 5=the easiest to use) and learnability (1=the most difficult to learn, 5=the easiest to learn) regarding to FootUI and finger touch. Participants were required to rate the above four indicators on 5-point Likert scales. After that, we took an interview with each participant about their interaction experience and got some implications about the application of FootUI.

4.4 Result and Discussion

4.4.1 Accuracy and Memorability of FootUI. We labeled the operating steps and errors appeared in the tasks and found that errors were mainly occurred in three ways: 1) The cursor was not moved to the specified position. 2) The algorithm classified the foot gestures incorrectly so that FootUI evoked incorrect smartphone operations. 3) The participant confused some gestures or asked the host for their forgotten gestures. We defined the number of the first type of error as pointing error, the number of the second type of error as classification error and the number of the third type of error as memory error. Accordingly, we defined that 1) pointing accuracy = $1-\frac{pointingerror}{pointingsteps}$, where pointing steps denotes steps that participants move cursor for pointing, 2) classification accuracy = $1-\frac{classificationerror}{discretesteps}$

Table 2: Interactive task and operations

Start the e-reading app	
FootUI	Point to the e-reading app with the cursor (right foot dynamic continuous gestures) ->Trigger (left foot tap)
Finger Touch	Touch the e-reading app
Listen to a phonetic reading of a book	
FootUI	Point to a phonetic book with the cursor (right foot dynamic continuous gestures) -> Trigger (left foot tap) -> Move the cursor to the start button (right foot dynamic continuous gestures) -> Trigger (left foot tap)
Finger Touch	Touch a phonetic book -> Touch the start button
Tune the volume	
FootUI	Evoke volume bar (left foot rotate left + right foot rotate right) -> Slide the volume bar (right foot heel rotation) OR Evoke shortcut menu (left foot double tap or pointing+trigger) -> point to the volume button with the cursor (right foot dynamic continuous gestures) -> Trigger (left foot tap) -> Slide the volume bar (right foot heel rotation)
Finger Touch	Press the volume side-button
Suspend the phonetic reading	
FootUI	Point to the suspend button (right foot dynamic continuous gestures) -> Trigger (left foot tap)
Finger Touch	Touch the suspend button
Read the title page of the e-book	
FootUI	Point to the e-book with the cursor (right foot dynamic continuous gestures) -> Trigger (left foot tap) -> switch screen left (both feet rotate left)/switch screen right (both feet rotate right) or scroll down (press the scroll button or drag down)/scroll up (press the scroll button or drag up)
Finger Touch	Touch the e-book -> scroll left -> scroll right or scroll down ->scroll up
Tune the screen brightness	
FootUI	Evoke brightness bar (both feet tap once) -> Slide the brightness bar (right foot heel rotation) OR Evoke shortcut menu (left foot double tap or pointing+trigger) -> point to the brightness button with the cursor (right foot dynamic continuous gestures) -> Trigger (left foot tap) -> Slide the brightness bar (right foot heel rotation)
Finger Touch	Pull down easy access menu -> drag the brightness bar
Take a screenshot	
FootUI	Perform screenshot gesture (Heel to Heel) OR Evoke shortcut menu (left foot double tap) -> Move the cursor to the screenshot button (right foot dynamic continuous gestures) -> Trigger (left foot tap)
Finger Touch	Press the combination of side-buttons
Go to home screen	
FootUI	Perform home gesture (Toe to Toe) OR Evoke shortcut menu (left foot double tap) -> Move the cursor to the home button (right foot dynamic continuous gestures) -> Trigger (left foot tap)
Finger Touch	Press the home button
Read the notice and get back	
FootUI	Perform notice/easy access gesture (left foot Kick up) -> Perform back gesture (left foot rotate left) OR Evoke shortcut menu (left foot double tap) -> Move the cursor to the notice button (right foot dynamic continuous gestures) -> Trigger (left foot tap)-> Move the cursor to the back button (right foot dynamic continuous gestures)->Trigger (left foot tap)
Finger Touch	Scroll down the screen from the top -> Scroll up
Get into app backstage and return to the e-reading app	
FootUI	Perform app switch gesture (feet close up) -> Point to the e-reading app with the cursor (right foot dynamic continuous gestures) ->Trigger (left foot tap)
Finger Touch	Touch the app switch button -> Touch the e-reading app

and 3) memory accuracy = 1- $\frac{memoryerror}{discretesteps}$, where discrete steps denotes the number of discrete foot gestures that participants performed. The mean pointing accuracy is 96.75% (SD=8.36), the mean classification accuracy is 84.65% (SD=8.14) and the mean memory accuracy is 92.97% (SD=6.59).

The pointing accuracy showed that FootUI can track the foot stably and users can point the object on the interface accurately. The classification error showed the robustness of the foot gesture recognition algorithm remains to be improved. Another reason for classification error was that participants' gestures were not

Table 3: Subjective evaluation results of FootUI (from 1 - not good to 5 - good)

Statement	FootUI	Finger touch
user satisfaction	3.63 (SD=0.86)	4.44 (SD=0.70)
fatigue	3.38 (SD=0.78)	4.44 (SD=1.11)
ease-of-use	3.56 (SD=1.00)	4.44 (SD=0.79)
learnability	3.75 (SD=1.09)	4.19 (SD=1.18)

standard. Some of their subconscious feet movements also evoked shortcuts by mistake. Memory accuracy showed high memorability of FootUI, in consideration of short learning time.

4.4.2 Efficiency of FootUI. The efficiency is measured as task completion time. We compared the task completion time and task completion rate between FootUI and finger touch for more intuitive evaluation. Since the loss of finger touch data of two able-bodied participants. We collected 160 (16*10=160) trials for FootUI and 120 trials for finger touch. With finger touch, participants were able to complete all of the tasks (100%). With FootUI, participants were able to complete 95.63% (153 of 160) of the trails. Some participants failed to finish the task due to forgetting some sub-tasks and error-prone. The mean task completion time of finger touch is 96.33s (SD=32.04) and 198.10s (SD=86.07) of FootUI. The task completion time was calculated from participants who completed all the sub-tasks with both finger touch and FootUI (data from 5 participants was removed).

4.4.3 Subjective Feedback.

User Experience. Table 3 showed the average ratings of FootUI and finger touch on four indicators. Wilcoxon signed-rank test revealed that finger touch got significant higher ratings on user satisfaction, fatigue and ease-of-use with $p<0.01$. However, there are no significant diffenrece between FootUI and finger touch on learnability ($p=0.08>0.05$), which shows FootUI is easy to learn.

The result of the interviews revealed that the foot pointing was fatiguing and the shortcut gestures didn't bring much fatigue. P5 said, "Moving the cursor with right foot brings some burden to right foot" and he also said "Shortcut gestures bring no fatigue". P6 and P9 also indicated that pointing with right foot continuously was the main source of the fatigue and the shortcut gestures wasn't fatiguing. Some participants said that the FootUI is easy to learn and use. P7 said "Although the requirement for foot postures was a little strict, I felt the using experience was good when I got familiar with it." P3 said "I think this is a good way to help the people with upper body motor impairments use smartphones conveniently. The interaction process was smooth and the learning cost was relatively low." P5, P7 and P11 all expressed that using FootUI for smartphone interaction was novel and interesting.

Two upper body motor-impaired participants also expressed positive attitude to FootUI and confirmed its usability. P15 (both-arm amputee) said "when I get closer to the phone for carefully reading, my foot can't touch it and your software really solve this problem."

Suitable Usage. Since the both-arm amputees can use smartphones in different postures and using smartphones when reclining on the bed is not the only usage scenario of them, P15 and P16 said FootUI can't replace their current interaction method that touching smartphone screens with feet. P16 (both-arm amputee) said "I think the people with cerebral palsy and those who just lost their arms due to accident may need this technique much more than me." They also expressed that FootUI can help them interact with smartphones especially when they recline on the bed and browse the phone. P15 (both-arm amputee) said "This is convenient. When reading a novel or watching a video, you don't have to look down at the phone all the time, and the operation is relatively simple." P16 (both-arm amputee) expressed that FootUI is really suitable for e-reading and video watching. she said "Now it feels that reading novels and watching TV is quite convenient. I can turn pages or do other things with a simple move, which is pretty good." Some able-bodied participants also proposed that FootUI can be useful to able-bodied people in such scenario. P2 said that she would like to use FootUI while watching videos and eating concurrently on the bed. P11 said "I think it is also an interesting experience for able-bodied people. For example, when I cast my phone to the wall, using foot may better than touching the phone."

5 CONCLUSION AND FUTURE WORK

We present FootUI, a novel interaction technique that enables people with upper body motor impairments but sounds lower limbs to use smartphones reclining on the bed. We designed the foot gesture set for FootUI and found that toes-based gestures are not suitable for this scenario. The corresponding gesture tracking and recognition algorithm achieved an average accuracy of 90.53% for 8 gesture classifications. The usability evaluation shows that FootUI is easy, efficient and interesting to use and can help the people with upper body motor impairments interact with smartphone more smoothly and comfortably in such scenario.

There are some limitations that we will address in future work, such as: (i) Because of the difficulty for recruiting upper body motor-impaired participants, we recruited some able-bodied participants to our user study and the number of our target users was small. We will recruit more upper body motor-impaired participants to evaluate FootUI in the future. (ii) The evaluation of FootUI is rough, we will choose an appropriate baseline and evaluate FootUI thoroughly.

REFERENCES

[1] Jason Alexander, Teng Han, William Judd, Pourang Irani, and Sriram Subramanian. 2012. Putting Your Best Foot Forward: Investigating Real-World Mappings for Foot-Based Gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1229–1238. https://doi.org/10.1145/2207676.2208575

[2] Lisa Anthony, YooJin Kim, and Leah Findlater. 2013. Analyzing User-Generated Youtube Videos to Understand Touchscreen Use by People with Motor Impairments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 1223–1232. https://doi.org/10.1145/2470654.2466158

[3] Y. L. Chi, S. K. Ong, M. L. Yuan, and A. Y. C. Nee. 2007. Wearable Interface for the Physical Disabled. In *Proceedings of the 1st International Convention on Rehabilitation Engineering & Assistive Technology: In Conjunction with 1st Tan Tock Seng Hospital Neurorehabilitation Meeting* (Singapore) *(i-CREATe '07)*. Association for Computing Machinery, New York, NY, USA, 28–32. https://doi.org/10.1145/1328491.1328500

[4] Jon Froehlich, Jacob O. Wobbrock, and Shaun K. Kane. 2007. Barrier Pointing: Using Physical Edges to Assist Target Acquisition on Mobile Device Touch Screens. In *Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility* (Tempe, Arizona, USA) *(Assets '07)*. Association for Computing Machinery, New York, NY, USA, 19–26. https://doi.org/10.1145/1296843.1296849

[5] Xiaozhu Hu, Jiting Wang, Weiwei Gao, Chun Yu, and Yuanchun Shi. 2021. FootUI: Assisting People with Upper Body Motor Impairments to Use Smartphones with Foot Gestures on the Bed. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 436, 7 pages. https://doi.org/10.1145/3411763.3451782

[6] Christopher Kwan, Isaac Paquette, John J. Magee, Paul Y. Lee, and Margrit Betke. 2011. Click Control: Improving Mouse Interaction for People with Motor Impairments. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility* (Dundee, Scotland, UK) *(ASSETS '11)*. Association for Computing Machinery, New York, NY, USA, 231–232. https://doi.org/10.1145/2049536.2049582

[7] Guanhong Liu, Yizheng Gu, Yiwen Yin, Chun Yu, Yuntao Wang, Haipeng Mi, and Yuanchun Shi. 2020. Keep the Phone in Your Pocket: Enabling Smartphone Operation with an IMU Ring for Visually Impaired People. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 2, Article 58 (June 2020), 23 pages. https://doi.org/10.1145/3397308

[8] Katsumi Minakata, John Paulin Hansen, I. Scott MacKenzie, Per Bækgaard, and Vijay Rajanna. 2019. Pointing by Gaze, Head, and Foot in a Head-Mounted Display. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications* (Denver, Colorado) *(ETRA '19)*. Association for Computing Machinery, New York, NY, USA, Article 69, 9 pages. https://doi.org/10.1145/3317956.3318150

[9] Kyle Montague, Vicki L. Hanson, and Andy Cobley. 2012. Designing for Individuals: Usable Touch-Screen Interaction through Shared User Models. In *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility* (Boulder, Colorado, USA) *(ASSETS '12)*. Association for Computing Machinery, New York, NY, USA, 151–158. https://doi.org/10.1145/2384916.2384943

[10] Kyle Montague, Hugo Nicolau, and Vicki L. Hanson. 2014. Motor-Impaired Touchscreen Interactions in the Wild. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility* (Rochester, New York, USA) *(ASSETS '14)*. Association for Computing Machinery, New York, NY, USA, 123–130. https://doi.org/10.1145/2661334.2661362

[11] Martez E. Mott, Radu-Daniel Vatavu, Shaun K. Kane, and Jacob O. Wobbrock. 2016. Smart Touch: Improving Touch Accuracy for People with Motor Impairments with Template Matching. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1934–1946. https://doi.org/10.1145/2858036.2858390

[12] Martez E. Mott and Jacob O. Wobbrock. 2019. Cluster Touch: Improving Touch Accuracy on Smartphones for People with Motor and Situational Impairments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300257

[13] Francis Quek, David McNeill, Robert Bryll, Susan Duncan, Xin-Feng Ma, Cemil Kirbas, Karl E. McCullough, and Rashid Ansari. 2002. Multimodal Human Discourse: Gesture and Speech. *ACM Trans. Comput.-Hum. Interact.* 9, 3 (Sept. 2002), 171–193. https://doi.org/10.1145/568513.568514

[14] Guarionex Salivia and Juan Pablo Hourcade. 2013. PointAssist: Assisting Individuals with Motor Impairments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 1213–1222. https://doi.org/10.1145/2470654.2466157

[15] Jeremy Scott, David Dearman, Koji Yatani, and Khai N. Truong. 2010. Sensing Foot Gestures from the Pocket. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA) *(UIST '10)*. Association for Computing Machinery, New York, NY, USA, 199–208. https://doi.org/10.1145/1866029.1866063

[16] Shari Trewin, Simeon Keates, and Karyn Moffatt. 2006. Developing Steady Clicks: A Method of Cursor Assistance for People with Motor Impairments. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility* (Portland, Oregon, USA) *(Assets '06)*. Association for Computing Machinery, New York, NY, USA, 26–33. https://doi.org/10.1145/1168987.1168993

[17] Eduardo Velloso, Dominik Schmidt, Jason Alexander, Hans Gellersen, and Andreas Bulling. 2015. The Feet in Human–Computer Interaction: A Survey of Foot-Based Interaction. *ACM Comput. Surv.* 48, 2, Article 21 (Sept. 2015), 35 pages. https://doi.org/10.1145/2816455

[18] Jacob Wobbrock. 2003. The Benefits of Physical Edges in Gesture-Making: Empirical Support for an Edge-Based Unistroke Alphabet. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI EA '03)*. Association for Computing Machinery, New York, NY, USA, 942–943. https://doi.org/10.1145/765891.766083

[19] J. O. Wobbrock, B.A. Myers, and S. E. Hudson. 2003. Exploring edge-based input techniques for handheld text entry. In *23rd International Conference on Distributed Computing Systems Workshops, 2003. Proceedings.*

[20] Jacob O. Wobbrock, Brad A. Myers, and John A. Kembel. 2003. EdgeWrite: A Stylus-Based Text Entry Method Designed for High Accuracy and Stability of Motion. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology* (Vancouver, Canada) *(UIST '03)*. Association for Computing Machinery, New York, NY, USA, 61–70. https://doi.org/10.1145/964696.964703