Doctoral Thesis

WearPut: Designing Dexterous Wearable Input based on the Characteristics of Human Finger Motions

Hyunjae Gil

Department of Biomedical Engineering

(Human Factors Engineering)

Ulsan National Institute of Science and Technology

2023

WearPut: Designing Dexterous Wearable Input based on the Characteristics of Human Finger Motions

Hyunjae Gil

Department of Biomedical Engineering

(Human Factors Engineering)

Ulsan National Institute of Science and Technology

WearPut: Designing Dexterous Wearable Input based on the Characteristics of Human Finger Motions

A thesis/dissertation submitted to
Ulsan National Institute of Science and Technology (UNIST)
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Hyunjae Gil

January 9th, 2023 of submission

Approved by

Ian Oakley

Advisor
Ian Oakley

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

WearPut: Designing Dexterous Wearable Input based on the Characteristics of Human Finger Motions

Hyunjae Gil

This certifies that the thesis/dissertation of Hyunjae Gil is approved.

January 9th, 2023 of submission

Signature

Ian Oakley

Advisor: Prof. Ian Oakley

Signature

Prof. Kyungho Lee

Signature

이재연

Prof. Jaeyeon Lee

Signature

Seongkook Heo

Prof. Seongkook Heo

Signature

Jin Ryong Kim

Prof. Jin Ryong Kim

Abstract

Powerful microchips for computing and networking allow a wide range of wearable devices to be miniaturized with high fidelity and availability. In particular, the commercially successful smartwatches placed on the wrist drive market growth by sharing the role of smartphones and health management. The emerging Head Mounted Displays (HMDs) for Augmented Reality (AR) and Virtual Reality (VR) also impact various application areas in video games, education, simulation, and productivity tools. However, these powerful wearables have challenges in interaction with the inevitably limited space for input and output due to the specialized form factors for fitting the body parts. To complement the constrained interaction experience, many wearable devices still rely on other large form factor devices (e.g., smartphones or hand-held controllers). Despite their usefulness, the additional devices for interaction can constrain the viability of wearable devices in many usage scenarios by tethering users' hands to the physical devices. This thesis argues that developing novel Human-Computer interaction techniques for the specialized wearable form factors is vital for wearables to be reliable standalone products.

This thesis seeks to address the issue of constrained interaction experience with novel interaction techniques by exploring finger motions during input for the specialized form factors of wearable devices. The several characteristics of the finger input motions are promising to enable increases in the expressiveness of input on the physically limited input space of wearable devices. First, the input techniques with fingers are prevalent on many large form factor devices (e.g., touchscreen or physical keyboard) due to fast and accurate performance and high familiarity. Second, many commercial wearable products provide built-in sensors (e.g., touchscreen or hand tracking system) to detect finger motions. This enables the implementation of novel interaction systems without any additional sensors or devices. Third, the specialized form factors of wearable devices can create unique input contexts while the fingers approach their locations, shapes, and components. Finally, the dexterity of fingers with a distinctive appearance, high degrees of freedom, and high sensitivity of joint angle perception have the potential to widen the range of input available with various movement features on the surface and in the air. Accordingly, the general claim of this thesis is that understanding how users move their fingers during input will enable increases in the expressiveness of the interaction techniques we can create for resource-limited wearable devices.

This thesis demonstrates the general claim by providing evidence in various wearable scenarios with smartwatches and HMDs. First, this thesis explored the comfort range of static and dynamic touch input with angles on the touchscreen of smartwatches. The results showed the specific comfort ranges on variations in fingers, finger regions, and poses due to the unique input context that the touching hand approaches a small and fixed touchscreen with a limited range of angles. Then, finger region-aware systems that recognize the flat and side of the finger were constructed based on the contact areas on the touchscreen to enhance the expressiveness of angle-based touch input. In the second scenario, this thesis revealed distinctive touch profiles of different fingers caused by the unique input context for the touchscreen of smartwatches. The results led to the implementation of finger identification systems for distinguishing two or three fingers. Two virtual keyboards with 12 and 16 keys showed the feasibility of touch-based finger identification that enables increases in the expressiveness of touch input techniques. In addition, this thesis supports the general claim with a range of wearable scenarios by exploring the finger input motions in the air. In the third scenario, this thesis investigated the motions of in-air finger stroking during unconstrained in-air typing for HMDs. The results of the observation study revealed details of in-air finger motions during fast sequential input, such as strategies, kinematics, correlated movements, inter-fingerstroke relationship, and individual in-air keys. The in-depth analysis led to a practical guideline for developing robust in-air typing systems with finger stroking. Lastly, this thesis examined the viable locations of in-air thumb touch input to the virtual targets above the palm. It was confirmed that fast and accurate sequential thumb touch can be achieved at a total of 8 key locations with the built-in hand tracking system in a commercial HMD. Final typing studies with a novel in-air thumb typing system verified increases in the expressiveness of virtual target selection on HMDs.

This thesis argues that the objective and subjective results and novel interaction techniques in various wearable scenarios support the general claim that understanding how users move their fingers during input will enable increases in the expressiveness of the interaction techniques we can create for resource-limited wearable devices. Finally, this thesis concludes with thesis contributions, design considerations, and the scope of future research works, for future researchers and developers to implement robust finger-based interaction systems on various types of wearable devices.

Contents

List of Figures

List of Tables

List of Publications

This thesis is based on the following peer-reviewed publications which I am the primary author.

- **Chapter III: Fingers and Angles: Exploring the Comfort of Touch Input on Smartwatches**

  Hyunjae Gil, Hongmin Kim, and Ian Oakley, "Fingers and angles: Exploring the comfort of touch input on smartwatches," Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., vol. 2, no. 4, dec 2018.

- **Chapter IV: TriTap: Identifying Finger Touches on Smartwatches**

  Hyunjae Gil, DoYoung Lee, Seunggyu Im, and Ian Oakley, "Tritap: Identifying finger touches on smartwatches," in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ser. CHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 3879–3890.

- **Chapter V: Characterizing In-Air Eyes-Free Typing Movements in VR**

  Hyunjae Gil, Yonghwan Shin, Hyungki Son, Inwook Hwang, Ian Oakley, and Jin Ryong Kim, "Characterizing in-air eyes-free typing movements in vr," in 26th ACM Symposium on Virtual Reality Software and Technology, ser. VRST '20. New York, NY, USA: Association for Computing Machinery, 2020.

- **Chapter VI: ThumbAir: In-Air Typing for Head Mounted Displays**

  Hyunjae Gil and Ian Oakley, "ThumbAir: In-Air Typing for Head Mounted Displays," Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., vol. 6, no. 4, dec 2022.

Three publications (Chapter III, IV, and VI) in this thesis have been worked with the supervision of Ian Oakley. In Chapter III, Hongmin Kim supported the last user study that evaluated potential applications. In Chapter IV, DoYoung Lee and Seunggyu Im supported the development process of study implementations and provided constructive feedback on studies and artifacts. One publication (Chapter V) in this thesis has been worked with the supervision of Jin Ryong Kim. In Chapter V, Hyungki Son conducted study. Also, Hyungki Son, Yonghwan Shin, and Inwook Hwang supported data processing and analysis. I appreciate all of the dedicated supports for the researches in this thesis.

In addition, I have contributed to other publications during my Masters-Ph.D combined program at the Department of Biomedical Engineering (Major. Human Factors Engineering) at Ulsan National Institute of Science and Technology (UNIST). The following publications are not included in this thesis.

- Hyunjae Gil, Hyungki Son, Jin Ryong Kim, and Ian Oakley, "Whiskers: Exploring the use of ultrasonic haptic cues on the face," in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, ser. CHI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–13.

- Tae-Heon Yang, Hyungki Son, Sangkyu Byeon, Hyunjae Gil, Inwook Hwang, Gwanghyun Jo, Seungmoon Choi, Sang-Youn Kim, and Jin Ryong Kim, "Magnetorheological fluid haptic shoes for walking in vr," IEEE Transactions on Haptics, vol. 14, no. 1, pp. 83–94, 2021.

- Tae-Heon Yang, Jin Ryong Kim, Hanbit Jin, Hyunjae Gil, Jeong-Hoi Koo, and Hye Jin Kim, "Recent advances and opportunities of active materials for haptic technologies in virtual and augmented reality," Advanced Functional Materials, vol. 31, no. 39, p. 2008831, 2021.

- Yatharth Singhal, Haokun Wang, Hyunjae Gil, and Jin Ryong Kim, "Mid-air thermo-tactile feedback using ultrasound haptic display," in Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology, ser. VRST '21. New York, NY, USA: Association for Computing Machinery, 2021.

- Eunjee Kim, Hyunjae Gil, Jiwon Ryu, Ian Oakley, and Gwanseob Shin, "The effect of peripheral cues on motion sickness mitigation when using a vr hmd in a car," in Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 66, no. 1. SAGE Publications Sage CA: Los Angeles, CA, 2022, pp. 1967–1971.

- Jiwan Kim and Hyunjae Gil, "Top-levi: Multi-user interactive system using acoustic levitation," in The Adjunct Publication of the 35th Annual ACM Symposium on User Interface Software and Technology. New York, NY, USA: Association for Computing Machinery, 2022.

I would like to thank all collaborators for supporting me to go through the various fields of Human-Computer Interaction.

I Introduction

1.1 Wearables

'Wearables' are worn objects, such as clothing or glasses, that contain computer technology and/or can connect to the internet [1]. Recently, a wide range of wearable devices are thriving in the market. By combining high fidelity with low cost, devices such as smartwatches, fitness trackers, smart clothes, smart jewellery, or Head Mounted Displays (HMDs) for Virtual Realtiy (VR) and Augmented Reality (AR) are proving to be invaluable tools to support diverse set of user activities.

As demands in the healthcare and Internet-of-Things (IoT) areas drive growth, the market size of wearables is predicted to grow to approximately USD 265 billion by 2026 with a CAGR of 18% [2]. Powerful microchips for computing and wireless networking are fundamental in enabling the market growth by allowing wearables to be miniaturized with high reliability and availability and placed in any body parts with flexible form factors. Especially, the growing demand for the smartwatches and HMDs is resulting in the launch of numerous commercial products.

Smartwatches

Smartwatches are the most mature and commercially successful wrist wearable. Recently, pairing smartwatches with smartphones became a popular way of sharing or subsidizing the role of smartphones. They allow users to quickly access the frequent tasks such as visual [3] and haptic notifications [4], sound alerts [5], or short text information [6], during other activities. Through simple manipulations on a touch screen, many useful and productive applications are available such as music [7], health management [8], mobile wallet applications [9], text message reply [10], or calling [11]. In addition, internal sensors measure user's behaviors while exercising, travelling or resting by tracking their journey [12], steps [13], consumed calories [14], or sleep patterns by measuring movements [15], breathing [16], and heart rate [17]. Contact with user's body enables continuous monitoring of bio-signals, such as heart rate [18], blood pressure [19], or blood sugar level [20], in daily life. Similar to smartwatches, fitness trackers worn on wrist have increased the scope of healthcare monitoring services. These devices primarily manage a fitness journey by providing exercise tips or statistics [21], and some health monitoring trackers enable patients to directly communicate bio-data to healthcare professionals remotely [22].

Head Mounted Displays

Another emerging wearable form factor is Head Mounted Displays (HMDs). Various commercial products target application areas such as video games [23], media [24], entertainment [25], medicine [26], education [27,28], simulation [29], training [30], or productivity tools [31]. There are two basic types of HMDs: those that focus on Augmented Reality (AR) and those that address Virtual Reality (VR) [32]. AR HMDs provide an interactive experience with virtual contents altering users' perception of the real world environment in real time through the technologies of optical or digital see-through displays, complex 3D trajectory tracing, and hand tracking [33]. AR HMDs have high potential to be applied in various scenarios from supporting daily life to specialized industry settings [34]. The virtual contents shown on AR HMDs can be used to present visual information about or manipulate the status of real objects for the purpose of entertainment, social networking, or education in daily life [35]. Due to the benefits of hands-free and overlays, AR HMDs are also useful in industrial/medical fields as they can support interactive visual guidance or data visualization [36]. Users commonly interact with virtual contents shown in AR with hand gestures, voice commands, or additional controllers [37]. VR HMDs provide a virtual environment that immerses users in their surroundings. VR HMDs commonly consist of a stereoscopic display, head motion tracking sensors, and stereo speakers [38]. Some VR HMDs include hand tracking [39] or eye-tracking sensors [40]. Many consumer VR HMDs are standalone devices that do not require additional equipment to track the position of the HMDs or their controllers [41]. They focus on ease-of-use and high quality visual presentation. Various advances in technology and digital platform market have contributed to the popularization of VR HMDs. Especially, a desire for more realistic and immersive experiences has lead the development of VR. For instance, haptic feedback can be delivered to the skin through the hand-held controller or various types of contact and non-contact haptic device [42]. Users can interact with virtual contents by using a touch surface or buttons on a hand-held controller [43–45], with bare hands via an embedded hand tracking system [45–47], or accessories in the form of gloves, shoes, guns, or treadmills [48–51].

These powerful wearables are capable of assuming various functions previously performed only by handheld devices such as smartphones. However, many wearables still rely on the smartphones for core functionality. This is because there remain substantial technical challenges to be achieve fully standalone wearable devices including the hardware size, battery life, and supported networking system. Besides these technical issues, many wearables expose core challenges in input and interaction - they provide highly limited and constrained interaction experience to their users [7, 52–54]. This is because their miniaturized and specialized form factors, designed to fit the body, inevitably feature very limited spaces and surfaces for input and output. This is one major reason why many wearable devices still rely on the other large form factor devices for input and interaction. For example, smartwatches are typically paired with smartphones and HMDs operated by hand-held controllers. Challenges in achieving effective input and output represent a major bottleneck restricting the viability of wearable devices for many meaningful usage scenarios. Accordingly, this thesis argues that, for wearables to be successful standalone products, it is currently imperative to develop novel Human-Computer interaction techniques specific to the wearable form factor.

1.2 Input on smartwatches

Smartwatches face the challenge of providing a constrained interaction experience to users due to its small form factor. First of all, the smartwatches occupy one hand, which restricts to reproduce the performance of touch input using two hands in large form factor devices [55]. In addition, they inevitably carry the physically small touchscreen for both input and output. The small touchscreen of smartwatches yields the relatively small size of graphics. The small graphics of text and icon require higher visual attention and cognitive workload to react by touching them with fingers [56]. The touch input on the small touchscreen also causes an issue of fat-finger that the relatively large size of finger tip rather than the target size results in incorrect input by occluding the visual feedback during the touch [57].

A common solution is enlarging targets to prevent the fat-finger issue, but it results in only few targets being displayed on the small screen. Many prior works have explored novel interaction methods to avoid this fat-finger issue by extending input spaces to the alternative areas near the device, such as touching the edge [58], strap [59], or crown [60], while others have used the finger/hand gestures [61, 62], around-device skin touch [63], or physical movement of smartwatches [64]. However, the additional sensors/wearables can disrupt smartwatches to be successful standalone products requiring a high portability. The novel input space different from where visual feedback appears can also confuse users to perform input accurately.

The dynamic touch methods widely used in large form factor devices (e.g., scrolling, switching, or zooming) also causes the fat-finger issues by occluding contents with the moving fingers on the small touchscreen. To complement it, researchers have explored to enrich the touch interaction by improving the expressivity of the static touch input, such as force-based touch [65] or sequences of tapping [66]. While they can be an alternative for the traditional gesture inputs, the novel interaction techniques require to be memorized and trained. Another approach for adopting the gestural touch input is a rotational gesture input at the edge of the touchscreen [67]. In spite of the usefulness, it can arouse an issue of discomfort during the input due to the unique input context of smartwatches. The touching fingers approach the fixed touchscreen on the opposite hand. It makes a specific touching pose that limits the range of comfort angle for the rotational inputs.

Unsurprisingly, the fat-finger issue on smartwatches constrains high bandwidth tasks that require a wide range of input available such as text entry. The availability of the high bandwidth tasks is essential for smartwatches to be successful standalone devices. Many prior works have proposed novel interaction techniques to achieve the high bandwidth tasks with additional finger-mounted wearables [68] or the sequences of tapping [69], swiping [52], or zooming [70] on the touchscreen. However, the additional sensors/wearables can be cumbersome or unreliable in many usage scenarios. The untraditional touch techniques can demand an extra training and more execution time with additional input steps. Therefore, this thesis argues that novel interaction techniques for achieving high bandwidth tasks need to consider both the prevalent input skills and accessible sensors in smartwatches.

1.3 Input on HMDs

The HMDs are relatively free from visual occlusion problem while the touch input on the surface of HMDs, since the display for an output space is separated with the input space. However, the small size of HMD devices constrains direct touch input on the hardware surfaces. Various touch methods have been proposed by using the side or front parts of HMDs based on back-of-device touch interaction [71], such as a character drawing [72], cursor-based input [73], or swipe [53]. Despite utilizing the limited spaces for input, they can cause an issue in memorability due to the complex sequences of coded tapping or swiping. The touch input around head is not visible, which is an inherent issue that requires eyes-free input. The touch gestures directly on the HMDs can also cause physical discomfort on nose by shaking the hardware [74]. In addition, the long exposure to the touch input at the head location causes the muscle fatigue known as a gorilla arm effect [75]. For these reasons, the small touch space on the surface of HMDs constrains the touch interaction to achieve the acceptable performance for the desired tasks in three dimensional space such as typing or drawing.

Instead of the touch input providing the limited interaction experience, a manipulation with hand-held controllers is a common method from simple to high bandwidth tasks in the three dimensional space of the HMDs, such as text entry [45] or 3D drawing [76]. The ray-casting methods with the controllers [43, 77] support to point and select the virtual objects even at a long distance [78]. The diverse input techniques with the physical joystick and buttons on the controller have been proposed by integrating with the ray-casting method [79]. In addition, the methods of directly colliding with the hand-held controller is useful to operate the virtual objects around the user [80]. However, the ray casting methods have an issue of occlusion that the other objects can screen the target in dense environments [44, 81]. The direct collision methods support to interact only with the close targets still having the occlusion issue. Also, the selection methods based on the pointing are inefficient to make a fast sequential inputs such as text entry [45, 82]. In addition to the interaction issues, the occupied hands tether users who use additional devices or do other activities while wearing the HMDs. Especially, in the case of wearing AR glasses in daily life, the use of additional input devices causes inconvenience to perform a main task with the assistance of AR contents.

Recently, many prior works have investigated how to extend the input space without holding additional input devices with hands, such as eye gaze [83], head movement [84], on-body touch [85], or bare-hand gesture [86]. Among the various approaches, the bare-hand interaction using an embedded or add-on hand tracking systems is emerging as an alternative of the hand-held controllers. It supports to directly manipulate the virtual objects or make the hand gestures with the projected virtual hands in mid-air. In an early stage of hand tracking system, users make a favorite use of the predefined hand gestures which trigger the simple short-cut commands [46], but each predefined gesture needs to be memorized. Other prior works have proposed novel bare-hand interaction methods using the state-of-the-art optical tracking systems to manipulate the virtual objects [87]. However, the commercial hand tracking systems are not mature enough to imitate the exact positions of actual hands in real time due to the tracking errors in the finger joints [39]. It is insufficient to track the subtle movement which is necessary to achieve the high bandwidth tasks with small and dense targets. In additional, the hand gestures can generate a muscle fatigue on the shoulder and arm [75], and meet an issue of social acceptability that users are willing to use them in public places [86].

Nonetheless of the early-stage hand tracking system, the bare-hand interaction is good fit for the various applications of HMDs in the three dimensional environments. It is also an accommodating interaction technique for users to induce the intuitive movements to manipulate the virtual objects as we do with real hands. This traditional manipulation methods with bare-hands can lower the entry barriers for users of all ages rather than novel input methods with the hand-held controllers. In addition, the dexterous input can facilitate high bandwidth tasks with high performance. Many prior works have proposed novel bare-hand interaction techniques for the HMDs, such as error-correction techniques [88] or recognition methods [47]. However, most of prior works used the high-end optical trackers, which limits to utilize the interaction techniques in the currently available HMDs. They also controlled the hand or finger movements to get the clear features for the hand gesture detection or recognition [47], but the users move their hands without any constraints for the gesture input in real usage scenarios. Accordingly, this thesis argues that novel bare-hand interaction techniques for high bandwidth tasks need to consider both the unconstrained finger movements and the available hand tracking systems in commercial HMDs.

1.4 Properties of Finger for input

The finger-based input techniques are prevalent on many large form factor devices such as smartphones or physical keyboards. It has already been verified that fingers can achieve fast and accurate performance for complex or meticulous interaction tasks. With the advantages of prevalent finger input methods, the characteristics of fingers in anatomy and movements have the potential to generate various unique input features for the specialized form factors of wearable devices. Fingers consist of segments with the joints and associated muscles. They have their special shape at each segment and distinct regions such as nails, fingertips, pads, or knuckles. In addition to the unique appearance, the fingers can move very freely with their joints [89]. They have a kinematic model which includes a total of 21 degrees of freedom. The index, middle, ring, and little finger have 4 degrees of freedom. They have one degree of freedom only for flexion and extension motion with the proximal interphalangeal (PIP) or distal interphalangeal (DIP) joints, and two degrees of freedom for flexion, extension, abduction, or adduction motion in the metacarpophalangeal (MCP) joint. Especially, the thumb has five degrees of freedom: two for each of the trapeziometacarpal (TM) and metacarpophalangeal (MCP) joints, and one for the interphalangeal (IP) joint. In terms of the joint angle perception, the Just Noticeable Difference (JND) is around 1.7° to 2.7° at the proximal interphalangeal (PIP) and metacarpophalangeal (MCP) joints [90]. Based on the dexterity of fingers, I argue that the potential input features that can be extracted from the characteristics of fingers in appearance, anatomy, and perception are promising to increase the expressiveness of traditional finger input techniques on wearable devices.

In terms of finger input on the surface, the prevalent touch input performs fast and accurate target selection on the large form factor devices. However, the form factor of smartwatches has a small display which limits the input bandwidth and also generates the interaction issues of fat-finger [57] during the touch. To address the inherent issues, many prior works have explored novel input modalities with properties of finger motions during the touch, such as the physical movements of smartwatches by the impact force [91], the force/pressure level on the surface [65], the impact sounds caused by different finger regions of the nail, fingertip, pad, or knuckle [92] or the placements of each finger while the touching finger approaches to the touchscreen [93]. However, unpredictable or uncontrollable input noises from the variability of force or sound can degrade the performance of the proposed input techniques with the low recognition performance in the actual use scenarios. Meanwhile, the capacitive sensors in touch-screen are reliable to accurately estimate finger motions on the surface without any undesired input noises. The finger motions with the information of contact area on the surface can improve the input expressivity of the touch, since the different anatomy and skin shape of each finger produce distinct touch profiles on the surface. Many prior works have focused on the properties of finger movements on the large form factor devices. The contact area during touch input enables the finger orientation estimation with the touch geometry [94] or the finger identification with the relative finger positions [95] or the fingerprint [96]. However, little work has explored the finger motions during touch on smartwatches. The touch motions on the small and fixed surface of smartwatches have the potential to generate unique input features which can increase the input bandwidth for smartwatches.

Meanwhile, the fingers with a high degree of freedom are able to shape a bunch of hand gestures for the input, such as meaningful gestures [86], pinch [97], or single finger touch [88]. However, the issues of immature hand tracking systems in the HMDs limit the use of skillful finger stroking that is prevalent on the physical keyboard [39]. The finger stroking is simple, and also easy to perform the sequential inputs efficiently. In spite of the high potential, it is usually adopted for simple selection since the immature hand tracking system in commercial HMDs is insufficient to track all fast and subtle finger movements during the sequential inputs [39]. Even, a lack of haptic for the key confirmation causes the tough controls on the finger movements that generate different kinematics from the finger stroking to the surface [98–100]. Prior works have explored novel interaction techniques by using error-correction techniques [88] or developing new hand gestures [86]. However, little work has explored characteristics of unconstrained in-air finger stroking. Previous study that explored the constrained in-air finger stroking have yet to reveal sufficient features to precisely detect or recognize the fast and complex finger stroking inputs [47].

In addition, the anatomy of the hand has a confined range of motion for the rotational movements of the finger and wrist in abduction/adduction and flexion/extension. The input that deviates from the comfort zone affects the ergonomic comfort that impacts both usability and performance. Therefore, the design of finger input techniques needs to consider comfort based on the properties of finger movements. Previous study has explored the comfort during the touch input on the surface of large form factor devices [101]. However, the unique input context of smartwatches can generate different comfort ranges of finger movements during the angle-based touch input on the touchscreen or bezel. The bare-hand touch input in the air is also associated with the range of finger motion. The interface for the direct touch needs to consider the comfortably reachable arrangement of multiple virtual targets [102, 103].

In sum, finger motions during input have the potential to generate various novel input features and modalities. In particular, understanding the finger motions for the specialized form factors of wearable devices can enable identify unique input features with the reliable built-in sensors. Accordingly, this thesis argues that the properties of finger motions during input are able to generate unique input features that can improve the expressiveness of the traditional finger input techniques by avoiding the interaction and technology issues of the resource-limited wearable devices.

1.5 Claim

Among various approaches to the input for wearables, I focus on exploring how to increase input bandwidth by improving the expressiveness of the traditional finger input techniques, such as tapping or swiping on the surface and finger stroking in the air. This approach enables the novel interaction techniques to retain the prevalent input skills with high familiarity and learnability. In addition, commercial wearables support the detection technologies of finger movements with built-in sensors such as the touch surface or hand tracking system. Within this space, the finger motions have the potential to generate a range of unique input features during input based on the high degree of freedom, complex anatomy, distinct skin shape, and sensitive joint angle perception. The inputs on the special form factor devices can provoke somewhat different finger movements due to the unique input contexts of wearables. Therefore, understanding the properties of finger motions while manipulating wearables is fundamentally required to reveal the unique input features. by overcoming the addressed interaction issues generated from the form factors of wearables. In particular, the unique input features can improve the input expressiveness of novel interaction techniques that can enable a high bandwidth task requiring a fast sequential input and a wide range of input available on the resource-limited wearable devices. Accordingly, this thesis addressed a general claim that ***understanding how users move their fingers during input will enable increases in the expressiveness of the interaction techniques we can create for resource-limited wearable devices***.

1.6 Outline, Contribution, and Conclusion

This thesis demonstrates the general claim with the diverse approaches to the input scenarios of both smartwatches and HMDs. This thesis starts with an overview of various prior works related to the study background and interaction issues of wearables. Then, the empirical studies examine the properties of finger movements during input both on the surface and in the air to derive the unique input features. Firstly, this thesis explores the comfort range of angle-based touch on the small touchscreen of smartwatches during the static and dynamic touch inputs. Then, this thesis investigates the classifiable touch profiles generated by the unique input context of smartwatches to build finger identification system. Secondly, in order to examine the general claim of this thesis on a range of wearable scenarios, this thesis analyzes the movements of unconstrained in-air finger stroking that can enable the fast sequential input tasks such as text entry. Then, this thesis specifically explores the properties of in-air thumb movements based on the comfort range of thumb motion, input performance, muscle fatigue and social acceptance to develop a novel text-entry system for the HMDs. Finally, this thesis provides the design considerations, limitations, and future works on using the properties of finger movements for developing novel interaction techniques on resource-limited wearables. More specifically, in **Chapter II**, this thesis reviews the background of input techniques smartwatches and HMDs through various prior works. The previous works have proposed novel interaction techniques to tackle the interaction issues caused from the wearables' form factors by exploring novel input modalities, developing detection technologies of input, or improving expressivity of existing input skills. The input techniques for smartwatches consist of the methods of touch on screen, touch on alternative parts of device, around-device gestures, and one-handed gestures. The input techniques for HMDs include the touch on the surface of device or body parts, the manipulation methods of hand-held devices, hands-free methods of eye gaze or head-based motion, and hand gesture methods on the surface or in the air. In particular, the review for the HMDs includes the subtle hand interaction techniques considering social acceptability for the usage scenarios in public space. The following paragraphs describe contributions of this thesis based on the empirical studies for exploring the properties of finger movements during the input on the surface and in the air.

Chapter III investigates comfort angles of touch input on the small touchscreen of smartwatches. The finger touch input in the smartwatches has a unique input context that the touching fingers approach the small touchscreen fixed on the opposite hand. The specific hand pose limits a range of approachable angles of the finger input due to the anatomy of the hand and arm. The limited angles can arouse a physically uncomfortable input when the static or dynamic touches based on the angle are out of comfort range. The discomfort also diminishes both the reliability and usability of input techniques in achieving the expected performance. Prior work has investigated the comfort input ranges with the large-scale display of smartphones using two hands [101]. However, I highlight that there is no data or guideline about the comfort of touch input for both the static and dynamic gestures on the smartwatches. To fill this research gap, investigation of the comfort of finger movement during the touch is firstly necessary to achieve a robust and reliable interaction technique. This thesis explores the comfort range of static and dynamic touch input with different finger poses through empirical studies. In addition, the raw touch images are used to develop region-aware classifiers that infer the contacted finger regions. Finally, guidelines and example applications for the angle-based input and region-aware interaction techniques with the specific comfort range are suggested based on the subjective feedback on the comfort and objective input performance.

Chapter IV introduces a finger identification system using finger touch profiles generated by the unique input context of smartwatches. The contact area of finger inputs on the touchscreen is a reliable feature with the capacitive touch sensors. The touch profiles obtained by the raw touch image include the information on finger contact distribution on the touchscreen [94]. The individual fingers can generate distinct touch profiles in the contact area due to the different appearance of fingers and the unique input context of the smartwatches where the touching hand approaches the fixed touchscreen on the opposite wrist. Prior works have explored a finger identification to access different functions, but they remain challenges in the practicality by using either additional implementations [68, 104–106]. The finger identification based on the touch profiles of fingers can fill this research gap without any additional implementations. This thesis conducts empirical studies to collect the touch profiles on the touchscreen of smartwatches through the tapping and swiping tasks. Then, this thesis proposes finger identification models based on the distinctive touch profiles of each finger in natural and exaggerated touch poses. The models enable the recognition of the finger that is responsible for the touch. As an application of the finger identification system, this thesis suggests the keyboard interfaces with a novel button design that allows different fingers to have different functions on a button. The keyboard for complex typing tasks shows increases in the expressiveness of touch input on the small touchscreen of smartwatches.

Chapter V explores the movements of in-air finger stroking during the typing on HMDs. The built-in Hand tracking systems in the commercially available HMDs have the potential to alternate the hand-held controller-based interaction that occupies the user's hands to bare-hand interaction. However, the immature hand tracking system limits the fast in-air finger stroking for high bandwidth tasks such as text entry. Also, a lack of haptic on the fingers causes the fingers to move differently in the air than on the surface. Prior work explored the finger motions during constrained in-air finger stroking [47], while the results have yet to reveal sufficient features to accurately detect and recognize unconstrained finger stroking for actual usage scenarios. Therefore, this thesis explores unconstrained in-air finger stroking during typing tasks. The empirical studies identify two typing strategies that affect in-air finger movements. Based on the strategies, this thesis analyzes the kinematics of in-air finger stroking (e.g., amplitude, direction, location, and speed of fingertip), the correlated movement of fingers, the inter-finger stroking relationship, and in-air key distributions. In the end, this thesis suggests a design recommendation for developing a text entry system using in-air finger stroking.

Chapter VI examines the movement of the thumb to develop an in-air thumb typing system for HMDs. The touch with the thumb is a prevalent input method on large form factor devices. This thesis highlights that the prevalent thumb input can enable efficient text entry on the HMDs. To avoid the hand-held controllers occupying the hands, this thesis uses the built-in hand tracking system that supports the bare-hand finger input. However, the immaturity of the hand tracking system seldom achieves the subtle manipulation of the small virtual targets near the hands. In this limitation, understanding the properties of thumb movements is necessary to arrange the virtual targets in the appropriate location. This thesis explores the properties of thumb movement during the thumb input on the small targets around the hands. The empirical studies draw the highly effective location of eight targets for the in-air thumb input based on the effective range of thumb movements. A layout selection process assigns the characters to the eight buttons. This thesis verifies the in-air thumb typing system with word repetition and phrase typing tasks. Through a comparison study with a baseline keyboard, this thesis examines the in-air thumb typing system with objective typing performance and subjective feedback for muscle fatigue and social acceptance.

The contributions of this thesis are threefold through the background reviews and a series of empirical studies for finger input motions on various scenarios: smartwatches and HMDs. Firstly, this thesis contributes to understanding the properties of finger movements that can enable increases in the expressiveness of finger inputs for various scenarios. Chapter III suggested the specific comfort ranges of angle-based touch inputs on the small touchscreen of smartwatches and also the region-aware input techniques using the contact areas. Chapter IV presented the distinctive touch profiles of each finger by the unique input context of smartwatches. Chapter V analyzed the characteristics of finger stroking during unconstrained in-air typing such as finger stroking strategies, kinematics, correlated movements, the interval between two sequential inputs, and individual in-air key analysis. Chapter VI presented the viable target locations of in-air thumb touch input on the HMDs. Also, this thesis provided detailed guidelines for these features to build novel finger input interfaces. Secondly, this thesis has an empirical contribution to data collection. Various empirical studies with natural or artificial conditions provided objective and subjective data of finder motions. For example, Chapter V provided 25,932 manually labeled finger motions to each in-air key and high-level information of in-air finger stroking through an observation study for unconstrained in-air typing, while Chapter VI provided all viable locations of thumb touch input above palm. The empirically collected data will be a ground-truth to design comfortable, accurate, and fast input techniques based on finger motions. The third contribution of this thesis is the artifacts of novel finger input interfaces that increase the expressiveness of finger input for resource-limited wearable devices. Two novel finger interaction techniques showed increases in the range of input available on the small touchscreen of smartwatches. Chapter III proposed potential applications with the integration of angle-based touch techniques and a finger region-aware system. Chapter IV introduced virtual keyboard interfaces based on the finger identification system that enabled the keyboards to have two or three letters on each key. This thesis also presented a novel in-air typing system for HMDs by improving the expressiveness of in-air thumb touch input. In Chapter VI, an in-air thumb typing system with an ambiguous keyboard layout and word prediction system showed fast and accurate typing performance with eight keys above the palm of both hands.

In conclusion, I argue that the contributions support the general claim of this thesis by examining various scenarios that the novel interaction techniques reflecting the properties of finger movements on smartwatches and HMDs can improve the input expressiveness on complex typing tasks for resource-limited smartwatches and HMDs.

II Related Work

This chapter presents an overview of prior input techniques and their sensing technologies for both smartwatches and HMDs to review a range of input scenarios on the surface and in the air. This review also describes the advantages and limitations of previous works considering the aspects of usability, practicality, social acceptability, and feasibility. The first section summarizes various prior approaches using hands to the interaction issues of popular touch input on smartwatches. This section starts by presenting the touch input techniques using the touchscreen of smartwatches such as sequences of touch gestures, the status of fingers, or additional finger-mounted wearables. Other touch techniques that extend the touch surface using the alternative parts of the device such as the bezel, strap, or crown are also described. Then, I introduce around-device input techniques for avoiding occlusion on the touchscreen. The around-device input includes hand or finger gestures or skin touch near the device with various recognition technologies. The last approach described in this section is the one-handed gesture input method that uses the occupied hand by a smartwatch to perform the hand gestures. The second section reviews various interaction techniques and input modalities for AR and VR HMDs. This section begins with a summary of input techniques using physical Hand-held devices that are popular on commercial HMDs. After this, I describe alternatives to hand-held devices. One approach is the touch input technique on the surface of HMDs and the body parts. Another approach is the hands-free input technique using head-based motion or eye gaze which is promising for interaction scenarios in uninstrumented environment. This section closes with a description of in-air hand/finger input techniques and also subtle manipulation methods that consider the social acceptability in public space scenarios.

2.1 Input Techniques on Smartwatches

Touch on the touchscreen

While the touch input is a prevalent and reliable input method for most smart devices, smartwatches are struggling for employing the traditional touch input on the physically limited space for input and output. To overcome the limitation, many researchers have explored how to improve the input expressivity of finger touch based on the already implemented touchscreen. First and foremost, smartwatches display a limited number of soft buttons due to the small touchscreen, which hinders high bandwidth tasks requiring many buttons such as a text entry. Many prior works have proposed various approaches to address this issue. For example, the multiple functions per button allow a large number of input options on the small space by retaining the effective size for touch. The input methods using the sequences of tap and swipe gestures to the different directions enabled the text entry with seven [107] or nine [52] buttons. Other works adopted the ambiguous button with several characters and interpreted the sequential taps [69, 108] or the swiping gestures [109] to the most likely word by auto-correction or prediction system. While the input techniques based on the tap-and-swipe and ambiguous buttons are potentially effective systems, they require more steps to finish the intended inputs. Huang *et al.* [110] assigned multiple functions per button to the different contact areas of left, right, and center of the finger pad, which can omit any additional inputs. Others have investigated the zooming interface to enlarge the soft keyboard with tiny buttons to the effective size for selection [70], or to display a pop-up window showing an enlarged touch region [111]. Despite a more familiar interface, it can demand additional swipes to shift the touching fingers to the target exactly.

The issue of fat-finger frequently occurs on the small touchscreen of smartwatches [57, 112]. The relatively large finger decreases the touch accuracy by occluding the target or visual feedback. To address this issue during the touch input, prior works have explored to improve the touch input expressivity by avoiding visual occlusion. Recognizing finger angles on the surface [113, 114] can be effective in broadening the input bandwidth of smartwatches using the input context of angles. Xiao *et al.* [94] proposed diverse applications by estimating the pitch and yaw of touching finger on the smartwatches, such as twist, pan, zoom, rotate, and 3D manipulation. Vogelsang *et al.* [115] discussed the design space for the finger angle input to present the interface elements by mapping with various use cases. In spite of improving the input expressivity, the angle-based input techniques need data refinement for the comfort and social acceptability. Another effective approach for augmenting the touch input is using a continuous pressure touch at the initial touch position [116, 117]. Yeo *et al.* [65] suggested the pressure input interface with three levels of pressure in eight directions. While showing the feasibility of pressure input on smartwatches with high recognition accuracy, it had a relatively long input time about 2.4 seconds. In addition to the bare-finger touch, other authors have implemented additional input devices for the touching finger to address the fat-finger problem. NanoStylus [118] used a small finger-mounted stylus that significantly reduces the error rate compared to the traditional finger touch. Su *et al.* [119] suggested an additional display on the nail of touching finger. The nail display reveals the occluded area and guides the touching finger to the target location. While the additional devices accomplished precise touch input on the small touchscreen, they may not fit the use cases of smartwatches by occupying the tip of the touching finger.

The benefit of smartwatches is quick access to messages and visual or haptic notifications while mobile or other activities. In line with the use cases, the short input time is vital to respond to them instantly. Eyes-free finger touch input is a good fit for this desire. A tactile landmark on the bezel can enable the eyes-free input [120]. Kubo *et al.* [121] used the bezel-to-bezel swipe gestures on the rectangular touchscreen by using the physical characteristics of screen shape as the landmark. It accomplished the quick gesture input due to the short distance between the two bezels. Another work showed that the bezel-to-bezel swipe gesture can also minimize a screen occlusion during the touch interaction through an investigation of contents visibility [122]. To achieve fast eyes-free input, others adopted the sequences of tapping [7] or the contact area of flat fingers [66] to command the shortcut menu by mapping with action metaphors. However, the eyes-free techniques based on swiping and tapping remain to command the shortcut menu, not to perform more complex tasks. In addition to fast input, on-the-go interaction in the actual use cases is a key consistent with the benefit of smartwatches. While mobile, the problems of physical disturbance [123] and increased cognitive load [124] frequently occur during the on-the-go interaction, which reduces the input performance more for the small targets [125, 126]. Singh *et al.* [127] investigated the impact of mobility on various touch input methods, and suggested new navigation techniques for mobile contexts by adopting tapping and flicking for zooming and panning. Ahn *et al.* [126] discussed more complex tasks while mobile. The HoldBoard system employed the watch holding motion with the sequences of two finger touches—tapping the target button with the index finger while touching the bezel with the thumb. While the prior works represent empirical results relating to mobility, the on-the-go interaction needs to consider other confound interaction problems such as the fat-finger issue.

Touch on the alternative parts of device

The finger touch on the small touchscreen remains the issue of screen occlusion during the touch. In addition, soft button for manipulation can obstruct the view of visual contents in the small output space. To completely avoid any visual possession for input, alternative parts of smartwatches have been examined as an input space, such as edge, strap, or crown. Among the various parts, the bezel, the edge of the smartwatch, is the most relevant part for input as it is located right around the screen. Oakley *et al.* [58] examined the viability of edge touches with capacitive sensors around the bezel by evaluating the targeting performance of single and spatial multi-touches. Other authors detected a single edge touch with the built-in Inertial Measurement Unit (IMU) [128]. Darbar *et al.* [129] investigated pressure-based input on the bezel to enhance the input expressivity, augmenting the touch to three pressure levels. Another approach is rotating the bezel to select the button with cursors [67]. It used multiple cursors based on the prefix system to minimize the rotating distance. In terms of eyes-free input using the bezel touch, Wong *et al.* [130] showed the feasibility of up to an eight-segment layout with a mean accuracy of over 93%. Despite avoiding the issues of visual occlusion, the bezel interface is somewhat circular smartwatch-centered design, which likely needs more investigation for rectangular smartwatches.

The strap is a part of smartwatches with a relatively large space for making input without visual occlusion. Perrault *et al.* [59] explored a design space of eyes-free touch input on the strap with simple pointing and sliding gestures. They discussed the strap gesture techniques more suitable for shortcuts. A more complex task was examined with the arrangement of multiple touch buttons along the strap [131]. During the task, the strap touch interface revealed an interaction issue that the screen is out of view by the rotated arm while touching the button away from the screen. Other authors have proposed strap interfaces having both input and output spaces. The ambient notification on the strap can enable the back-of-device interaction with precise target selection [132]. Lyons *et al.* [133] and Klamka *et al.* [134] demonstrated the feasibility of back-of-device interaction by implementing strap interfaces consisting of multiple displays that facilitate all strap to be input and output spaces. Instead of adopting most areas of the strap for input, Ahn *et al.* [135] augmented the strap input space with pressure levels. However, since the interactive surfaces on the strap have a different dimension from the touchscreen and their location is away from the screen, the strap interaction can provide a constrained experience for subtle manipulation requiring sophisticated mapping of small and complex visual contents.

The crown is a traditional input part on regular watches. Many commercial smartwatches adopt the crown as a physical input channel to manipulate digital content. Brule *et al.* [136] investigated the design space of physical rotary inputs using the crown. The elicitation studies for the combinations of physical rotary inputs of crown and bezel resulted in various eyes-free gestures such as swipe, pinch, and return. Another approach is a tilting crown [60] that extends the input space of the crown from bi-directional rotations to a joystick including tilt, rotation, and press. The tilting crown showed fast input performance on a circular layout. While successfully broadening the input bandwidth of the crown, the crown interaction needs more investigation on mapping with digital contents on the touchscreen.

Around-device input

A promising approach to finger occlusion on the small touchscreen is an around-device gesture input with the opposite hand of the watch-wearing hand. Kim *et al.* [61] examined the design space of around-device gestures such as straight and rotational hand movements above the watch. The prototype with proximity sensors showed the feasibility of recognition for simple gestures. Abracadabra [137] employed magnetic sensors to recognize around-device finger gestures that manipulate a cursor for the digital contents on the screen. It contributes to extending the interactive space to a much larger in-air space than the limited input surface of smartwatches. zSense [138] explored a recognition of different static and dynamic finger gestures around the device by using infrared emitters and receivers. The authors examined the spatial configurations of emitters for different form factor wearables including one-axis linear, two-axis linear, and angular displacements. Other authors introduced the eliciting around-device gestures by mapping with various smartwatch tasks [139]. A set of user-defined gestures was suggested by considering the mental model, social acceptability, and preference of the gestures. Meanwhile, Han *et al.* [140] proposed an around-device interaction technique using continuous gestures from the touch on the screen. The continuous gestures based on single or multi touches extend the interactive space around the device as using a large touchscreen. Magnetips [141] integrated haptic feedback on fingertip to around-device gesture. It investigated the design space of haptic guidance for eyes-free interaction or behind-arm and back-of-device interactions which has visual occlusion. The prior works show the feasibility of around-device gestures for resource limited smartwatches, while they require additional implementations to accurately recognize the gestures and the lack of haptic feedback in air can cause inaccurate input. Future works need more investigation on accurately recognizing diverse gestures based on the user-defined gesture set [139].

Meanwhile, many researchers have explored around-device touch input on the skin to extend the input surface. During the on-skin input, a benefit of haptic feedback on the fingertip can enable discrete inputs with key confirmation. The preference and social acceptability of on-body input for the forearm record relatively high scores rather than other body parts [142]. Various input modalities for the on-skin input on the forearm have been explored to map users' mental models to on-skin gestures such as touch, grab, pull, press, or squeeze [143]. In a proof of concept of the on-skin input, prior works examined the feasibility through handwriting input on the back of the hand using the magnetic field on a flat surface [144] or virtual button input on hand using a shoulder-worn projector and depth camera [145]. In terms of recognition technology with wrist worn system, Harrison *et al.* [63] characterized a property of skin during the finger tap to the skin. The authors implemented a bio-acoustic sensing system into the armband to recognize the touch location by analyzing mechanical vibrations from the finger tap on the skin. They also investigated the design space of the system with the accuracy performance on different locations at the forearm and validated the system in mobile scenarios. To achieve the on-skin input, various technical approaches with wrist worn form factor have been investigated such as using single or multiple ultrasonic rangefinders [146, 147], infrared reflective sensors [148, 149], inertial sensor and microphone [150, 151], acoustic sensor [152], or depth camera [153]. Especially, Zhang *et al.* [154] used an electrical signal propagating through the body. The system achieved high accuracy in detecting the location of skin touch input with a high resolution. While the authors note that the system only with small electrodes under the smartwatch is relatively independent from the environmental noise, the system still requires a signal emitting ring on the touching finger as an additional wearable device. In other prior work, Zhou *et al.* [155] integrated the signal transmitter with receiver electrodes under the smartwatch. It has a relatively low resolution and accuracy rather to the system with a separated signal emitter, while showing the feasibility of a small implemented system under a smartwatch for recognizing simple touch gestures. Visual feedback on skin with projector modules has been investigated to extend the interactive region by guiding on-skin touch input with virtual buttons [147, 156]. The technology of skin touch input is becoming a sophisticated interface by achieving high resolution and miniaturization. Further investigation on the integration of the touchscreen and extended interactive space on the skin can enable overcoming the confined interaction due to the resource limited smartwatches.

One-handed gesture input

The one-handed gesture is using only the occupied hand to perform input for the smartwatches, which can completely avoid any occlusion issue and extend the input space within the hand. During the one-handed input, the opposite hand is free for other activities. Among various one-handed gesture input methods, a wrist-tilting input is simple and similar to the operation of the joystick [157, 158]. The movements of the hand with the dexterity of wrist joints include flexion, extension, radioulnar deviation, and rotation. The dexterous wrist is able to discrete the amount of wrist tilt for target selection [159]. Various prior works have examined sensing technologies to adopt the wrist-tilting input technique for smartwatches. By detecting the physical device movement with the built-in inertial sensor in the smartwatch, Guo *et al*. [64] investigated the selection methods of the angular menu according to wrist rotation degrees and different tilt levels. The array of proximity sensors around the wrist enables the estimation of wrist tilt direction to control angular cursor [160] or two-dimensional gesture drawing [161] on the screen. Other authors have accessed bio-signals from the body such as photoplethysmogram (PPG) with motion sensor [162] or electromyography (EMG) [163, 164], while the EMG sensing requires additional wearable device attached to the large muscles of the forearm away from the clock. Strohmeier *et al.* [165] proposed an implanted interface detecting the position of the magnet implanted in hand to make input on the smartwatch. The prior works adopting the one-handed gesture techniques based on the wrist angles have shown the feasibility of replacing the opposite hand gestures. However, the angle-based inputs have a limit to providing subtle interaction for the small output space of smartwatches. In addition, the wrist tilting gesture can cause confusion in manipulating the digital content on smartwatches' screen which has a different plane of motion.

The one-handed finger gesture for triggering shortcut commands is preferable to quickly respond to the notifications from smartwatches. The finger gestures can achieve eyes-free input that is suitable for the use scenarios of smartwatches. In addition, the popular gestures are less obtrusive during the input.

To adopt this input method for smartwatches, many researchers have explored finger gesture recognition systems. Wen *et al.* [62] showed the classification of five popular finger gestures using the built-in motion sensors in the smartwatch. In another work, both motion and audio data from the built-in sensors were used to improve gesture classification by gathering more information about the gestures [166]. In spite of using the built-in sensors, the motion data with specific window size can have a latency to classify the gestures, and the noise while mobile can interrupt accurate classification. Other prior works have developed additional sensors to detect the movement of fingers such as ultrasonic-based rangefinders [167], infrared cameras with LED and camera [168], pressure sensors [169], or photo reflectors [170] around the wrist. The systems can identify the finger gestures with real-time signal processing, but require additional attachments to the smartwatch. Han *et al.* [171] used an array of microphones to detect non-vocal acoustics during the finger gestures, since some finger gestures have their own unique sound such as snapping or rubbing. The non-vocal acoustics can provide a rich interaction experience in situations where the environmental context matches well with the use of finger gestures. Meanwhile, other authors have explored the bio-signals from the wrist area where the smartwatch is located. They detect the skin or muscle deformation during the movement of fingers for generating the finger gestures, such as electrical impedance tomography [172], infrared transmission and reflection [173], ultrasound image [174], or PPG-based heart rate [175]. The examined bio-signal sensors are applicable to the form factor of smartwatches due to the sensing locations. Especially, PPG-based interaction is a promising approach since some commercial smartwatches include the PPG sensor for healthcare purposes. Zhang *et al.* [175] showed the feasibility of using PPG data to identify finger gestures by investigating essential parameters. Despite of the effective one-handed finger gestures on shortcut menu, high bandwidth tasks require a number of input keys with subtle interaction techniques.

The fine-grained finger gesture input with one hand enables micro interaction to conduct high bandwidth tasks on the small touchscreen of smartwatches such as pointing, drawing, writing, or text entry. To achieve it, the basic requirement is the continuous tracking of finger gesture input. Kienzle *et al.* [176] suggested a ring-type input device with an infrared proximity sensor and gyroscope to facilitate the continuous finger input on the surface. Through a pointing performance, this system showed the feasibility of continuously tracing the position of the finger to provide two-dimensional input. Other authors introduced deformation-based input for smartwatches with the soft hemisphere surface worn on the finger [177]. The light-sensitive photodiodes measure the reflected amount of light to detect surface deformation by pressing, shearing, or pinching gestures. The sensors support a high resolution on the optical force-sensitive deformations to conduct two-dimensional pointing input on the screen of the smartwatch. While the ring-type devices have a small and familiar appearance as finger accessories, the additional ring wearable device can be cumbersome to wear. To achieve the micro interaction by avoiding the extra devices, Zhang *et al.* [178] used only the built-in inertial sensors in the smartwatch without any additional sensors to accomplish the finger writing on the surface. The system maps the continuous motion signals by finger movements into the word for providing word-level recognition. In spite of recognizing the micro finger gestures for the high bandwidth task, it relies on the flat physical surface which tethers the user. Meanwhile, Loclair *et al.* [179] introduced a micro interaction technique with a thumb-to-finger interface for the watch-form factor devices. The authors suggested eyes-free scenarios of thumb-to-finger input with tactile feedback generated by the landmarks on the fingers. Huang *et al.* [180] explored the feasibility of the thumb-to-finger interface for smartwatches by rating the physical comfort of each finger segment during the thumb touch input. The first and second segments of the index and middle fingers recorded a high average of comfort. Through the letter-drawing tasks on the comfort segments, the authors noted that the interaction in the comfort regions can boost user performance. Other researchers have examined the feasibility of complex tasks with this one-handed thumb-to-finger interaction such as the letter drawing [181] by using fingers as an input surface, number pad [182], or text entry [183] by assigning keys to the finger segments. A thumb with the highest degree of freedom among the fingers is a useful input tool to accomplish the one-handed micro interaction. More investigation on the recognition methods of thumb movements can increase the feasibility of the thumb-to-finger interface with small built-in implementation in smartwatches.

2.2 Input Techniques on HMDs

Hand-held devices

Hand-held types of input devices are popular and familiar, since the traditional tools for making, writing, or manipulating require hands. In addition, most computing devices adopt hand-held input devices such as controllers, mouse, or smartphones. Many commercialized HMDs have also adopted hand-held controllers due to high efficiency and familiarity. A ray casting with the hand-held controllers is the most popular input technique to point, select, or manipulate the virtual objects by intersecting them [43, 77]. The smartphone-based ray casting methods have been explored to manipulate three-dimensional contents in air [184–186]. However, the ray casting methods have an occlusion issue in that the target can be obscured behind other objects in the dense target environments, while the selecting performance for the single target is fast and accurate [44]. To address this issue, many researchers have suggested novel selecting techniques such as the in-depth [44] or flexible [81] ray casting methods. In addition, the ray casting methods show low performance in selecting the small virtual objects at long distance [187]. Shoemaker *et al.* [78] investigated the distance-independence technique by using the relative cursor movements with hand-held controller-based gestures. Jones *et al.* [188] and Jackson *et al.* [189] used the angular tilting input of controllers. While the authors noted that the relative cursor movements with gestures can also support eyes-free input, the moving or tilting gestures showed lower performance in execution time. Speicher *et al.* [45] and Xu *et al.* [82] showed that the ray casting method with hand-held controllers is powerful to conduct the sequential inputs to the two-dimensional targets in the virtual environments. On the other hand, the close targets near the HMDs can be selected with the direct object collision methods using the hand-held controllers [45, 80]. However, the collision methods with the controllers can have similar occlusion issues to the lay casting methods and a constrained interaction only for the close targets.

The hand-held controllers for the HMDs support the prevalent thumb touch input on the touchpad near the hand like smartphones. However, the form factor of controllers limited the size of the touchpad. In this limitation, many prior works have investigated how to efficiently conduct the touch input for the virtual contents. The circular shape of the touchpad is a good fit for mapping with circular interfaces. Jiang *et al.* [79] investigated the tap resolution on the edge of the touchpad. The layout with six buttons showed the highest performance on the small touchpad of the controllers. To achieve a complex manipulation that requires a larger number of inputs, various approaches have been proposed such as designing the ambiguous button having multiple keys [79], the sequences of coded swipe gestures [190], or the multiple continuous cursors [191]. Other authors have explored indirect touch input methods using smartphones as hand-held controllers. It can bring the benefits of a reliable touchscreen on the large form factor device to the resource-limited HMDs on input. In order to use the occluded smartphones in the immersive environments of VR HMDs, prior works tracked the location of hovering thumb above the touchscreen [192] or projected the smartphones into the virtual scenes [193, 194]. The smartphones can also extend the input space for the AR HMDs with a constrained field of view. Zhu *et al.* [195] proposed a bi-directional interaction technique to elicit the touch gestures for the manipulation of three-dimensional contents on AR HMDs. The touch gesture [196] and pressure [197] on smartphones have been explored to achieve the eyes-free selection for the text. Despite the benefits of the prevalent thumb touch with hand-held controllers, the prior works are mostly limited to two-dimensional manipulation in virtual environments. In addition to the thumb touch, the contact area of other fingers along the hand grip can be a promising feature to extend the input space or expressiveness for the HMDs.

The physical buttons or sticks on the hand-held devices can provide a strong key confirmation with haptic feedback of landmarks. The accustomed types of input devices allow users to have high learnability. Yu *et al.* [198] proposed a rotational input method using two thumb-sticks. The thumbsticks generate the combination of two rotational movements in a circular interface. A mouse is also a popular tool to manipulate two or three-dimensional content. To adopt the mouse in the VR environments, Zhou *et al.* [199] explored a three-dimensional pointing technique for the VR HMDs with a typical desktop mouse. The depth-adaptive cursor estimates the intended selection with the user's viewpoint and the location of the cursor and target. The authors considered the display environment of VR HMDs such as the diplopia problem from stereoscopic views [200], perspective problem from the viewpoint [201], and sensitivity problem from the small targets [202]. Other authors suggested tool-based interaction techniques for specialized use cases, such as grasping [203, 204], or touching [205] the virtual objects. These tools provide a specific interaction experience with haptic feedback, corresponding to the purpose. The hand-held input devices become developing to provide a rich interaction experience with virtual objects through various input modalities and feedback. However, the hand-held controllers tether users' hands by not allowing other tasks during the input. In particular, AR Assistant systems may restrict the use of hand-held input devices for HMDs in the usage scenario of military or surgery.

Touch on the surface of HMDs

The hardware of HMDs includes a surface where finger touch input is available like other smart devices. However, the hardware location on the head due to their basic working principles needs eyes-free for the touch input. In addition, the AR HMDs have a limited size of the surface at the side parts because the see-through display occupies the front part of the hardware. To address these constraints, various prior works have investigated the eyes-free touch techniques by improving the input expressivity for the limited size of the surface on AR HMDs. Grossman *et al.* [53] explored the swipe gestures on the small side touchpad of smart glass. The authors measured the swipe performance in eight straight or diagonal directions. The shape of the surface that's longer horizontal yields more accurate and fast performance in straight swiping gestures. Based on the sequences of two swipes, they evaluated a text entry task that requires multiple keys. For the touch technique on the side part of the hardware, other authors presented the text entry system with one-dimensional handwriting by coding the characters with the different combinations of sequential swipe gestures [72]. The sequences of swipe gestures can improve the touch expressivity on the small input surface, while the coded gestures need to be memorized to reduce the error rate and the relatively long execution time. Instead of the use of coded sequential swipe gestures, Ahn *et al.* [73] investigated the combination of eye gaze and swipe gesture on the side pad. The input consists of an eye cursor and swiping gesture to select each of the three layouts and eight keys respectively. The results showed that the eye gaze assistance can reduce the execution time compared to the gesture-only input methods by reducing the touch input steps. However, the system requires a visual interface for the key layout selection. Islam *et al.* [54] explored the different contact areas of the side touchpad during the eyes-free tapping gestures with one or two fingers. The authors note that tapping is faster touch input method with a short execution time than swiping for the surface. In spite of the fast input performance, they still faced the issue of memorability in the coded tapping gestures.

Meanwhile, the VR HMDs have a relatively large surface compared to the AR HMDs, since they can use the front side of hardware for the touch input due to the inner lens and displays. Kato *et al.* [71] presented the HMD's cardboard design to customize the input space with conductive ink. The modified case can support rotating gestures on the side parts or swiping gestures on the side and front parts. In various prior works, the back-of-device interaction techniques have been explored due to the large front side of VR HMDs as an input space. An attachable magnet clip on the front side of cardboard enables interaction to manipulate virtual contents [206]. The authors explored the touch resolution of the clip by evaluating the back-of-device performance of cursor pointing with the target selection task. The results showed the feasibility of back-of-device touch interaction with the sixteen distinguishable points. Gugenheimer *et al.* [207] investigated the touch interaction on the side and front parts of VR HMD. The authors noted that users can conduct the eyes-free touch input on the surface around the head by relying on the proprioceptive senses that humans can know the position of their own body parts. Through the performance evaluation of selection tasks with different sizes of targets, they found that the closer the touch point is to the center of the face, the more accurate it is. In addition, the touch input on the side part of HMD recorded lower performance due to the different mental mapping with the virtual contents which are located perpendicular to the touch positions. Despite the feasibility of a touch interface on the surface of HMDs, the heavy weight of touch panels on the HMDs can generate a usability issue to wear for a long time. To complement this issue, Gao *et al.* [208] suggested the two-dimensional touchpad with conductive ink for VR interaction. The lightweight touch interface with the paper and ink can be customized for a purpose. It supports continuous tracking of the touch with multi-touch.

Many prior works have shown the feasibility of an eyes-free touch interface on the surface of HMDs. However, the two-dimensional touch input on the surface can provide a limited interaction experience in manipulating the three-dimensional virtual contents of the HMDs [209]. While the sequences of touch gestures on the small surface can generate the issue of memorability, they can be efficient to trigger shortcut menus once familiar with them by long usage. In terms of usability, the location of hardware away from the hands disturbs the finger touch on the surface of the main hardware with high muscle fatigue at arm [53, 207, 210]. The touch input can generate physical discomfort on the nose supporting the HMDs [74]. More investigations on the issues of usability and social acceptability are needed to build practical and reliable interaction techniques while the touch on the surface. In addition, the Hardware being miniaturized to reduce the weight can scale down the surface in the future. The other parts of HMDs need to be investigated to extend the limited touch input space on the HMDs, such as the whole temples around ears or head bands.

On-body input

HMDs, exceptional technology, are becoming developed to be a daily product. Non-handheld interactions have been explored to avoid additional encumbered devices in the usage scenario of HMDs. In the context of HMDs, non-handheld and non-touch interactions are preferred to the handheld interaction [86]. In particular, on-body input is the most preferred method in non-handheld interaction techniques [86]. The on-body input is also a promising approach that extends input space to the body area. The proprioceptive sense with haptic on the skin can enable eyes-free input during the on-body touch [85]. On-Hand input has high accessibility and social acceptability with the small movements for input [142]. Gustafson *et al.* [85] showed the feasibility of on-hand touch input based on the result that the passive tactile sensing on the palm guides the touching finger during the eyes-free input task. Other authors have explored the performance of on-hand touch input for the HMDs through vision-based detection technologies with the wrist form factor. Wang *et al.* [211] investigated the comfort region to make touch gestures on the palm area. The authors suggested guidelines for on-hand gesture detection, such as the same starting points to draw different gestures due to proprioception. Another work applied the on-hand touch input to a text entry by optimizing a QWERTY keyboard layout on the divided palm sections [212]. The on-hand text entry was preferred by participants and faster in typing speed than the touchpad-based virtual keyboard of commercial smartglasses. The forearm is also a good candidate for on-body touch input as having a flat surface for on-body touch input. A touch task on the forearm showed the feasibility of eyes-free interaction with a resolution of five or six points, while the touch accuracy was highest at landmarks such as wrist and elbow joints [146]. As the skin is deformable, input modalities of the forearm skin have been investigated to improve input expressiveness. The prior works elicited a user-defined set of skin gestures such as press, grab, pull, scratch, shear, squeeze, and twist [143]. In addition, other authors integrated the various touch gestures on the forearm with the virtual widget of HMDs by mapping with the input contexts [213, 214].

The hardware of HMDs has a small surface of the temples for touch interaction on the device. Hand-to-face input can extend the input space to a relatively large area around the device. The passive tactile feedback on face skin and proprioceptive sense facilitate eyes-free input [215]. In addition, hand-to-face gestures are often used in everyday life by having the meanings such as the casual gestures of thinking or boring [216], which are socially acceptable for the hand-to-face inputs in the public [74]. Serrano *et al.* [74] explored the design space of the touch gesture inputs around the face with mobile tasks. The results of the elicitation study showed that the cheeks and forehead are the preferred areas for the gestures. Lee *et al.* [217] investigated the hand-to-face input using different thumb areas for the chin and cheek. The input strategy on the face facilitates covering the subtle movements of the touching thumb behind the fingers. Meanwhile, the built-in sensors in HMDs are close to accessing the hand-to-face input events. Prior works have examined the detection methods with optical sensors placed on the HMDs. The detecting systems used the fact that the hand-to-face input generates a skin deformation on cheek [218] or nose [219]. However, the universal gestures of hand-to-face input are discrete to fit into the shortcut menu trigger, not continuous inputs. Future works need to investigate both the performance and social acceptability of continuous hand-to-face inputs for high bandwidth tasks such as text entry. In addition, since direct touch on the face is not preferred for those wearing makeup, mid-air interaction around the face would be an alternative. The ears that support the temples of HMDs also have a surface with various landmarks that enable eyes-free input. The social acceptability of touch gestures on the ears is relatively high compared to other parts of the face [217]. The input devices worn on ears can also be a relatively unencumbered and socially acceptable due to the common use of ear accessories. Lissermann *et al.* [220] investigated the design space of various touch gestures on the ear regions such as multi-touch, swipe, and grasp. Lee *et al.* [217] examined the performance of touch and swipe gestures on the ears. The authors noted that the cursor manipulation on the ear skin, which is not as smooth as the touchscreen, may be challenging. Besides the hand-to-face interaction, researchers have explored touch inputs for the HMDs from various body parts such as belt [221] or thigh [222] at standing or sitting postures respectively.

Hands-free input

An interaction scenarios in an uninstrumented environment constrain the hand input for interacting with HMDs [223]. Other activities can occupy hands while wearing HMDs. As an alternative, the body parts with a high degree of freedom and sensitivity in movement can operate the HMDs instead of hands. A promising body part is a head that meets these conditions. The HMDs can also access the data of the head movements with built-in IMU sensors. Yi *et al.* [224] examined the design space of head movement through an observation study using the built-in motion sensors in AR HMDs. The authors suggested the guidelines to make a recognition system of head-based gestures while various statuses such as sitting, standing, walking, and running. Augusto *et al.* [84] applied the head-based gestures to the input modality of AR HMDs by matching the head motions with the moving targets in orbital and rhomboidal trajectories. The scenario of controlling surrounding objects with AR HMDs showed the feasibility of head-based input in smart environments. Other authors conducted an elicitation study to obtain a set of user-defined head gestures for operating the AR HMDs [225]. Based on the results, they introduced various applications with a recognition system of the selected head gestures. Various prior works have explored improving the bandwidth of head-based input with a high bandwidth task such as text entry [45, 82]. The simple dwell input, holding the head on the target key, enables the sequential inputs while it is slower than the combination of head gesture and button tapping [226]. Other authors examined an effective layout for the head gestures to improve the input performance [227]. Xu *et al.* [228] introduced a circular keyboard layout with a language prediction model to improve the efficiency of head-based input. The dynamic location of the suggested word from the prediction model reduces the moving distance of the head while typing.

While the head-based motion is impactful on the hands-free input by using the built-in sensors in the HMDs, another promising approach is using eye gaze. Recently, commercial HMDs support eye tracking with built-in or plug-in eye trackers. Interaction techniques using eye gaze input have been explored to achieve a hands-free manipulation [229,230]. However, the eye gaze interaction has an inherent problem of Midas touch [231] that eyes move consciously and also unconsciously during input. Many prior works have investigated how to process unconscious eye movements during input. One approach is a dwell-based gaze pointing [232]. Park *et al.* [233] adopted the dwell time selection for manipulating the virtual contents of AR HMDs, while the authors noted that the Midas touch problem remains during the dwell-based gaze input. The relatively long dwell time of around 300ms to 1000ms also provides a constrained interaction experience [234]. Other authors analyzed the specific eye gaze patterns with the Midas touch problem during the input scenario of AR HMDs [235]. They used representative gaze patterns to trigger input. Meanwhile, the gaze gestures show better performance than the dwell-based gaze input in navigation task [236]. For the gaze gesture input for HMDs, Delamare *et al.* [237] proposed an eye gaze guiding system with virtual objects. The authors designed the input interface with dynamic targets by considering the patterns of eye movements during the exploration and return phases. Other authors investigated a design space of gaze-depth gestures in the three-dimensional environment of HMDs, based on the properties of HMD's displays and human depth perception [238]. Chen *et al.* [239] suggested an eye gaze tracking method based on the user's intention to improve the performance of gaze gesture recognition on HMDs. Another approach of eye gaze input is a method of smooth pursuits that eyes follow a moving target along the orbital trajectories to trigger an input [83, 240]. In virtual environments, the trajectory size influences the performance of smooth pursuit eye gaze input, while the distance and size of the target have a minor impact on the performance [241]. Sidenmark *et al.* [242] proposed a selection method for the smooth pursuit eye movements following the outline of the occluded virtual objects in dense virtual environments. The hands-free gaze-based selection method complements the occlusion issue of conventional ray-casting selection with less movement.

In-air Hand/Finger input

The hand can generate dexterous gestural inputs with proprioceptive sense [207], many degrees of freedom [89], and high sensitivity of movement [90]. The advanced technologies of computer vision support the real time hand tracking system that enables the in-air hand gesture inputs. Recently, the commercialized HMDs have a built-in hand tracking system to enable bare-hand interaction in the virtual environments. In line with the development in the hand tracking technology, many prior works have explored the in-air hand or finger gestures to extend the limited input space of HMDs. Aigner *et al.* [243] explored the preferences of in-air hand gesture types. The authors noted that the meaning of the hand gestures influences the usage of gesture types for manipulation. Various prior works elicited the user-defined hand gestures for manipulating the virtual contents on the HMDs. The small in-air hand gestures near torso were preferred due to the social acceptability in the AR scenario [86], while users were prone to form large hand gestures with path to directly interact with virtual objects in the immersive environment of VR HMDs [46]. Other authors showed the feasibility of multi-finger gestural input by characterizing the performance of individual fingers and their co-activation during the gestural input [244]. To achieve the fast inputs with simple hand gesture, many researchers have explored indirect hand gesture input using a pinch with the thumb and index finger. Markussen *et al.* [97] investigated a per-word level text entry using the in-air pinch gesture. The cursor on the virtual keyboard completes word typing by tracing the letters of word while the pinch gesture. Lin *et al.* [245] used the pinch gesture to interact with surrounding objects while wearing the AR HMD. The pinch gesture for the drag and drop manipulate the AR menu overlays aligned with physical objects. Other authors integrated the pinch gesture with a depth-based interface [246]. The pinch gesture displays different outputs depending on the relative depth position of the hand, which can take up only a small portion of the AR display.

Another promising approach for the in-air finger input is a collision-based touch in that fingers directly intersect with virtual targets in the air. Many researchers have explored how to improve the performance of this direct touch input in speed and accuracy. Dudley *et al.* [88] proposed an input adjustment technique after the initial touch. It can reduce the unnecessary input step for error correction. The authors also explored the cursor design on the virtual keyboard to minimize the input error of direct touch to the virtual target. Other authors explored spatial awareness during direct touch input [102]. They designed virtual interfaces that magnify or protrude the target near the touching finger to improve accuracy and speed. Brasier *et al.* [247] investigated an indirect touch input for AR HMDs by touching a virtual pad separated from the output space. Through the in-air pointing tasks, the indirect touch showed a similar performance to the direct touch to the interface. Sun *et al.* [103] detects the in-air input moment by monitoring the ultrasound phrase pattern around the HMDs. After the detection, a depth camera estimates the location of the in-air finger touch on the virtual keyboard. While the in-air direct touch input to the virtual objects is simple and easy to use, it is a distance-dependence selection method that virtual targets need to be located close to the user. In addition, it showed low performance for the sequential inputs [45] compared to other traditional input methods, since the in-air tapping gesture input using one finger is based on the hunt-and-peck style that can have a limited input performance in speed.

An emerging input technique with bare hand is in-air finger stroking with ten fingers that enables fast sequential input such as a text entry. Various prior works have investigated the movements of in-air finger stroking to detect and recognize the moving finger. While traditional finger stroking is prevalent in physical keyboards, the lack of haptic feedback during the in-air finger stroke may generate different finger movements. Yi *et al.* [47] addressed this issue by analyzing the kinematics of in-air finger stroking such as amplitude, velocity, and duration in both flexion and extension phases. Also, the authors investigated the correlated movements of other fingers. Other authors suggested the changes of joint angle as a feasible feature for recognizing the in-air finger stroking [248], while they estimated the angle changes during in-air finger stroking without any data collection of actual human typing behavior. Dudley *et al.* [87] compared various in-air finger typing methods on the virtual keyboard. In the contrast to the typing performance on a physical keyboard, the results showed a lower typing speed and higher error rate in in-air typing with ten fingers than in in-air typing with two fingers. The sequential inputs with more fingers may disturb accurate touch input for the small virtual targets. Also, the correlated movements of other fingers during the finger stroking may cause unintended finger movement to make input errors. Foy *et al.* [249] addressed this issue by suggesting the detection strategies for unintended finger stroking that is co-activated with the actual finger stroking.

Subtle finger input

While the mid-air finger gesture enables fast sequential inputs, the relatively large hand gestures in mid-air can cause arm fatigue known as Gorilla arm effect [75]. In addition, large mid-air motions can have low social acceptance in the public space. The results of elicitation studies for the hand gestures on HMDs showed that subtle hand interaction is the most preferred input method due to the concerns with social acceptance [86]. Many prior works have explored the subtle finger gestures with additional wearable mounted on the finger. Ogata *et al.* [250] proposed a ring type of wearable to detect the subtle gestures of finger rotation and bending. The IR reflection sensor in the ring detects finger gestures by observing the distance between the skin and the sensor. Other authors used a ring with electromagnetic transmitter coil [251]. A wristband with sensor coils estimates the five-degree-of-freedom pose of the ring to enable fine-grained finger interaction such as drawing letters or manipulating buttons in the air. Chen *et al.* [252] explored the subtle interaction with a finger-mounted wearable device. The device using a magnetometer estimates the distance from a magnet mounted on a different finger in order to track the three-dimensional finger movement around the palm.

Another approach for the subtle hand gesture is a thumb-to-finger interaction that the thumb touches the area of other fingers on the same hand. The fingertip pad has been investigated to allow subtle interaction with small finger movements. The small movements can be hidden by hand during input. Also, the touch on the fingertip pad enables eyes-free input with passive haptic feedback on both the fingertip and touching finger. Chan *et al.* [253] explored the touch performance on the index fingertip pad by implementing a nail-mounted device with a hall sensor grid of nine points. Through the target selection tasks, the results showed that the dexterity of the finger enables the precise two-dimensional touch input on the small fingertip pad even while mobile. Other authors investigated the touch resolution of fingertip pad [254] to build a text entry system. The authors discussed that the two-by-three grid layout has less overlap for each touch point and its spatial awareness for identifying the touch location is sufficient to enable eyes-free input. The side of the finger is also approachable by the thumb. Kim *et al.* [255] proposed a text entry method by using a small touchpad mounted on the side of the index finger, while the touch resolution is confined to the two-by-three grid. Boldu *et al.* [256] adopted a ring type wearable to facilitate the gestural touch input on the side of the index finger while athletic activity. The users preferred the simple swipe gestures towards the fingertip and palm while standing, walking, and running.

In addition to touching one finger, the thumb-to-finger interaction technique for the four fingers has been explored to enable an eyes-free subtle input on a relatively large touch area with multiple skin landmarks. Many prior works have adopted the glove form factor to detect the tap or gesture inputs on the fingers. While the additional glove for the input can be cumbersome to users, it is a promising input channel to simultaneously support high-fidelity haptic feedback and precise hand tracking while wearing the HMDs. Kuester *et al.* [257] suggested a glove form factor for a thumb-contact interaction by other fingers. The fingers with the electrode can touch the different areas of the thumb to input different functions. Peshock *et al.* [258] introduced a conductive glove that enables the thumb to make touch on the area of nails and finger segments. Other authors investigated the performance of force-based thumb-to-finger interaction [259]. The force sensors attached to each finger segment detect three levels of force input that extend the input space of one hand. Although the gloves with sensing modules precisely identify the touch inputs, they may disturb the movements of fingers by the bulky setup. On the other hand, Whitmire *et al.* [260] introduced a touch-sensitive glove with resistive fabrics. The reconfigurable glove detects the touch gesture and two-level pressure inputs for HMDs. The authors suggested a text entry system by assigning the characters to the finger segments. For the finger-mounted sensor type, Tsai *et al.* [261] used two IMU sensors to estimate the thumb touch location on the other fingers. Lee *et al.* [262] proposed a nail-mounted wearable with the design space of on-nail touch input. The nail-mounted wearable supports the subtle touch input on the side and tip of each nail. Meanwhile, the recent HMDs with the built-in or plug-in hand tracking system can enable thumb-to-finger interaction with bare hands. Prior works have adopted additional RGB [263] or depth [264] cameras to see the feasibility of bare-hand thumb-to-finger interaction with the precise touch performance. For the thumb-to-finger interaction on HMDs, Pratorius *et al.* [265] suggested a wrist-worn hand tracking system with low-cost camera modules that captures the image of fingers at the moment when the thumb touches the joints or fingertip. Soliman *et al.* [266] used a depth camera mounted on the shoulder to detect discrete and continuous thumb-to-finger input. The high-quality optical trackers have been used to demonstrate the bare-hand thumb-to-finger input techniques [267].

2.3 Summary

This chapter summarized the previous works of input techniques on smartwatches and HMDs by discussing their advantages and limitations with the raised interaction challenges in the usage scenarios of wearable devices. Various interaction techniques with state-of-the-art sensing technologies for enhancing the widely-used methods or developing novel input modalities show the feasibility of enabling various input scenarios for the wearables such as on-device, non-contact, eyes-free, hands-free, on-body, and bare-hand input methods. I conclude this chapter with the potentiality that finger-based input techniques with high dexterity are applicable to enable high bandwidth tasks requiring fast or subtle manipulation such as drawing, sketching, modeling, and typing. In the following chapters, I will present the development process of novel interaction techniques enabling the increases in input expressiveness based on the properties of finger movements. The development process includes rationales for the novel interaction systems, empirical studies for the data collection and system evaluation, and discussion for design guidelines, limitations, and future works. The diverse wearable input scenarios on the small touchscreen of smartwatches and the in-air space around HMDs demonstrate the general claim of this thesis.

III Fingers and Angles: Exploring the Comfort of Touch Input on Smartwatches

3.1 Properties of Finger Movements in Comfort Angle of Touch Input

The first scenario supports the general claim of this thesis by understanding the comfort range of angle-based touch input on the small touchscreen of smartwatches. The unique input context of smartwatches would limit the touch angles that require the touching finger to reach over the watch and twist back to the touchscreen. In addition to this inherent limitation, the angle-based touches would generate different levels of comfort according to the touching poses, since the viable angles depend on the range of motion in flexion/extension and abduction/adduction of wrists and fingers. The discomfort during the touch gestures will affect the usability and also performance. To address this comfort issue on smartwatches, two empirical studies explored the comfortable range of static and dynamic angle-based touch input on variations in fingers, regions, poses, and angles. Also, using the raw touch images from the two studies, the classifiers were constructed to recognize the finger region during angle-based touch input. Based on the results of two empirical studies, the data-based design guideline and potential interaction techniques were recommended to show the feasibility of angle-based touch input and region-aware interaction techniques that can enable increases in the expressiveness on the smartwatches.

3.2 Abstract

Smartwatches present a unique touch input context: small, fixed to one wrist and approachable from a limited range of angles by the touching hand. Techniques to expand their input expressivity often involve variations in how a watch must be touched, such as with different fingers, poses or from specific angles. While objective performance with such systems is commonly reported, subjective qualities such as comfort remain overlooked. We argue that techniques that involve uncomfortable input will be of limited value and contribute the first data on the comfort of input on smartwatches via two studies that combine subjective ratings of comfort with objective performance data. We examine both static and dynamic touches and three finger poses. Based on the study results, we contribute a set of design recommendations for comfortable, effective smartwatch input. We close by instantiating the recommendations in interface prototypes that we evaluate in a final qualitative study.

3.3 Introduction

Wrist worn wearables such as smartwatches have a unique input context. Their mounting point on the forearm provides benefits: a ready visual availability [83] and an easily adjustable device position that makes them suitable for novel input styles such as arm gestures [62]. On the other hand, their small touch and display surfaces present substantial challenges for input techniques, such as multi-touch, that have been instrumental in the popularization of larger scale mobile devices such as smartphones. Researchers have begun to explore the opportunity this represents with proposals for how touch interaction can be customized for the constraints of the watch form factor. Ideas include the use of pairs of fingers in temporal [7] or spatial [268] patterns, or distinguishing amongst touches based on properties such as contact area shape [66] or the finger issuing the touch [68]. The goals of this work are typically to improve expressiveness and/or speed of input.

However, little attention has been paid to how physically comfortable these techniques are to use. We argue this issue is particularly relevant for interaction on smartwatches as they are operated with a specific and highly limited pose between the arm wearing the watch and the arm/finger touching the screen. Many of the techniques authors have designed imply substantial variations in this pose that may, in turn, result in reduced comfort. This is most clear in the growing body of work that deals with yaw rotation of touches [94], or techniques based on the angle derived from the shape of a single finger touch [66] or that between a pair of finger touches [268]. In these examples, users perform tasks such as menu selection, parameter setting or shortcut activation by specifying angles in ranges of between 120° and 180°, spans that imply substantial movements of one or both hands or arms. They also achieve these manipulations using touches issued by a range of different finger areas such as the flat [94] or side [66] of the thumb, index or middle fingers and the combination of index and middle finger [68] and/or index and thumb [268]. While reports on the accuracy, speed and reliability of these systems are staples in this literature, we highlight the fact that there is no data or commentary about how comfortable they are to use. Specifically, in these representative examples, comfort is not mentioned and even reports of high level data such as NASA TLX [269] or BORG-10 [270] are absent. This paper argues that the comfort of interacting with a smartwatch will vary substantially depending on both the angles users need specify and the finger area or areas they use to do so and seeks to gather data to elucidate this issue.

Comfort is a key issue in the design of touchscreen input techniques [271]. Regardless of how rapidly or accurately an input action can be performed, one that results in discomfort will be neglected over more comfortable alternatives. While prior work on larger form-factor devices suggests this is particularly true of touches specifying angle information [101], no work to date has examined the comfort of input of smartwatches. Accordingly, this paper gathers and contributes baseline data on the comfort of input on smartwatches from a pair of studies that combine and contrast traditional metrics of time and accuracy with ratings of the comfort of touches. These experiments were designed to directly complement prior work reporting objective aspects of input performance – they focus on representative input tasks. Specifically, we consider eight different angles and three different finger regions: the flat of the index finger, the side of the index finger and the tips of index and middle fingers. The first study involves static touches – atomic touches and releases of the screen that indicate a single angle. This type of touch appears in prior smartwatch input techniques, such as Lafreniere *et al.* [268]'s description of pairs of variously angled simultaneous taps on a 3x3 grid of targets covering a smartwatch screen. The results highlight viable, comfortable ranges for angular input for each of the three finger regions. Building on these findings, a follow up study captures the comfort and performance of dynamic touches involving on screen rotations from one input angle to another, a form of input that has been proposed in, for example, Xiao *et al.* [94]'s description of yaw input. This paper complements this existing work with its focus on the comfort of input. The paper closes by contributing practical recommendations for input on smartwatches in terms of the angles and finger regions that users can comfortably and effectively use. We showcase the value of these recommendations by designing interaction techniques that instantiate them and capturing user reactions to the applications, and interaction techniques they feature, in a final qualitative study.

3.4 Related Work

The small size and worn context of smartwatches presents new challenges and opportunities for the design of input and interaction techniques. Approaches are as diverse as adapting existing primitives such as tap [111], swipe [121] or multi-tap [7] to small screens, extending input spaces to the skin areas surrounding a device [147] or leveraging different aspects of the worn context of a watch. For example, a watch's loosely anchored attachment to the body enables it to be pushed or twisted, providing a rich space for input [65, 272], and its proximity to the hand can enable various forms of gestural input, what Kerber *et al.* [163] term "same-side" interactions, such as detecting finger motions [62] or simple tilts of the arm [162] or wrist [160]. Wearables are also inherently available for use in wide range of settings and researchers are being to characterize user performance while mobile or encumbered [124] and to present design guidance to better enable the use of watches in these distracted and complicated scenarios [127].

Comfort is an important aspect of wearable device use and a valuable lens though which new input techniques can be examined. A practical definition revolves around movements that can be readily achieved – for example in Le *et al.* [273], "comfortable areas" are defined as regions that can be reached by the fingers, during single handed use of a smartphone, without hand posture or grip changes. Based on this definition, comfort and input involving variations in the angle between the sensor surface and the touching finger are intrinsically linked. Hoggan *et al.* [274], in one of the relatively few studies to systematically examine human performance of multi-touch rotations, introduce a notion of "ergonomic failure" to denote uncomfortable situations in which a rotation cannot be performed due to the fingers jamming against one another. Hoggan *et al.* [274] use a tablet sized touch surface, input with the index finger and thumb, initial touch angles in fairly extreme ranges (0-120°) and fixed 90° rotations. They report ergonomic failure led to abandonment of 39.7% of trials. However, despite acknowledging the importance of comfort in this interaction, they provide no data on how comfortable participants felt any of the tasks they actually completed were – we argue that prior to failure, there is likely to be a spectrum of more and less comfortable actions.

Other authors have also sought to formalize performance of general multi-touch input, including rotational tasks. Nguyen *et al.* [275], for example, explore performance of combined rotation and translation operations on a multi-touch table using a Fitts law [276] study design. Using a similar tabletop setup, Zhao *et al.* [277] examine the full set of translation, rotation and scaling operations with the objective of creating a model of human performance for multi-touch input. Both articles also discuss the strategies and approaches users adopt to combine different types of transformation into single on-screen operations. However, we note that neither of these studies comments on the subjective qualities of the tasks their participants complete – their focus remains solely on production of objective models of performance. In closely related work, Voelker *et al.* [278] explore performance of rotational tasks in tangible and tabletop settings, concluding that grasping tangible, physical controllers can boost performance by 20%.

Numerous authors have also proposed rotation input using single fingers [94, 113], but the focus has typically been on system rather than user performance; there are few accounts of objective usage data, let alone subjective experience. One notable exception is Mayer et al.'s recent work on the ergonomics of single finger rotation input on tabletops [101] and mobile devices [279]. For both settings, this work considers touches with the finger tip from four pitch and 16 yaw finger angles (spanning a full 360) and captures subjective ratings of the feasibility of these angular touches. Among other analyses, these are used to define a comfort zone for touch input with each hand. On the tabletop, these are a 135° region where the touching finger can make contact with the screen while remaining aligned approximately parallel to the arm. In the mobile setting, these regions are larger: 180°. This is likely due to participants moving the touch surface (on a mobile device) during the course of their input. This work highlights the importance of comfort in performing input that involves different angles on a touch surface – it argues that input that is not comfortable is infeasible.

The work in the current paper extends these ideas in two key ways. First, it moves from a tabletop or mobile device scenario to that of a wrist wearable, a physically distinct input setting. Prior work on the biomechanical aspects of touch input, although not directly considering a wearable scenario, has highlighted the fact that touch input performance fundamentally differs depending on device form-factor and body posture [280] – different devices and postures engage different muscle groups, enable different movement ranges and result in different performance. The unique constraints of touch input on wearables, with small touch surfaces attached to the body, mean that findings from other input scenarios are unlikely to be directly applicable: new studies of performance in wearable settings, such as the one described in this work, are required. Secondly, this paper considers three different touch input techniques for specifying input angles: touches by the side and flat of the index finger and those by the pair of index and middle fingers. In this way, this paper complements prior work by improving our understanding of the comfort of touch input on a new form factor of wearable devices and with a larger set of finger regions.

3.5 Sensing Rotations

All work in this paper was implemented on a Sony Smartwatch 3, a device with a 28.25mm square screen and a bevel of between 5 and 7mm. Following prior authors [94, 281, 282], and in order to capture rotations reliably, we adapted open source modifications to the Android kernel [283] to report raw touch data from the device's seven by seven sensor capacitive grid. Our implementation ran at 80Hz, sufficient to provide a fluid response. We calculated ellipses through a process of flood filling to isolate individual touch areas and image moments to derive their description: centroid, orientation, the size of major and minor axes and eccentricity. We thresholded sensor data at 25% of its maximum value and, unlike prior work [94], did not apply any gamma corrections to the touch image. Both the system's similarity to those deployed in prior studies [94] and extensive iterative testing during development indicate this configuration recorded input angles relatively accurately – this approach leads to errors of approximately 10°, as reported by Xiao *et al.* [94] in their description of yaw detection accuracy for flat (low pitch) touches. We argue this level of accuracy is both realistic (e.g. supported by standard touchscreen sensor arrangements) and sufficient to support our empirical objectives – to elicit particular angular inputs in order to assess their comfort.

Figure 1: Sony Smartwatch 3 capturing raw touch input; data shown in callout at top-left of screen. The three images show the finger regions used in this work. Left image shows a touch with the *flat* finger region, center with the *side* and right with *pair*.

We acknowledge that more advanced algorithms have been proposed to extract finger angle from touch screen input patterns, such as the deep learning approach proposed by Mayer *et al.* [284], but we argue that while these techniques can boost angular accuracy, their relatively high computational overhead makes them unsuitable for interactive tasks on resource constrained devices such as smartwatches. Furthermore, the benefits of these advanced algorithms are not clear for the scenario in this paper – while Xiao *et al.* [94] report genuine data from a watch, Mayer *et al.* [284] use a larger touch surface and note that their results may not generalize to small screens where finger contact regions may be truncated at sensor edges. Furthermore, while Mayer *et al.* [284]'s detailed critique highlights improvements in accuracy over Xiao *et al.* [94] yaw accuracy, it is not clear these apply to the situation in which fingers are flat on the touch-screen surface (low-pitch). Performance of Xiao *et al.* [94]'s heuristic algorithm is optimal in this setting and variations in reporting in Mayer *et al.* [284] make it hard to assess whether their algorithm offers improvements when fingers are flat. This combination of suitability for use in an interactive study (rather than the purely measurement scenario in Mayer *et al.* [284]) and proven performance in a wearable setting motivated our choice of Xiao *et al.* [94]'s implementation for the current work. Figure 1 shows the watch with a callout visualizing the sensor data.

3.6 Static Angle Study

We first captured performance and comfort data during production of static angles using three common finger regions: the *flat* or *side* of the index finger and the *pair* of index and middle fingers – these are shown in Figure 1. The flat and side of the index finger were selected as they are the only two ways on-screen yaw angles can be specified by a single digit. They have also been explored in prior work [94, 101, 283]. We opted to study index plus middle over the alternative pair of index plus thumb

Figure 2: Left image shows eight angular targets and smartwatch used in the static study. Targets are subsequently referred to by the degree angles 30° through 240°. Right image shows the smartwatch with the study interface worn by a user about to select the 120° target with the side of their finger.

as it has also been specifically studied in prior work [68]. We also selected a range of input angles to consider based on those proposed by prior work [101] and common sense plausibility – we excluded angles that would require a user initiate touches from the far side of arm wearing the watch as these would involve the touching hand reaching over the wrist wearing the watch and twisting back to touch the screen. This would be an inherently uncomfortable arrangement and indeed, prior work suggests that users may find this kind of task impossible to perform [101]. To select angles, we assumed the watch was always mounted on the left wrist and defined a range of 210°, spanning the clock position numerals between one and eight and aligned to the axes of the watch. Within this range, we considered eight angles, spaced at 30° clock position intervals. Following [94] we label these angles in degrees and sequentially in the clockwise direction with a zero point at the top: between 30° and 240°. Figure 2 shows these angular targets on the Sony smartwatch used in the studies.

The study followed a fully balanced repeated measures design for the three finger region conditions. Before each condition started, the experimenter explained the finger region to be used, demonstrating this if required, and emphasized the focus on comfort and that, within the constraints of the task, participants were free to touch the watch any way they liked. This included moving both arms, as well as the wrist and finger joints of hand touching the screen. One consequence of this approach is that, for pair touches, participants could use the index finger to mark either the edge or center of a specified angle: input for the opposing angles of 60° and 240°, for example, could be specified with an effectively

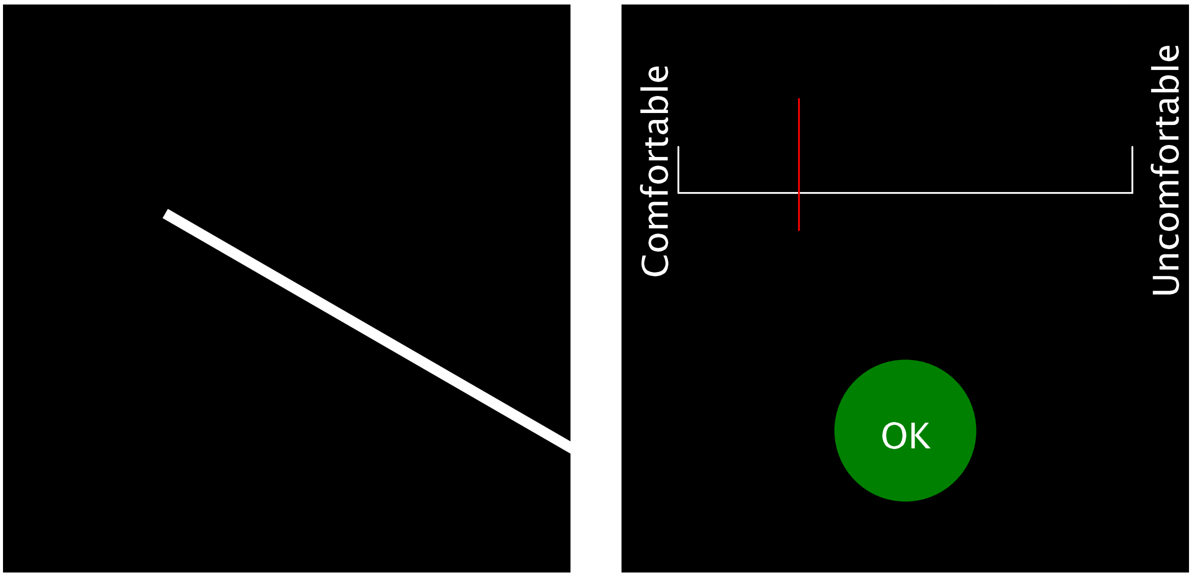


Figure 3: Screen shots of watch interface from static study. Left image shows angle instruction for touching the target at 120° in the form of a simple white line. Right image shows the comfort scale used to enter ratings. To enter a rating, participants positioned the red line on the slider and then clicked OK.

identical pair of touches. Although an experimenter was present in the room throughout, the study did not include any external checks on participants' touches – the task was simple and consistent and we relied on participants to follow the finger region instructions faithfully.

Within each condition, participants completed two blocks of eight randomly ordered trial-sets, one for each angle considered in the study. Each trial-set was composed of five repetitions of the same angle followed by entering a rating about the experience on a single item continuous, unsegmented 100-point scale from comfortable (0) to uncomfortable (100), similar in structure to that used in Mayer *et al.* [101]'s closely related prior study. Participants in the current study were asked to rate "how comfortable the input was to perform". This question, and the scale labels, differs from that used by [101] (a question about feasibility and an explicit option to indicate an input was not feasible) as our core concern relates to comfort, and as such, our study design already excluded extreme angles that participants would likely perceive as impossible. The inclusion of five repetitions before each rating was intended to increase the reliability of the data (as it would be derived from a sustained experience), while the temporally separated repetition of ratings for identical angles in the two blocks provided a way to verify each participants' consistency.

Eighteen participants completed the study (three in each possible order). They were recruited online from the local student population, screened for right-handedness and compensated with 10 USD in local currency. Nine were male and nine were female and they had a mean age of 21.6 (SD 1.8). They reported high levels of experience with touch screen smartphones (4.8/5) but low levels of experience with smartwatches and other wearables (1.25/5). In total the study resulted in 864 ratings derived from 4320 individual user inputs (18 participants by 3 finger regions by 2 blocks by 8 angles by 5 repetitions).

The structure of each trial was designed to solicit natural movements and poses – this stands in contrast to prior work on angle input which has tended to capture screen touches from controlled and somewhat artificial finger poses achieved by, for example, requiring participants to place their fingers in plastic guides [94, 101]. We argue these restricted settings are inappropriate for work focused on the comfort of wearable input; we need assess behavior in less artificial and constrained situations. Accordingly, we borrow from Lafreniere *et al.* [268] and had participants stand throughout the study. Furthermore, each trial in the study first instructed the participant to lower the arm wearing the watch to a vertical position with the hand over the thigh (detected via the device's accelerometer). After completing this, the arm was raised and a message requested them to tap the screen. A fixation spot was then displayed for 500ms, followed by a simple white line indicating the required angle. For each trial, we logged *preparation time* (from the presentation of an instruction until the first screen touch), *touch time* (the duration of a screen touch) and the absolute *accuracy* of the angle data at the temporal center of the touch [279]. The accuracy was calculated as the absolute difference between the target angle and either the orientation of one touch, or the angle specified by a pair of touches. We did not judge the success or failure of trials based on the accuracy. Finally, we also recorded all raw sensor data. After a trial-set was complete, participants entered a *comfort rating* using a slider shown on the watch. Study instructions are pictured in Figure 3.

Results

Figure 4 shows comfort (left), time (center) and accuracy (right) data per finger region and input angle. We first assessed the consistency of participants' ratings by correlating data from the two blocks in each condition. This led to a mean correlation of 0.78 (SD 0.11), indicating a strong relationship. This suggests that we can be confident in the validity of the data and that participants' subjectively experienced comfort was consistent throughout the study. Beyond this check, we analyzed each measure with a three by eight repeated measures ANOVA and *post-hoc* t-tests. We incorporated Greenhouse-Geisser corrections for sphericity violations and Bonferroni corrections to prevent alpha inflation in *post-hoc* testing. Based on longstanding arguments that this is appropriate, we opted to treat the data from the comfort rating scale as parametric [285]. At this stage in the analysis, we chose not to conduct

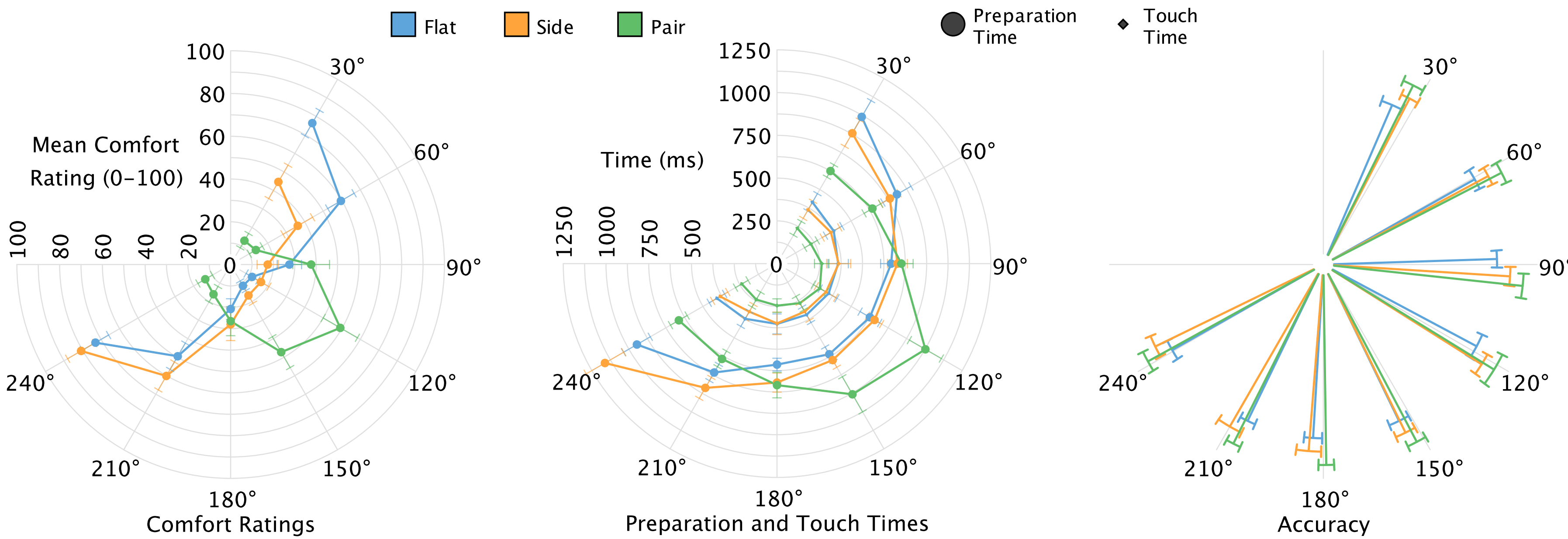


Figure 4: Data from the static study: Comfort Ratings (left) on scale from Comfortable (0) to Uncomfortable (100); Preparation and Task Time (center) and; Angular Accuracy (right). Bars show standard deviation.

post-hoc tests on the angle variable, as these would be aggregated over the three finger regions. Table 1 shows the statistical results.

Discussion

While the study showed significant differences in all variables, the comfort ratings led to the strongest effect sizes. This suggests this measure captures an important quality of behavior that should not be overlooked in the design of touch interfaces for wearables. The interaction effect between region and angle, clearly illustrated in Figure 4 (left), is strongest. It indicates that the comfort of specific angles varied with the finger regions used. We interpret this to mean that each pose has a different viable input range. To explore what these ranges might be, we calculated the overall mean comfort rating recorded in the study (37%) and performed three sets of Bonferroni adjusted pair-wise comparisons for the angle data from each finger region. This generated a large number of significant differences, but we note the key data is the Cohen's d effect sizes shown Table 2. These range from small to extremely large (large is typically defined as 0.8), reflecting the relative consistency of scores for each angle/finger region pair and the substantial differences between angles.

While these data could be interpreted in many ways, we make recommendations by thresholding data to be under the study mean and considering the effect size compared to the angle(s) with the best (lowest) comfort rating for each finger region (see Figure 4, left). Applying a threshold of 1.5 for effect size differences leads to selecting comfortable side touches as those between 60° and 180°, flat touches between 90° and 180° and pair touches from 0° to 90° or 180° to 270°. This dual viable range for pair touches is due to the fact that they involve approaching the watch from above, rather than from one side (as with the other two regions). As such, there is no practical difference between the opposing angles – a touch at 30° is effectively the same as one at 210°. This fact likely contributed to the significantly lower ratings that pair touches achieved in the main effect of region. We note that while data from the flat finger-region is broadly aligned with prior work on the comfort of input on tabletops [101], data for the other finger regions differs considerably, highlighting the value of exploring these alternative types of input. We speculate that the increased range in side over flat (the region from 60° to 90°) is due to participants' ability to use finger flexion during side input – bending the finger enables comfortable input over a larger set of angles. Similarly, we suggest the prominent increase in the discomfort of pair input around 120° is due to participants' need to to twist their wrist in order to align their index and middle fingers to make touches at this angle.

There are several other notable outcomes. Firstly, the significant interaction effects in the time and

Table 1: ANOVA and post-hoc test results in static study.

Measure	*Variable(s)*	*Outcomes*			
Comfort	Region by Angle	F(4.2,70.7) = 56.8	$p < 0.001$	$\eta_p^2 = 0.77$	–
	Region	F(2,34) = 16.71	$p < 0.001$	$\eta_p^2 = 0.5$	*Post-hoc*: pair<side and flat (p<=0.002)
	Angle	F(3.4,57.8) = 48.3	$p < 0.001$	$\eta_p^2 = 0.74$	N/A
Preparation Time	Region by Angle	F(3.3,55.6) = 26	$p < 0.001$	$\eta_p^2 = 0.6$	–
	Region	F(2,34) = 1.94	$p = 0.16$	$\eta_p^2 = 0.1$	–
	Angle	F(3.9,66.6) = 28.8,	$p < 0.001$	$\eta_p^2 = 0.63$	N/A
Touch Time	Region by Angle	F(5.2,70.7) = 88.18	$p < 0.001$	$\eta_p^2 = 0.22$	–
	Region	F(2,34) = 16.32	$p < 0.001$	$\eta_p^2 = 0.49$	*Post-hoc*: pair<side and flat (p<=0.003)
	Angle	F(3.4,57.8) = 48.3	$p < 0.001$	$\eta_p^2 = 0.19$	N/A
Accuracy	Region by Angle	F (14, 238) = 5.41	$p < 0.001$	$\eta_p^2 = 0.24$	–
	Region	F(2,34) = 0.87	$p = 0.43$	$\eta_p^2 = 0.05$	–
	Angle	F(4,68.3) = 1.92	$p = 0.072$	$\eta_p^2 = 0.1$	N/A

Table 2: Cohen's d effect sizes for post-hoc pairwise comparisons on the comfort ratings for angles from 30° to 240° in the static study. One asterisk signifies a significant difference at $p<0.05$, two asterisks at $p<0.01$. All comparisons incorporate Bonferroni corrections.

Angle	*60°*			*90°*			*120°*			*150°*			*180°*			*210°*			*240°*		
	Flat	*Side*	*Pair*	*Flat*	*Side*	*Pair*	*Flat*	*Side*	*Pair*	*Flat*	*Side*	*Pair*	*Flat*	*Side*	*Pair*	*Flat*	*Side*	*Pair*	*Flat*	*Side*	*Pair*
30°	0.85**	0.35	0.07	2.75**	1.38**	1.4**	4.56**	1.5**	2.49**	4.84**	1.49**	1.99**	3.37**	0.71	0.9	1.36**	0.66	0.28	0.17	1.57**	0.08
60°	gray!10	gray!10	gray!10	1.77**	1.03*	1.41*	3.29**	1.14*	2.53**	3.49**	1.13**	2.01**	2.3**	0.38	0.88	0.5	1.09**	0.22	0.67**	2.06**	0.01
90°		gray!5		gray!5	gray!5	gray!5	1.26*	0.07	0.85	1.36**	0.05	0.4	0.45	0.6	0.51	1.22*	2.42**	1.19**	2.52**	3.66**	1.35*
120°		gray!10			gray!10		gray!10	gray!10	gray!10	0.01	0.03	0.48	0.82	0.69	1.44**	2.6**	2.62**	2.26**	4.22**	3.93**	2.43**
150°		gray!5			gray!5			gray!5		gray!5	gray!5	gray!5	0.89*	0.68	0.97**	2.76**	2.61**	1.76**	4.48**	3.93**	1.93**
180°		gray!10			gray!10			gray!10			gray!10		gray!10	gray!10	gray!10	1.71**	1.51**	0.67	3.1**	2.51**	0.84
210°		gray!5			gray!5			gray!5			gray!5			gray!5		gray!5	gray!5	gray!5	1.17**	0.98**	0.2

accuracy variables serves to reinforce the primary conclusion that specific angles were more or less challenging depending on the finger regions used. We also point out that the three significant main effects of angle likely reflect the fact that two conditions (flat and side) yielded relatively similar results, while pair data is almost opposite – this unbalanced split suggests these main effects can be discounted. Beyond this, touch times were significantly lower with the pair touches than with either flat (by 121ms) or side (by 99ms) touches, replicating prior descriptions of this kind of touch [101]. This effect may be due to participants' greater familiarity with two finger taps, or something intrinsic about these larger touches. Although they show few differences, the accuracy data are also useful for deriving viable angle target sizes: overall absolute mean and standard deviation were 3.7° and 2.47°. This suggests that angular targets should be approximately 20° wide to support easy acquisition – the mean plus two to three times the standard deviation in each angular direction. We note that the similarity of the accuracy measures provides a limited validation that the simple algorithm we use to assess angle in this work [94] is sufficient for our objectives. Specifically, assuming the human ability to specify angles in flat and side poses does not differ, the close similarity of the accuracy data for these two types of touch suggests that the algorithm is as accurate for side touches as for the flat touches on which it has previously been validated.

It is also worth discussing relationships among the different measures in the study – in particular, we are interested in assessing whether or not the comfort ratings genuinely capture unique aspects of user behavior or experience, or whether they simply replicate outcomes from more traditional objective data. To explore this issue, we correlated the full set of individual comfort ratings against both timing measures and the absolute accuracy data. As can been inferred from Figure 4, the results show some links: preparation time was moderately [286] correlated with comfort ($r=0.546$, $p<0.001$), but relationships with touch time ($r=0.066$, $p=0.17$) and absolute error ($r=0.025$, $p=0.603$) were non-significant. These outcomes clearly indicate that comfort is independent from accuracy (corresponding to error rate in more conventional studies) and time on screen during our static touch task. The similarity to preparation time is likely due to more awkward and uncomfortable poses requiring longer to produce. The fact the relationship is only moderate in strength, together with the higher effect sizes for comfort (and a visual inspection of Figure 4) suggest participants were able to maintain relatively high objective performance levels under conditions of moderate to substantial discomfort – mean preparation times exceed 1000ms in only a single pose/angle combination. A second implication from this data it is non-trivial to infer comfort from objective measures. In the current study, the only measure showing links is one that captures performance before screen contact – data that is easy to capture in a controlled lab study, but hard to observe in more naturalistic settings. These findings are important as, we argue, users are unlikely to adopt real systems that are uncomfortable to use, even if objective performance remains high. In this way, we argue that the results from this study highlight and support the value of explicitly measuring comfort.

Finally, we also examined the raw data from this study. Histograms on the eccentricity and major and minor axis lengths of the moments suggested that flat and side touches were highly distinctive – basically flat touches were considerably wider than side touches. In order to explore the feasibility of distinguishing between these two cases, we extracted a single centrally timed moment from each touch in the flat and side conditions and constructed a logistic regression model on this data using a ten-fold cross validation process in Weka [287]. This yielded an accuracy of 95.9%, a figure we suggest is high enough to make these distinguishing between side and flat touches a practical option.

3.7 Dynamic Angle Study

In order to further explore the differences between finger regions, we conducted a follow-up study of dynamic touches that involve rotations from an initial to a final angle. This type of input is a prominent feature of multi-touch systems [288] and has been proposed as a modality for single finger input on tabletops [113], mobile phones [114] and smartwatches [94]. While performance studies of multi-touch rotation exist [274], the majority of literature relating to single finger input is concerned with the accuracy with which angle can be determined by the sensing system [94, 114], or with interface design proposals. We are aware of no prior work that examines performance with the side of the finger, or which explicitly examines the comfort of the movements. This study sought to fill these gaps.

Eighteen new participants completed this study (11 male, 7 female, mean age 22.3 (SD 2.4)) recruited and screened as in the static study and with similar experience of smartphones (4.9/5) and wearables (1.6/5). For each of the three regions, we selected a 120° angular range centered on the most comfortable region identified in the static touch study. These were 60° to 180° (side), 90° to 210° (flat) and 150° to 270° (pair). We again spaced targets at 30° and each participant completed each region condition in a fully balanced design. The study design sought to explore key properties of targeting movements: distance and size (or width). While there are many paradigms studying these variables, we opted to build on prior work on rotation input [275] and use a Fitts law inspired study design. This is a well-validated experimental paradigm that confers numerous benefits including a formalism for selecting appropriate target properties (size/width) to study, guidance for designing, balancing and repeating trials, and detailed analysis procedures. Applying these perspectives to targeting study designs can increase the reliability and validity of the final outcomes. Based on this approach, trials for each finger region involved an exhaustive combination of all five start points to all four end points, leading to four movement distances in two directions (clockwise and anti-clockwise). We defined the angular width of the start target as 20° and included two possible angular widths for the end target: 5° and 15°. The two end target sizes and four distances led to eight Index of Difficulty (ID) values, spanning 1.58 to 4.64 and typical of those used in prior studies of rotation input [275]. The start width was set derived from data in the first study: the mean absolute accuracy plus 2.5 times the standard deviation in each angular direction. This target size was intended to provide a moderately challenging input task.

However, the study procedures deviated from Fitts law recommendations [276] in several major ways. Firstly, we focused on input in a natural setting - we maintained the study setup with standing participants who lower and raise the watch prior to each trial. This broke up the trials and increased the study duration. In total, for each region condition, we ran three blocks of randomly ordered trials involving all combinations of start-angle, end-angle and width. The first block was discarded as practice and, although the study took approximately one hour to complete, this led to too few trials to adjust for accuracy and calculate IDe [276]. Secondly, the largest angular distances (120°) could only be performed between one start and end point, whereas the smallest distances (30°) could be performed between any pair, leading to unequal numbers of trials for each ID value. To deal with these deviations, we explicitly logged trial correctness and had participants repeat erroneous trials, thus supporting separate analyses of time and error data.

The primary independent variable in the study was finger region. We also considered rotation direction (clockwise/counter-clockwise) and, in non-Fitts analyses, final target size. For each trial, we measured *preparation time* and *touch time*, as in the first study, as well as errors on both the initial touch and final release, respectively termed *initial errors* and *final errors*. Errors were defined as touch angles outside the specified start or end target regions. Based on these data, we calculated *throughput* [276] and also logged all raw sensor data. Finally, in order to assess comfort, participants filled a short paper questionnaire at the end of each region condition; the lack of sustained trial repetitions precluded capture of repeated comfort ratings as in the static study. However, participants answered the same question used in the static angle study: they rated both clockwise and counter-clockwise rotations using a comfortable-uncomfortable scale. This was also labeled identically (comfortable to uncomfortable) to the scale in the static study and, again similarly, no units were marked. However, in order to facilitate subsequent scoring by an experimenter, it was divided into 20 equal segments. We opted for a paper questionnaire due to the limited amount of data collected and to provide a non-watch based activity to break up study conditions. In total, we retained 4320 correct trials (18 participants x 3 finger regions x 2 blocks x 5 start-angles x 4 end angles x 2 widths) and 108 comfort ratings (18 participants x 3 finger regions x 2 directions) for analysis.

Instructions for the study were designed to maximize visibility on the small watch screen in all three finger region conditions. A circle was drawn full screen and the direction of the required rotation indicated by four arrows in the corners. The initial angle was always marked with a line across the circle, while the final angle and width were shown by symmetric wedges. During movement, the line acted as a cursor, following a user's rotation. The task was to move the line into the wedges. In the flat and side conditions, a colored highlight indicated that users should leave the tip of the line uncovered in order to receive feedback as to the current angle. In the pair condition, two small circles indicated where to touch to ensure the center potion of the line provided a similar response. We note these differences are an inevitable result of using different finger regions on a tiny screen. Figure 5 shows these interfaces.

Results

Data were analyzed using procedures similar to the static study: repeated measures ANOVA and post-hoc tests incorporating any necessary adjustments for sphericity violations and applying Bonferroni confidence intervals adjustments in *post-hoc* testing. Table 3 summarizes these outcomes for all variables. Data from the comfort ratings (not charted) showed no significant differences and an overall mean (11.08, SD 4.9) in the mid-range of the 0-20 scale used. Data per finger region were: flat (M: 10.11, SD: 4.91); side (M: 11.02, SD: 4.49) and; pair (M: 12.11, SD: 5.2). This suggests that the selection of angular zones to match finger regions was successful: participants experienced similar levels of comfort in each condition. Furthermore, the direction of movement (CW: 11.46, CCW: 10.7) did not impact comfort. However, the overall rating (55%) was higher than in the first study, suggesting that the dynamic task was found to be, on the whole, moderately less comfortable that the static task in the first study. As such, input techniques that rely on static touches may be preferred over those that require dynamic motion.

Analysis of input time showed more diversity. Time data are shown in Figure 6. All main effects were significant for preparation time, indicating that counter-clockwise motions towards large targets with side and pair finger regions led to optimal performance. We note the magnitude of some of these temporal differences (e.g. in the size variable) are quite small and may have limited real world impact. In contrast to the static study, touch time data revealed that side touches were significantly faster than pair or flat touches. Target size also strongly influenced performance. The main effect of region was borne out by generating Fitts' law regression models and calculating throughput for each region and direction. Direction did not yield significant variations, but side (mean of 1.79 bits/second, SD 0.06) significantly improved over flat (1.61, 0.07) and pair (1.62, 0.08). We note that, as we did not include

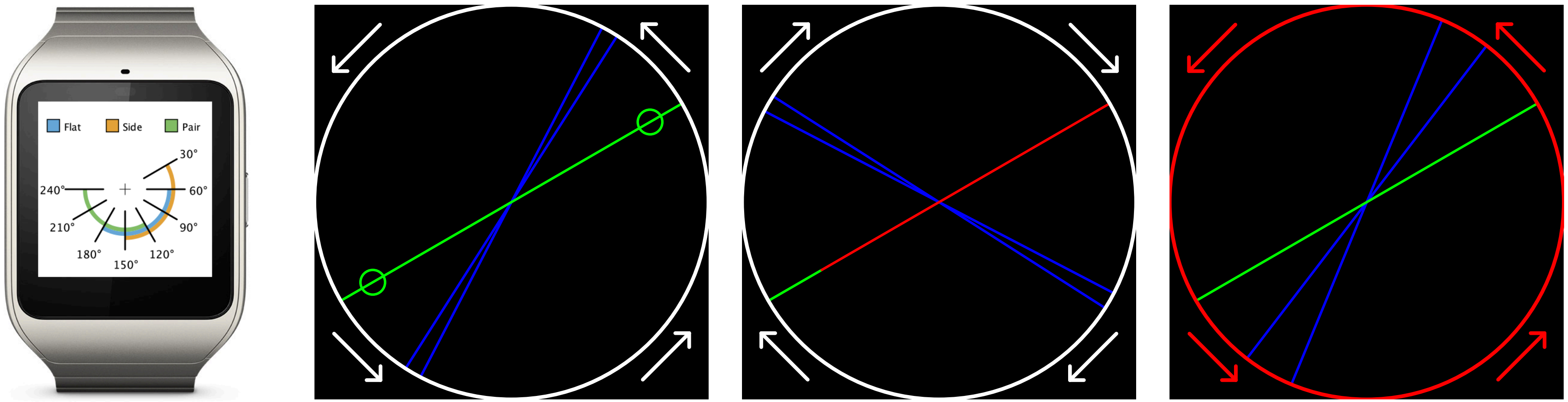


Figure 5: Left image shows the 120° angle ranges used in the three finger-region conditions in the dynamic study. Right three images show screen shots of the watch interface from dynamic study. The large circle and arrows in the corners indicate the rotational direction for the trial. The circle and arrows are white at the start of a trial (center two images) and turn red when the screen is touched (right image). In each trial two targets are shown – the initial target users must touch at the start of a trial, marked by a single red/green line, and the final target, marked by a symmetric blue wedge. After touching the initial target. the single red/green line serves as a cursor and the user's task is to rotate it so that it is within the blue wedge. Highlighting on the initial target indicates how it should be touched. Left-center shows two small circles, used to show appropriate finger positions for the pair condition while the large red portion of the line in the right-center image shows the touch angle for flat/side conditions.

Table 3: ANOVA and post-hoc test results in dynamic study. Data from non-significant interaction effects are not presented.

Measure	*Variable(s)*	*Outcomes*			
Preparation Time	Region by Direction	F(2,34) = 5.9	$p < 0.006$	$\eta_p^2 = 0.26$	–
	Region	F(1.4,23.7) = 9.24	$p < 0.003$	$\eta_p^2 = 0.32$	*Post-hoc*: flat>side and pair (p<=0.009)
	Direction	F(1,17) = 27.45	$p < 0.001$	$\eta_p^2 = 0.62$	–
	Size	F(1,17) = 10.06	$p < 0.001$	$\eta_p^2 = 0.37$	–
Touch Time	Region	F(2,34) = 5.96	$p < 0.006$	$\eta_p^2 = 0.26$	*Post-hoc*: side<flat and pair (p<=0.046)
	Direction	F(1,17) = 0.65	$p = 0.43$	$\eta_p^2 = 0.04$	–
	Size	F(1, 17) = 440.7	$p < 0.001$	$\eta_p^2 = 0.96$	–
Throughput	Region	F (2,34) = 8.22	$p < 0.001$	$\eta_p^2 = 0.33$	*Post-hoc*: side>flat and pair (p<=0.026)
	Direction	F(1,17) = 0.37	$p = 0.85$	$\eta_p^2 = 0.002$	–
Initial Errors	Region by Direction	F(2,34) = 16.3	$p < 0.001$	$\eta_p^2 = 0.49$	–
	Region	F(2,34) = 30.7	$p < 0.001$	$\eta_p^2 = 0.64$	*Post-hoc*: pair<side and flat (p<=0.001)
	Direction	F(1,17) = 6.2	$p < 0.023$	$\eta_p^2 = 0.27$	–
	Size	F(1,17) = 3.51	$p = 0.078$	$\eta_p^2 = 0.17$	–
Final Errors	Region by Size	F(2,34) = 3.81	$p < 0.032$	$\eta_p^2 = 0.18$	–
	Region	F(1,17) = 0.72	$p < 0.5$	$\eta_p^2 = 0.04$	–
	Direction	F(1,17) = 0.94	$p < 0.35$	$\eta_p^2 = 0.05$	–
	Size	F(1,17) = 82.58	$p < 0.001$	$\eta_p^2 = 0.83$	–

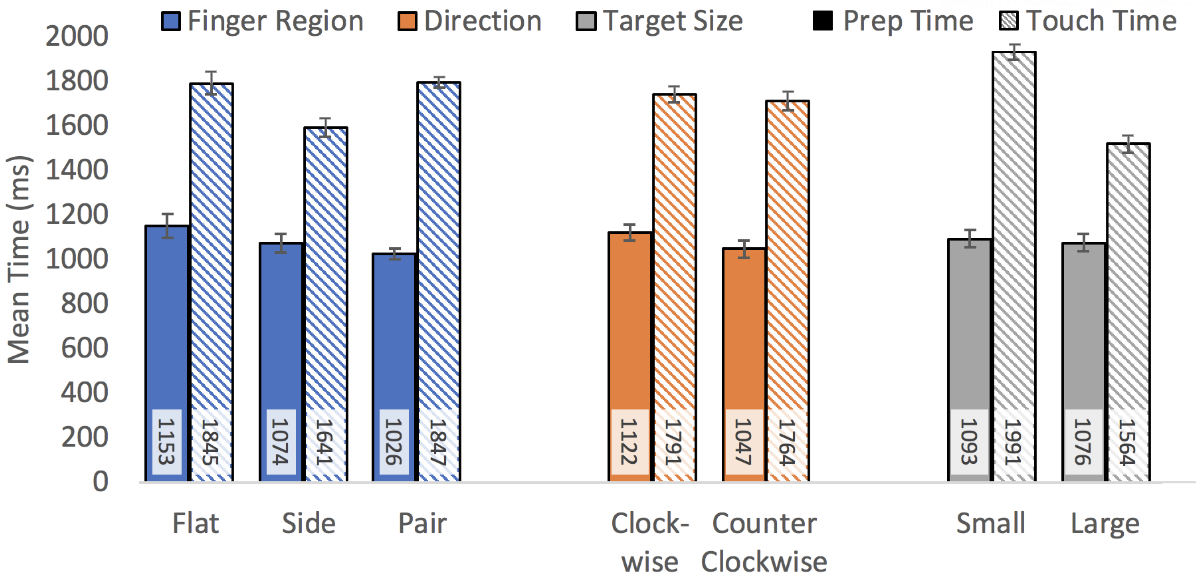


Figure 6: Preparation and Touch time data from dynamic study organized by three main effects. Bars show standard error.

accuracy adjustments in these models (i.e., we used ID not IDe), they simply represent temporal peak performance, rather than more nuanced figures that incorporate an error term. In contrast to the data from the static study we conclude that, for the 120° angular input areas we studied, touches using the side finger region provide an optimal combination of fast preparation and touch times for dynamic tasks. A likely reason for this is that participants could use finger flexion [168] in addition to hand and wrist movements during input.

Error rates are shown in Figure 7. The dominant result from the initial touch error data was that this was lowest in the pair condition, although still somewhat high (6.2%). While this indicates that pair touches are more accurate, it also suggests that larger initial targets are needed in general. It is also higher than predicted based on the accuracy data in the first study – success rates of 97% could be expected with targets set to 2.5 times SD of recorded touch accuracy. This may indicate that preparing to complete the dynamic task influenced the accuracy of participants' initial touches – they may have commenced rotational movements prior to screen touches due to a focus on the final steps of the task rather than the initial ones. Counter-clockwise rotations also led to significantly lower initial errors than clockwise rotations, indicating that the initial stages of this movement can be achieved more accurately.

Final error data flesh out these findings. We note that before analyzing these data, we removed trials involving very short touches (more than 3.5 standard deviations from the mean) as these were likely

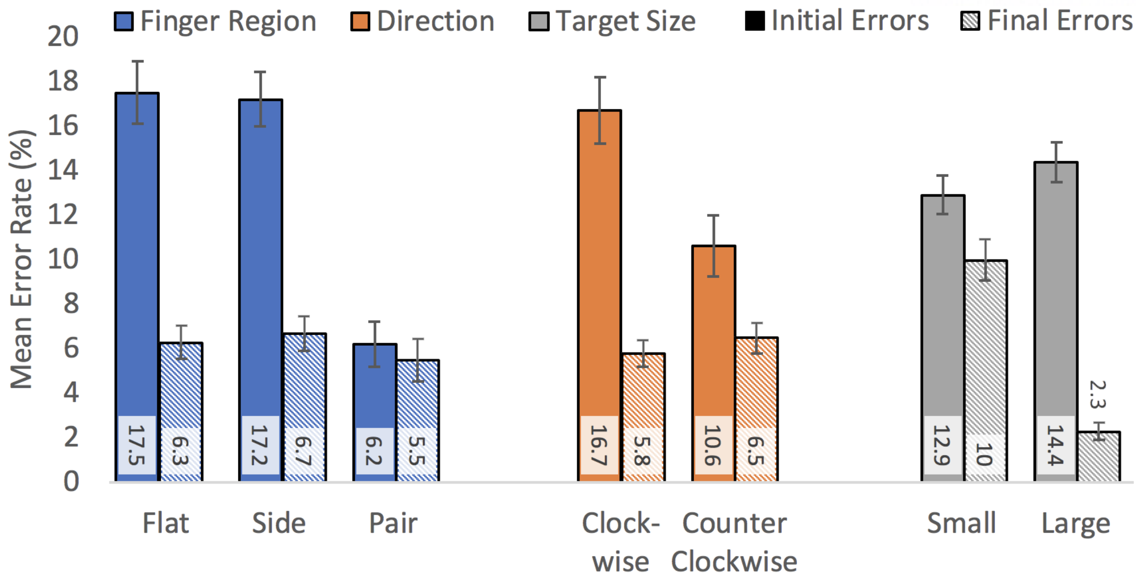


Figure 7: Errors in initial and final touches in dynamic study organized by three main effects. Bars show standard error.

due to unintentional artifacts of the experimental procedure rather than genuine attempts to complete the trials. It total, this accounted for 1.43% of total trials, evenly distributed over the different finger region conditions. The strong main effect of size dominated the final error results, revealing a very substantial difference between the high error rates with small 5° targets and very low rates with the 15° targets. In fact, even though the large final targets were smaller than the initial targets, they recorded much lower error rates, most likely due to the fact that performance was guided by a cursor.

Finally, in order to validate the logistic regression model that distinguished between flat and side touches in the static study, we tested it with data from the dynamic study. We once again extracted a single moment from the center of each flat and side trial, and used this data as a test set for the previously generated Weka model. This led to an accuracy of 95.6%, almost precisely the same as in the static study. This strongly suggests that touches with flat and side finger regions can be reliably distinguished from each other across all users and a range of different tasks.

Discussion

The goal of this study was to apply the outcomes from the static angle study, in terms of the comfortable input ranges, in a more complex setting. The primary result from this study is that the flat comfort ratings suggest these recommendations are valid - participants did not rate any of the input conditions in the study as more or less comfortable than any other. This is a positive result that suggests that future researchers and designers will be able to apply the comfortable angular input ranges and finger region combinations we propose to help them create novel smartwatch input techniques. We also note that these ranges not only differ between the different finger regions, but also from those identified in prior work [101, 279] on other platforms. This highlights the importance of assessing comfort for different device form factors.

A second key finding relates to temporal aspects of performance. In contrast to the static study, participants performed most rapidly with input via the the side of the finger – they showed a 10% – 11% increase in throughput during correctly performed trials, most likely due to the use of finger flexion during input. This result suggests that interaction techniques that rely on input via the side of the finger (e.g., [283]) may be particularly suitable for the watch form factor. In contrast to other devices, the small size and wrist mounted location of a smartwatch may make it uniquely accessible to touches with the side of the a finger. Its also worth contrasting rotational input performance in this study with prior work. While differences in methods and setting make direct comparisons impossible, the 2.4s–3.2s mean task completion times in the current study are broadly similar to those reported in prior work, such as the 2.5s–3s reported for two finger rotations of between 60° and 120° degrees on tablets by Hoggan *et al.* [274] and Voelker *et al.* [278]'s report of approximately 1.8s movement times for two finger rotations of between 50° and 150°. This latter result broadly matches the mean 1.44s touch time in the pair condition for targets that are between 30° and 120° degrees distant. These comparisons suggest that, despite their small size, rotation input is achievable on smartwatches at similar speeds to those attained on much larger input surfaces.

Beyond these points, we discuss the remainder of the study results in combination with those from the static angle study and in the form of recommendations for design. These are presented in the following section.

3.8 Design Recommendations

The studies support a range of recommendations for the design of smartwatch interfaces based on both the objective and subjective data. We collect these by theme below:

Input ranges. The comfort ratings from the first study show viable angles for input on smartwatches using three different finger regions. Based on differences from the most comfortable angles, and a threshold approximating the mean overall comfort rating, the side region is viable between with angles from 60° to 180°, the flat between 90° and 180° and two finger touches of the index and middle fingers between 180° and 270° (and/or 0° and 90°). We suggest that interface designs that require poses outside of these ranges are not likely to be well received by users. The objective data back this up: in the first study, preparation times increase in a moderately similar pattern to comfort, suggesting it is takes more time to adopt a pose suitable for input outside of these comfortable ranges. Data from the dynamic study, where comfort (and to some extent preparation time) are less variable, support these conclusions – designing for comfortable ranges of motion will increase usability of input techniques. We note that these comfortable input ranges extend those proposed in prior work on tabletops [101] and mobile devices [279] by considering a range of finger regions and focusing on smartwatch input – a context where both the device and touching finger are able to move.

Target sizes. The studies provide a rich characterization of pitch angle target sizes for smartwatch input based on standard touchscreen sensing technology. We consider three scenarios: a feedback-free absolute task in the static study, a variant of this task including only correctness feedback for initial touches in the dynamic study and a cursor driven task for final touches in the dynamic study. We can make a number of recommendations. For initial touches, a minimum target angular width would be 20°, as recommended in the static study. If tasks involve subsequent on-screen movements, larger targets of 30° or more are recommended. For tasks involving selection on release and a cursor indicating current angle, sizes can be lower: the 15° targets in the dynamic study led near error-free performance. While these figures represent safe minimums, it is worth noting that more accurate finger angle detection technologies (e.g. with less than the approximately 10° of error expected with our system [94]) may enable effective use of smaller targets.

Static vs Dynamic. We noted a range of performance variations between the static and dynamic studies. While we did not perform a direct comparison, mean percentage comfort ratings are better for static touches, even though we considered a broader range of angles. This suggests that participants found the dynamic movements more generally uncomfortable than the static touches: these simpler inputs should be used when possible. Objective performance also differed between these two scenarios. Specifically, we note a switch in touch time data. Optimal static touches were performed with the index and middle fingers, while optimal dynamic touches were performed with the side of the index finger. This suggests that different finger regions may be more and less suitable for these different types of input.

Distinctive touches. While distinguishing between paired and single touches is a trivial and commonplace part of modern interfaces, the data reported in this paper indicates we can also distinguish between flat and side finger regions. Using a simple regression model, we attained a high degree of accuracy

(~95%) that was valid across different participants and tasks. While authors have previously proposed distinguishing between tip and flat of the finger [66] or thumb [112], this new distinction develops this idea further and can be explored in future interface designs.

Finger Flexion. Finally, we note that finger flexion, previously proposed for in-air input scenarios [168, 244] may also be a viable and understudied modality for input on touchscreen smartwatches. In the static study, compared to the two other conditions studied, the side region offered a modestly wider range of comfortable input angles. In the dynamic study, it combined low task preparation times with low task execution times and correspondingly increased throughput. Error rates matched those from the more commonly studied flat finger touches [94] throughput. We suggest these benefits are likely due to the use of flexion and that future researchers and designers may be able to use this finger region to develop effective and comfortable input techniques for small devices.

3.9 Interaction Techniques and Applications

In this final section of this paper, we present the design of smartwatch interaction techniques and applications based on the data we present, and human-centered recommendations we derive, in this article. These largely adapt prior ideas with a particular focus on the side and flat touches, as these are less well studied than pair touches. We generated the following set of different input techniques.

Angle Menus. Menus which assign options to different angular ranges (like a pie menu) are a commonly proposed interface primitive for angular input [113], providing easy access to commands and placing limited memory load on users. In a *static* configuration, feedback about menu options should be displayed continuously (i.e., including when a user is not touching the screen), targets should be at least 30° wide and triggered on finger release in order to support adjustment of touches prior to selection. *Dynamic* angle menus occupy similar regions, but are not shown by default – they are summoned by a long touch (or dwell) with either flat or side finger regions. Subsequently, a user can make angular movements and they show highlighting or a cursor depicting the current selection. They support a greater target density (down to a minimum size of 15°) as they are intended to be used only with interactive feedback.

Shortcuts. A variant on static angle menus, shortcuts do not involve any graphical cues and are triggered only by a rapid touch and release of an angled finger on the screen. They enable rapid execution of a small number of commands based on quick angled touches. As with static angle menus, 30° targets are a minimum; due to the lack of graphical feedback inherent in the technique, larger 45° target sizes are recommended.

Sliders. Rotation input is well suited to setting continuous or analogue parameters such as volumes, times or percentages [94]. We contribute to this idea by specifying comfortable 90°-120° ranges for input with different finger regions: 60° to 180° for side touches, 90° to 180° for flat touches and between 180° and 270° (and/or 0° and 90°) for two finger touches of the index and middle fingers. Data from the dynamic study also provides insight into the accuracy of final angles in this kind of input: reliable input can be achieved with a target size of 15°, and a drop in accuracy (to ~10% error rate) will result from 5° target sizes.

Flex–Swipes. Reflecting the potential of using finger flexion input on watch screens, we propose a novel variation on swipes. In contrast to commonplace straight movements, Flex-Swipes involve a screen touch with the side finger region followed by a rapid contraction or extension of the finger. This creates a unique input event composed of a large and rapidly rotating touch area in either clockwise or anti-clockwise directions. We propose this primitive as a rapid technique suitable for interface actions such as mode changes (e.g. like flipping pages) or large-scale, fixed quantity rotations, such as changing screen or content orientation by 90°.

Region–Aware Rotation. The distinctiveness of flat and side finger regions can enable different types of rotation input, allowing users to specify different parameters with by using different finger regions in the same interface. For example, a user might control the brightness or hue of smart lighting with continuous sliders in same user interface by accessing these functions with different finger regions.

Example Applications

We integrated these ideas into three example applications implemented on the same Sony smartwatches used in the studies. Figure 8 shows these prototypes. The first application was for viewing and light-weight editing of photographs. In addition to traditional touches to issue commands and swipes to navigate between images, Flex-Swipes in clockwise and counter clockwise directions resulted in matching 90° image rotations. A dynamic angle menu on the top and left of the watch could be summoned with a prolonged (>500ms dwell) side region touch. The menu featured four items, each occupying 18° of angular space and representing an image filter. After selection, the filters were adjusted using a flat touch to control a slider shown at the top of the screen.

A second prototype implemented Static Angle Menus in the context of a messaging app. A list of

Figure 8: Demo applications. Left shows a user pointing to reply in a static angle menu in a message app using a flat touch. Center-left shows rotations of a side touch adjusting zoom in a map. Center-right shows selection of a brightness filter from a dynamic angle menu in photo app and right shows adjusting the brightness filter with a subsequent flat touch.

messages was displayed in the center/bottom-right of the screen, with a slim menu showing three options along both the left and top edges. Messages could be opened with a tap. Flat touches were used to select among reply, reply-all and forward commands along the top menu, while side touches could select from new, done and delete commands on the left. Rapid touches and releases led to immediate selection, or users could maintain their touch and make adjustments to the angle for high precision selection. Flex-Swipes were implemented as an alternative mechanism for rapidly issuing reply (clockwise) and forward (counter-clockwise) commands.

In the final map app, a canvas was displayed that could be translated with standard single finger drags. Region-Aware Rotation was used to access different functions. Rotations with the flat of the finger controlled orientation, while rotations with the side adjusted zoom. The map app also implemented three Shortcuts triggered by horizontal and diagonal side touches and vertical flat touches. The horizontal touch launched a search bar at the bottom of the screen (the most likely touch location) and the vertical touch a left side menu bar. Finally, the diagonal touch toggled between satellite and street view.

Application Study

To validate these designs, and make an assessment of whether the recommendations proposed in this paper can help developers create viable, comfortable interaction techniques for smartwatches based on angled or rotational input, we conducted a final qualitative study with 10 participants (six male, four female, mean age 23 (SD 2.04), all right-handed). These individuals had not participated in either prior study described in this paper and self-rated themselves as highly familiar with touch screen input (5/7), but not wearables (1.4/7). The study was based on procedures previously deployed to assess reactions to novel wearable interaction techniques [7, 289] and involved an experimenter performing a scripted demonstration of the input primitives available in each application, followed by the participants donning the watch, trying these out for themselves and commenting liberally.

We also asked a series of semi-structured interview questions about the usability attributes of each application that sought to ensure consistent coverage of issues including comfort, efficiency, learnability and memorability. We closed the study session by asking participants to comment on three general issues that appear across the whole set of techniques. These were: the contrast between static and dynamic rotations; the potential of rotation input to reduce the need for large on-screen targets and; the mapping of different inputs to different finger regions. The study took an average of 37 minutes to complete. The goal of this study was to complement the numerical, empirical ratings from the first two studies with direct capture of participants' reactions to the final designs and input techniques. We analyzed data in the study by recording all sessions and transcribing the comments verbatim. We then divided the transcriptions into meaningful semantic units (in the vast majority of cases, sentences) and performed an affinity process to cluster these comments into themes and categories. The semi-structured interview process, as it covered a specific set of topics and issues, provided an initial framing for these issues. We present key outcomes in terms of *comfort*, *performance* and *interaction design* below.

Comfort. We were particularly interested in comments relating to perceived comfort. These were positive, sometimes surprisingly so – for example, after using the mail app, P3 reported expecting the rotational "gestures will [induce] fatigue", but found this was not the case. Similarly, in closing comments, P2 noted that input expressivity is "surprisingly expanded" simply by "using different sides of [the] finger" via "tilting [the] wrist". Designing the inputs based on the findings from the initial studies helped achieve this. Expanding on his/her comments on the mail app, P2 indicated that activating the menu on the left of the screen with the side of the finger and top menu with the flat "matched [] expectation[s]", while P6 remarked that "diagonal touch with the side" of the finger (pointing from bottom right to top left) was "the easiest input" and should be assigned to very common functions. P6 qualified this comment by suggesting that rotations with the flat of the finger led to more "fatigue than other gestures" as the whole hand needed to be rotated. In more negative comments, P3 expressed reservations about input with the side of the finger due to the potential for "spoiling [his/her] nail" and its painting and decorations, while P1 noted that traditional single finger taps are "very comfortable" and questioned why rotational input is needed.

Performance. Participants were also generally positive about the input techniques in terms of their performance and effectiveness. Ten participants (across all three apps) remarked on the value of increasing the range of input available on a tiny screen. For example, after using the map app, P5 explained the rotational input lowered the need for a "messy screen with too many touchable menus" while still maintaining fast and convenient access to commands. P3 found the inputs in the mail app to be "easily expected and executed" and P10 saw value in using the flex-swipes as a generalized shortcut, suggesting they would be particularly "useful if [] customized". Two participants expressed doubts about the ability to perform rotational input in truly mobile scenarios – it was not "valid while moving" (P7). Learnability and memorability were closely related issues: most participants (15 mentions over the three apps) reported that although they needed "some time to be familiar" (P2) the rotational inputs were "not hard to learn" in "only a few minutes" (P3) in part due to their "consistency" (P10), "predictability" (P4) and, in the map app, the clear match between "menu bars and touch directions" (P2). There were concerns about whether older users could learn and memorize the inputs (5 participants), two participants doubted the discoverability of the techniques ("how can a user predict these gestures?" (P6)) and 9 participants, over all three apps, remarked on the departure of the techniques from standard two finger multi-touch operations, even though such traditional inputs were recognized as perhaps not "suitable for this tiny screen" (P6).

Interaction Design. Finally, participants also made various comments related to distinguishing or confusing the different inputs and their graphical presentation on the screen. In the gallery app, three participants felt that that "dividing the finger face into tool selection and [applying effects] is a very good idea" (P5) as they could be "naturally skilled" (P5) at differentiating between the two modes from the offset. Five participants felt the two rotational gestures in the map app readily supported input – P9 went so far as to claim that single finger rotational input was "easier than interacting with a smartphone map app". Some of the icons also supported the gestures: in the mail app, the four participants reported expecting that the "rapid flex-swipe input would be connected to reply and forward shortcut[s]" (P9) because of the similarity of these motions to traditional arrow icons for these commands. On the other hand, some participants also reported confusion between the different input modes, such as the two rotational inputs in the map app (3 participants), or the speed threshold that should be exceeded to trigger flex-swipes (5 participants). This was in part due to the fact that the app designs somewhat intentionally lacked consistency and differed "in each application" (P8), so needed to be memorized separately. Unsurprisingly, participants identified that "apps need to be made with consistent input gestures" (P1). Finally, just two users remarked on the fat finger problem [57] and only for the map app. P8 put it as the "screen of smartwatch is too tiny. I want to see how it is rotated". This highlights the fact that rotational input occupies large portions of a watch screen, potentially exacerbating the inevitable occlusion of content caused by the touching finger.

Discussion

In sum, data from these participants suggests our guidelines can yield input techniques that are both comfortable and effective. Participants reported pleasant surprises and few complaints about the input primitives. We take this as an endorsement of the design recommendations presented in this paper. More generally, participants also responded positively to the increased expressiveness enabled by the techniques and anticipated that using them would be relatively easy to learn and remember, so long as they are deployed across a system in a coherent and consistent manner.

3.10 Limitations

A number of limitations impact this work. Firstly, we considered a narrow definition of comfort focusing on specifying yaw angles. Extending this to include other forms of input including pitch [101] or roll [290] is a logical next step. Secondly, the studies also took place in a lab setting and assessing the validity of the data in a more realistic real world environment involving activities such as walking or interaction while encumbered [124] would be a valuable extension to the current findings. Thirdly, while we use an established, practical, feasible and real-time system for detecting finger angles on smartwatches, based on Xiao *et al.* [94], there are a number of weaknesses to this approach. These include assuming that Xiao *et al.* [94]'s existing characterization of input performance holds for the current study. We note that while the close similarity between the smartwatch systems (in terms of touch sensor size, controller and underlying kernel modifications) suggest this is a reasonable assumption, we do also apply their algorithms to a new setting: input with the side rather that the flat of the finger. While the absence of differences in accuracy data in the static angles study suggest this algorithm performs well in this setting, future work should formally assess this. We also note that more advanced algorithms for touch angle sensing have been recently proposed [284]; determining if these offer improvements in our study setting is a next step for this work. This issue is particularly pertinent as the algorithm we employed has a known error in accuracy of detecting touches (with near flat finger pitch) of approximately 10° [94]. This means the results and design guidance we employ are valid only for systems that perform at this level of input accuracy. While it is represents a realistic, real time platform for developing interactive systems, a more accurate sensing algorithm (e.g., Mayer *et al.* [284]) may impact some of conclusions relating to target sizes reported in this article - a more accurate sensor may enable targets to be shrunk. Finally, reflecting the different study designs, comfort was assessed in different ways in each of the three studies in this paper. This variability in measurement instrument may serve to confound the conclusions in this paper, and the use of different techniques reflects a lack of well-validated instruments or approaches in this area. Future work should seek to establish robust standards for evaluating issues such as comfort of interaction techniques.

3.11 Conclusions

Comfort is an important quality of any input technique. This paper notes that many recently proposed interaction techniques for smartwatches imply a broad range of postures, poses and arrangements of the fingers, hands and arms. The comfort of these kinds of activity is an unknown quantity. This paper contributes initial data on this issue via two lab studies. The first gathers data from simple screen touches and releases using three finger regions (or poses) and eight input angles. Comfort data show strong variations with the different regions, clearly highlighting that user experiences vary considerably depending on both the finger regions used to make input and the angle of the input being made. Building on these findings, a second study looks at dynamic touches that move from one angle to another in the optimal ranges identified in the first study. Results show flat comfort ratings, and highlight variations in objective performance. The data from these studies can be used to reflect on the design of prior smartwatch input techniques and to inform future schemes. The paper closes by contributing design recommendations and interface examples that showcase how this could be done and by assessing how users react to these designs in a final study. The results indicate that the design recommendations support the creation of comfortable, effective interactions. We believe the data and design recommendations that form the core contributions of this paper will help researchers and designers create more objectively and subjectively usable smartwatch input techniques in the future.

IV TriTap: Identifying Finger Touches on Smartwatches

4.1 Properties of Finger Movements in Finger Identification

The second scenario supports the general claim of this thesis by understanding the touch profiles on the small touchscreen of smartwatches to develop the finger identification system. The finger identification is a promising approach to increase the expressiveness of touch input on small surfaces by assigning different functions to each finger. This scenario begins with the assumption that the unique input context of smartwatches, in which the touching finger has a limited range of approachable angles during the touch input to the fixed touchscreen on the opposite hand, will generate the distinctive touch profiles of each finger. The empirical studies examined the distinctive touch profiles for the thumb, index, and middle fingers. Based on the touch profiles of each finger, the finger identification systems were developed to classify the two (thumb and index fingers) or three fingers (thumb, index, and middle fingers). The finger classifiers were evaluated with two types of virtual keyboards: Di-Type and Tri-Type with two and three letters on each key, respectively. Finally, through the recommendations and potential application designs, it was discussed how to apply this finger identification system based on the touch profiles to designing interfaces requiring increases in touch input expressiveness on the small touch surfaces.

4.2 Abstract

The small screens of smartwatches provide limited space for input tasks. Finger identification is a promising technique to address this problem by associating different functions with different fingers. However, current technologies for finger identification are unavailable or unsuitable for smartwatches. To address this problem, this paper observes that normal smartwatch use takes places with a relatively static pose between the two hands. In this situation, we argue that the touch and angle profiles generated by different fingers on a standard smartwatch touch screen will differ sufficiently to support reliable identification. The viability of this idea is explored in two studies that capture touches in natural and exaggerated poses during tapping and swiping tasks. Machine learning models report accuracies of up to 93% and 98% respectively, figures that are sufficient for many common interaction tasks. Furthermore, the exaggerated poses show modest costs (in terms of time/errors) compared to the natural touches. We conclude by presenting examples and discussing how interaction designs using finger identification can be adapted to the smartwatch form factor.

4.3 Introduction

Interaction with wearable devices such as smartwatches is a highly constrained experience. Small screens and touch surfaces provide few opportunities to create powerful, expressive interfaces using the conventional tap and swipe input popularized on larger devices such as smartphones. This problem is both fundamental, in that users want small, useful, discrete wearable devices, and interesting, in that it raises substantial new challenges for researchers and designers working in the space. Indeed, as more powerful and mature wearables come to market, increasing attention in the HCI community is being devoted to interacting with them in non-traditional ways, such as using their on-board inertial sensors to detect physical movements of the watch [65, 91] or gestural movements of the hand and fingers [62]. Commercial work in this area includes the Apple Watch (www.apple.com/watch) and its use of pressure input [291] and its crown, a novel physical controller.

While these approaches show considerable promise, the touch screens of smart watches remain powerful high fidelity sensors with a direct connection to the primary display surfaces of the device. As such, researchers are also exploring how to maximize the value of touch input on tiny screens through techniques such as diverse as tapping gestures [7], multi-touch menu systems [268] and inferring touch properties such as finger angle [94]. Within this space, one promising technique for increasing the expressiveness of touch input is finger identification [292]. At heart, this is a simple idea: if a system can process screen touches to disambiguate which of a user's fingers is responsible for a touch, then different functions or operations can be assigned to each finger. A range of prior work on this topic has discussed the interaction and application scenarios enabled by this technique in contexts as diverse as physical buttons [291], tablets [293] and tabletop computers [294].

However, while technologies to achieve finger identification have been achieved in fixed [96] or large format devices [295], they remain challenging to implement in small or mobile devices. In fact, most current work on this topic on wearables or mobiles simply instruments the touching fingers [68] to support system development or assumes instructions as to which finger should be used will be faithfully followed during studies [292]. While these approaches are effective for pursuing purely application design or empirical goals, they are also somewhat impractical – it is unlikely that real users will commit to wearing sensors or markers on their fingers simply to interact with another wearable device. Existing research to classify touches on small devices does exist, such as Harrison *et al.*'s [92] use of touch impact sounds to identify hand regions such as the fingertip, pad, nail or knuckle. However, little work has examined how we might use the properties of a touch to distinguish between fingers, rather than finger regions, on a small wearable device.

This paper aims to fill this gap by building on recent research showing it is possible to extract finger angles from the raw touch image generated by a standard capacitive touch screen [94]. Within this space, the contributions of this paper are threefold. Firstly, it proposes the idea that, in the constrained input poses available on a smartwatch (i.e. fixed to the wrist and approachable by the touching hand from a very limited range of angles) the touch profiles and angles generated by a user's fingers may by sufficiently distinctive to support reliable finger identification. Secondly, it contributes an empirical investigation of the validity of this idea: two studies of common smartwatch input techniques capture touch in both natural input conditions and those in which participants are instructed to exaggerate the angles of their finger touches. Our analysis describes the touches and discusses the efficiency and accuracy of user input and the recognition rate for finger identification in these scenarios. Thirdly, this paper contributes a discussion of how finger identification could be realistically implemented in current wearables and the types of interaction it could support. This discussion is showcased with application examples including two- and three-finger keyboard designs (and limited validations) and tricons, smartwatch application icons customized for use with finger identification technology.

4.4 Related work

Smartwatches and other wearable devices are powerful computational tools packaged with highly limited input and output capabilities. Authors are responding to the design opportunity this represents by proposing techniques to enhance and enrich interactions. The scope is broad, spanning topics as diverse as sensing input [296] or producing output on the skin [147], integrating touch sensitive surfaces into alternative parts of a device such as the strap [59] or edge [58] and utilizing body sensing techniques such as EMG [297] or tomography [172] to infer user actions. In contrast, finger identification has received limited attention. Gupta and Balakrishnan provide the most thorough exploration of the potential of this technique on smartwatches in their study of DualKey [68], a keyboard in which letters are clustered in horizontal pairs. Tapping on a pair with the index finger selects the leftmost letter, while tapping with the middle finger selects the rightmost letter. This effectively doubles target sizes and Gupta's comprehensive study highlights the resulting performance benefits. However, their system remains a lab prototype as it relies on a cumbersome finger mounted distance sensor to disambiguate touches by one finger from those by the other.

Finger identification has received more attention on larger platforms. Attracted by the simplicity of the technique there is a relatively large body of research that can be broadly categorized into technical systems for recognizing fingers or user centered investigations into the design of [292], or performance with [293], the technique. In terms of designs, most proposals are variations on the idea that specific fingers can be used to access different functions. For example, one might trigger a context menu [92], or provide easy access to commonplace commands such as cut, copy and paste. In terms of evaluations, studies have catalogued user performance with different fingers during tapping and dragging tasks on, for example, tablet computers [293]. Roy *et al.* [292] provide a laudably comprehensive review this literature. However, we note that while this work represents a substantial design and empirical resource relating to interaction with systems capable of identifying touching fingers, few articles have considered the specific form factor of smartwatches: small touch screens mounted on one wrist. Data and design guidance for larger devices likely needs updating or refinement for small screens.

Techniques for identifying the finger making a touch come in several forms. In the simplest, the finger is instrumented with a marker [104, 106] or sensor [68]. Another approach has been to infer individual finger locations in the context of multiple simultaneous touches based on the physical constraints imposed by the shape of the hand [294]. This technique has limited potential on the small screens of smartwatches. The Beats system [7], for example, accepts a pair of simultaneous touches on a watch and associates the left-most one with the index finger and the rightmost one with the middle finger to form sequences of tap gestures. However, all further position information is discarded, limiting the scope of the system for general purpose interaction. Other approaches for identifying fingers involve sophisticated touch surfaces capable of, for example, detecting the finger prints of the touching fingers [96, 105], or rely on advanced visual tracking systems positioned either under [298] or over a screen. As Gupta and Balakrishnan [68] note, none of these techniques is available or suitable for use with the current touchscreen technology available in smartwatches.

To address this technological lack, this paper explores whether the data reported by a standard capacitive touch screen in terms of finger contact area and finger angle [94] is sufficient to identify the touching finger. Many prior authors have recognized the value that can be gained by using finger contact area as an input modality. For example, Wang *et al.* [113] discuss the elliptical nature of touches on a tabletop and how this can be used to create interaction techniques such as ray based pointing. Other authors have applied these ideas to mobiles. Boring *et al.* [112] discuss how changes to the profile of a thumb touch on a phone can be used to shift between interface modes while Rogers *et al.* [114] show that tracking finger angle can improve pointing performance and create interfaces that automatically adjust for finger occlusions or scroll through menu options based on pitch. Recently Xiao *et al.* [94] discuss the accuracy with which single finger orientations can be inferred using the raw data from a standard watch

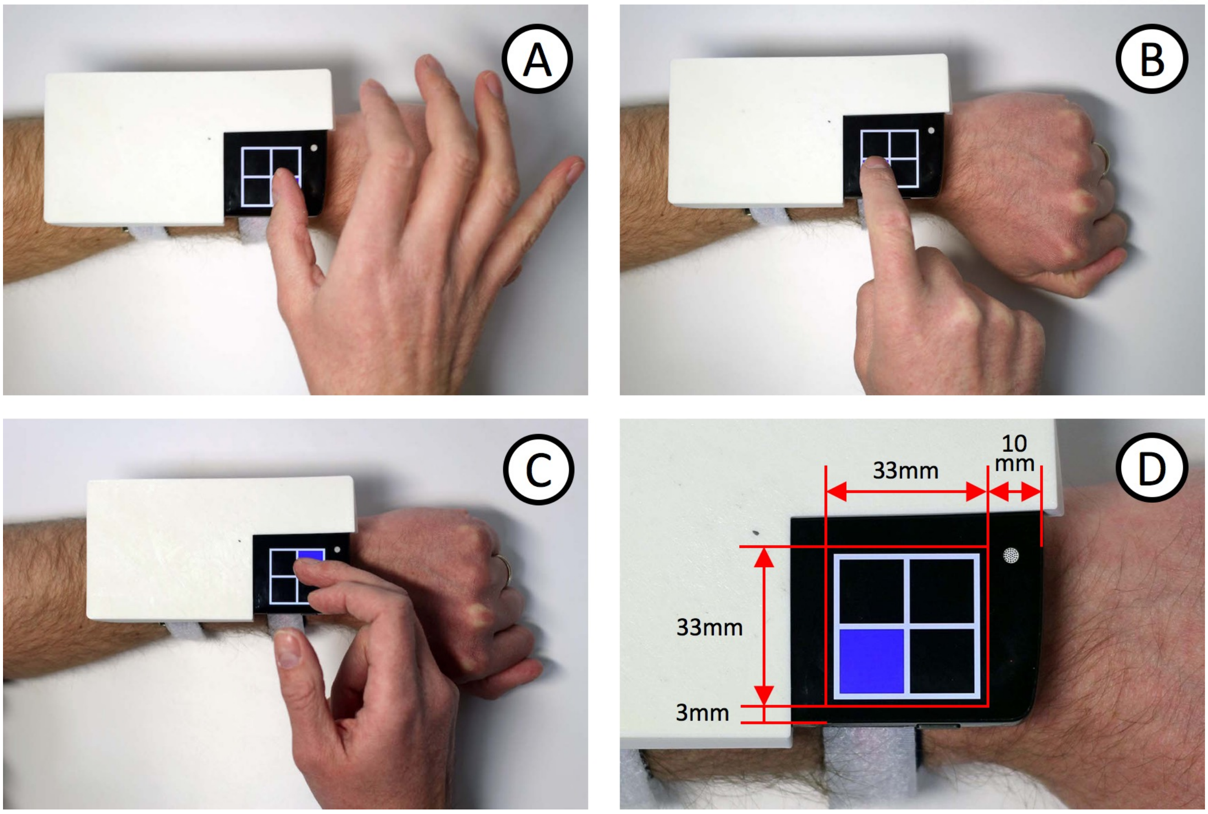


Figure 9: Experimental setup: wrist mounted phone showing four target tapping condition and exaggerated touches with thumb (A), index (B) and middle (C) fingers. Close up including annotations of sensor region and bevel sizes (D).

or phone touch screen. The goal of this paper is to build on these findings and ideas to create a robust finger identification system that operates with data from currently available touch screen technology.

4.5 Performance Study

The main goal of these studies was to explore the viability of using touch contact area to recognize fingers during interaction with a smartwatch. We considered three fingers: thumb, index and middle, as prior research has suggested that ring and pinky fingers are rarely deployed by users in touch input tasks on smartwatches [66] and, indeed, perform relatively poorly in situations where they are used [293]. In line with prior work, we conducted two studies to cover two common forms of touch screen input: tapping, or selecting targets by touching the screen, and swiping, or making rapid stroke gestures in cardinal directions.

We also considered two input conditions: natural and exaggerated. In the natural conditions, participants performed input tasks in any way they were comfortable. In the exaggerated conditions, partici-

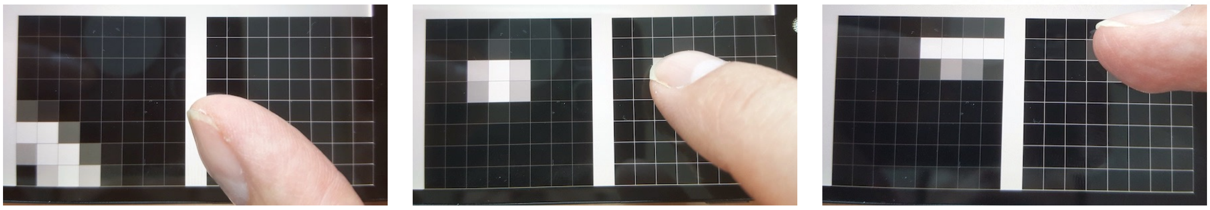


Figure 10: Three touches on the 8x8 33mm square touch sensor grid on the Nexus 5 phone. The thumb (left), index (center) and middle (right) fingers are touching the screen in the exaggerated poses. The left region of each image shows the sensor data generated by the touch on the right.

pants were instructed to make input by using the left side of their thumbs, the tip of their index fingers and the right side of the middle fingers. These poses are shown in Figure 9 (on our phone based study prototype) and were selected to maximize the distinctiveness of the touch contact profiles of the different fingers while also remaining convenient and comfortable in the context of smartwatch use. The goal of including both natural and exaggerated conditions was twofold. First, it enabled us to investigate the feasibility of inferring the touching finger during natural interaction. Second, it allowed us to characterize both the benefits of requiring participants to use a specific pose for input (in terms of finger recognition performance) and the costs of doing so (in terms of time, errors or comfort and workload). The following sections describe these studies in detail.

System

A wide variety of mechanisms have been deployed to gain access to detailed information about touch contact regions. These include using the standard reporting methods in the operating system [112], using specialized hardware such as fingerprint scanners [96], constructing bespoke sensor grids [66] and modifying the touch drivers on existing smart devices [94, 281]. We followed this latter approach as it provides high fidelity data while relying on commonplace and relatively high performance sensing hardware. Using a commodity device also supports our objective of exploring whether reliable finger identification is possible using current technology. However, as the main goals of this work are empirical, we opted to simplify the development process (in terms of better documentation and easier access to features such as network connectivity and storage) by using a region of an Android smartphone as a surrogate for a smartwatch. We note this is a common approach [111] and that touch sensors used in both classes of device are reported to be very similar [94].

As with prior authors [281], we modified the Android kernel to poll the touch screen driver for the touch image – the raw sensor data recorded by each capacitive electrode. This data varies with

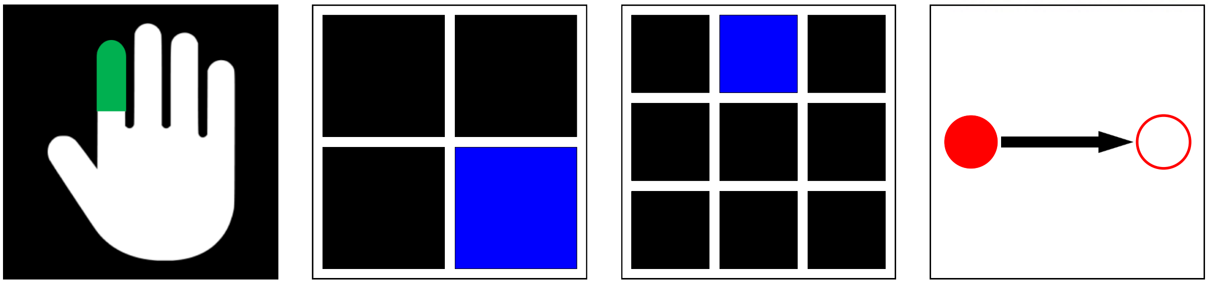


Figure 11: Study instructions. Left: the icon used to index what finger should touch the screen (index in the example). Center: the 2x2 and 3x3 button grids, each with one target highlighted in blue. Right: an example swipe instruction to drag the red target to the white one. All were shown on the 33mm touch screen area used in the studies.

the proximity of each sensor to a touching finger to form a greyscale intensity image that captures the finger-screen contact area and, to a lesser extent, the finger regions directly above the screen [296]. Our implementation ran on a Google Nexus 5 Smartphone and captured 16-bit touch intensity data in an eight by eight sensor grid covering an area 33mm square in the top right corner of the phone at 33Hz. During the initial studies, we processed this data using flood fill based blob detection and ellipse fitting [66] to derive a centroid for each screen touch. The modified kernel source files and example applications are available for download (https://github.com/UNIST-Interactions/tritap). Figure 10 shows the system in operation.

Methods

Twenty participants (mean age 22, nine female) completed the tapping study while nine (mean age 21, five female) completed the swiping study. They were recruited from the local student body via online methods and word of mouth and received 10 USD as compensation. We screened for right-handedness. Over both studies, participants rated themselves as familiar with smartphones (mean 4.5/5) and touch-screens (mean 4.6/5) but not wearables (mean 1.4/5). No participant completed both studies.

In both studies participants rested their left arms comfortably on a desk top in front of them and wore the smartphone strapped along their wrist using two watchbands attached to the back of the phone with Velcro. The top right corner of the phone was adjusted to be in a typical location for a watch: center of the wrist, just back from the hand. It also had the smallest possible bevels at its base (3mm) and right edge (10mm). A 3D printed cover obscured the rest of the phone and served to indicate the study touch area to participants. Figure 9 shows this setup, including annotations showing bevel sizes. All content in the study was shown in the 33mm square region used for data capture and each trial took the same form.

First participants tapped the screen to start. A hand graphic highlighting which finger to use was then presented for 1000ms, followed by the experimental trial. This took the form of a grid of targets in the tapping study and a single target and direction in the swiping study. The instructions are illustrated in Figure 11. As with prior work [292], we did not independently verify if participants used the requested finger in each trial. The task is simple and prior work suggests the compliance rate will be very high.

In both studies all participants completed natural input conditions prior to exaggerated conditions. This unbalanced repeated measures design ensured that the instructions given about poses in the exaggerated conditions did not impact the touches recorded in the natural conditions. Both studies were composed of sequences of identical trial blocks. In the tapping study three were three blocks, the first of which was treated as practice and discarded. In the shorter swiping study, there were five blocks, the first two of which were treated as practice and discarded. Within each block in both studies trials were delivered in a random order and participants were required to repeat error trials.

In the tapping study, two target sizes were used: a 2x2 grid and a 3x3 grid, corresponding to targets of 13mm and 8.25mm square. There was a 3mm border around the targets and an inter-target spacing of 1mm. Each block of trials was composed of a set of trials in the 2x2 grid and a set of trials in the 3x3 grid. The order of the sets was balanced among participants. In both sets, participants were required to complete six trials per target, three with each finger. In total, 6240 trials were retained for analysis (20 participants x 2 conditions x 2 blocks x (9+4) targets x 3 fingers x 2 repetitions). In the swiping study, each block consisted of a single set composed of two repetitions of each of the three fingers completing a stroke from one on-screen target to another in each of the four cardinal directions. The required stroke distance was always 2cm. In total, 1296 trials were analyzed (9 participants x 2 conditions x 3 blocks x 3 fingers x 2 repetitions x 4 directions). Stroke direction was not treated as an independent variable.

For each trial we recorded: the preparation time, the span between the start of each trial and the first touch to the screen; the touch time, or period in which the finger was in contact with the screen; the error rate in terms of successful completion of the requested interface operation (e.g. selecting a button) and; the stream of 8x8 raw sensor data. In order to acquire reliable data, all touches were required to generate at least three packets of sensor data – given the system's 33Hz update rate, this meant 90ms of touch time. In trials when the user touched the screen for less than that time, they were required to repeat the trial, and no data or targeting error was recorded. The NASA TLX was used to capture workload after the natural and exaggerated conditions were completed in both studies.

Performance Results

Initial analysis of the data focused on the fundamentals of performance: time, errors and workload. Specifically, we report on time and error data per finger in the natural condition in order to contrast this smartwatch data with that derived from related studies on larger form factors such as tablets [292, 293]. We also compare data in the natural condition with that in the exaggerated condition to understand the impact of requiring the user to adopt specific finger poses. Finally, in the tap study, we also examine the differences between the small and larger targets. All analyses, except where otherwise mentioned, were factorial RM ANOVA incorporating Greenhouse-Geisser corrections to adjust for sphericity violations and followed-up, if required, by post-hoc t-tests incorporating Bonferroni corrections. We also report effect size for ANOVA results as partial-eta squared (η_G^2). For brevity, only significant results at p<0.05 are reported.

Figure 12 shows the preparation and touch time data from the tap study. Preparation time showed two significant interactions: number of targets by finger (F (2, 38) = 4.094, p<0.05, η_G^2=0.177) and condition by finger (F (2, 38) = 40.44, p<0.001, η_G^2=0.68). All main effects were also significant: condition (F (1, 19) = 20.067, p<0.001, η_G^2=0.521), number of targets (F (1, 19) = 111.427, p<0.001, η_G^2=0.854) and finger (F (1.46, 27.8) = 58.5461, p<0.001, η_G^2=0.775). Interpreting these results in terms of the three strongest effects, we can say that selections of the smaller targets required more preparation time and that this effect was stronger in the exaggerated condition and specifically with the thumb and middle fingers. The touch time data showed fewer differences, but a similar story. Only the condition by finger interaction (F (1.049, 19.93) = 15.899, p=0.001, η_G^2=0.456) and main effects of condition (F (1, 19) = 10.404, p<0.01, η_G^2=0.354) and finger (F (1.055, 20.042) = 15.912, p=0.001, η_G^2=0.456) attained significance. This suggests that the interaction is the key effect in this case and the difference can be simply explained by the increased touch time in the thumb and middle finger trials in the exaggerated condition.

Figure 13 shows preparation and touch time for the swipe study. All preparation time and touch time comparisons were significant. For preparation time the figures are: interaction (F (2, 16) = 21.706, p=0.001, η_G^2=0.732) and main effects of condition (F (1, 8) = 23.186, p=0.001, η_G^2=0.743) and finger (F (2, 16) = 26.627, p=0.001, η_G^2=0.769). For touch time, these data are: interaction (F (1.27, 9.734) = 32.687, p=0.001, η_G^2=0.803) and main effects of condition (F (1, 8) = 21.151, p=0.002, η_G^2=0.726) and finger (F (1.147, 9.177) = 31.213, p=0.001, η_G^2=0.796). The interactions are again the dominant effects and these data reinforce the findings from the swipe study that the use of exaggerated poses for the thumb and middle fingers negatively impacted task completion times.

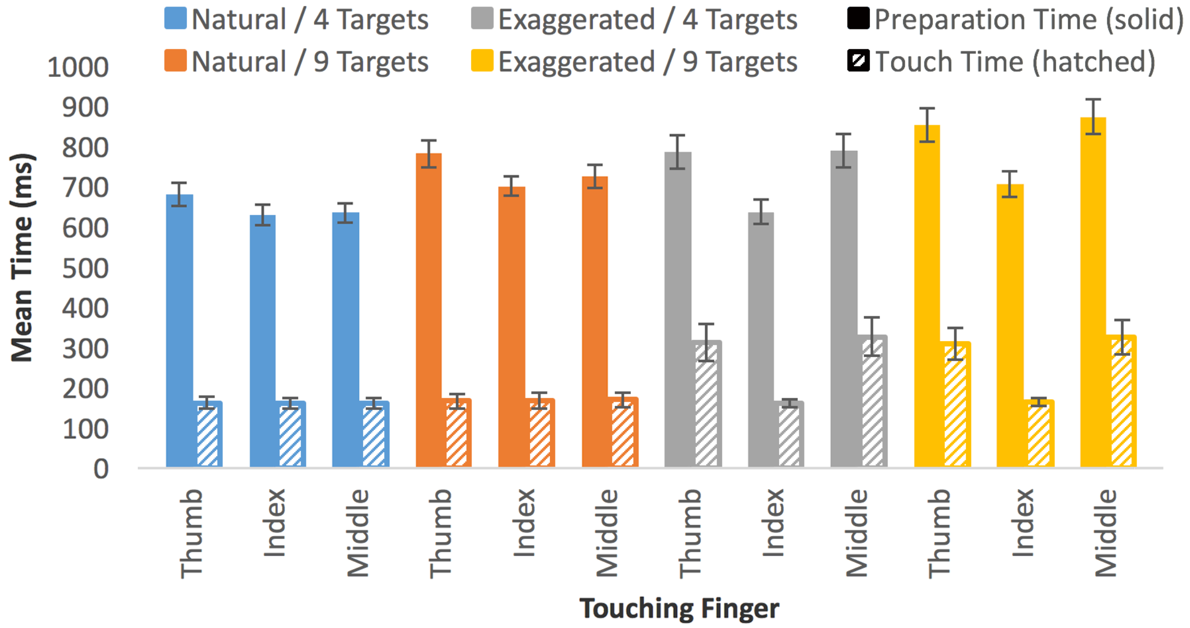


Figure 12: Mean Preparation and touch times from the tap study by finger and condition. Bars show standard error.

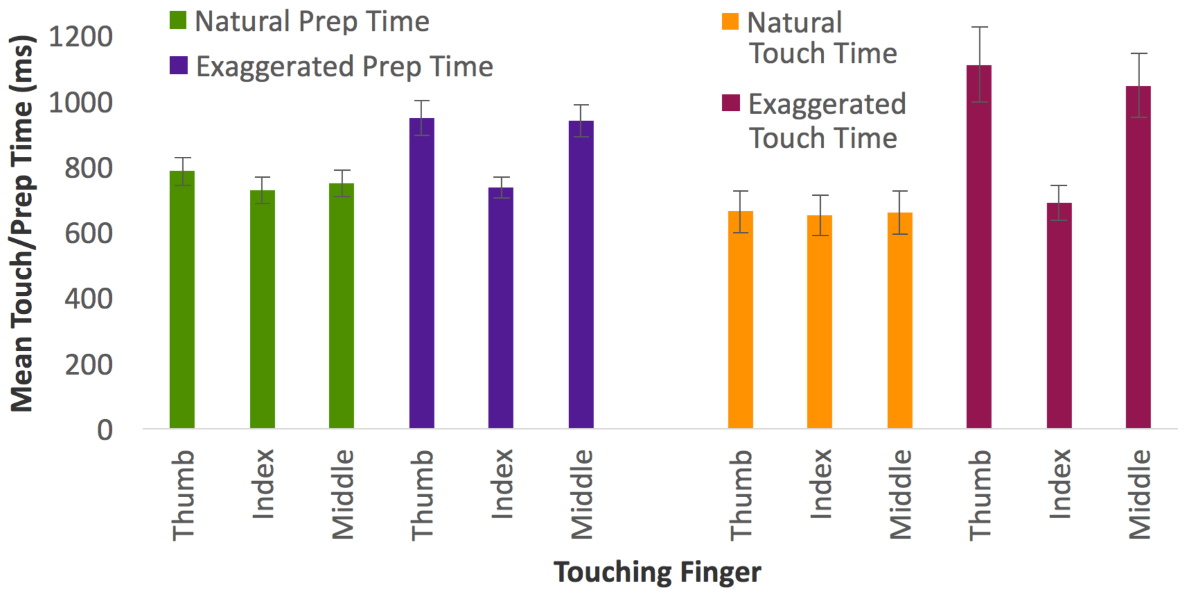


Figure 13: Mean Preparation and touch times from the swipe study by finger and condition. Bars show std. error.

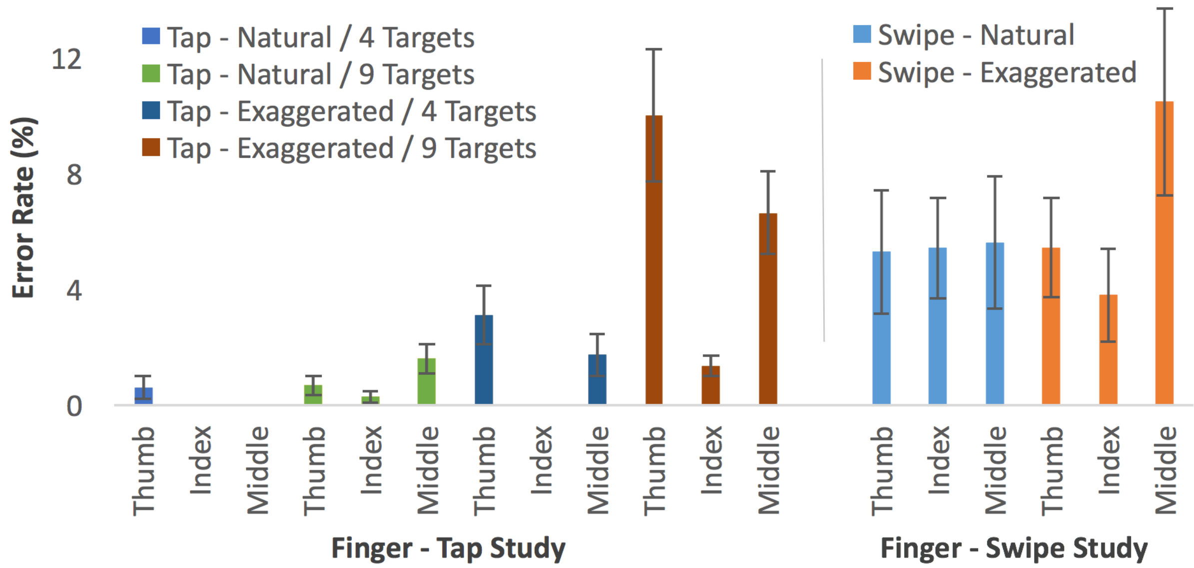


Figure 14: Error rates from both tap and swipe studies by finger and condition. Bars show standard error.

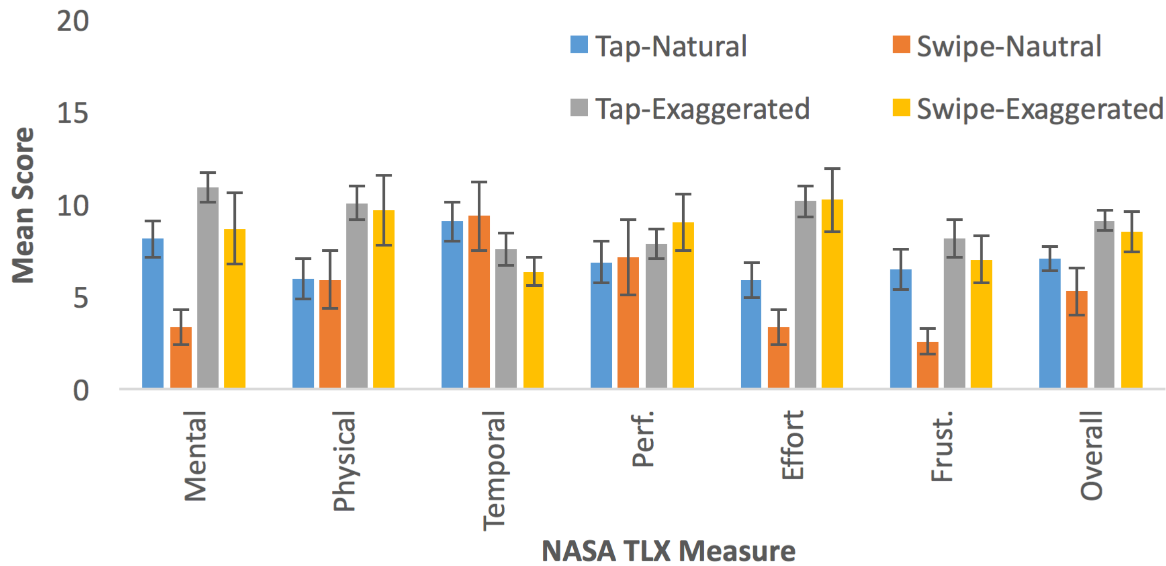


Figure 15: Mean TLX workload scores from both tap and swipe studies by condition. Bars show standard error.

Error rates are shown for both studies are shown in Figure 14. Errors in the tap study did not feature normal distributions – the mode for all bar one combination of conditions was zero. As such, we analyzed these data with three separate Friedman tests, one for each variable. All three returned significant results: finger ($\chi^2(2) = 12.5$ p = 0.002), number of targets ($\chi^2(2) = 20.0$ p < 0.001) and condition ($\chi^2(2) = 10.889$ p = 0.001). Follow-up Wilcoxon tests indicated that the index finger resulted in significantly lower error rates than the thumb (Z = -3.057, p = 0.002) and middle finger (Z = -3.7, p < 0.000). Beyond confirming the additional challenge of smaller targets, these results also indicate that tap performance is optimal with the index finger and natural input condition. Although we were not able to examine interactions, the chart suggests that these effects are largely due to the spike in error rates with the thumb and middle finger when completing trials with the smaller targets. Error rates at other times remain relatively low. In contrast to these variations, error data in the swipe study were fairly flat. They were also somewhat higher, most likely due to the fact the compound dragging task was more challenging, and distributed more normally. As such, we analyzed them with a single three-way RM ANOVA. However, no comparisons in the swipe study reached significance at the p<0.05 level.

Finally, TLX data are shown in Figure 15. In the interests of brevity, we describe overall workload as a representative measure. This data was analyzed with matched pairs t-tests to contrast performance in the natural and exaggerated conditions in both studies. The results indicate that the natural conditions received lower ratings of workload than the exaggerated conditions (both p<0.001), mirroring the trend suggested in the chart.

Classifier Results

In order to build a finger classifier for the touch images, we first selected a single touch image from the temporal center of the data for each trial. This is because finger touch profiles at the start and end of a touch (the moment a finger touches or release the screen) may vary substantially from those during the middle portion of a touch [113], when the finger is fully in contact with the screen surface. We wanted to exclude these transient data points. We then generated ellipses for all these touches using both blob tracking [66] and image moments [94] approaches. Ellipses were defined as angle, major and minor length and eccentricity. In the tap condition, we also recorded the x and y center with respect to the current target, while in the stroke condition we just logged the raw center position. Visual inspection revealed the image moments led to ellipses that better matched the raw data, most likely due to the fact that the blob tracking approach thresholds the image to black and white rather than considering it as a greyscale image. Following prior authors [94, 281] we also applied several gamma corrections to the

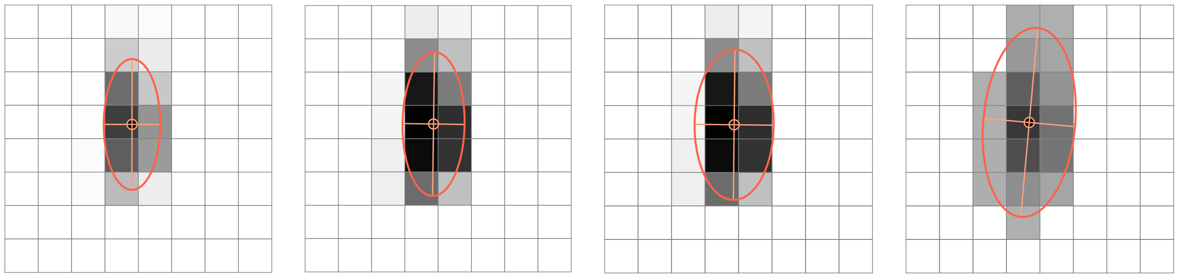


Figure 16: The four touch images used for classification. Left: raw-image, left-center: power3, right-center: thresholded-power3 and right: thresholded-log. Red ellipses are calculated from each figure's image moments.

image to enhance the ellipses, although we found different parameters more effective. Specifically, we created three adjusted images: transformed to the power three; thresholded at 5% of the maximum reported data value then transformed to the power three and; thresholded and log transformed. Figure 16 depicts an example of the four touch images generated and Figure 17 shows the mean ellipses derived from all the raw images for both natural and exaggerated touches in both the tap and swipe studies.

We used this data to construct recognizers for the touching finger using Weka [287]. All recognizers were built using a ten-fold cross validation process and Random Tree or Random Forest decision trees. We selected these techniques as they are mature and relatively quick to execute (so suitable for small devices). In the following description figures and statistics are included for clarity, but we note that Table 4 summarizes all the content reported in terms of recognizers, datasets, attributes and results.

Visual inspection of the raw touch data indicates touches in the exaggerated condition are highly distinctive, while those in the natural condition show substantial overlap. Accordingly, we first constructed static recognizers based on all data from the exaggerated conditions in each study. For the exaggerated tap data, class-wise histograms showed the attributes of eccentricity and orientation had high discriminatory power. We used these attributes to construct a simple three level Random Tree and achieved a mean accuracy of 98%. However, applying the same approach to the swipe data led to a lower mean accuracy: 92.3%. In addition, class-wise performance varied considerably (Kappa: 0.88), with the middle finger at 96.3% accuracy and the index finger at 86%. This suggests that the dynamic touches in the swipe study are harder to classify than simple taps. In order to increase performance, we constructed a 10 tree Random Forest using the full description of the ellipses from the raw data set: position, size, angle and eccentricity. This attained a mean accuracy of 97.7% (Kappa: 0.96). We believe these figures are sufficiently high to reliably identify fingers if users are instructed to use exaggerated touches on a

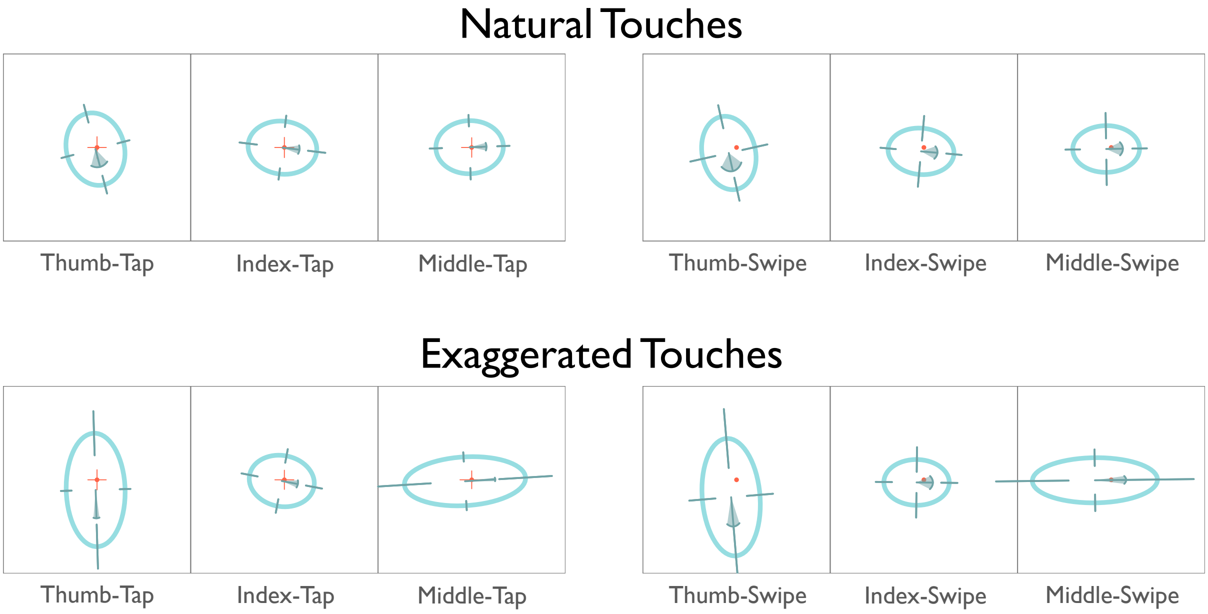


Figure 17: Mean ellipses from the raw touch image for each study and finger. Standard deviations for angle and major/minor axis length are shown via the arc and bars. The red dot marks center of the target and displacement from the center of the ellipse marks the mean center position; red bars show its standard deviation.

Table 4: Results from the machine learning models constructed to analyze touch shape during tap and swipe.

Recognizer	***Attributes***	***Touch Image(s)***	***Condition / Data Set***	***Accuracy***				***Kappa***
				Mean	**Thumb**	**Index**	**Middle**	
Random Tree (Static, 3 deep)	Eccentricity, Orientation	Normal	Tap-Exaggerated	98.6%	98.5%	96.4%	99.1%	0.97
			Swipe-Exaggerated	92.3%	94.5%	86%	96.3%	0.88
			Tap-Natural	61.9%	82%	12.6%	91.1%	0.43
			Swipe-Natural	55.3%	80.6%	71%	14%	0.32
Random Forest (Static, 10 trees)	Eccentricity, Orientation, Major, Minor, CenterX, CenterY	Normal	Swipe-Exaggerated	97.7%	98.6%	99.1%	95.3%	0.97
			Tap-Natural	67%	85.2%	50.6%	65.3%	0.50
			Swipe-Natural	62.7%	80.1%	53.3%	54.7%	0.44
Random Forest (Static, 100 trees)	Eccentricity, Orientation, Major, Minor, CenterX, CenterY	Normal, Power3, Thresholded-Power3, Log	Tap-Natural	68%	88.1%	53.8%	62.1%	0.52
			Swipe-Natural	65.4%	85.6%	55.6%	54.7%	0.48
			Tap-Natural - Long	70.6%	90.7%	55.6%	66.3%	0.56
Random Forest (Per-user, 10 trees)	Eccentricity, Orientation, Major, Minor, CenterX, CenterY	Normal	Tap-Natural (mean result)	79.4%	93.4%	70.4%	74.7%	0.69
			Swipe-Natural (mean result)	72.1%	83.4%	65.3%	67.2%	0.58

smartwatch.

Unsurprisingly, applying these relatively simplistic approaches to data from the natural conditions, where we expect both a greater diversity of touches and a less distinctive set of features, resulted in lower accuracies. To boost accuracy, we first created static models using all ellipse attributes from our four data sets, an approach similar to Xiao *et al.* [94], and increased the number of Random Forest trees to 100. For the tap data, this led to an overall accuracy of 68%. Class-wise performance was split (Kappa: 0.52) with the thumb at 88.1% and the index and middle fingers showing lower performance (53.8% and 62.1%, respectively). Following Wang *et al.*'s [113] observation that touch profiles on screen are time varying, we also examined performance with this recognizer on the subset of the tap data extracted from longer trials – those that recorded at least 150ms, or five packets of data. In total this was 1568 trials (50.2% of the original set), spread evenly over the three finger classes. This led to a modest improvement in the mean recognizer performance to 70.6%. Finally, we explored the impact of individual differences on performance by generating separate per-user models (using a 10-tree random forest on all data from the normal touch image) yielding a mean accuracy of 79.4%, again with best performance for the thumb (93.4%) and lower performance for the index (70.3%) and middle (74.4%) fingers. Applying these same approaches to the swipe data led to lower figures: the mean accuracy of the per-user models was 72.1%, maintaining the trend in which the thumb is more distinctive (83.8%) than the index (65.3%) and middle (67.2%) fingers.

Discussion

Mean performance in the tap study was relatively fast and accurate compared to prior work documenting performance on smartwatches [111]: overall mean task time was 947ms and error rates were 2.1%. The more complex swipe actions, which took the form of a drag between two targets, took longer and yielded more errors: 1617ms and 6%, figures that are again consistent with prior work [293]. It is informative to compare this data with Goguey *et al.*'s [293] and Roy *et al.*'s [292] recent examinations of different finger input on tablets. Differences in study design and objectives make direct time comparisons challenging (Roy terminates time measurement on first screen contact, while Goguey's work examines the span of a pair of screen contacts), but both articles report performance variations for different fingers. Data on thumb, index and middle during tapping and stroking reveals temporal and error data follow the V-shape also observed in the current study: performance is optimal with the index finger. Error rates for tap in these articles are reported in the range of 1.9%-3.2%, figures consistent with all bar the smallest targets in the exaggerated input condition. These comparisons suggest that performance in the current study was typical and serves to confirm that findings on finger input reported in prior work on tablets also applies to the smartwatch form factor.

Beyond establishing this baseline, the main objective of capturing time/error data in the current studies was to explore the costs incurred by requiring users make specific finger poses during interaction. These costs are clear: in the tap study, preparation times edge upwards (from 704ms to 824ms) when users have to make touches in specific poses; touch times nearly double (from 167ms to 318ms) as they actually make these contacts. In the swipe study, time data show similar trends. Error data tells a somewhat more complex story, with data remaining relatively flat when tapping large targets and during swipes and spiking dramatically in selection tasks with smaller targets. This suggests that finger identification techniques based on touch contact profiles might be best applied to tasks involving coarse-grained targeting actions. Finally, TLX data confirm these variations caused participants to feel increased levels of workload in the exaggerated conditions.

Despite these costs, we note that performance remains within acceptable levels in exaggerated study conditions: mean task times of 1142ms (tap) and 2022 (swipe) and error rates of 0-5%. Furthermore, variations in workload (and other data) need be treated with caution due to the fact that the natural condition always preceded the exaggerated condition. While this design prevented the exaggerated instructions from biasing natural behavior, it may mean that fatigue is artificially inflating differences. Evidence to support this idea comes from the fact that the lowest workload scores were recorded in the first condition administered in the more complex but shorter swipe study. We also note that workload levels remain generally low – from 3-10 out of 20 – across the whole study. This suggests participants never felt the tasks to be high demand. In sum, while the costs of using specific poses for finger identification are clear, their magnitude is relatively limited. While these costs may make the technique unsuitable for highly frequent interactions (like the repeated and prolonged tasks in the studies), we argue they likely remain acceptable for the more sporadic use scenarios that are more typical of genuine device operation in the real world.

The flipside of documenting these costs to performance is an exploration of the benefits of exaggerated finger poses for finger recognition. These are powerful. Recognizers in the exaggerated conditions are simple and high accuracy (98%). This suggests that touches made under these constraints could be effectively deployed in real interfaces. Results for recognizers in the normal conditions are less clear cut. Although the 72.1% (swipe) to 79.4% (tap) mean accuracies achieved in models created for each user show promise, they are clearly insufficient for deployment in realistic interfaces. Performance with static models covering the whole participant population is worse yet. However, examining the class-wise error rates suggests the picture is more nuanced. Specifically, the thumb can be identified relatively reliably (up to 93.4% with the per-user models) while the index and middle finger remain challenging to distinguish (70.4%-74.7%). Furthermore, the confusion matrix (not pictured) from the per-user tapping models reveals that thumb and index finger taps are cross-classified only 2.8% (thumb as index) and 3.1% (index as thumb) of the time. This indicates that it may be possible to create reliable finger identification systems using our approach if only the thumb and index are considered. As such, we note that while it would be infeasible to use our system to directly enable interfaces such as Gupta and Balakrishnan's [68] index plus middle finger keyboard, it might be possible to adapt these interfaces to leverage the optimal performance of our recognizers via, for example, the use the thumb, or by requiring exaggerated touches. This observation highlights a critical point: the design of finger recognition interfaces needs be informed by the properties of the underlying recognition system.

The remainder of this paper explores this issue: given the user performance constraints captured and recognizer accuracies documented, what interface designs are useful, effective and feasible? We investigate this issue in two ways. Firstly, inspired by the DualKey system [68], we implement and study two finger-identification powered keyboards. The goal of this work is to understand real world performance with the static recognizers and input modes documented and proposed in this paper. Rather than just relying on the data from the studies, we use this complex task to push the boundaries of the system and observe how recognizer and user performance instantiate and interact in a more realistic task incorporating, for example, immediate graphical feedback relating to the outcomes of a user's actions. We close by providing a more general discussion that consolidates all the work in the paper into practical recommendations and design examples.

4.6 Di-type and Tri-type Keyboards

We created two smartwatch virtual keyboards that use static natural (100 tree random forest) and static exaggerated (random tree) tapping models generated from the data captured in the first tapping study. For natural taps, we created Di-Type, a dual finger design, with two letters marked on every key. Tapping with the thumb recorded the leftmost letter and tapping with the index or middle finger recorded the rightmost letter. This reflects the fact that the natural model recorded higher accuracy for the thumb and lower scores for the index and middle fingers. For the exaggerated touches, we placed three letters on each key to create Tri-Type. The leftmost key was activated by the thumb, the center one by the index and the right one by the middle finger. This keyboard requires participants to mimic the touch poses used in the exaggerated conditions of the studies. The keyboards were both ordered alphabetically to facilitate novice users in the task of locating letters. Where possible we also used common key arrange-

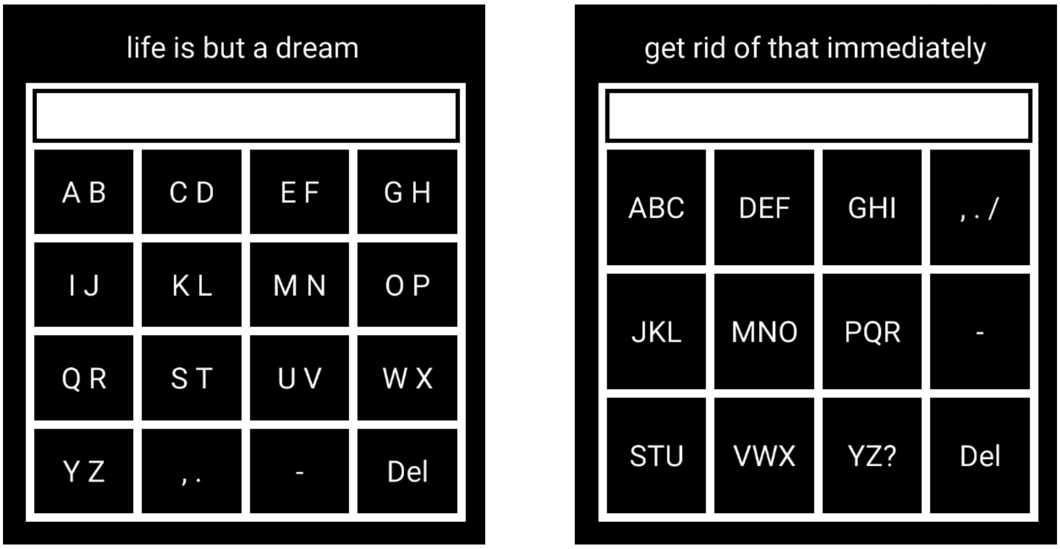


Figure 18: Di-Type (left) and Tri-Type (right) keyboards. Thumb taps selects the leftmost key. In Di-Type, the index or middle finger select the right key. In Tri-Type, index selects the center key and middle the right key.

ments, such as the 3x3 arrangement of letter triples in Tri-Type (see Figure 18). Both keyboards were 33 x 33mm in size. They featured a 2.1mm border, an inter-key spacing of 0.6 mm and a text display bar at the top with a height of 4mm. The buttons divided up the remaining space equally: 6.75mm x 5.75mm and 6.75mm x 7.85mm, respectively, for Di-Type and Tri-Type. They are both shown in Figure 18.

We performed a limited evaluation of these systems with 11 participants (mean age 21, six female). All participants entered 30 randomly selected sentences from Mackenzie *et al*'s [299] phrase set using both keyboards. As with the earlier studies participants always used the natural Di-Type system first. The first 15 phrases entered with both keyboards were considered practice and not analyzed. During their initial use of the keyboards, participants were encouraged to explore the keyboard, the novel finger-identification input scenario and the finger recognition process freely for up to 30 minutes. In the natural condition, they were not given formal instructions on how to touch the keyboard, but an experimenter did demonstrate how to use the keyboard if requested. In the exaggerated condition, the finger poses were demonstrated to participants. In total the experiment took approximately 90 minutes per participant and each was compensated with 15 USD in local currency.

The goals of this study were more focused on validating the recognition performance than on text entry performance. As such we logged raw Words Per Minute (WPM) to support a basic comparison with prior work [68] and asked participants not to correct any errors. The primary measures were then calculated from the text streams. We classified each character as either correct, or as a wrong-key error (meaning the wrong keyboard key had been selected) or a wrong-finger error (meaning that the wrong finger had been used or recognized). If participants entered additional or insufficient characters in a string, this was treated as an error and they were required to enter another string in order to complete the study. They could also tap the top of the keyboard to cancel a trial at any time.

Results and Discussion

Over the course of the study a total of 8164 characters were entered and retained for analysis. Participants achieved a mean of 8.08 (SD 0.78) raw WPM with Di-Type and 7.53 (SD 0.87) raw WPM with Tri-Type, figures that a paired t-test revealed were not significantly different from one another. These figures are somewhat slower than those recorded in the initial sessions of Gupta and Balakrishnan's [68] DualKey – they report mean WPMs of around 10.8. There are many possible explanations for this. One is that the finger recognition system used in the studies took longer for users to operate because of, for example, its reliance on the thumb, or its use of three fingers or specific poses. We also note that while our participants were engaged in study at an institution whose language of instruction is English, none were native readers of Latin characters. A dedicated comparison study with an implementation of both systems would be required to understand the cause of these differences. Instead, we note that the WPM figures indicate participants were able to type using both systems at a reasonable speed on a tiny screen.

More interesting are the error results. First, we recorded a mean of 1.45 (SD 1:87, median 1) sentences with an incorrect number of characters. Participants also cancelled entry processes when they observed extra characters – on average 4.13 times, a distribution skewed by one participant who frequently performed this behavior (SD: 8.2, median: 1). We did not analyze these trials further, but their presence does indicate that participants did at times cancel tasks on noticing errors involving entry of extra characters, potentially skewing the data towards more successful trials. Given the relatively infrequent occurrence of this behavior, we do not believe it exerted a strong effect on the study. Wrong-key errors were low thorough the study, running at means of 1.7% (SD: 1.2%) with Di-Type and 1% (SD: 0.9%) with Tri-Type. A matched t-test revealed these figures were not significantly different. This indicates that participants were able to select the small keys used in the keyboards with a very high degree of accuracy and regardless of the use of the exaggerated touch pose. Wrong-finger errors were more commonplace. Means per participant were 8.3% (SD 4.4%) with Di-Type and 10% (SD 3.8%) with Tri-Type. We note these figures include genuine input mistakes – situations when a user actually tapped with the wrong finger. A paired t-test revealed no difference in the wrong-finger rate between the keyboards.

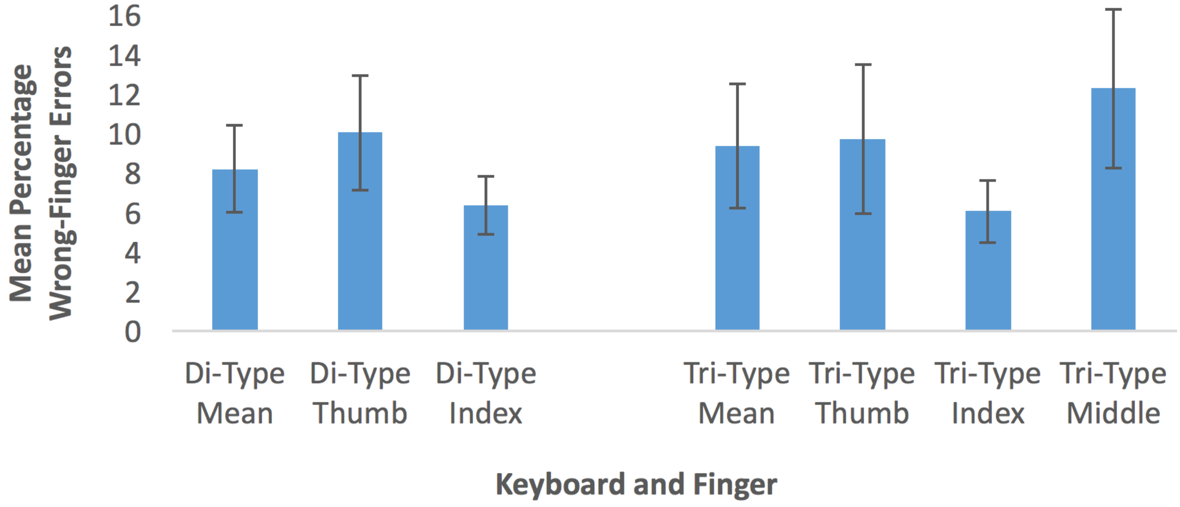


Figure 19: Mean percentage wrong-finger errors (by key) in Di-Type & Tri-Type keyboards. Includes mean and per finger data for both keyboards. Bars show Std. Dev.

Figure 19 shows the overall mean wrong-finger rate (per key) for each keyboard and finger. There are substantial variations among the fingers, with the index finger being recognized most accurately. Looking at the data in detail, we observed a disproportionate number of thumb wrong-finger errors in the bottom row of keys. In the single Di-Type letter key on the bottom row, thumb input logged a 42% wrong-finger rate. In Tri-Type, this ran at a mean of 24% wrong-finger errors for the bottom three keys. This suggests it was highly challenging to correctly identify the thumb at the bottom of the screen, an effect that was likely not observed in the main studies due to their larger targets. If errors from the bottom row are removed overall wrong-finger rates for the thumb drop from 10% to 7.3% for Di-Type and, more substantially, from 9.7% to 2.3% for Tri-Type. This problem impacted participants: in post-study comments, they indicated a preference for Di-Type due to a combination of the fewer fingers required to operate it and problems acquiring bottom row targets with Tri-Type.

In sum, this study shows our approach to finger identification has considerable promise. Static finger identification models generated from the tapping data of one set of participants enabled a second set to reasonably successfully and rapidly enter text, a challenging input task. Although mean accuracy for each keyboard was in the range of 90%-92% (see Figure 19) and clearly lagged behind systems that use dedicated hardware [68], we note that these figures incorporate actual user mistakes and also expect that redesigning the keyboard to avoid trouble spots such as the bottom of the screen and integrating customized per-user recognizers can improve performance in the future.

4.7 Recommendations and Designs

Moving beyond this validation, the user and recognizer performance data captured in the studies are rich enough to support a range of practical recommendations about how finger identification using capacitive touch profiles could be best applied to designing interfaces on smartwatches. We break these down into key themes.

Fingers

Prior authors have documented how performance with different fingers varies in touch tasks [292, 293]. The current work confirms this is also true on smartwatches. However, recognizer performance also impacts design choices. The work in this paper clearly indicates that taps with the thumb are distinct from other taps – relatively high recognition accuracy could be achieved with no prior instructions and no changes in task performance and workload. Therefore, any system requiring only two fingers should first consider a design that discriminates between the thumb and other digits as the most practical and comfortable approach. Furthermore, results (and stated preferences) from our Di-Type and Tri-Type prototypes seem to suggest that less is more and that systems should use as few fingers as possible in order to achieve their objectives. Operating a system with two fingers is easier for users to deal with than three.

Targets

In the main studies in this paper, performance with thumb and middle finger input decreased with smaller targets; index finger input was unaffected. This trend was particularly prominent with the exaggerated poses. This suggests that finger identification technology can be most effectively applied to relatively large targets. On smartwatches, with their limited screen space, this may serve to restrict effective systems to specific types of content such as application icons, or continuously available actions such as a back function. While this effect was not evident in the final typing study, we did observe that recognition performance dropped substantially during thumb touches on the bottom of the screen – another risk for small targets is that proximity to screen edges may mean that full touch contact areas cannot be accurately captured. If small targets are used (and the typing study suggests that may be viable), then they should be situated away from the bottom and right edges of a watch screen.

Actions

Slow and simple touches can be more accurately recognized. Specifically, in the natural tapping study, the fingers making touches over 150ms were classified 2.8% more reliably the full set of touches. Due

Figure 20: Example finger identification enabled interfaces. Left shows two tricons, icons that respond differently to each finger. Center left shows a middle finger touch that would start an exercise routine on an activity tracker. Center-right depicts a typical music player interface. Right shows how a touch with the thumb can bring up a menu to switch between tracks (swipe left/right) or adjust volume (swipe up down).

to the length of touches captured in that study (mean of 150ms for the index finger) it was not possible to explore whether longer touches achieved greater gains. However, we also note that more complex actions, even if they are prolonged (such as the movements in the swipe study), will likely result in greater variability in touch contact area and lower classification accuracy. To be reliable, finger identification systems should therefore rely on techniques such as dwell thresholds before triggering classification processes. This kind of technique may also serve to lower false positive rates – touches under a certain duration are all treated as the default regardless of the touch contact area. Dwell thresholds could also be used to combine finger identification with more complex input techniques like swipe – classification could take place during the dwell and a subsequent movement could then further specify input.

Building on these discussions we present two interactions designs that adapt themes for interaction design with finger identification systems presented in prior work. The first of these is tricons, an idea that relates to the multiple finger icons discussed by, for example Roy *et al.* [292]. Tricons enhance applications icons by providing multiple points of entry. For example, a fitness application could be opened as normal using a regular tap, have an exercise routine start with a middle finger tap and open settings with a thumb tap. Equally, a clock could access alarm, timer or main functions depending on the tapping finger. This kind of icon matches our design recommendations as they are relatively large, accessed sporadically and usually situated away from the extreme edges of the screen. Figure 20 (left) shows two possible tricon designs. The second design is a context menu, similar to those proposed by Harrison *et al.* [92]. We envisaged this design for a music player and operating as follows. A user calls up the menu with an easily recognizable thumb tap that is held against the screen. After a short dwell, a context menu appears around the thumb and subsequent horizontal swipes navigate between tracks while vertical movements adjust volume (Figure 20, right). All normal controls of the smartwatch that respond to regular touches are unaffected. Rather than as fully novel contributions, we present these designs as customized versions of existing concepts that fit the capabilities of the functional finger recognition system proposed in this paper.

4.8 Discussion and Conclusion

This paper contributes the idea of recognizing the finger touching a smartwatch from the profile it generates on the device's capacitive touch screen. It also contributes data that explores the tradeoffs between finger recognition accuracy and user input performance in natural and exaggeratedly posed touches and validates its approach in a challenging text entry task. Taken together this work is a first characterization of the feasibility of finger identification using standard smartwatch touch screens and a comprehensive investigation of the practical limitations of the technique – its strengths, weaknesses and how these impact what can be built with it.

There are a number of interesting future avenues for research. This paper focused on the recognition of fingers from single frames of touch data. Exploring features than span entire screen contacts is an obvious next step. A prerequisite for achieving this is likely a faster sensor – the 33Hz sensor data used in this work resulted in rapid touches leaving few records. More empirical work is also required in terms of data capture – the lab studies in this paper suffer from typical issues of ecological validity. While we did focus on a common pose (seated at a desk), wearables are clearly used much more diversely. Contexts such as use on public transport or during discrete operation under a desk or on a lap are valid and worth studying. To do so will require porting the system to a genuine smartwatch (as in [94]) and this is a clear next step for this project. We also note that the work in this paper deals with sensing finger touch profile and angle – it may be possible for users to operate the system simply by angling their fingers appropriately to form different shapes [66, 112]. Exploring how these closely related input modalities could complement one another would be another interesting next step for this work.

In conclusion, finger identification is a simple, effective input technique that can yield many benefits on wearables. This paper provides a first examination of how it might be enabled using standard capacitive touch screen technology. We believe the ideas, techniques and recommendations we present can guide designers and developers as they introduce finger identification into real devices.

V Characterizing In-Air Eyes-Free Typing Movements in VR

5.1 Properties of Finger movements in In-air Finger Stroking

The third scenario supports the general claim of this thesis by understanding the properties of finger movements during fast sequential input with in-air finger stroking gestures. Enabling fast sequential input using in-air finger stroking can allow already prevalent finger stroking skills on the surface as in a physical keyboard to be applied to resource-limited wearable devices such as HMDs. To understand the finger motions in the air, the empirical studies explored various perspectives of natural in-air finger stroking during unconstrained typing tasks such as strategy and kinematics of in-air finger stroking, correlation movements of other fingers, the time interval between sequential finger stroking inputs, and characteristics of the individual in-air key. In sum, the various analysis showed that the unconstrained in-air finger stroking complicates the detection of intended finger inputs due to the faster and shorter finger movements with more interleaved and correlated motions. Finally, the design considerations and recommendations were described for future researchers and developers to achieve accurate detection for the in-air finger stroking that can improve the input expressiveness for the resource-limited wearables.

5.2 Abstract

We empirically explore fundamental requirements for achieving VR in-air typing by observing the unconstrained eyes-free in-air typing of touch typists. We show that unconstrained typing movements differ substantively from previously observed constrained in-air typing movements and introduce a novel binary categorization of typing strategies: typists who use finger movements alone (FINGER) and those who combine finger movement with gross hand movement (HAND). We examine properties of finger kinematics, correlated movement of fingers, interrelation in consecutive key-strokes, and 3D distribution of key-stroke movements. We report that, compared to constrained typing, unconstrained typing generates shorter (49 mm) and faster (764 mm/s) key-strokes with a high correlation of finger movement and that the HAND strategy group exhibits more dynamic key-strokes. We discuss how these findings can inform the design of future in-air typing systems.

5.3 Introduction

Virtual Reality (VR) is undergoing a renaissance: the emergence of high-fidelity, low cost Head Mounted Displays (HMDs) is transforming it from the province of the lab to that of the living room. VR now impacts a very broad range of application areas from media consumption through gaming [300] to simulation [29] and expressivity [301] or productivity [302] tools. As its reach spreads, more emphasis is being placed on the interactive aspects of VR - the majority of headsets ship with dedicated input controllers or advanced bare hand tracking systems. However, such systems still struggle to effectively support many fundamental input tasks, such as text-entry. We argue text-entry is of growing relevance to VR. Gamers need chat with peers, or perform administrative tasks like logging into accounts or configuring settings. In more productivity-oriented domains, tasks such as file manipulation, communication or browsing the Internet will all frequently require text input. Removing headsets to perform these tasks is, at best, laborious and unappealing. Reflecting this perspective, a number of text input systems for VR have been proposed. A simple approach is to rely on existing VR controllers (e.g. Cutie Keys, Punch Keyboard, etc), but input bandwidth is typically low. Physical keyboards can also be tracked and integrated into a virtual scene, superimposing the real on the virtual [303]. While this can offer good performance, it tethers a user to a single physical space [304].

Freehand text input can solve this problem - users simply type in the air. However, systems that implement virtual keyboards based on finger interaction are slow and cumbersome to use [88, 305] – hitting targets in mid-air is not the same as hitting targets on a real keyboard. A more promising approach is to capture high fidelity finger movements during in-air typing [47, 244, 306]. In this way, users can rely on their motor system [307] to type at high speed [308]. Furthermore, as no keyboard needs to be presented, users are free to move around and the VR scene remains clutter free.

While this idea is appealing, it remains difficult to implement – typing finger movements in mid-air are rapid and complex. Prior work targeting this space has tended to simplify the problem. For example, ATK [47] achieves reasonable recognition performance using a probabilistic tap detection algorithm based on the height of the stroking finger; to ensure each tap is clearly performed, users are instructed to issue a sequence of controlled, temporally separated motions with single fingers. We refer to this as *constrained* in-air typing and note that the movements it is based are highly simplified compared to the interleaved bi-manual activity that characterizes real-world typing. Here, we define *unconstrained* in-air typing as the uninstructed typing behaviors that users exhibit during in-air typing and which reflect their real-world typing behaviors. To achieve truly eyes-free typing in mid-air with unconstrained finger kinematics, it is important to fully understand how people naturally behave without any instruction when they are engaged in mid-air typing. This allows us to define fundamental requirements and guidelines for designing in-air typing systems in VR. To achieve this objective, we collect 25,932 in-air finger stroke traces of unconstrained typing in VR through an empirical study. We characterize this data and identify key properties and features including: different typing strategies among the users; finger kinematics; correlated movement of fingers; interrelation in consecutive keystrokes; and 3D end-point distribution. This analysis provides quantitative data that can support the development of in-air typing systems for VR based on unconstrained finger movement. This data is important as relaxing the constraints on typing motions will increase the difficulty of accurately classifying the keys they indicate. Only by understanding and detailing the behaviors involved in unconstrained in-air typing will we enable input that is rapid, accurate, fluid, and eyes-free.

The contributions of this paper are 1) empirical data characterizing unconstrained behavior during eyes-free in-air typing and contrasting it with prior data from studies of constrained in-air typing, 2) a description of how typing strategy impacts this data, 3) the development of features describing interrelated finger strokes that can support improved segmentation and recognition of a single finger stroke, and 4) design considerations for building in-air typing system and feasibility of finger classification based on our findings. Throughout the paper we contrast the data we report with existing data captured in constrained in air typing settings (ATK [47]).

5.4 related work

Text Entry on HMDs

Authors have proposed a variety of approaches to text entry during HMD use including touchscreen typing via integrating a large input surface into an HMD [207]. Others have proposed systems based on handheld controller motions that describe text-entry gestures [188] or control a ray-based pointer that intersects a keyboard [78]. While these systems effectively support typing, their efficiency is lacking: performance ranges from 10 WPM or less [188, 207] through between 10 to 20 WPM [78]. A potentially more effective approach is to integrate physical keyboards into virtual scenes. Walker et al. [303] exemplify this approach. This work implemented a HMD-based text entry system that tracks a real keyboard. The VR system includes a virtual keyboard assistant to provide visual feedback inside the virtual scene. Performance reached 43.7 WPM with 2.7% error rate. Similarly, McGill et al. [309] report that typists on a physical keyboard can achieve 38.5 WPM with 1.07% uncorrected error rate while wearing a HMD. We note that while blending keyboards into the virtual scene provides high typing performance, it restricts VR simulations in other important ways: users are tethered to the keyboard and/or desk.

Ten Finger Typing on Flat Surfaces

The design of in-air typing systems can also be informed by the substantial literature on unconstrained text entry on flat surfaces such as mobile phones and tablets. This typically involves examining naturalistic typing patterns in order to propose the design of future touchscreen keyboards [310, 311]. For example, Findlater *et al.* [312] analyzed twenty typists' touch contact points and hand contours and reported that individual typists exhibit spatially consistent key press distributions. They also showed that typists could achieve up to 58.5 WPM mean typing speed without visual cues showing keyboard layout. In follow-up work, they showed that personalized ten-finger touchscreen typing models can improve both typing speed and subjective experience [313]. Haptic feedback has also been shown to be important in touch typing on flat keyboards. Kim et al. [314–316], for example, used piezoelectric discs under each key to create specific feedback for each keypress. Their results showed significant improvements to typing performance when haptic and auditory feedback were included.

In Air Typing in VR

Freehand typing techniques have the potential to combine two beneficial properties: input that is high bandwidth input and also not tethered to a physical device. Numerous authors have proposed systems to explore this potential. ARKB [305] relies on markers on the fingers to implement this approach. A more complex approach is to recognize finger strokes based on highly consistent movement schemes such as touch typing. TiTAN [306] attempts this, showing up speeds of up to 9.4 WPM with 10 fingers. ATK [47], a freehand-based mid-air typing system based on a Leap Motion's 3D hand tracking data, also implements this approach. It recognizes ten-finger typing by adopting a probabilistic tap detection algorithm and augmented version of Goodman's input correction model [317]. ATK achieves up to 29.2 WPM of typing speed after practice, showcasing the strong potential of this approach. Despite this performance, we note that the typing task of ATK involved "clearly performed tap[s]" on an "horizontal imaginary keyboard". The relatively low WPM, compared to real world touch typing, reflects the controlled nature of these typing movements and enabled the authors to use simple vertical finger motions to reliably classify stroking fingers and tapped keys. This enabled the authors to develop a final system with a word-level accuracy of 99.7%.

5.5 study: Understanding In-Air Typing

The goal of this study was to characterize unconstrained finger kinematics and behaviors during in-air typing to provide insights and implications that can inform the design of in-air typing systems. To achieve this goal, we explored and analyzed the typing behaviors and strategies of mid-air typists in order to investigate the time at which a single stroke is executed, the finger that is issuing it and the location it is intended to indicate from a rapid, complex and interleaved sequence of unconstrained finger motion. Throughout our study, we compare our analysis with ATK [47], a closely related study that discusses the features needed to support in-air typing based on constrained finger motions—participants in ATK were instructed to clearly tap on each key with temporally separated individual finger movements. Our study involves an in-air typing task completed by touch typists. We collect a large number of unconstrained finger movement traces and characterize a range of key features from this data including finger kinematics, correlated movement of fingers, inter-relation in consecutive keystrokes, 3D end-point distribution, and typing strategies among the typists. This analysis provides quantitative engineering specifications and design insights that can support the development of in-air typing systems for VR that integrate rapid and fluid typing experiences based on unconstrained finger movement.

Participants

Sixteen typists (female=5; M=30.06; SD=7.76, two left-handed) were recruited for this study. Prior to the main experiment, we confirmed their typing performance and finger to key mapping with a physical keyboard using TextTest [318], a text entry evaluation tool. Mean typing speed was 62.68 words per minute (WPM) (SD=12.1) with 0.8% of uncorrected error rate using a physical keyboard. They were paid for their participation with a $3 coffee voucher.

Experimental Setup

Figure 21 shows the experimental setup and a virtual scene of typing test. Participants were asked to take a seat, wear a HMD device (Oculus Rift CV1) and reach out their hands above a Leap Motion finger tracking device (version 3.2.0+45899) to start the typing experiment. The Leap Motion has a deviation of below 0.2 mm (static) and 1.2 mm (dynamic) between a desired 3D position and the measured positions [319]. The Leap Motion device was placed 41 cm above from the floor using a tripod. The virtual scene was implemented using Unity3D (version 5.6.0.3f) displayed through the binocular displays in the HMD.

In the virtual scene, two hand skeletons were rendered from the Leap Motion to provide a visualiza-

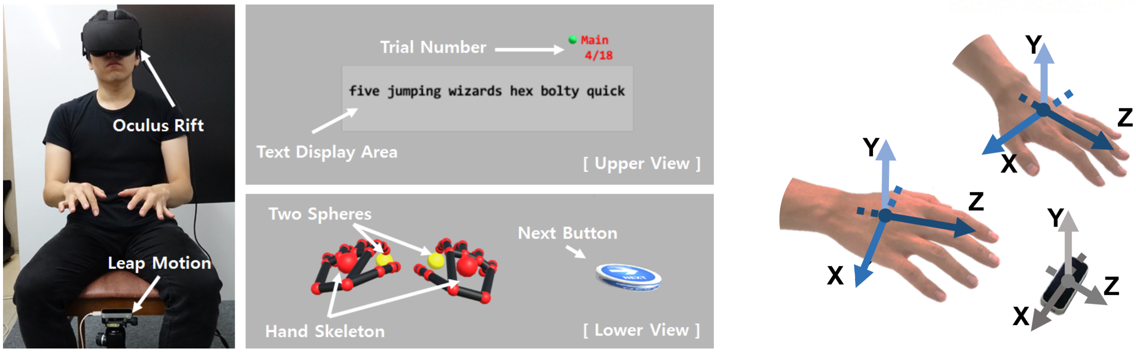


Figure 21: Experimental hardware setup with participant performing typing task (left) and typing test in a virtual scene (middle). Relative coordinate used palm position as the origin (blue axes) and absolute coordinate used the common origin of Leap Motion (gray axes) in 3D.

tion of the typing fingers. Two spheres were presented as indicators of initial finger positions that are similar to indicators on the F and J keys on physical keyboards. The two virtual spheres were placed 15 cm above the Leap Motion and the distance between them was 9 cm. We derived this spacing by starting with 6 cm, approximately the spacing between the two indicators of F and J on physical keyboards, and iteratively testing to find the most natural and comfortable distance. Subjective assessment by experimenters suggested that 9 cm was suitable for in-air typing. Participants were able to adjust the height of the virtual spheres to achieve a comfortable pose. Starting from the two spheres is important as they act as reference points for our measures. A next button with a diameter of 15 cm was placed 33 cm to the right of two indicators. This was used in order to move to subsequent trials. A text display area was placed in the visual field at 90 cm above the two spheres, approximately eye-height. This displayed the phrases to type. The current trial number was also displayed right above the text display area to communicate progress through the study.

There was no further visual content (e.g., a bounding box for hand placement, a visual keyboard layout) in the scene. In addition, visual feedback on finger strokes was not provided as the purpose of this study was to observe unconstrained finger motions. These choices were explicit. We argue it would be impossible to observe unconstrained typing finger movements while displaying a visual keyboard layout for several reasons. Firstly, people may look at the layout while making keystrokes. This visual constraint will further generate eye-gaze shifts between text and keyboard areas, reducing typing speed. As our goal is to characterize in-air eyes-free typing movements and most typists spend minimal time looking at the keys on a physical keyboard, providing minimum visual cues is the best way to elicit natural high speed typing behavior. Secondly, following [312, 313] we believe that typing behavior captured without a visually displayed keyboard provides ground truth about users' expected locations for keys that can be used to inform future virtual keyboard designs (e.g. key positions, sizes). Displaying a virtual keyboard will inevitably compromise this behavior. Finally, recent work on touchscreen keyboards demonstrates rapid performance without displayed keyboards [320]. Furthermore, we provided no auditory or haptic feedback to participants and also no constraints or guidance as to the location of the backspace key. By combining these unconstrained conditions and limited visual cues, we hoped to solicit unconstrained in-air typing motions.

All joint position data from finger stroking movements was captured in real time from the Leap Motion by C# scripts. The data includes 3D coordinate positions of all the joints between phalanx bones in thumb, index, middle, ring, and pinky fingers of both hands. We also collected the center position of each palm. This raw data describing finger stroking motions for all participants was then analyzed to build a ground-truth model for performance of in-air finger stroking for eyes-free typing in VR. We use two different 3D coordinate systems in our descriptions: absolute and relative (see Figure 21). Absolute is a global position measurement with its origin at the Leap Motion reference frame. In the relative coordinate system, the current palm position is defined as the origin.

The study was composed of four blocks of 20 trials. The first block was considered practice and the remaining three used for data collection. Each trial contained one phrase and the first two trials in all blocks were also treated as additional practice (as in [312]). A randomly selected phrase from either Mackenzie's phrase set [299] or one of 6 pangrams were used for each trial. In each block, the 6 pangrams were interleaved with phrases from Mackenzie's set to increase the occurrence rate of rarely used characters.

Procedure

A typing test with a physical keyboard was conducted prior to the main study. During this time, we closely monitored participants' behavior with a web camera to ensure their fingers follow the standard finger-to-key mapping. In the beginning of each study block, participants were asked to take a seat, wear a VR headset, and reach out their hands above the Leap Motion. They first completed an informal training session to get used to the input scenario, then began the first trial in main study by placing their index fingers into the two spheres. The spheres then disappeared, a text phrase was displayed and the color of hand skeletons turned white to indicate the start of a trial. Participants then typed the phrase as fast and accurately as possible. We recommended they hit backspace whenever they felt they made errors. When they felt that they completed the trial, they were instructed to press the next button to go on to the next trial, which again started by requiring they place their index fingers into the spheres to begin. Participants were requested to take a break after each block to minimize fatigue.

Data Processing

We collected and analyzed a total of 25,932 labeled finger strokes. The traces, containing data from the Leap Motion API, include positions, directions and velocity for all fingers and both hands. We use these to derive finger stroke amplitude, finger stroke start and end positions, finger movement directions, and also movement of the palms. Data analysis procedures were as follows. First, we processed the data to create ground truth typing behavior. We parsed the raw data and loaded it into a Unity3D based typing sequence visualiser. This system loads the 3D position data of all fingers' stroking movements and provides a visual environment to browse 3D depictions of typing actions and manually label each action with the appropriate character (e.g. the character that should have been typed). This procedure is inspired by that used in ATK [47] and was achieved through observation of finger acceleration and movement profiles to ascertain the 3D endpoint of each finger stroke. This time consuming process was required since there is currently no existing model of in-air finger stroking during unconstrained eyes-free typing.

After end-points were identified, we traced backwards to determine stroke start-points. Prior to the end-point the finger is moving at speed, so we simply categorized start points at the origin of these essentially ballistic motions – the most proximate local speed minima. To perform this task accurately, we used a relative coordinate frame (Figure 21) and applied linear interpolation to fill the gaps in the velocity data and a rolling average with a window size of five to eliminate noise. After segmenting all strokes, our subsequent analysis used this raw but delimited data. The final data set includes information of all 3D positions, directions, keystroke amplitudes, velocities, and accelerations of finger strokes for all ten fingers and two palms. We also measured typing speed in WPM and finger-level accuracy as the ratio of incorrect finger strokes (defined as use the wrong finger) to the total number of finger strokes.

5.6 General description: in-air typing

Typing Speed

Participants achieved up to 78.3% of their typing speed on a physical keyboard during in-air typing: 49.1 WPM in-air vs. 62.7 WPM keyboard. It is obvious that in-air typing is slower than physical keyboard typing. A number of factors likely contribute to this. Firstly, the lack of haptic feedback on finger strokes resulted in relatively deep stroking motions (49 mm, SD = 12 mm), which may reduce speed [314]. Secondly, prior work has suggested that fast typists keep their hands still and slower typists move their

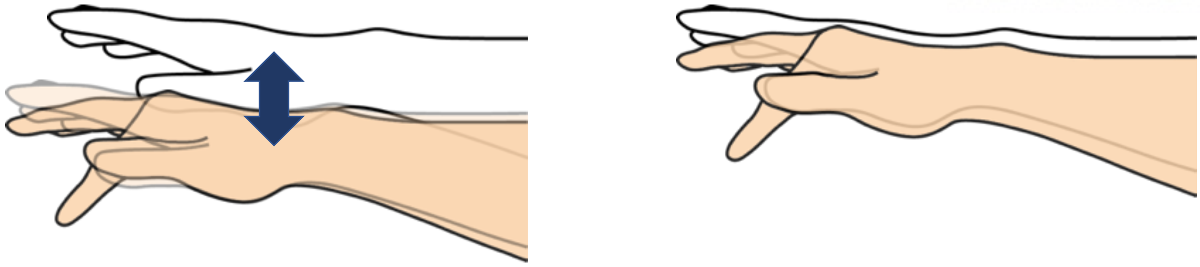


Figure 22: Typing strategy: Participants either move their entire hand and finger together (HAND strategy, left) or only finger (FINGER strategy, right) for single finger-stroke

hands more [321]. We observed larger hand motion (16 to 31mm) compared to prior studies of typing on keyboards—these additional motions likely slowed participants down. Despite these factors, WPM that participants achieved suggests that unconstrained in-air typing speeds in VR may be relatively high in optimal conditions. It contrasts strongly with the speeds reported in prior work—ATK: 29.2 WPM and TiTAN: 9.4 WPM.

Typing Strategy: HAND vs. FINGER

Global hand movement referring to motions of the entire hand expressed in absolute 3D coordinates, has been shown to support identification of different typing strategies during use of a physical keyboard [321]. We specifically investigate this issue in unconstrained in-air typing as the lack of explicit guidance as to key locations and feedback during key presses may increase the amount of hand movement that occurs. We therefore assumed that in-air typing would exhibit variations in typing strategies based on the different patterns of global hand movement reported in prior work.

Results & Discussion

Two distinct in-air typing strategies were observed during the study. Some typists tend to stroke the keys by moving their entire hand and finger together (HAND strategy) while other typists use only their fingers to reach keys, minimizing their hand movements (FINGER strategy). Figure 22 shows the two strategies. We examined the travel distance of the hand during finger strokes to categorize participants into the two strategies. For this, we considered a palm movement as a movement of the entire hand and calculated the position change in absolute 3D coordinates between the start and end of each finger-stroke. With a feature of palm movement, we adopted a K-means clustering method to cluster each participant into one of two strategies: HAND (N=6) and FINGER (N=10). The centroid of each cluster was 37.2 mm in HAND strategy and 14.1 mm in FINGER strategy, respectively. This indicates that HAND strategy group moved their hand greater distances than FINGER strategy group during the finger-stroking. We confirmed the clear distinction between two strategies through visual inspection.

As increased dynamic palm movement would affect the properties of finger-stroke generally, we identify typing strategy as a critical factor to understand unconstrained in-air typing behavior. Following typical processes, we will compare two groups and discuss how their typing behaviors and tendencies in each group affect other components in designing and developing in-air keyboards.

5.7 Finger Kinematics

We examined finger kinematics to understand the structure and characteristics of finger-strokes in terms of their duration, amplitude, travel distance, and max velocity for both hands and all four fingers: index, middle, ring, and pinky. Each finger-stroke includes 'flexion' and 'extension' phases referring to, respectively, the fingertip displacement from the start point of the stroke to when the finger at its most bent and then back to its position when the finger is most fully stretched. We also measured the amplitude (finger stroke depth relative to the depth at the start of the stroke) and the max velocity of a fingertip in relative 3D coordinates. Finally, we measured the travel distance representing the total movement of the fingertips in the air in absolute 3D coordinates. We argue that the features showing statistical differences represent distinct kinematic behaviors that can be adopted for classifying fingers or strokes in the future.

Results & Discussion

Table 5 shows the kinematic features of fingers during flexion and extension phases, including duration, amplitude, travel distance, and max velocity for each finger. Table 6 shows the same kinematic features in each typing strategy group. Figure 23 provides a graphical illustration of the finger kinematics. Examining this data reveals that unconstrained typing shows faster finger velocities (764mm/s, SD=214), shorter finger-stroke times (322ms, SD=25), and lower amplitudes (49mm, SD=12) than the figures reported for ATK's constrained typing, respectively: 623mm/s, SD=262; 496ms, SD=170 and; 64mm, SD=24. Basically, in unconstrained typing, fingers move shorter distances more rapidly and, thus, reach their terminal destinations more quickly. These differences suggest that typists in our study performed with natural typing motions—with short, rapid finger movements. In Table 7, we compare stroke features in more detail. We argue that the short and fast finger strokes in unconstrained typing will increase the difficulty of detecting finger strokes accurately compared to the more constrained movements used in systems such as ATK. These qualities also alter the features that may be most salient. While times and distances traveled may be less useful, due to the fact they are smaller in magnitude, the faster finger

Table 5: Mean (SD) of kinematic features of fingers during flexion and extension phases, including duration, amplitude, travel distance, and max velocity for each finger. One-way ANOVA results show the main effect of the finger.

		INDEX	*MIDDLE*	*RING*	*PINKY*	*F*	*p-value*
Overall Duration (ms)		331 (30)	316 (27)	316 (22)	327 (27)	F(3,60)=1.241	0.303
Flexion	***Duration (ms)***	143.3 (12.0)	138.0 (10.93)	140.7 (9.9)	144.2 (11.6)	F(3,60)=1.006	0.397
	Travel Distance (mm)	69.7 (18.6)	62.3 (22.2)	63.4 (21.2)	62.8 (20.4)	F(3,60)=0.446	0.721
	Amplitude (mm)	54.0 (13.3)	47.6 (14.6)	48.9 (13.6)	47.3 (13.5)	F(3,60)=0.810	0.493
	Velocity (mm/s)	829.7 (227.4)	741.8 (256.1)	762.2 (215.6)	722.1 (238.8)	F(3,60)=0.635	0.595
Extension	***Duration (ms)***	187.2 (19.06)	177.86 (17.93)	175.6 (13.5)	182.6 (18.14)	F(3,60)=1.419	0.246
	Travel Distance (mm)	61.8 (16.1)	50.4 (14.1)	49.6 (12.45)	52.2 (13.0)	F(3,60)=2.818	0.0466
	Amplitude (mm)	45.8 (10.9)	39.0 (10.8)	38.8 (10.5)	41.9 (11.7)	F(3,60)=1.404	0.250
	Velocity (mm/s)	658.2 (179.06)	564.58 (232.8)	607.1 (225.9)	613.03 (272.2)	F(3,60)=0.445	0.722

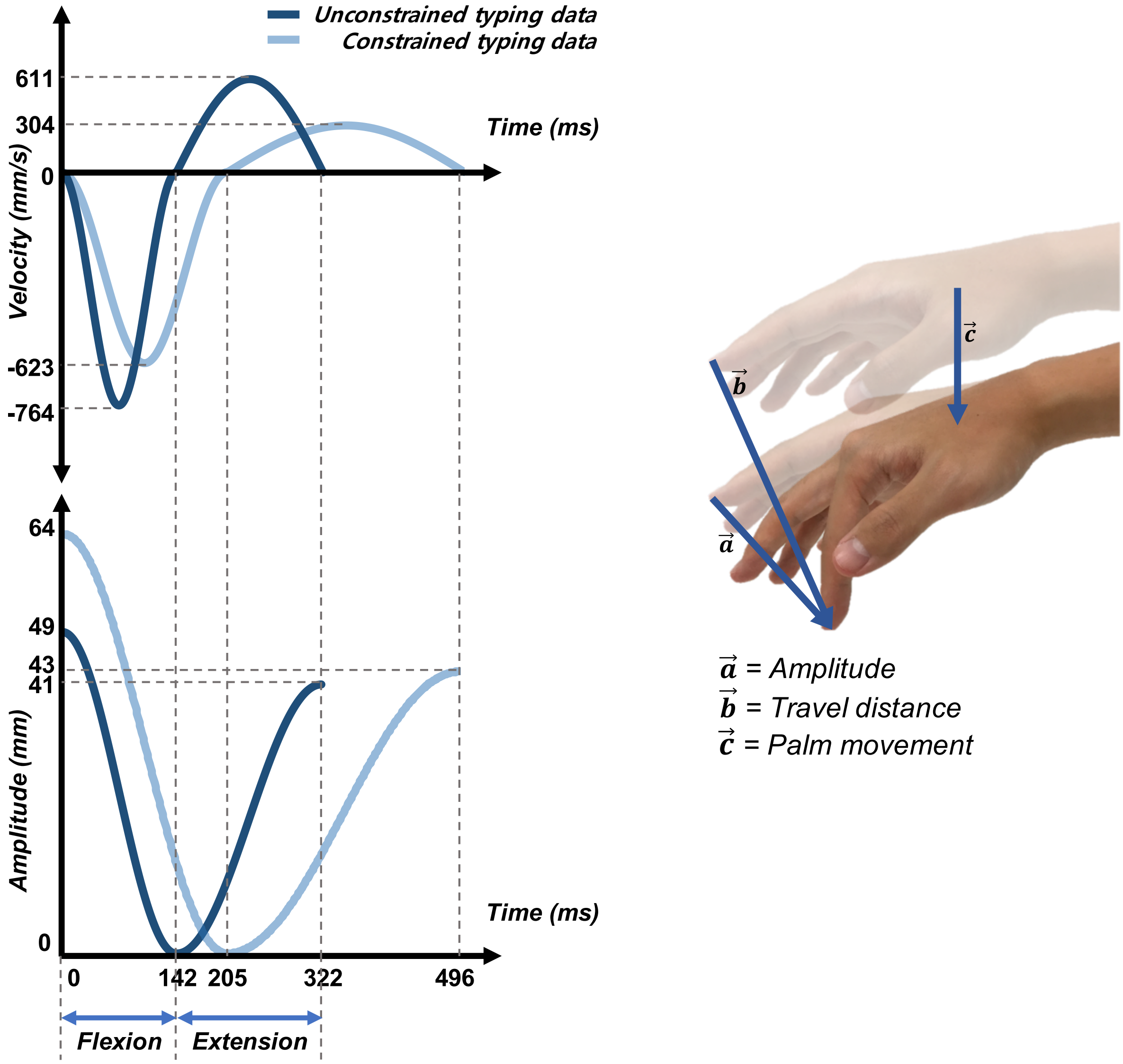


Figure 23: Finger kinematics (left). Dark blue indicates unconstrained typing from current study and light blue indicates constrained typing from ATK. Left-top shows velocity and left-bottom shows amplitude. Features of finger stroke (right); amplitude, travel distance and palm movement.

Table 6: Mean (SD) of kinematic features of fingers during flexion and extension phases in each strategy group. A Mann-Whitney U test was applied for statistical analysis.

		FINGER (n=10)	*HAND (n=6)*	*Mean difference*	*p-value*
Overall duration (ms)		314.53 (19.96)	335.46 (28.49)	-22.94	0.1471
Flexion	***Duration (ms)***	138.25 (9.64)	147.1 (8.49)	-8.76	0.0727
	Travel Distance (mm)	51.56 (8.93)	86.22 (11.78)	-0.034	**0.0002**
	Amplitude (mm)	43.84 (10.13)	58.84 (10.55)	-0.02	**0.016**
	Velocity (mm/s)	672.22 (177.93)	916.86 (188.92)	-254.14	**0.0225**
Extension	***Duration (ms)***	176.28 (10.89)	188.36 (20.39)	-12.11	0.3132
	Travel Distance (mm)	45.33 (7.54)	67.17 (5.68)	-0.02	**0.001**
	Amplitude (mm)	38.57 (9.15)	46.06 (10.4)	-0.008	0.1806
	Velocity (mm/s)	572.29 (225.84)	674.78 (163.53)	-123.15	0.0934

Table 7: Comparison between constrained (ATK) and unconstrained typing in duration, amplitude, and velocity.

	Flexion		***Extension***	
	ATK	***Current Study***	***ATK***	***Current Study***
Durtaion (ms)	205 (81)	141 (10)	291 (148)	181 (16)
Amplitude (mm)	64 (24)	49 (12)	43 (26)	41 (10)
Velocity (mm/s)	623 (262)	764 (214)	304 (136)	611 (205)

velocities observed suggests this feature may be particularly important.

Data were further analyzed using one-way ANOVA on the variable of the finger. Tukey's test for Post-hoc testing was applied to reveal the differences among fingers. We found a significant main effect in travel distance during the extension phase ($F(3,60)=2{,}818$, $p=0.0466$). Post-hoc testing confirmed that the index finger moved longer distances than all other fingers ($p<0.001$) in the extension phase. There is no main effect on other features. However, we noticed that the index finger tended to move faster (velocity) and further (amplitude) than the other fingers. These variations are likely because each index finger is responsible for six keys, rather than three or fewer keys for other fingers, and this greater diversity requires larger but faster movements.

Table 6 shows the mean of kinematic features of fingers during flexion and extension phases in each typing strategy group. We performed a Mann-Whitney U test for a non-manipulation variable [322] with unpaired samples and unbalanced sample size. The HAND group showed faster velocity ($p=0.0225$) with longer travel distance ($p=0.0002$) and amplitude ($p=0.016$) in flexion phase and longer distance ($p=0.001$) in extension phase than the FINGER group; their hands are more mobile during the typing task. Our analysis indicates that the HAND group makes more substantial typing motion than the FINGER group, which can be a more clear trigger for keystroke detection. However, the larger finger motion in the HAND group could generate higher correlated movements in other fingers (passive fingers), possibly yielding the recognizer more complicated. We will discuss the correlated movement of fingers in Section 5.8.

An interesting finding is that, in the flexion phase, we observed a relatively long travel distance of 65 mm (SD=19.9) (similar to the amplitude of 64 mm in constrained typing). This may be due to the addition of hand movements since the mean distance covered by the palm is 23.0 mm (SD=16.1) and 23.5 mm (SD=11.5) for left and right hands, respectively. Since finger stroking in unconstrained typing leads to these marked movements of the palms, we suggest the travel distance may be particularly salient as a key feature for finger stroke detection in unconstrained typing.

Another interesting finding is the recovery ratio of movement amplitude in the flexion phase over that in the extension phase. We found that unconstrained typing achieved an 84% recovery ratio (flexion: 49mm and extension: 41mm) while constrained typing is reported to reach a ratio of 67% (flexion: 64mm and extension: 43mm) (see Figure 23). We interpret this as indicating that, in an unconstrained typing scenario, the previous finger stroke is quickly returned to its original position to yield smoother

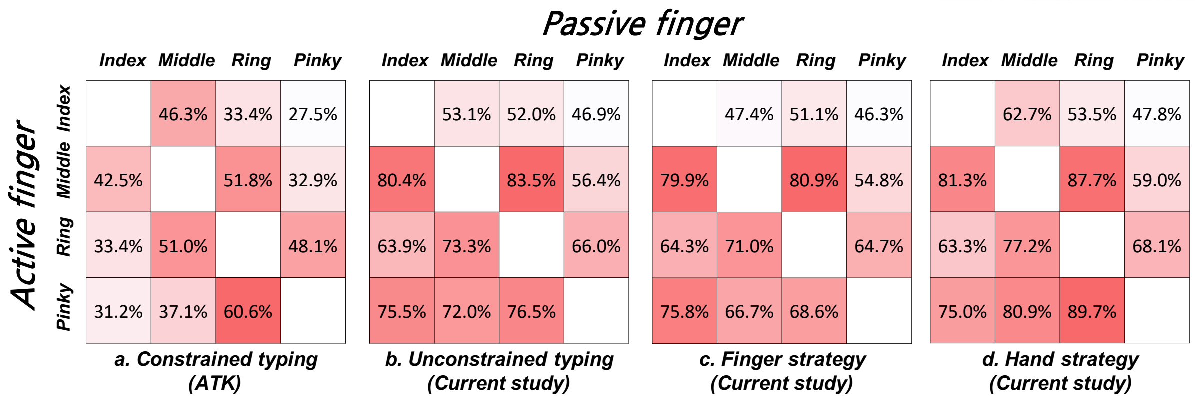


Figure 24: Amplitude ratio (AR) between active and passive fingers. a) Constrained typing (ATK, b) Unconstrained typing, c) Finger strategy, and d) Hand strategy.

consecutive keystrokes. This leads to a higher recovery ratio. In contrast, the recovery ratio was lower in constrained typing since it emphasizes clearly individuating keystrokes rather than continuous typing. We speculate that the higher recovery ratio may help support clearer segmentation of finger strokes and recognition of the finger.

5.8 Correlated Movement of Fingers

Human fingers do not move alone—due to the arrangement of muscles and tendons in the palm, hand, and wrist, intentional movements of a single finger inevitably lead to unintended, but correlated, movements in other fingers. An accurate understanding of this correlated movement has been identified as an important factor to support accurate finger classification in in-air typing [47]. To analyze the correlated movement among the fingers in our data, we use the*Amplitude Ratio (AR)*. This is the ratio between the active finger's amplitudes, the one making the intended stroke, and all other fingers, which are termed passive fingers. Correlated movement among the fingers is calculated as: $AmplitudeRatio(AR) = (Amplitude_{passive\ finger}/Amplitude_{active\ finger}) \times 100\%$.

Results & Discussion

The AR of unconstrained typing (46.9-83.5%) (Figure 24-b) was higher than that of constrained typing in ATK (27.5-60.6%) (Figure 24-a). This suggests that unconstrained typing exhibits a higher correlation of finger movements. This is likely due to the fact that, in ATK, participants were asked to perform a clear tap to capture the gesture of a finger stroke and see their hands during typing (i.e. they were not wearing an HMD). We argue that participants in the ATK data set made substantial effort to produce clear individuated input, resulting in reduced amounts of passive finger movements. In contrast, in the unconstrained task in the current study, participants exhibited greater movements of passive fingers due to study instructions requesting them to freely "type" in the absence of any cues indicating keyboard layout.

We note that the higher levels of correlated movement in unconstrained in-air typing may pose challenges for accurate finger classification—during any given key-press, more fingers were moving further. This is a particular problem for the HAND group (see Figure 24-d) as they showed stronger correlated movement (47.8-89.7%) than the FINGER group (46.3%-80.9%). We note this difference may be due to passive finger motions that occurred due to forces applied as an incidental but inevitable result of using more dynamic whole hand movements during typing. Regardless, the reduced finger individuation we observed during keystrokes indicates that relying on stroke amplitude for finger classification, as in ATK's constrained typing system, would likely lead to poor results.

5.9 Inter-keystroke Relationship

It is essential to obtain a precise tapping movement for each finger stroke to achieve stable finger-stroke detection and finger classification. The prevalence of either overlap or very short intervals between strokes makes this hard—it reduces the clarity of stroke movements and adds uncertainty to measures of initial finger position. Furthermore, errors in one stroke can cascade into follow-up strokes. Thus, characterizing typical time interval between endpoints of consecutive finger strokes can support improved finger-stroke detection and finger classification. We examined this 'inter-keystroke interval' [313] over three key variables: 1) hand combination between previous and current hands (either same or different), 2) typing strategy (either HAND or FINGER) and 3) digraph frequency (either conventional or other). This last factor refers to the how often letter pairs co-occur—we use it to contrast performance with highly practiced pairs and less well-practiced pairs. The typical set included twelve digraphs for both same (in, er, on, re, at, es, ea, io, ou, ar, as, ve) and different (th, he, an, nd, ha, en, of, nt, ti, to, le, is) hands [323]. We excluded digraphs stroked with the same finger. We calculated overlap time between consecutive finger strokes as *"Overlap time = Keystroke duration - Inter-keystroke interval"*.

Results & Discussion

Table 8 shows inter-keystroke interval for all three variables. Data were analyzed using the statistical procedures similar to those used in prior analyses: two-way ANOVA followed by post-hoc tests for variables of hand combination and digraph frequency. The grand mean of the inter-keystroke interval was 279.1 ms (SD = 73.8 ms). Given that the average stroke extension duration is 181 ms (see Table 7),

Table 8: Inter-key press timing with interrelation variables. Data were analyzed using Mann-Whitney U test for the unpaired and unbalanced samples.

Inter-keystroke interval (ms)		*Overall*	*FINGER (n=10)*	*HAND (n=6)*	*Mean difference*	*p-value*
Grand mean		279.1 (73.8)	239.5 (44.4)	345.2 (66.3)	-87.87	0.0047
Same-hand	***Normal***	321.6 (89.6)	278.4 (58.0)	393.7 (89.9)	-108.03	0.011
	Digraph	271.8 (75.6)	230.4 (44.3)	340.6 (67.3)	-108.83	0.003
Cross-hand	***Normal***	290.1 (71.6)	252.3 (42.4)	353.0 (67.3)	-89.22	0.0017
	Digraph	233.1 (65.5)	196.8 (42.6)	293.7 (50.6)	-90.08	0.003

this reveals that keystrokes are overlapped by an average of 23.8%, or 43 ms. In other words, flexion during a keypress starts well before the extension of the prior keypress is complete. In terms of hand combination, the inter-keystroke interval for different hands (M=261.6 ms, SD=67.5) was not significantly difference with that for the same hand (M=296.7 ms, SD=81.3): ($F(1,60)=3.40$, $p = 0.07$).

Common digraphs (252.4 ms, SD=69.5) showed significantly reduced inter-keystroke intervals compared to other digraphs (305.8 ms, SD=79.8): ($F(1,60)=7.88$, $p=0.006$). Typists' habitual use of these key pairs on physical keyboards leads to very rapid key stroking behaviors. The relatively short intervals for digraphs could require a faster finger-stroke recognition system or digraph gesture recognition system in in-air typing.

Data were further analyzed using Mann-Whitney U test on the variable of the typing strategy. HAND strategy group (345.2 ms, SD=66.3 ms) showed longer inter-keystroke intervals than FINGER strategy group (239.5 ms, SD=44.4). This was further confirmed with a Mann-Whitney U test that typing strategy was statistically significant factor for inter-keystroke press timing. We argue that dynamic hand movement in HAND group leads to longer inter-keystroke intervals. Due to the shorter inter-keystroke intervals, FINGER group caused longer overlap time between consecutive finger strokes than that of HAND group. The relatively short keystroke duration (314.5 ms) and long overlap time (75.0 ms) of the FINGER group could yield the finger stroke detection more complicated.

5.10 Individual in-air keys

We further analyzed finger stroke amplitude in terms of in-air finger travel. Figure 25 shows the amplitude and maximum velocity of active fingers from their initial state to their fully pressed point for each key. In addition, we examined the end point of participants' fully extended key strokes in absolute 3D coordinates. We averaged absolute coordinate positions and directions of the active fingers at the

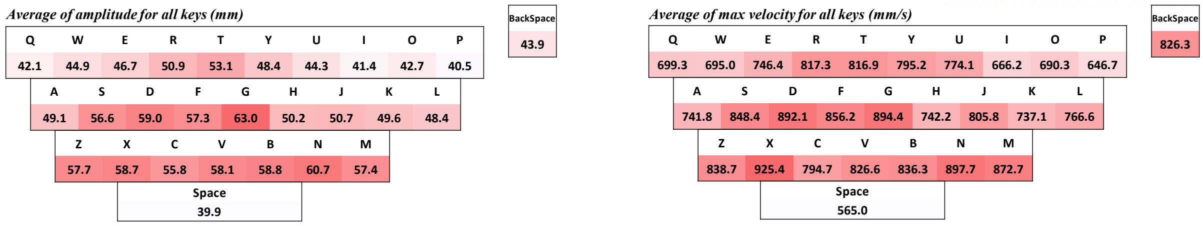


Figure 25: Amplitude (left) and max velocity (right) for each in-air key; towards the bottom of the keyboard, both features showed faster and deeper values.

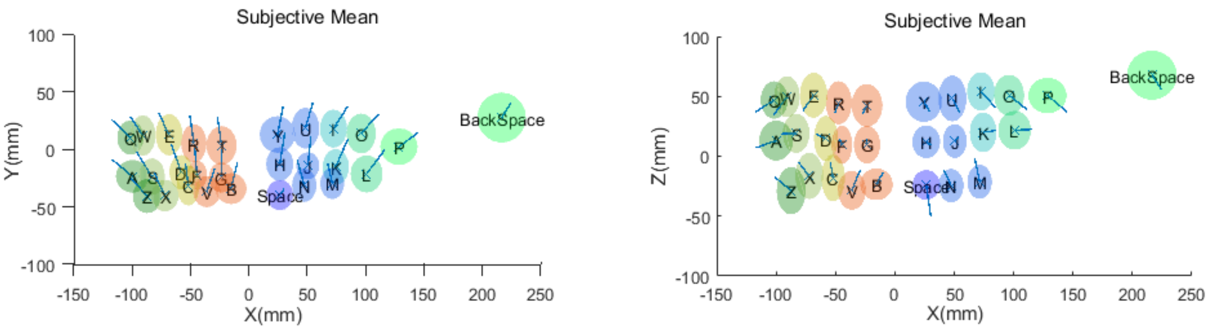


Figure 26: Averaged positions (cross marker) and directions (line) of fingertip when each of "virtual" keys is fully pressed (end point) in absolute coordinate. SD of positions for each participant were averaged and represented as lengths of axes on ellipsoids. Each ellipsoid was colored to show the dominant finger used to stroke the corresponding key.

endpoint of a key press for each subject. Figure 26 shows ‘xy’ and ‘xz’ projections of the 3D positions and directions.

Results & Discussion

Amplitude and velocity for each in-air key. In Figure 25, in terms of keyboard rows (i.e. top, middle, and bottom rows), the active finger moved further to keys located in bottom (M=58.2mm, SD=1.5) than top (M=45.5mm, SD=4.2mm) and middle (M=54.8mm, SD=5.3) rows. The average maximum velocity also varied among rows: 734.7mm/s (SD=63.4), 809.4mm/s (SD=65.1) and 856.0mm/s (SD=45.0) for top, middle, and bottom rows, respectively. This data shows that finger-strokes became faster and deeper towards the bottom of the keyboard. This may indicate that participants envisaged their fingers hovering near the top of the keyboard—as such, shorter, slower strokes would be sufficient to select keys there, while longer, faster strokes were required for keys on the center and bottom rows. These variations may help support accurate key classification, the final recognition step in an in-air typing system.

3D End-point Distribution. The deviations of fingertip end-positions were smaller in x-axis (M=12.26 mm, SD=2.38) than those in y-axis (M=15.67 mm, SD=2.09) and z-axis (M=16.67 mm, SD=2.16). Vertical and sagittal movements are essential to perform a finger stroke and change rows on the keyboard. Meanwhile, we do not need to move our fingers laterally except the left index finger, which covers two columns and the right pinky, which tends to press the backspace key. This observation is supported by the large position deviations of the backspace key: 21.03, 21.93, and 20.55 mm, in x, y, z, respectively. The averaged positions shown in Figure 26 are also relatively well aligned to the layout of a real keyboard. The mean distance between F and J keys (indicators in physical keyboards) was 97.1 mm, little different from the fixed 90 mm enforced at the beginning of each trial. While the y-positions of the most keys can be distinguished by their rows, the spacebar tended to be positioned in line with the keys on the bottom row. The locations of the air-keys were spread over approximately 250 mm in the x-axis and 100 mm in the z-axis, slightly exaggerated compared to typical physical keyboard sizes (200 mm by 70 mm). This larger horizontal size may have been due to greater initial horizontal hand spacing—F and J are typically separated by about 60 mm on a physical keyboard, rather than the 90 mm used in this study. The larger vertical size may simply reflect participants, possibly intentionally, use of exaggerated motions.

Participants' virtual keyboards were also conceptualized to be highly slanted in the z-axis, with a mean slope of 54.9°. This is much higher than the 0-10° slants common in physical keyboards. We suggest this is due to the fact the finger must be stretched forward to press a key in the top row, and this action reduces the depth of vertical movement that can be achieved by the finger joints. When the participants perform a finger stroke to the keys towards the center of the keyboard (around the F and J keys), the directions of fingertip strokes were close to vertical. Stroke directions were increasingly rotated towards the center with greater distance from the center. The space bar was the only exception from this tendency. It was typically pressed by the thumb with a stroke closely aligned to the forward direction of the participants.

5.11 Feasibility of Finger Classification

To examine the feasibility of finger classification in unconstrained typing, we constructed a new classifier, including promising features from our analysis. The feature set for the classifier consisted of the absolute and relative 3D position of endpoints, maximum velocity and amplitude of the fingers at the endpoint. We applied this feature set to a RandomForest classifier with 10-fold cross-validation process using data from both all users and also in a per-user arrangement to assess the impact of individual dif-

Table 9: Accuracy of finger classification with *RandomForest* using features from unconstrained typing system.

Finger	***Index***		***Middle***		***Ring***		***Pinky***		***Total***	
Hand	***L***	***R***	***L***	***R***	***L***	***R***	***L***	***R***	***L***	***R***
Accuracy of static (%)	85.5	90.2	79.7	74.9	77.8	79.6	88.1	91.5	83.3	84.0
Accuracy of per-user (%)	89.2 (6.5)	90.7 (3.9)	83.3 (8.4)	80.3 (9.1)	83.1 (10.2)	82.4 (8.8)	90.8 (6.0)	92.2 (8.2)	86.9 (7.0)	86.0 (5.6)

ferences.

As shown in Table 5, the static classifier achieved 83.3% and 84.0% finger classification accuracy for the left and right hands. The average accuracy of the per-user classifiers was 86.9% (SD=7) and 86.0% (SD=5.6). The highest accuracy among classifiers was 92.2% (right pinky), and the lowest accuracy was 80.3% (right middle). It is noticeable that the index and pinky fingers achieved higher accuracy of finger classification than middle and ring fingers. It is probably because features from finger movements in pinky and index fingers are distinct from their neighbor fingers, whereas middle and ring fingers are relatively not.

5.12 General Discussion

This work investigates the fundamental requirements for developing in-air keyboards for unconstrained in-air eyes-free typing scenario in VR. It captures data in an unconstrained typing scenario and contrasts this with prior reports of data in constrained settings: ATK [47]. In general, we show that unconstrained typing involves faster, shorter and more interleaved and inter-correlated motions than studied in prior work. This means that recognition systems that have been successfully deployed in the past may not be applicable to real-world scenario. Here, we summarize the key findings of unconstrained in-air typing and discuss how they can impact the design of in-air keyboard system.

Typing speed is faster in unconstrained in-air typing. Our participants typed at 78.3% of their typing speed with a physical keyboard in unconstrained in-air eyes-free typing. This is a significant improvement since constrained in-air typing occurs at 44.4% of a user's typing speed on a physical keyboard. This improvement is due to the fact that unconstrained typing generates shorter (49 mm) and faster (764 mm/s) keystrokes with a high correlation of finger movement, yielding increased velocity, shorter duration, and overlapping fingers.

Finger stroke recognition is complex. Our analysis showed that there are multiple factors to consider for recognizing a finger stroke. We claim that increased velocity with a shorter amplitude of finger stroke in unconstrained in-air typing can add more difficulties in detecting a finger stroke. We further claim that a higher correlation of finger movements in unconstrained typing can lead to a higher ratio of false detection. In addition, the shorter inter-keystroke interval with longer overlap time and increased degree-of-freedom in finger movements will bring more complexity for finger stroke recognition. Furthermore, the typing strategy (HAND/FINGER) should be considered to optimize the recognizer in order to improve the recognition rate. More apparent and distinct features should be identified in order to detect and recognize a series of finger strokes accurately in unconstrained in-air eyes-free typing. This will be discussed in the following subsection.

Haptics can play a significant role. We argue that adding haptic feedback can enhance the in-air typing performance - the presence of both tactile feelings of the keyboard layout and confirmatory clicks will significantly improve the typing speed. In fact, several works have already demonstrated the benefits of adding haptic cues to virtual typing systems on, for example, touchscreens [98–100]. In fact, a mid-air haptics display using focused ultrasound waves [324] can provide the haptic cues in mid-air for each finger stroke [325]. We suggest that mid-air haptic feedback can provide a confirmation for each finger stroke, and this may lead to reduced finger movements in both active and passive fingers, yielding decreases in finger travel distance and increases in typing speed.

Takeaways

Based on a detailed characterization of unconstrained in-air eyes-free typing, we believe that researchers can adopt the analysis from our study as a ground-truth to construct an in-air typing system. To move towards this goal, we present the following data-based recommendations to achieve accurate recognition of unconstrained in-air typing, including time components, finger stroke detection, finger/key classification, and in-air keyboard layout.

Time components for in-air typing system. Given that the average finger flexion duration is 141 ms, the recognition processing should be completed within this time period. The recognizer should be able to detect the initial moment that triggers the finger stroke, retrieve all the features during the flexion finger motion, feed those features into the finger/key classification model, and determine the correct key with proper cues (i.e. visual, audio, haptic, etc).

An inter-keystroke interval is another important time factor to distinguish the sequential finger strokes during a rapid in-air typing. The overlapping between the previous and current finger strokes can lead to an unstable acquisition of necessary features for both finger strokes. We noticed that overlap time takes about 30% of the flexion and 24% of the extension, and this may cause additional false triggers. The overlap time becomes even larger with cross-hand digraphs. One possible solution that we suggest is to use the overlapping finger movements in the extension phase as a feature for the recognizer. Since the recovery movement of the previous keystroke in the extension phase can affect the movement of a current keystroke in flexion phase, the finger movements can be a good option to be a feature to improve the accuracy.

Typing strategy is an important factor. Typing strategy, split between HAND and FINGER in this work, is a key to understanding unconstrained in-air eyes-free typing behavior. The HAND typists tend to stroke the keys by moving their entire hands while the FINGER typists use only their fingers to reach keys. Our analysis revealed that the HAND strategy group moved all their fingers further than the FINGER strategy group. These variations in motor control behavior are likely transferred, and possible amplified, from participant's typing patterns on physical keyboards. On a physical keyboard, the range of movements for a given keystroke is constrained by the physical relationship between fingers, hands and keys. In mid-air, this relationship is relaxed, likely resulting in more diverse behaviors. In addition, the lack of physical terminators for motion (i.e., the physical travel depth of actual keys) and the haptic feedback associated with such impacts, likely contributed to the universal extensions to keystroke length we observed. Regardless, these variations in strategy will increase the challenges associated with accurate detection of finger strokes and, indeed, likely demand approaches specific to each strategy. For example. the HAND strategy group, showed stronger correlated movement among the fingers. This will increase the difficulty of accurate key-press detection. On the other hand, the FINGER strategy group exhibited shorter inter-keystroke intervals and greater overlap between strokes. This will also present challenges for finger stroke classification, albeit relating to the close temporal proximity of the strokes rather the increased quantities of unintentional finger motion. Based on this analysis, we argue that it will be necessary to model typing strategy to achieve accurate in-air typing systems based on unconstrained finger motions.

Finger stroke detection. Detecting a finger stroke in unconstrained in-air typing should consider other factors as features besides simply amplitudes in finger flexion. Since the unconstrained typing generates relatively lower amplitudes, the irrelevant finger movements from passive fingers can add more false triggers. We argue that the velocity of finger stroke can be one of the key features. We noticed that the velocity of the finger stroke was faster (764ms/s) than those from irrelevant movements in passive fingers. In addition, the palm movement can be an another important feature for finger stroke detection because the palm often moves before the finger stroke, indicating that a finger is about to be pressed.

Finger/Key classification. Unlike constrained in-air typing, it is difficult to use the amplitude alone as a feature in unconstrained in-air eyes-free typing due to its high amplitude ratio (AR) between active and passive fingers (46.9%-83.5%). Alternatively, other features can be adopted together with the amplitude. These include acceleration of finger stroke, direction of fingertip, angle of finger flexion, and 3D end-position.

In-air keyboard geometric features. Participants envisage an in-air keyboard to be large and slanted. Specifically, our data suggest 250 mm by 100 mm at a slope of 54.9°. In-air key depth (derived from stroke endpoints) increases from the top row (45.5mm, SD=4.2) to the bottom row (58.2mm, SD=1.5). We derive recommended sizes for in-air keys based on the standard deviation of fingertip end-positions over all keys (see Figure 26): 12.26mm (SD=2.38), 15.67mm (SD=2.09), and 16.67mm (SD=2.16) in x, y, and z (depth), respectively. The backspace key occupies a relatively large region (21.03mm, 21.93mm, and 20.55mm in x, y, z, respectively) compared to its size on a physical keyboard.

Typing is a fundamental input task across a wide range of device form factors. We believe in-air typing is therefore highly relevant to VR scenarios [88, 326]. Understanding finger movements during the unconstrained in-air typing will be essential for future designs. We contribute a discussion of the features that can be used to accurately recognize user input in this setting. This includes differences from prior work (e.g. correlated movement of fingers has weak discriminatory power) and a set of specific data and recommendations for how to detect finger strokes (duration during flexion and extension, inter-keystroke interval, amplitude, velocity) and recognize stroking fingers (amplitude ratio between active and passive fingers, end-point, layout of in-air keyboard-size, depth and skewed). This data is of direct use for researchers seeking to build in-air typing systems. Immediate future work is the validation of our current feature set through further empirical studies with a new set of typists. This new data will allow us to validate and refine our data, analysis and conclusions. A larger sample will also increase confidence in our measures and may suggest new features, in particular if further approaches to in-air typing strategies are uncovered. In addition, future work should apply approaches from bio-mechanics, or models of human movement, such as Fitts' law [327], to in-air finger typing movements. These models may provide additional insight into the relationships between users finger motions and the in-air keys they intend to select.

5.13 Conclusions

In summary, we explored the properties of unconstrained in-air eyes-free typing to determine the feasibility of, and requirements for, development of a real world in-air typing system. We also contrasted our data with that captured in a constrained typing setting. We contribute practical observations about basic finger kinematics, typing strategies, correlated movement of fingers, 3D endpoint distribution, and interrelation features of consecutive finger strokes that can support stable finger stroke detection and finger classification. We close by discussing design considerations for developing real-world in-air typing system based on our findings. These contributions can promote the development of more effective, practice in-air eyes-free typing systems in the future.

VI ThumbAir: In-Air Typing for Head Mounted Displays

6.1 Properties of Finger Movement in In-air Thumb Typing

The fourth scenario supports the general claim of this thesis by understanding the properties of thumb movements to develop an in-air typing system enabling fast sequential inputs with high comfort and social acceptability. The thumb with the highest degrees of freedom among fingers can have a large design space for generating in-air finger gestures above the palm. The thumb touch skill is also prevalent on many smart devices. However, the performance on the surface will not generalize to the in-air interaction experience due to the unique attributes of in-air thumb motions such as the specific hand pose, the lack of haptic feedback, and the more complex thumb movements on large design space. In this scenario, the two empirical studies examined viable locations of in-air thumb touch input to select the final key locations. The in-air thumb typing system was developed through a computational design process to assign the characters to the final keys. The final two studies evaluated the typing performance of in-air thumb typing system. The two studies also accessed the perceived exertion and social acceptability on the experience of in-air typing with the designed thumb motions.

6.2 Abstract

Typing while wearing a standalone Head Mounted Display (HMD)—systems without external input devices or sensors to support text entry—is hard. To address this issue, prior work has used external trackers to monitor finger movements to support in-air typing on virtual keyboards. While performance has been promising, current systems are practically infeasible: finger movements may be visually occluded from inside-out HMD based tracking systems or, otherwise, awkward and uncomfortable to perform. To address these issues, this paper explores an alternative approach. Taking inspiration from the prevalence of thumb-typing on mobile phones, we describe four studies exploring, defining and validating the performance of ThumbAir, an in-air thumb-typing system implemented on a commercial HMD. The first study explores viable target locations, ultimately recommending eight targets sites. The second study collects performance data for taps on pairs of these targets to both inform the design of a target selection procedure and also support a computational design process to select a keyboard layout. The final two studies validate the selected keyboard layout in word repetition and phrase entry tasks, ultimately achieving final WPMs of 27.1 and 13.73. Qualitative data captured in the final study indicate that the discreet movements required to operate ThumbAir, in comparison to the larger scale finger and hand motions used in a baseline design from prior work, lead to reduced levels of perceived exertion and physical demand and are rated as acceptable for use in a wider range of social situations.

6.3 Introduction

Augmented Reality (AR) and Virtual Reality (VR) Head Mounted Displays (HMDs) are rapidly developing. Recent models feature expansive high resolution, high refresh rate screens capable of providing rich experiences, while also remaining sufficiently comfortable and lightweight for prolonged use. Both end-user application areas, such as casual gaming and entertainment [328], and professional applications, such as training [329] or work and task support [330], are blossoming [331]. For AR, many of most compelling use cases involve relatively uncontrolled settings such as to enhance experiences as visitors wander around a museum [332], to support equipment maintenance tasks in the field [333], to enhance educational activities in a classroom [334] or in various accessibility scenarios [335]. In tandem with this growth, HMD sensing capabilities are also increasing, with current standalone devices supporting not only accurate and responsive controller tracking but also state-of-the-art bare hand tracking [41]. However, while these systems are impressive, they still struggle to support key high bandwidth digital input tasks, such as text entry—a study of Microsoft's HoloLens default keyboard interface, for example, reported a mean WPM of just 5.86 [88]. This is problematic as text entry is an important input modality required for a very wide range of common, generic and everyday tasks such as entering identification information and authentication credentials, composing written chat/messages, accessing or searching online resources and dealing with file and system management tasks.

Reflecting the importance of text entry tasks, a wide variety of work has explored how rapid and reliable text entry can be achieved on both VR and AR HMDs. The dominant approach mimics the traditional computer typing experience—a virtual keyboard is presented to users, who must then select the desired keys in sequence. However, specific implementations of this basic design are diverse. Many rely on tracked hand held controllers and use the series of intersections between a cursor [82, 226] or ray [45, 78] and keyboard buttons generated by either a set of discrete taps [191, 195, 336] or a continuous stroke [79, 198] to specify characters. While systems of this sort benefit from the high accuracy of controller based tracking, text entry performance (in the range of 13.6 to 24.73 WPM for the representative examples cited above) may be artificially limited by their reliance of cursor-based input. In addition, using controllers makes these designs unsuitable for many AR (or casual VR) settings in which holding and managing input peripherals is either impossible (e.g., during hands-busy maintenance) or impractical (e.g., during a prolonged museum visit). Projects that eschew controllers and rely on bare hand tracking have the potential to address these limitations—to attain speeds approaching those of keyboard typing while also enabling users' hands to remain unencumbered. The existing performance data is promising. Using a high end marker based optical tracking system, Dudley *et al.* [87], for example, achieved mean WPMs of 42.1 for two-finger in-air typing and 34.5 for 10-finger in-air typing on a virtual keyboard. ATK [47] explored an alternative approach and combined a relatively low fidelity consumer level finger tracker positioned directly under the fingers with the simpler task of making individual finger strokes according to a touch typing scheme. In ATK, no virtual keyboard is present; rather intended key entries are derived solely from finger motions. This approach places fewer demands on tracking fidelity and the authors report relatively rapid text entry speeds of 29 WPM.

While these systems highlight the potential of achieving rapid unencumbered text entry on HMDs, they are not currently practical. The systems reported by Dudley *et al.* [87] and Yi *et al.* [47] rely on external and/or high performance trackers able to capture the hands and fingers from angles, and with a fidelity, that is not currently achievable on HMDs. Indeed, due to occlusion issues, it is unlikely that the currently dominant inside-out approach to hand tracking on HMDs will lead to reliable monitoring of the kind of downward finger strokes used in these systems: the movements are simply too small, rapid and liable to be obscured by the hand. We argue for exploring alternative approaches to bare hand text entry on HMDs and take inspiration for this from a prevalent tactic on mobile devices: two-thumb text entry. This widely studied technique [337–340] involves a palms-up two-handed grip on a mobile device and the interleaved use of both thumbs to strike keys. It is prevalent—used by up to 82% of people [341]—and data from both lab [55] and field [341] studies indicate it results in high text entry speeds of, respectively, 50.1 WPM and 38 WPM, the peak performance rates reported in both these studies. In addition, the hands-up pose provides a clear view of the thumbs from cameras mounted on a HMD. Despite this advantage, the popularity of the technique and the high performance its users achieve, we are not aware of prior work implementing thumb typing for HMDs.

This paper seeks to rectify this omission and design, develop and evaluate a thumb typing interface for the bare hand tracking system integrated into a commercially available HMD (an Oculus Quest [41]). To achieve this we present four studies. In the first, we capture performance with a large set of targets displayed in a volume above the palm. We use the results to characterize in-air thumb targeting motions and, in particular, to identify an appropriate surface on which to position keyboard keys. Inspired by common thumb-based input devices, such as game controllers, we select a design based on four keys per thumb for further study. A second study captures performance on all possible pairs of these keys—a sequential input task mimicking the interleaved input that occurs during typing. We use this data to design a accurate key selection procedure and an ambiguous keyboard layout that balances maintaining QWERTY-similarity [342] with achieving reasonable word disambiguation performance [161]. A third study evaluates this layout in both a word repetition text entry task, intended to simulate expert performance, and a more realistic phrase entry task. We report final WPMs of 27.1 and 13.73 and final error rates of 3.31% (Uncorrected Error Rate) and 10.1% (Corrected Error Rate) for word repetition and phrase entry, respectively. We complement these results with a final study that compares novice performance with ThumbAir against a baseline freehand text entry technique that relies on index finger taps to an in-air QWERTY keyboard [45]. The results suggest that while initial trials with ThumbAir, when users are familiarizing themselves with the layout, may be slower than this baseline, this effect rapidly diminishes: after entering just 15 phrases mean WPMs (at 11.24 and 12.09) did not significantly differ between the conditions. In addition, ThumbAir led to reduced levels of perceived exertion ("gorilla arm" [75]) and was viewed as more discreet and social acceptable than the baseline.

The contributions of this paper are two-fold. Firstly, the design of a bi-manual, unencumbered thumb-based text entry system for a commercially available HMD. This data driven process includes a description of the viable volume for locating thumb targets during bare hand HMD input tasks, a characterization of performance during a sequential input task with a selected subset of targets located in this viable volume, the design of a system to accurately detect target selections and the final computationally designed keyboard layout. The second contribution is two evaluations of the keyboard design and layout in repetition (once) and phrase based (twice) text entry tasks. These studies demonstrate that our keyboard design enables users to achieve rapid, accurate, comfortable and socially acceptable bare hand text entry using the tracking systems available in a current consumer HMD.

6.4 Related Work

Text-entry on HMDs

The challenge of typing on AR/VR HMDs has provoked the design of numerous text entry systems. As many headsets ship with a hand-held controller, early approaches sought to re-purpose the input capabilities of this device. However, as such devices are predominantly oriented towards pointing input, performance of the resultant systems was somewhat limited. For example, Jones *et al.* [188] proposed accelerometer-based gesture typing and achieved 5.4 WPM while Shoemaker *et al.* [78] used a ray-casting technique and achieved performance levels of 10.1-14.5 WPM. Error rates in both these projects remained high. More recent projects have sought to integrate traditional text entry devices and touch typing input styles into HMD scenarios, most commonly by combining a real physical keyboard with a virtual counterpart [303, 309, 343–345]. While this taps into users' physical typing skills and supports rapid text entry (WPMs of between 26.3 and 43.7), it essentially tethers the user to the location of the physical keyboard, typically a desk. One way of avoiding this limitation is to sense natural typing behaviors against arbitrary surfaces. Zhang *et al.* [346], for example, used a pair of commercial wearable devices composed of five inter-linked motion sensing finger-rings to track finger impacts on a range of surfaces. Participants operating this system achieved very rapid text entry speeds (of 70.6 WPM) by leveraging their existing typing skills. While this is impressive, it requires users to don two additional wearable devices. Recognizing that managing this additional equipment may be impractical for many HMD users (and use scenarios), other authors have designed systems that rely on sensors or surfaces built-in to the HMD itself. Numerous modalities have been proposed, such as head rotations (13.24 WPM for expert users at the end of a longitudinal study [228]), head gestures (24.73 WPM [226]) or eye-gaze assisted touch input (11.05 WPM [73]). While performance in these systems can be rapid, they suffer from the disadvantage that they dominate visual attention—in contrast to regular typing activity on a keyboard or touch screen, users must visually attend to the input surface, a typically undesirable property that precludes focusing on the written text, or anything else, while typing.

In-air Typing on HMDs

One approach that has the potential to achieve a more familiar experience is in-air typing. A very wide variety of designs have been proposed in this space. For example, Gupta *et al.* [347] captured the orientation of a finger ring to make input on a novel rotary keyboard layout, ultimately achieving expert performance levels of 14 WPM. A more typical design involves a QWERTY-like virtual keyboard—a grid of targets that a user strikes with one [88, 103], two [45, 87], or more fingers [47, 87, 306] to enter text. While such systems can be effective, yielding between 9.8 [45] and 42.1 WPM [87], they typically do so by either relying on high performance external optical trackers [87, 306] or cumbersome and potentially uncomfortable hand poses that clearly expose the tapping fingers to an HMD based sensor. For example, Sun *et al.* [103] describe a system based on index finger strokes that take place directly in front on an HMD's cameras—short stabbing motions towards the face at eye height. A more common design has been to present a virtual keyboard in front of a user, within ready reach, and require they strike the keys with one [88] or both [45, 87] of their index fingers. While this pose offers advantages, such as its resemblance to typing on a vertically mounted touchscreen, it requires maintenance of a hands-raised posture that is widely acknowledged to cause cumulative arm fatigue (or gorilla arm [75, 348]), a problem that that may ultimately restrict its practicality. In addition, while limited work has examined the social acceptability of HMD input methods [217], numerous authors have remarked that a pose entailing the arms raised and directly in front of the body may be considered awkward and undesirable [349]. As such, users may be reluctant to adopt it in a wide range of public and semi-public settings. These concerns may additionally serve to constrain the practicality of such approaches and highlight a need to develop in-air text entry methods that rely on input actions that are less strenuous and more discreet.

One approach to meet these requirements has been to capture performance with in-air typing tasks in the absence of explicit keyboards. The aim in such systems is to capture natural typing finger motions and infer intended targets. While the results suggest this can be highly effective—from 29.2 WPM in a working system [47] up to a possible maximum of 49.1 WPM if user performance is fully unconstrained [350]—the high fidelity of finger tracking required again necessitates the use of external trackers. As such, while in-air typing is a promising approach to HMD based text entry we note that current solutions are not yet practical. They are either dependant on external trackers, meaning they cannot be used standalone, or require uncomfortable, and potentially strenuous, artificial poses to ensure clear input. As such, we highlight a need for more research to identify viable in-air typing schemes that can work with current inside-out HMD based tracking systems and enable users to adopt physically comfortable and socially acceptable input poses.

Thumb Typing on HMDs

One potential solution is to use thumb input. While we are not aware of any prior implementations of in-air thumb typing for HMDs, numerous authors have explored the idea using physical controllers. For example, PizzaText [198] used the thumbsticks of a game controller to achieve a mean WPM of 13.77 while systems based on thumb-trackpads [79], trackpads-plus-hover [191], or a held smartphone [194] show peak performance levels of between 13.57 and 29.91 (after extended training). Fully wearable systems are also common. Designs include drawing gestures over the skin of the fingers [180] and thumb-to-finger taps [183] on the finger phalanxes [260] or nails [262]. While high levels of performance are frequently reported (e.g., up to 16 WPM in phrase typing task [260] and 31.3 WPM in word repetition task [262]), such systems invariably rely on elaborate wearable hardware such as touch sensing gloves [259, 351] or finger nails [262]. As such, while they highlight the strong potential of thumb based text entry for HMDs, these systems are lab prototypes that cannot be implemented on current commercial platforms.

Inspired by prior work on both in-air input and thumb-typing, the remainder of this paper seeks to implement a system that combines them to achieve a comfortable, discreet, and efficient HMD based text entry system. A key additional goal is to design a practical system that can be implemented for the inside-out tracker on a commercial HMD. The remainder of this paper describes our steps towards these goals.

6.5 Platform and Environment

All work in this paper was conducted using an Oculus Quest VR HMD. We note VR HMDs are often used to design and prototype interactions intended for both AR and VR scenarios [226, 267, 352] due to the greater maturity of the VR product category. This device features an advanced "inside-out" bare hand tracking system capable of capturing finger and thumb movements without additional hardware [41]. More specifically, it features a camera-based multi-stage hand tracking pipeline that implements processes of hand detection, hand keypoint identification, and model-based tracking [39]. It has an average finger joint angle error of 9.6° and an average temporal delay of 38ms [39]. In all studies, we configured both the frame rate of HMD display and the hand tracking system to be 60Hz. All studies and applications were developed using Unity3D and the Oculus Quest SDK. The size of virtual hand used in all work reported in this paper was the default size specified in the Oculus Quest SDK. Virtual hands did not vary in size by participant. To best support this device, we minimized variability in tracking performance due to environmental issues by conducting all work in the same laboratory room and ensuring it was always lit solely by constant fluorescent lighting. Windows were blocked by blackout curtains. Finally, we note that all studies reported in this paper were approved by the local IRB and conducted in full compliance with all national and institutional rules and recommendations relating to social distancing.

6.6 Range of Motion Study

Key location is a critical aspect of a keyboard design—a comfortable, effective keyboard is composed of keys in easy and ready reach. While numerous authors have conducted studies to establish appropriate key locations in scenarios such as touch typing on a tablet [337], thumb typing on an index finger [353] or even during in-air touch typing [47], no prior work has explored this issue for in-air thumb typing. There is reason to suspect performance from other settings will generalize to an in-air experience—the specific open-hand, palm-up we use, the complex articulation of thumb movement and the lack haptic feedback inherent during in-air input will likely create a unique performance profile [354]. To address this omission, we therefore conducted a study to establish thumb input performance over a large set of in-air targets. The goals are to improve our understanding of in-air thumb targeting motions and support the selection of a subset of target locations that support fast, accurate and reliable user performance.

Methods

Participants. Sixteen participants (mean age 23.8 (SD=3.7), eight female, eight male) completed this study. They were screened for right-handedness. They rated their familiarity with computers (5.0/5.0) and smartphones (5.0/5.0) as high, and with HMDs (1.94/5.0, SD=0.93) as relatively low. The study took approximately 50 minutes to complete and participants were compensated with the equivalent of 15 USD in local currency.

Study Design. The study was descriptive in design: all participants completed a single condition—tapping in-air targets with their thumbs—in which we sought to characterize performance. Participants were required to hold their palm up, with thumbs clearly visible to the HMD tracking system. We arranged a grid of targets above each palm with the goal of densely populating the full range of locations comfortably in reach of the thumbs. Through a process of subjective experimentation, we first defined a viable volume for these targets with respect to both center of the palm and the fixed virtual hand size, as reported by the Oculus SDK. This volume ultimately covered a range across the palm that spanned the region underneath the ring, middle and index fingers, and a range along the palm from the base of the thumb to the base of the fingers. Due to discomfort experienced when trying to actually touch the palm, we ensured the lowest targets were situated 2cm above the palm and defined the limit for highest targets based on an assessment of the vertical reach of the thumb above the palm. Based on these constraints we ultimately selected a volume that was 6cm by 4.75cm by 4.75cm, centered 3.875 cm above the palm, precisely centered across the palm and offset 1.25cm forward, towards the the finger tips. We then selected a typical target shape (sphere), size (1cm diameter) and spacing (0.25cm) and uniformly populated this volume. This resulted in a five by four by four grid of targets positioned relative to the center of the each palm: 80 targets per hand, 160 in total. In addition to this layout, we specified a start point for each thumb motion. This was 2.5cm beyond the outer edge of the grid of targets, aligned with the grid's center. The targets and thumb starting position are illustrated in Figure 27.

Each trial in the study started with a blue highlight over the left or right thumb start point. After moving their corresponding thumb to highlighted location, participants were required to dwell for 0.5 seconds, during which time the hand model was greyed out. The hand model was then returned to skin color, one of the grid targets was displayed in red and the participant's task was to move their thumb quickly and accurately to intersect it. No distractor targets were shown during this task. Trials were arranged in blocks of 160: one unique occurrence of each target. Trials in each block were delivered in a random order and the study consisted of six blocks, the first of which was discarded as practice. In this way, each we retained 12,800 trials (16 participants by 5 blocks by 160 trials) over the study. Each trial timed-out after a maximum of three seconds and we logged the success rate (whether or not the participant touched the target), the touch time (measured from initial display of the target until first contact) and thumb velocity during this period. Additionally, we recorded global changes in wrist

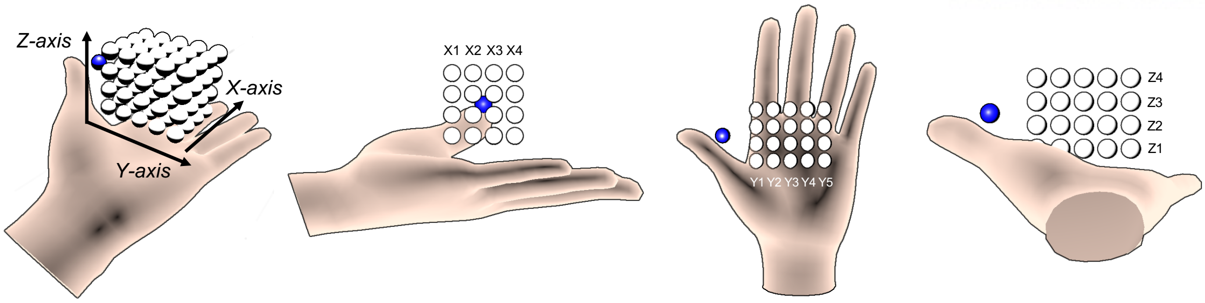


Figure 27: Target locations with x, y, and z axes marked in the range of motion study. The blue sphere is thumb starting point.

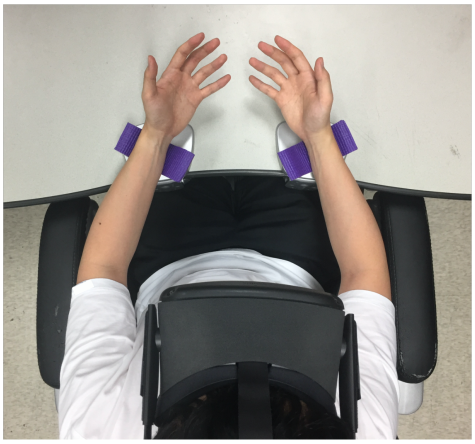


Figure 28: Participant in the range of motion study. Participants selected a comfortable palm-up hand pose and used wrist supports (in blue) to avoid fatigue.

angle in order to estimate difficulty—the intuition being that increased wrist motions correspond to more challenging targeting movements [355].

Procedure. The study took place in an empty office lit solely by uniform fluorescent lighting. Participants were seated in front of a desk. They first read study instructions, and completed consent and basic demographics forms. They then viewed a video clip depicting the study procedures and had the opportunity to ask any questions to an experimenter. After indicating they clearly understood the study task, participants adjusted the position of two wrist supporters mounted on the desk to achieve what they considered to be a comfortable palm-up hand pose (see Figure 28). We used these wrist supporters to minimize the impact of "gorilla arm" [75], or upper limb fatigue, during the course of this quite intensive, repetitive and prolonged study. Next, participants donned and adjusted the HMD, then returned their wrists to the desk-mounted supports and began the study trials. There was an enforced rest of three minutes between blocks. At the end of the study, an experimenter measured their thumb size.

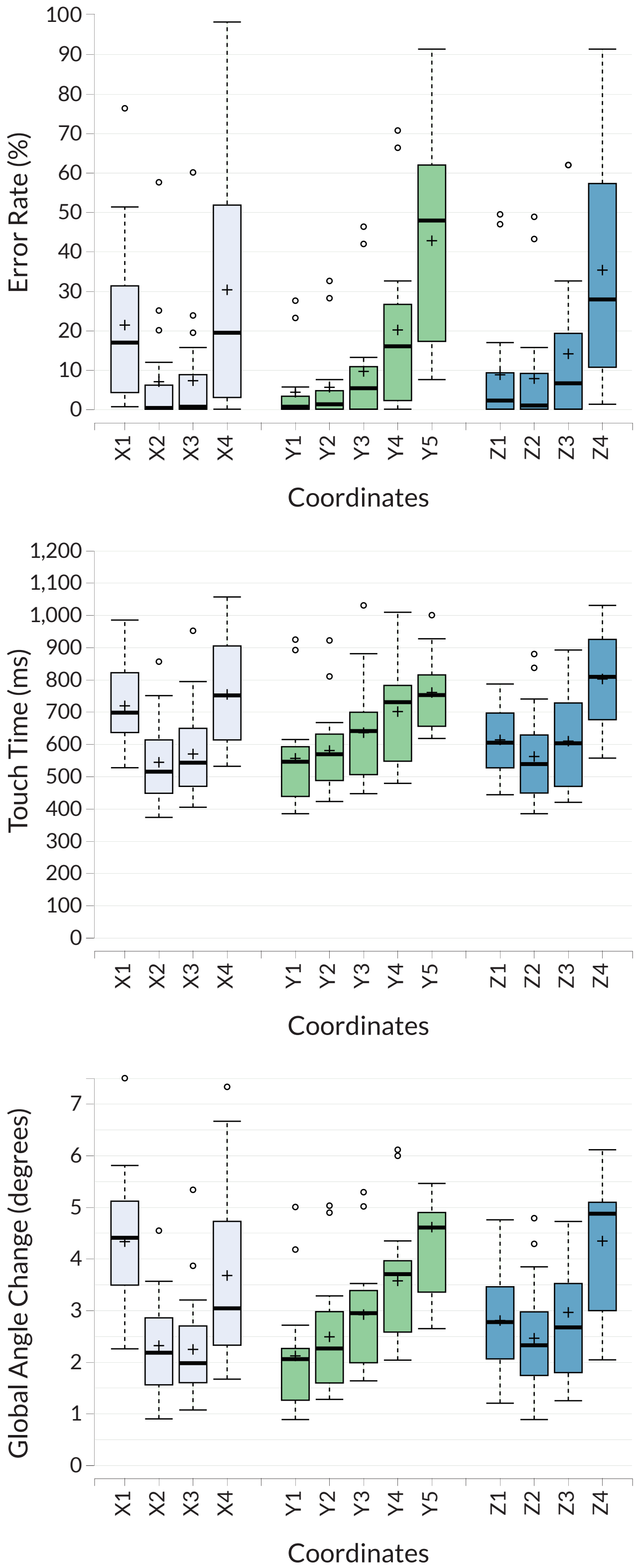


Figure 29: Error rate, touch time, and global changes in wrist angle in wrist for range of thumb motion study. Data is combined from both left and right hands.

Results and Discussion

We first examined mean data for time, errors and wrist angle change. We conducted paired t-tests for each measure to examine performance variations due to the handedness. No differences between dominant and non-dominant hands were observed (the lowest p value was 0.225), so we combined data between hands for all subsequent analysis. The resultant mean data for these measures are shown in Figure 29, organized to depict variations along targets in the x, y, and z axes. The overall mean success rate per target was 83.6% (SD 23.3%), indicating that many targets were hard to reach. This was expected: the target set was designed to clearly document the viable range of thumb motion by extending beyond it into more challenging territory. For those target selections that were successful, the mean touch time was 646ms (SD 179ms). This indicates that the in-air tapping task could be executed relatively rapidly. For example, prior thumb targeting data reports targeting times of between 560ms [356] on a mobile phone and 1500ms [357] on the fingertips. This high performance was enabled by a rapid (and not plotted, as it was highly uniform) mean thumb velocity of 0.138m/s (SD=0.028m/s) during targeting motions. This suggests that one factor supporting the rapid performance we observe was that participants took advantage of the purely virtual nature of the targets to approach at high speed, pass through the targets and decelerate in the space beyond them. We note that while this strategy is effective in the single target task studied here, it may be less useful in a more complex task, such as typing, that involves a sequence of selections. In such a task, an extended targeting motion for one target may make a subsequent targeting motion slower. In addition, if targets are densely arranged, such a strategy may result in intersections with multiple targets. This result suggests that an in-air thumb keyboard will need to be carefully design to prevent such inadvertent activations. We opted not to analyze this data statistically, as formally establishing the presence of performance variations would not serve our objectives. Finally, we note that the measure of wrist angle change (Mean = 3.14°, SD=1.66°), which was intended to capture subtler variations in the difficulty of different targeting motions, closely followed the trends in the time and error data (with Pearson correlations of between 0.831 and 0.995). This suggests these traditional metrics were able to effectively capture key trends in user performance.

We also examined the relationship between the thumb length (the distance from the end of thumb to the tip of thumb) and time and error data. The mean thumb length among our participants was 5.59cm (SD=0.39cm, Min=4.9cm, Max=6.4cm). However, Pearson correlations indicated that variations in thumb length were not linked to either success rate ($r = -0.223$, $p = 0.406$) or touch time ($r = 0.325$, $p = 0.219$). This suggests that participants with different thumb lengths were able to use the system with equal effectiveness.

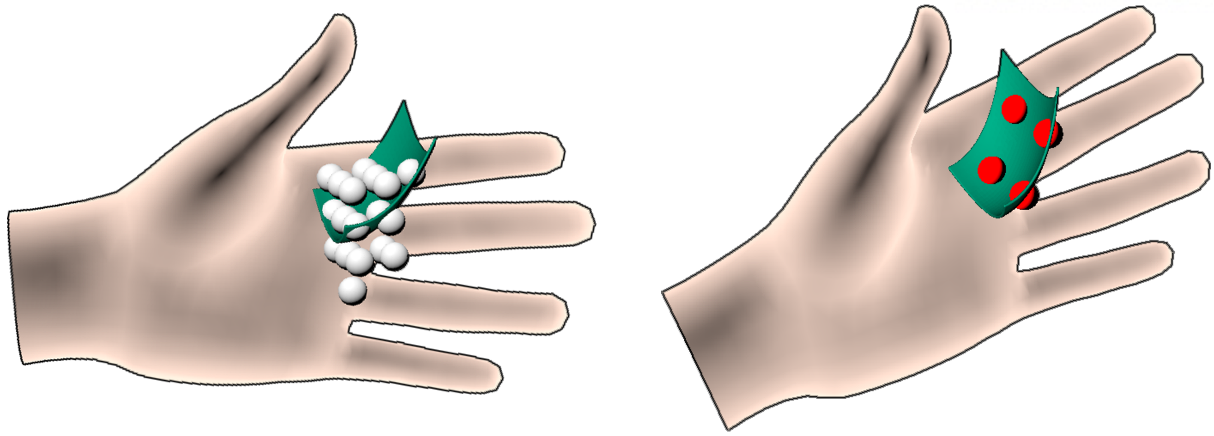


Figure 30: Surface (shown in green) derived from the 18 selected targets (shown in white, left image) in the range of motion study. The key locations (shown in red, right image) used in bigram study.

Selecting Target Locations

To determine appropriate target locations, we first created a single *performance score* for each target. This was based on normalized scores for success rate, touch time, and wrist angle; thumb velocity data was excluded as it showed highly limited variability. Data for each of these three metrics was first inverted to ensure larger scores corresponded to more desirable performance (i.e., higher success rate, lower times, and reduced hand rotation). The *performance score* for each target was then calculated as the magnitude of the vector composed of these three features. We then selected the top 20 scoring targets for each hand (25%) and then, as we intend to support bi-manual input, filtered these by those that appear in both hand's data sets. Ultimately, we retained 36 targets, or 18 on each hand. These are shown in Figure 30 and their coordinates, and the normalized scores recorded for all metrics, are reported in Appendix 6.12. They were arranged in a tapered wedge aligned with the movement of the thumb tip as it rotates around the metacarpal joint. In this set, targets above the center or ulnar (little finger) side of the palm were physically occluded by those on the radial (thumb) side—to reach targets over the ulnar region, the thumb would need to pass through targets above the radial region. As such, many of these target locations were incompatible with one another.

In order to establish a more limited set of possibilities for final target locations that would avoid issues of physical occlusion, we fit a paraboloid to the positions of the full set of targets (R^2=0.154), aligned and clipped so that it faced the thumb of each hand, maintained rectilinear edges and did not extend beyond the range of the 18 original targets. This surface specifies an appropriate set of locations on which to situate targets that will be rapidly and accurately in reach. We located candidate targets on this surface by considering our design goals of supporting quick, error (and inadvertent activation) free targeting performance on an ambiguous keyboard. As performance with larger numbers of targets will inevitably reduce performance we opted to locate four targets on each hand or a total of eight over both hands—eight targets are sufficient to support a wide range of ambiguous keyboard designs [183, 262]. We located these four targets by the simple expedient of dividing the paraboloid surface into four quarters, and placing a target at the center of each quarter. We note this four target design bares similarity to common input devices such as directional thumb joy-pads. We believe the familiarity of this arrangement may help to reduce novelty effects with our system. In addition, we validated these target locations with respect to the reported sensing resolution of the Quest HMD used in this work. Measured from the trapeziometacarpal joint of the thumb, the angular gaps between the four targets range from 17.0° to 35.2°, comfortably exceeding the Quest's 9.6° joint angle accuracy [39]. This suggests the performance of the hand tracking system used in this work should not overly impact or constrain the performance participants can achieve when selecting these targets.

6.7 Bigram Study

We conducted a second study to complement the data captured in our first study. It had two objectives. First, it captured data in a sequential input task. This is important because ambiguous keyboards [160, 259, 262, 358, 359] feature relatively few keys (in the range of between 5 and 10), each of which is mapped to multiple characters. They take advantage of the fact that the vast majority of possible character sequences do not represent valid words to create accurate and effective text entry systems—although individual key entries may be highly ambiguous, the vast majority of words can be precisely specified by appropriate sequences of ambiguous selections. However, in order to map characters to keys, a process known as the letter assignment problem [360], in arrangements that support good user performance, data about single input events, such as that gathered in first study reported in this paper, is insufficient. Rather, typing is a continuous task and data about performance of sequential input actions—how users select targets one after the other—is required. This study aimed to capture such data. In addition, we also sought to characterise the distribution of thumb motions made during targeting. Data of this sort has previously been captured for touch screen keyboards and used for a range of purposes, such as understanding typing behaviors [312] and optimizing or customizing [313, 361] key locations and sizes. We captured such data in our study to support the similar objectives: to use it to adjust our target locations and/or target selection process in order to improve user performance.

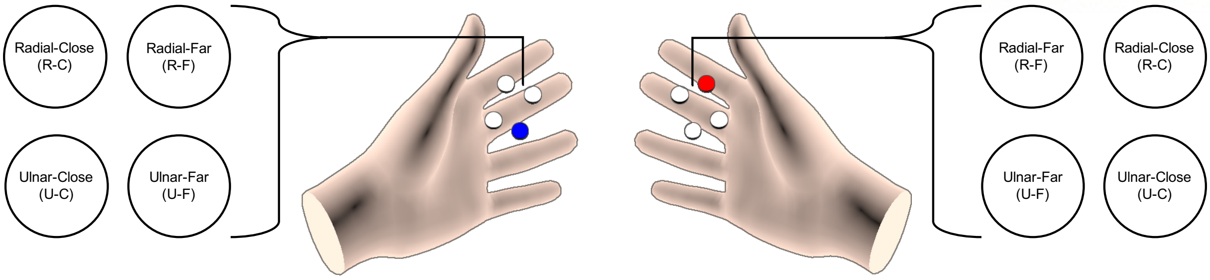


Figure 31: Study task in the Bigram study showing both hands and all eight keys. Participants were required to touch red target first and blue target second. Annotations indicate naming conventions for the targets and were not shown during the study.

Methods

Participants. Sixteen new participants (right-handed, mean age 25.8 (SD=4.1), seven female, nine male) completed this study. They again rated themselves as highly familiar with computers (5.0/5.0) and smartphones (5.0/5.0) but with relatively low familiarity with HMDs (1.94/5.0, SD=0.93). The study took approximately 50 minutes to complete and participants were compensated with the equivalent of 15 USD in local currency.

Study Design. The study again featured a single condition. It had the goal of characterizing user performance in the task of sequential in-air thumb taps. The candidate layout selected from the results of the first study features eight 1cm diameter targets, so this study used the full set of 64 possible sequential pair-touches (see Figure 31). The study was again designed with six blocks, each of which featured a single randomly ordered occurrence of each possible pair-touch, with the first two blocks discarded as practice.

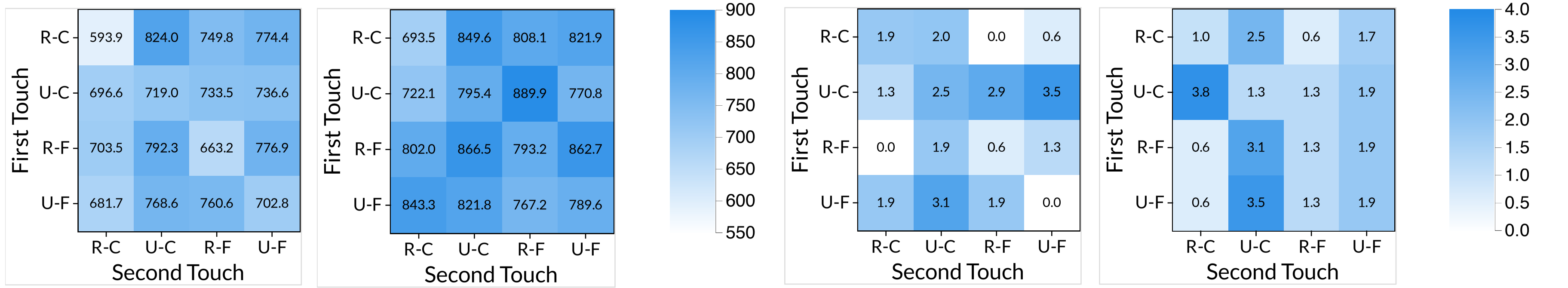
To characterize thumb motions, we opted for the simple expedient of logging the area of all thumb contact with the paraboloid surface during each targeting task. This approach is directly analogous to closely related prior work on touch screen typing [312, 313] which logs all finger contact points during key presses. To achieve this we created a grid of 0.15cm diameter transparent spherical colliders (15x16) distributed evenly over the paraboloid surface and spaced at 0.15cm intervals. Effectively, this formed a solid mesh of separate colliders. As the visually displayed targets were spherical, we offset this collider mesh towards the thumb and along the paraboloid's normal vector by 0.5cm—the radius of the visually presented targets. This ensured that we captured data from even shallow contact with the displayed targets. In addition, to log full data from each touch we considered each touch to start and end when the thumb made and released contact with the collider grid—a style of input analogous to interacting with a touch screen. To ensure that we captured the most diverse set of valid thumb motions toward each target we treated any touch that contacted the correct target to be correct, irrespective of whether or not it came into contact with other targets beforehand or afterwards. As such errors were recorded only if participants failed to make contact with one or both of the correct targets in a trial in the correct order. In addition, we recorded a timeout if no contact with any target occurred after five seconds. To ensure we logged a complete set of data for each target-pair, error and timeout trials were repeated in a new random order. In this way, the study retained data from 4096 successful trials (4 blocks by 64 pairs by 16 participants), with each trial composed of a pair of target selections.

Trials in the study were broadly similar to the those in the first study. Each trial started by tapping a start button, following by a fixation period during which time the hand model and all eight targets were greyed out. After two seconds, the hand model was re-colored and two targets were highlighted - red for the first touch and blue for the second touch. If the first and second touch were due on the same key, participants were required to tap this key twice. When touching one of the targets, feedback was provided by applying a green highlight.

Procedure. The procedure broadly followed the first study. It was completed in the same environment, and participants completed similar consent and demographic forms. Instructions were again read and exemplified via a video. To increase ecological validity, and due to the reduced fatigue expected during the somewhat shorter data collection period in this study, participants did not use the wrist supporters in this study. Instead, they adopted a comfortable hands-free, palm-up, in-air posture of their own choosing. Prior to the main study, participant's completed an informal practice session (max five minutes) in which the eight keys were shown and they were able to practice touching them with their thumbs. Participants then began the main study. A break of three minutes was enforced between each block and participants were additionally able to rest between each trial if needed. After completing the study, an experimenter measured their thumb size.

Results and Discussion

We recorded a total of 4223 trials, including 4096 successful trials and 127 failure trials in which the participant did not touch the specified targets in the specified order. From this set of trial failures, we excluded the 26 trials involving selection of a first target on the wrong hand—these failures likely represented confusion with regards to the trial instructions rather than performance of the actual study task. Similarly, we excluded the 13 timeout trials, thus retaining a total 88 errors for analysis. Based on this data set, we first examined the impact of participant thumb length (Mean 6.01cm, SD=0.47cm,



(a) Bigram-time within hand (ms)

(b) Bigram-time between hand (ms)

(c) Error rate within hand (%)

(d) Error rate between hand (%)

Figure 32: Bigram-time and error rate divided according to trials in which both targets were on the same hand (within hand) and those in which they were on opposite hands (between hands). Targets are described by their position: R(adial) or U(lnar) and F(ar) or C(lose). See Figure 31 for an illustration of this mapping

Min=5.5cm, Max=7.0cm) on performance using Pearson correlations, finding no relationship with either bigram-time (p=0.465 with r=0.197) or error rate (p=0.365 with r=0.243). As in the first study, participants' performance was not linked to the size of their physical thumbs.

We then plotted the data—Figure 32 shows confusion matrices showing the mean bi-gram time and error rate for all possible pairs of touches. We first analysed this data by conducting two-way RM ANOVAs on the measures of bigram-time and error rate using the independent variables of first-touch (left or right hand) and second-touch (same or different hand). These tests sought to establish whether there were performance variations due to handedness or the possibility of interleaving the target selections—of overlapping input in trials in which it was bi-manually split between the hands. As with the range of motion study, we did not detect any differences in performance between dominant and non-dominant hands in either time or errors, suggesting that handedness did not exert an effect. This result is positive—users should be able to use a bi-manual system effectively from the outset. Based on this result we combined data from both the hands during all subsequent analysis. However, while there were no differences in terms of errors, we did observe a single significant difference for the second-touch variable: a main effect of bigram-time (F(1, 60)=4.1, p=0.047). This indicates that two touches on the same hand showed were performed faster (730ms, SD=134ms) than a pair of touches composed of a single touch to each hand (806ms, SD=162ms). This result suggests that participants were unable to effectively interleave input between their hands. A potential explanation for this is that they may have needed to direct visual attention to each individual target selection task and faced challenges with shifting visual attention to a new target on different hand in a short period of time—due to a lack of familiarity with the in-air input task, participants may have visually monitored each thumb movement, in effect placing a central processing limit on performance.

The overall error rate in this study was a mean of 1.67% (SD=1.47%) over all target pairs. This high level of performance indicates all targets were in easy reach of the thumb. While this figure is inarguably low, we note it also applies an unrealistic decision criteria for determining target selection: any contact with the specified target, irrespective of whether it proceeds or succeeds contact with another target, is considered to be a successful trial. Use of this criteria enabled collection of the most diverse set of valid targeting motions possible. However, it cannot be implemented in a realistic system as intended targets are not known in advance. Accordingly, we opted to explore error rates with more realistic and unambiguous decision criteria. Specifically, we examined error rates based on treating the initially touched target as the intended selection and those on treating the last touched target as the intended selection—conceptually equivalent to touch-down and touch-up events. The error rates calcu-

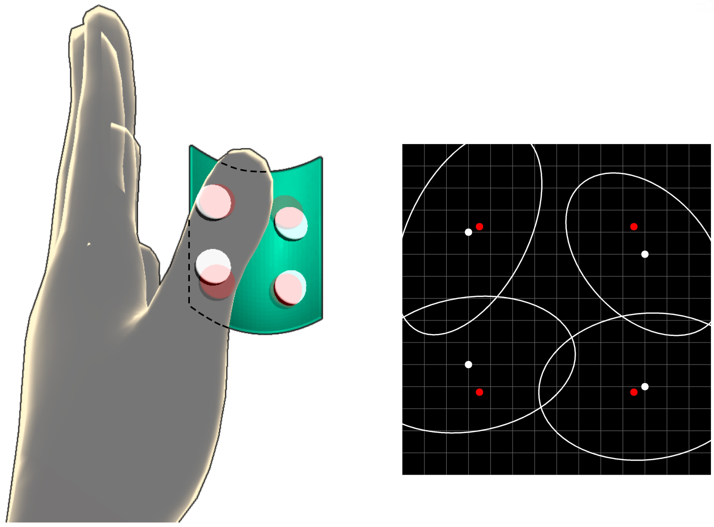


Figure 33: Four white keys for text entry study are the refined target locations (left hand) with the mean target centroid positions for each key. Four red keys are the previous target locations used in bigram study.

lated according to these criteria were substantially elevated at means of 9.01% (SD=6.17%) and 20.89% (SD=9.54%), respectively. These increases are likely due to a number of factors. Firstly, physical occlusion of the intended target by other targets. For example, of ulnar targets by radial ones. Such occlusion would result in the thumb path intersecting an unintended target on its way to or from the intended target. Secondly, overshooting an intended target and subsequently colliding with an unintended target. A final factor is doubtless the definition of errors used in the study. As we did not penalize contact with non-targets, participants did not explicitly seek to minimize these collisions. While a study applying a stricter error criteria would likely reduce the prevalence of unintended collisions we note that it would also represent a fundamentally more difficult task, in which the burden of avoiding unintentional collisions is placed entirely on the user. This would likely lead to elevated errors and time data. Based on these observations, we determined that developing a technique to accurately and unambiguously detect intended target selections from user's thumb motions is a key next step for this work.

To address this problem, we turned to the touch profile contact data recorded from the collider mesh. As we previously noted no performance differences dues to handedness, we first collapsed data between the hands. From each trial we then aggregated data from all frames in which the thumb made contact with the mesh, ultimately creating a 15x16 pixel binary image (a *collision map*) representing the full set of these intersections. We then explored the design of machine learning classifiers capable of determining the intended target from the collision maps. To build these models we constructed neural networks using PyTorch. We selected neural network models as they can process both raw image data and other forms of numerical data (e.g., summary statistics derived from the images). In addition PyTorch neural networks can run naively in Unity and on the Quest headset when converted to an Open Neural Network Exchange (ONNX) format. To construct classifiers we used Leave-One-Out cross validation (LOOCV), an approach in which each participant's data serves as the test set for a model built on all other participant's data. This reduces the risk of over-fitting. Other aspects of the models and training procedures were customized to match the features we explored.

Specifically, we examined performance with two feature sets. First, we calculated image moments for each collision map. We used these to extract the centroids [283, 312] of the touched regions. We designed a simple model for this data comprised of an input layer, a hidden layer, and a softmax output layer, all fully-connected using linear transformation. To support multi-label classification, we used Adaptive Moment Estimation (ADAM) as an optimizer and Cross-Entropy-Loss as the loss function. When training the model, we avoided under or over-fitting by monitoring changing trends in training and validation loss per epoch. This model ultimately attained a LOOCV classification accuracy of 97.6% (SD=1.8%). Second, we used the raw collision maps. We first down-sampled these using linear interpolation to create a 3x4 grey-scale image. We then constructed a simple model composed of one layer for linear transformation. We used Stochastic Gradient Descent (SGD) as an optimizer and Cross-Entropy-Loss as a loss function. This model achieved an LOOCV accuracy of 98.2% (SD=1.5%).

Results from both classifiers are strong, reporting accuracies close to the original 98.33% recorded in the study. This suggests that contact with non-intended targets is, in general, brief and/or slight and does not preclude reliable inference of intended targets from thumb motions. Based on these results, we made two revisions to the target layout and selection process used in further studies in this paper. First, we opted to use the collision maps and a version of the image-based classifier built using the full set of user data to determine selected targets. This is due to modestly higher accuracy and lower variability in the LOOCV results compared to the centroid based classifier. One implication of this choice is that target selections are calculated on release of the collider grid (an event equivalent to finger-up on a touch screen) rather than in relation to contact with visually displayed targets. Second, we refined the location of the visual targets. We achieved this by calculating all ellipses from the collision maps for each target and using these to calculate the mean target centroid position. These positions represent, on aggregate, the locations participant's actually moved towards during trials. To reflect these motions, we placed the final target locations at each of these mean positions. This process follows prior work identifying optimal key locations for touchscreen typing [313]. The revised target locations are illustrated in Figure 33 and reported in Appendix 6.12.

In sum, this study recorded mean bigram-times and error rates that that were were stable across both hands and showed relatively minor advantages for uni-manual input over bi-manual input. While error rates are somewhat high using simple (touch-down/touch-up) selection criteria, we show how simple classifiers can be used to dramatically improve these. It is worth contextualizing this data. In a broadly similar study involving pairs of thumb (and finger) taps to the nails of the same hand, Lee *et al.* [262] report mean bigram-times of approximately 550ms and error rates of approximately 4.5%. While our times are somewhat elevated compared to this prior study (at 768ms, SD=74ms), error rates from our classifiers are notably reduced. We argue the results are strong enough to validate our refined target locations and target selection procedures and support further investigation into their ability to support effective in-air thumb typing.

6.8 Keyboard Layout Selection

We used the data from the Bigram study to inform a keyboard layout design process. The goal was to map characters to the eight keys in our system in such a way that they can support rapid, accurate, unambiguous and familiar input. We achieved this through three mechanisms. First we specified a limited set of layouts that retain a strong similarity to QWERTY. Second, we defined four metrics and calculated these for each candidate layout. Next we reviewed the layouts and filtered them based the minimum and maximum number of characters assigned to each key. We then reviewed their performance on all metrics and selected a balanced candidate with a strong performance profile for further study. These processes are described and defined in the following sections.

Layout Constraints

We restricted the character layouts to closely resemble QWERTY. We achieved this by considering our bi-manual eight key layout as two rows of four keys. The top row of QWERTY characters was mapped to the top four keys, the bottom row of QWERTY was mapped to the bottom four keys and the middle row of QWERTY was assigned to either top or bottom keys. We also restricted column assignments by stipulating each row of QWERTY characters was divided as equally as possible over the four key columns—there could be either two or three characters per key for the top and middle QWERTY rows and one or two keys for the bottom QWERTY row. We then laid out all possible character to key assignments, following the QWERTY layout and respecting the possible sequences of character to key counts. For example, 'q' and 'w' could only be assigned to the top left key while 'e' could be assigned to either the top-left key, or the next key over. We determined all possible key assignments according to

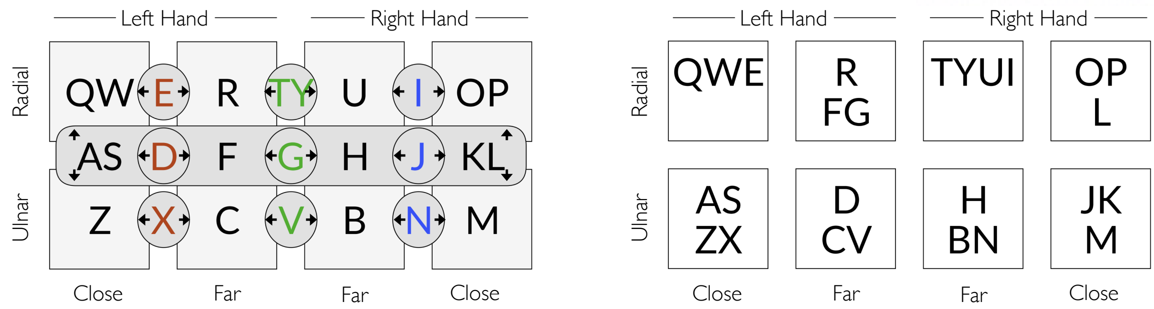


Figure 34: Keyboard layouts. Left shows the character to key mappings considered in the layout selection process. The eight keys are shown as grey squares, overlaid with the characters that could be assigned to them. Characters in black have a fixed key column. Characters in red could appear on the first or second column of keys (left hand); in green on the second or third columns (between the hands) and in blue between the third and fourth columns (right hand). In addition, while top row and bottom row QWERTY characters always appeared on, respectively, the top and bottom row of keys, all characters on the middle QWERTY row could appear on either the top or bottom row of keys. This limited set of variations was designed to ensure all layouts considered retain a close similarity to QWERTY. Right shows the final key layout selected for further study.

these constraints (illustrated shown in Figure 34, left) and then calculated all possible layouts: 524288 in total. Following prior work dealing with such restricted sets, we opted to calculate performance metrics for all layouts [353]. The number of candidate layouts is sufficiently small that a computational optimization process [259, 262] is not necessary.

Metrics

Speed. We combined our empirically captured bigram entry times with word frequencies from Norvig *et al.* [362] to calculate the text entry speed for a given layout using the standard quadratic formulation used for the letter assignment problem [360]. This calculates, for a given layout, a single measure that expresses how quickly text can be entered.

Accuracy. Accuracy was defined similarly to speed. We used the success rate of each bigram entry, together with Norvig *et al.* [362]'s word frequencies, to calculate a measure of the accuracy of text entry on each layout.

Qwerty Similarity. We defined a metric to model the extent to which layouts resembled the rows and columns in standard QWERTY. Row similarity was defined as the Levenshtein distance between the

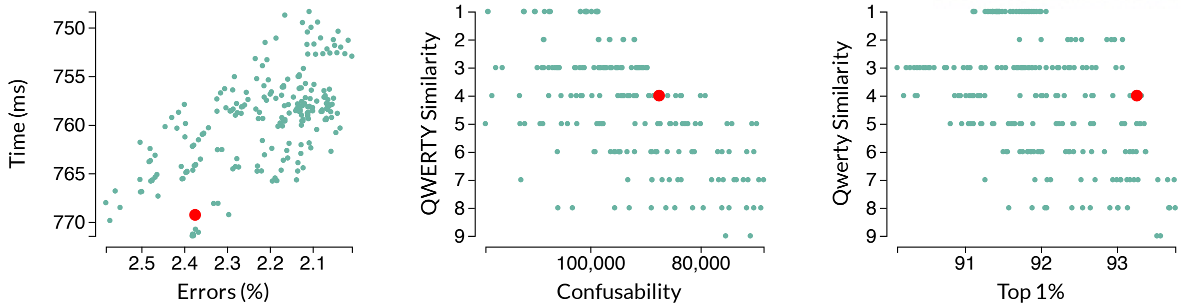


Figure 35: Layout scores for metrics of time/error (left), QWERTY similarity/confusability (center) and QWERTY similarity/Top 1% (right). Scales adjusted to show more desirable scores to the top-right. To facilitate illustration we show only the non-dominated solutions (i.e., the Pareto sets). Scores for the selected layout are manually highlighted in red. Note that due to the highly limited range of values calculated for time and error metrics (left), we did not consider them during layout selection.

characters in each row of the QWERTY layout and the order in which those characters were assigned to the keys in the layout. Column difference examined the characters in each QWERTY column (e.g., "QAZ") and calculated how many different key columns these were assigned to in each layout. We subtracted one from the value for each column to attain zero for a column arrangement than matched QWERTY. We summed these two measures to create a final metric that expressed how distant character assignment in a layout was from the rows and columns in QWERTY.

Confusability. A key measure for ambiguous keyboard layouts is how unambiguous they are—how uniquely sequences of key presses will accurately specify intended words. To estimate this, we calculated Lesher *et al.* [358]'s confusability matrices. We closely followed a recent implementation for confusability matrices [262], which we briefly review here. Confusability matrices assess ambiguity by calculating, for a given text prediction algorithm, the frequency with which all pairs of letters are mistakenly selected for each other in a given text corpus. The relative confusability of a particular key layout can then be estimated simply by summing, for each key, the matrix cells for the characters assigned to it. We created confusability matrices using Lesher *et al.* [358]'s k-gram algorithm (implemented via a dictionary of the thirty thousand most common words [363]) and a 322210 word corpus formed by combination of three mobile text entry data sets [364–366]. The confusability scores generated from these matrices are quick to calculate and effectively capture the relatively ambiguity of each layout.

Filtering and Selection

In order to facilitate layout selection, we explored methods to filter them. We first removed layouts with less than three or greater than four characters assigned to any key. The intuition here was that a relatively even distribution of characters would likely yield strong layout candidates (i.e., by avoid overloaded keys with many characters and high ambiguity) while also helping ensure the geometric arrangements of QWERTY were maintained. This led to a set of 10782 layouts. We then reviewed the raw data for metrics. Speed and accuracy showed highly limited variability (748-780ms and 2.0%-2.8%, respectively) over the full set of layouts; Pareto sets for these data are shown in Figure 35. The small ranges for these two variables likely reflects the fact that we examined a relatively small subset of possible layouts that all closely resembled QWERTY and, thus, all showed similar empirical performance profiles. As such we judged these two variables to have limited impact and did not consider them during layout selection. Due to challenges in contextualizing the confusability scores (68597-156939), we additionally calculated Gong *et al.* [160]'s disambiguation scores for all layouts. These scores express, for a given text corpus, set of word frequencies, number of character inputs, number of predicted words and key layout, the mean frequency with which all words in the corpus will be correctly predicted. This is a very practical measurement of how a given layout will behave—it estimates, as a probability, how often a user's inputs will result in prediction of the word they intend. We calculated disambiguation scores, assuming input of three characters, for most frequently predicted word (Top 1%). Figure 35 shows the distribution of scores in the non-dominated set of layouts for the metrics of QWERTY similarity, confusability and Top 1% scores. We note Top 1% scores were strongly correlated with the original confusability scores ($r(10782) = -0.748$). This indicates the measures assessed similar properties.

Based on these data, we then selected a balanced candidate attaining a good performance profile across QWERTY similarity, confusability and Top 1% score. Our goal was to support a familiar text entry experience without sacrificing the ability of the layout to unambiguously specify words. The character to key mappings in this layout are shown in Figure 34 (right) while its scores for the metrics of QWERTY similarity, confusability and Top 1% are illustrated with red highlights in Figure 35. It is worth contextualizing the Top 1% score. Notably, the whole range for this metric is elevated compared to typical examples in the literature. In Gong *et al.* [160]'s WrisText, for example, the Top 1% score for the layout they develop is 85.9%; for the layout selected here it is 93.3%, an increase which we believe will lead to markedly better word suggestion performance. In addition, we also calculated the Top 2% score for our proposed layout—the proportion of intended words that would be returned within the first two suggestions—and found this to be 98.1%. This suggests that a word selection interface featuring just two options would likely be highly effective when combined with our proposed layout. Based on these relatively strong results, we moved forward to further studies evaluating our key layout during actual text entry tasks.

6.9 Typing study: evaluating the performance of ThumbAir

We conducted a study to evaluate the performance of our in-air thumb typing system. During this study we used the target locations, target selection procedure and collider mesh based target classifier discussed and defined in the second study. To provide a rounded assessment we targeted capture of both an approximation of expert performance (via a word repetition task) and novice performance (via a traditional phrase entry task).

Methods

Participants. Fourteen new participants (right-handed, mean age 24.5 (SD=3.3), six female, eight male) completed this study. All were non-native English speakers enrolled in a full time English language degree program. They self-rated their familiarity with computers (5.0/5.0) and smartphones (5.0/5.0) as high, but they had passing experience with HMDs (1.86/5.0, SD=0.53). They took approximately 70 minutes to complete this study and each received the equivalent of 20 USD in local currency.

Study Design. This study was designed to assess typing performance with the keyboard and key layout designed based on data from prior studies reported in this paper. To do this, we included two stages. Participants first completed a word repetition task [262,367] designed to mimic expert performance. We followed closely related prior work [262, 367, 368] by using a 20 item word set ("the, and, you, that, is, in, of, know, not, they, get, have, were, are, bit, quick, fox, jumps, lazy, on") which includes all English letters and approximates monogram and bigram frequencies. Participants were presented with each word in this set in a random order and were required to enter it seven times. There was no support for correcting entered text. We recorded Words Per Minute (WPM) [369] from all entered words and calculated Minimum String Distance (MSD) error rate [370] at the character level. In this stage of the study, we collected 1960 typed words (composed of 6468 characters): seven repetitions by 20 words by 14 participants.

The second stage of the study was a phrase typing task. This more realistic typing task better reflects novice performance with the system. In this task, participants were required to correctly type a total of 40 phrases (1061 characters plus 240 spaces). We used this instruction to simulate a careful typing experience in which accuracy is emphasized over speed. Half of phrases consisted of selected ones from Mackenzie *et al.* [299]'s widely used set. In order to cover all English letters, the rest of phrases were composed of two pangrams repeated ten times ("the quick brown fox jumped over the lazy dog" and "pack my box with five dozen liquor jugs"), an approached borrowed from closely related prior work optimizing touch screen typing [312]. Participants were able to use basic error correction in the form of character deletion. We implemented space and delete actions via the simple gesture of touching the collider surface with the index fingers: left for delete and right for space. These actions involved simple finger flexion and did not require gross movements of the hand or wrist. In this task we logged WPM, CER (Corrected Error Rate at the character level) [370], and MSD error rate [371]. We collected 560 typed phrases (40 phrases by 14 participants) in this phase of the study.

The typing system in both study phases was identical and is shown in Figure 36. As previously, the basic scene featured the hands and the eight keyboard keys. In addition, we added a transparent object indicating the location of the collider mesh—the touch surface on which the keys were situated. This was important as the classifier based target selection system used in the study (see Section 6.7) was based on touching and releasing this mesh, rather than the actual keys. We used this surface to convey this functionality to participants. The characters assigned to each key were shown directly over it. Each word repetition or phrase task began with display of a large (10 cm) target directly in front of the participant. During this period, the keyboard keys and hands were shown in grey. After tapping this initial target, the trial word or phrase was shown and the hand and keys were re-colored. Text entry could then begin. Key selection was triggered by releasing the thumb from the collider surface and the selected target was then shown by briefly highlighting it in green. In each trial, the word or phrase to be typed was shown in green directly in front of participants at a height of 70cm; entered text was shown in emerald green underneath this. Feedback for entered text was predictive: for each typed word, we showed the most common word or word prefix based on the currently selected sequence of keys. However, we did not implement feedback or commands to view or navigate predicted word sets. Rather we manually adjusted word frequency data to ensure all words in our test sets were displayed in response to entry of the full set of characters they were composed of. This simplified the experimental system and task and matched our intention to focus on text entry performance in the form of entering a sequence of predefined characters (i.e., a phrase from Mackenzie *et al.* [299]'s set) rather than the design and use of a word prediction system and interface. We note that while this manipulation ensured that correct task performance could be achieved by selecting the full and correct key sequence required to enter each word, it also provided typical dynamic word prediction feedback (including display of non-target word

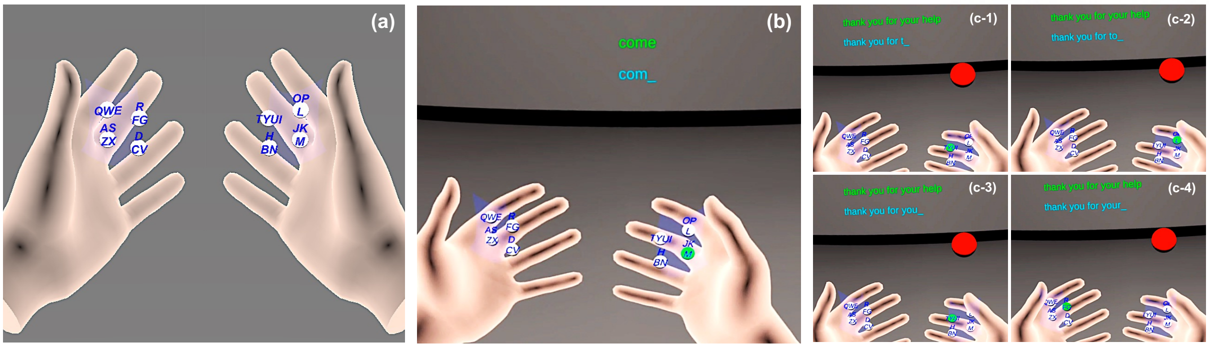


Figure 36: ThumbAir keyboard layout (a), screenshot showing the word repetition task (b), and screenshot showing the phrase typing task (c). In (b) and (c), the trial word or phrase is displayed in green with typed text in emerald green directly underneath. In addition, (c) shows feedback from the word prediction system—although the target word is "your", a higher probability word stem ("to") is displayed after selection of the first two correct keys (top row, right). Characters assigned to each key were always orientated towards the HMD to ensure they could be clearly seen by participants. Key selection was triggered by releasing thumb from surface. The selected target was highlighted in green.

stems that have high probability) during typing (see Figure 36-c).

Procedure. The study followed similar procedures to the prior studies: it took place in same room, and followed similar instruction delivery and enrolment processes, including the use of a video to exemplify instructions and inclusion of a practice session. This session initially mimicked practice in the bigram study and involved free tapping of the thumb targets for a maximum of five minutes. Participants were then exposed to the full text entry system and provided with an explanation of the qwerty-like keyboard mapping and operation of the predictive text feedback. They then practiced the full word repetition task with five words: "my, see, run, come, there". After completing this practice session, the word repetition task began. After completing all 20 words, there was a five minute break. Participants were then introduced to the space and delete gestures and completed a practice session involving typing three phrases (not overlapped with the set in the study) from Mackenzie *et al.* [299]'s set. They then completed the phrase entry section of the study. As with the bigram study, participants held their arms in a comfortable, unsupported pose throughout the study.

Results and Discussion

In the word repetition stage of the study, participants completed a total of 1960 trials. Figure 37 shows the overall mean WPMs for the entered words plotted by repetition and participant. It shows a very substantial improvement between first (12.30 WPM) and second (21.5 WPM) repetition and a more gradual upward trend thereafter. This suggests the task was relatively easy to pick up—after just a single experience. In line with prior work, we estimate expert performance as the mean of the final three repetitions: 27.1 WPM. This is on par with closely related work employing word repetition tasks during thumb to finger typing [262]. The fact performance during in-air thumb typing can match that attained in systems based on directly instrumenting the hands to precisely sense physical contact is a positive result. It suggests that current HMD trackers, when combined with careful keyboard design, are of sufficient quality to support expert text entry at rates previously achievable only by instrumenting the hands and fingers with dedicated worn sensors. Figure 37 shows the MSD error rate gradually decreased over the full set of seven repetitions: from 5.1% in the first repetition to 2.8% in the final repetition. The mean of the final three repetitions is 3.31%. This suggests that, as participants grew more familiar with the task, they were able to achieve speed improvements while maintaining or improving accuracy. We note such downward trends in error rate are common in word repetition tasks [262].

In the phrase typing task, we collected data from a total of 560 trials. We excluded two trials which featured fewer than 75% of the typed phrase from further analysis, leaving data from 558 phrase entries for analysis. Figure 38 shows WPM, CER and MSD error rate data from these trials. Performance differed substantially from that in the word repetition stage. Perhaps most clearly, it was slower. The mean WPM was 12.27WPM (the mean of the final five trials: 13.73WPM), less than half that achieved in the word repetition task (27.1WPM). This suggests that the phrase task achieved our objective of capturing novice performance levels—participants were not able to fully automate their input in the phrase task and resorted to the more traditional hunt-and-peck style typing behaviors common to inexperienced users. However, there is also evidence that they benefited in this task from our work to maintain the familiar QWERTY layout. WPMs over sequentially ordered trials do not show signs of an elbow point—a linear regression on trial sequence and WPM reveals a steady upward trend (slope=0.09, R^2=75.6%). This suggests participants were successfully and incrementally adapting their knowledge of QWERTY to the new format of our layout. Based on the performance levels achieved in the word repetition task and the steady slope we observe here, we suggest further practice would result in further WPM improvements. The WPM data also stacks up well to that reported in prior systems. For example, Fashimpaur *et al.* [267]'s PinchType achieves 12.54 in thumb to finger typing, but requires a high performance external optical tracking system. The benefits conveyed by physical feedback are also clear however. In a thumb typing task on a VR controller surface, Son *et al.* [191] report WPMs of up to 20.56. We suggest that the inherent lack of haptic feedback may increase the challenges of in-air typing for novice users [314].

While WPMs showed signs of steady improvement, error data remained relatively flat over all tri-

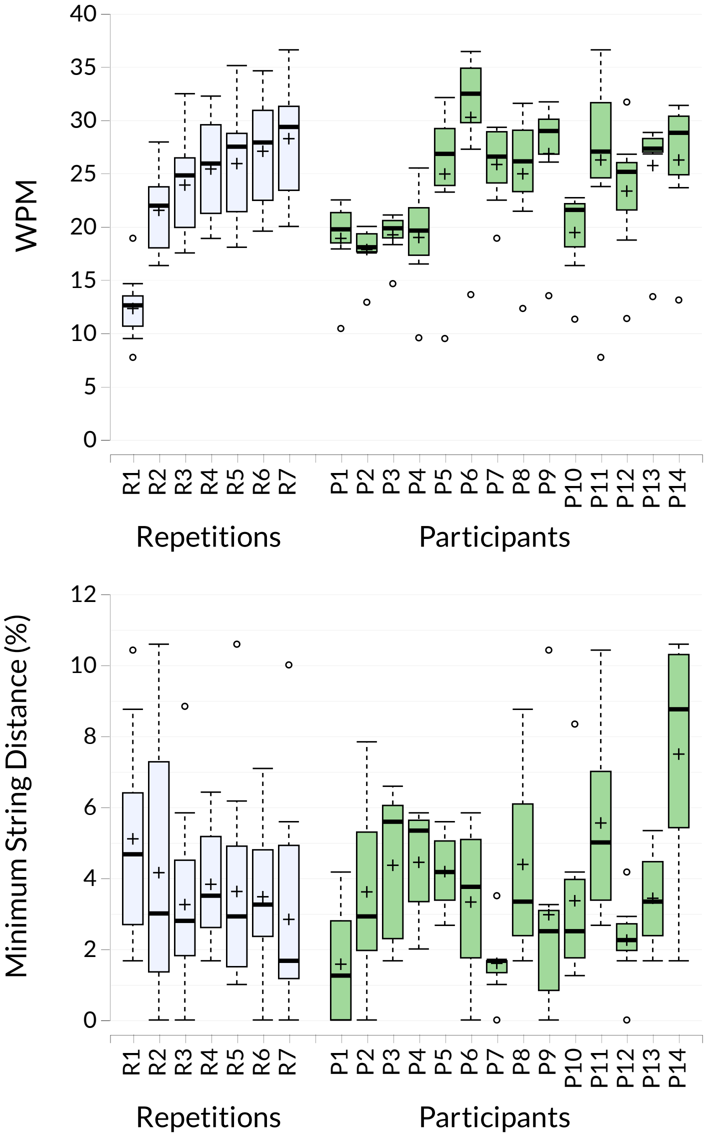


Figure 37: WPM and MSD error rate in the word repetition task. Mean WPM was 27.1 WPM over the last three repetitions while MSD error rate was 3.31% and gradually decreased over the full set of seven repetitions. Per participant data is included to highlight variability in the word repetition task across different users.

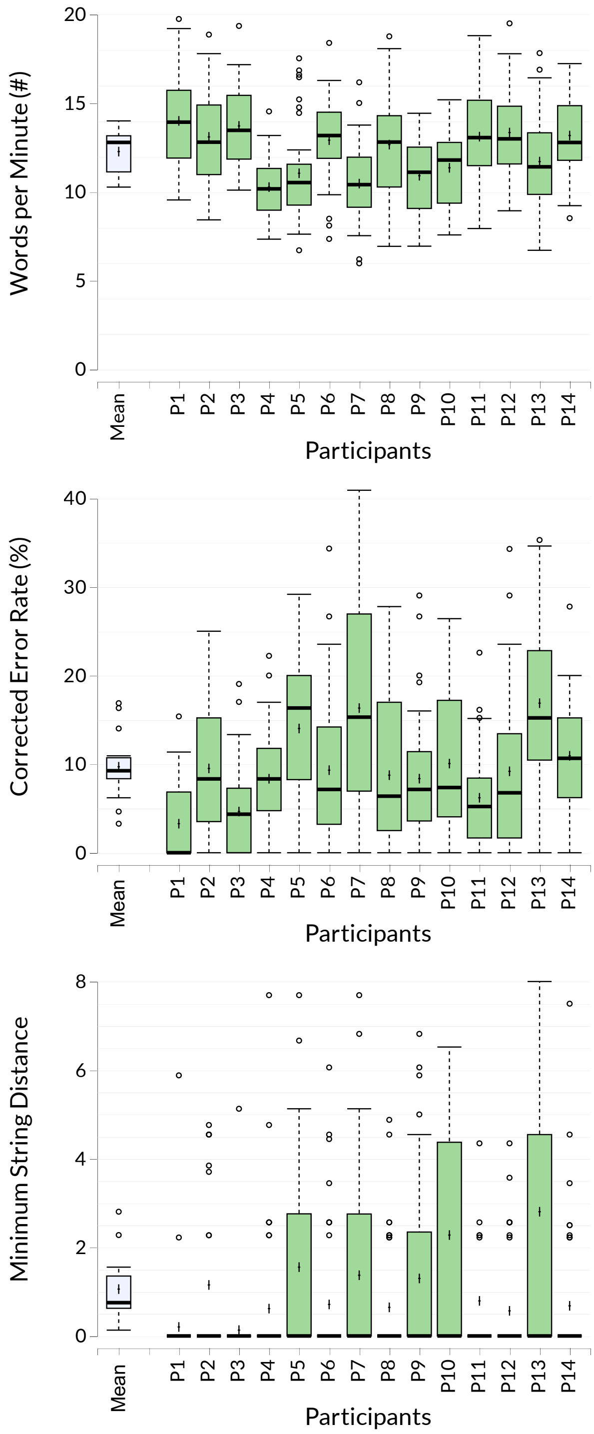


Figure 38: WPM, CER, and MSD error rate in the phrase typing task. Mean WPMs, CER and MSD figures over the whole data were, respectively, 12.27, 9.7% and 1.06%. In addition means from the final five phrases were, respectively, 13.73, 10.1% and 1.15%. Per participant data is included to highlight variability in the phrase typing task across different users.

als, showing no substantial relationship with increased experience in the task (R^2=0.077). In terms of corrected error rate, mean performance was 9.70% (SD=3.92%) and the mean of the final five trials was 10.1%. This is a typical score for CER in novice typing performance. Prior work reports 9.75% CER [191] in text entry on handheld controllers and 13.3% CER [267] with a system based on in-air hand tracking. Furthermore, the uncorrected error rate, which we calculated as the MSD error rate, was low over all the trials (1.06%, SD=0.76%), with many participants recording a median of zero. In addition the mean over the final five trials was 1.15%. This indicates that participants performed the phrase typing task carefully throughout the study, correcting the vast majority of errors they committed.

Beyond these analysis of the two typing tasks, we also explore two further fundamental issues. First, as in prior studies, we examined the relationship between the thumb length and performance in both tasks using Pearson correlation. Thumb lengths in this study were an average of 5.81cm (SD=0.34cm, min=5.0cm, max=6.3cm). Variations in thumb length were not significantly related to any of the metrics from either study task. We once again that conclude thumb length did not impact performance: users can operate our system effectively, regardless of their hand size. Second, we explored the key level accuracy, defined as the proportion of times participants selected the correct key during the phrase typing task, to gain a more detailed understanding of how our target selection classifier performed in practice. This measure combines incorrect target selections due to cognitive errors (i.e., choosing to select the wrong key) with performative (or classifier) errors in which a user attempts to select the correct key, but fails and selects another key instead. Mean key level accuracy was 92.6% (SD=2.4%), or 5.6% lower than the accuracy reported during the LOOCV procedures used to assess the classifiers developed in the second study (98.2%). While some of these additional errors are doubtless due to failures in the classifier, we suggest the majority are likely cognitive in nature and due to, for example, problems interpreting feedback from the word prediction system (which may display non-target word prefixes during partial word entry). Evidence to support this assertion comes from closely related prior work. Examinations of mean key level accuracy for, respectively, two-thumb text entry systems on soft tablet keyboards [337], miniature physical keyboards [339], and soft smartphone keyboards [55] are 94.8%, 93.9% and 89.2%. Our data fall squarely in the middle of this range, suggesting they may represent typical rates for thumb typing in general. This analysis, combined with the generally good performance recorded throughout the study, suggest our classifiers performed more or less as expected during this study. Indeed, the results of this study arguably serve as an effective validation of the performance of our key classifier as it includes both a new participant group and a more complex and naturalistic task.

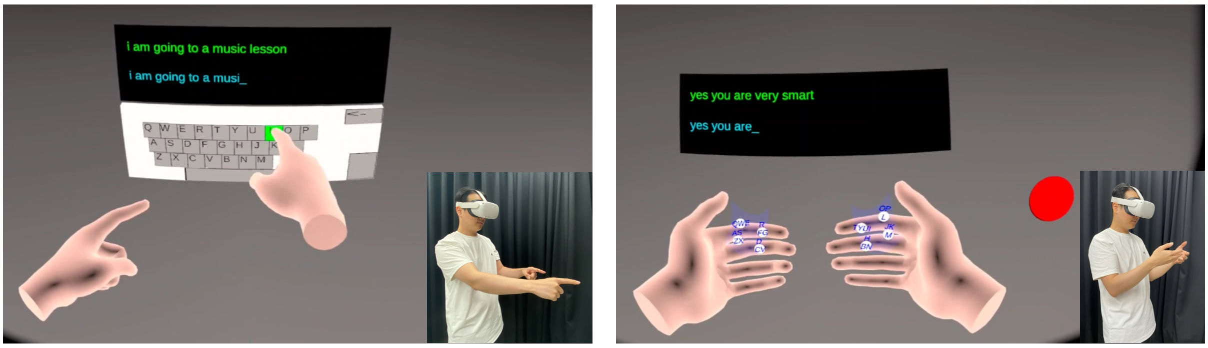


Figure 39: Baseline (left) and ThumbAir (right) interfaces in the comparison study. Participants stood throughout the study.

6.10 Comparison Study: Comparing ThumbAir against a Baseline

Building on the results of the typing study, we conducted a follow-up study that sought to contextualize and ground the performance data we report for ThumbAir. We achieved this by exploring typing performance in a more realistic pose (standing), by capturing salient aspects of participants' experiences and opinions via standard exertion, workload and social acceptability questionnaires as well as through short interviews and, finally, by contrasting ThumbAir's performance against that achieved in a baseline in-air typing system implemented following a design that has been frequently proposed in prior work [45, 88]. This design requires a user to use their the index fingers to strike keys on a QWERTY layout keyboard positioned directly in front of their body. We opted to compare ThumbAir against this design due to both its prevalence in the literature, and the equivalence of its design goals to our own: this design also enables unencumbered, internally tracked, in-air typing input.

Methods

Participants. We recruited twelve new participants (right-handed, mean age 20.33 (SD=2.64), five female, seven male). They self-rated their familiarity with computers (4.75/5.0), smartphones (5.0/5.0), and thumb typing on smartphones (5.0/5.0) as high, but they had relatively low familiarity with HMDs (1.92/5.0, SD=0.51) and very limited experience with typing on HMDs (1.0/5.0). They were all non-native English speakers enrolled in a full time English language degree program. They took approximately 60 minutes to complete this study and each received the equivalent of 15 USD in local currency.

Materials and Study Design. This study contrasted two conditions: ThumbAir and a Baseline. ThumbAir was configured as in the typing study (see Section 6.8), while the Baseline was implemented following a frequently proposed and studied in-air keyboard design [45, 88]. It is illustrated in Figure 39 (left) and involves a QWERTY keyboard located in the space immediately in front of a user; text is entered by striking keys with either index finger. Following recommendations in the literature, we situated this keyboard 35cm below and 50cm forward of the HMD [45, 88]. Keys were 22mm square [88]. We included space and delete keys on the right side of the keyboard, provided a green highlight when contact with a key was made and implemented a simple debounce routine that blocked multiple key selections within 200ms of each other. For both ThumbAir and the Baseline system, we displayed the instructions (i.e., the phrases to type) and entered text just above the input surface—see Figure 39.

All participants completed both conditions in a fully balanced repeated measures arrangement. Each condition was composed of two sequential sessions, a manipulation we included to enable us to examine short term learning rates. Each session featured 15 randomly selected phrases from Mackenzie and Soukoreff's phrase set [299]. We filtered the phrases to ensure they only contain words that appear in ThumbAir's dictionary. In this way, in total, we collected 720 phrases (2 conditions by 2 sessions by 15 phrases by 12 participants). Metrics were both objective (WPM, CER) and also subjective: for each condition we captured standard measures of perceived exertion (BORG CR10 [372]) and workload (NASA TLX [269]). We also used a questionnaire for social acceptability [349] that asks participants to report on the places (e.g., home, street) in which they would be willing to perform a input task and individuals/groups (e.g., alone, partner, colleagues) they would be willing to perform that task in front of. Finally, we asked participants for their comments and opinions on each of the text entry systems.

Procedure. The study used the same setting and equipment, and followed broadly similar procedures, to prior studies reported in this paper. Participants first read study instructions and watched a video demonstrating both typing systems. They did not complete any practice sessions; we sought to assess genuine novice performance. In addition, they conducted all tasks standing, a realistic HMD use posture for a very broad range of scenarios, and one that has previously been used to study performance in the design that inspired our Baseline system [45]. They next completed two sequential sessions, separated by a break of three minutes, for one of the typing systems. They then filled in study questionnaires and gave qualitative comments about their experience. After another five minute break, they completed the same process with the other typing system.

Results and Discussion

We captured 720 typed phrases in this study; five phrases produced by two participants in the Baseline condition were truncated, most likely by inadvertent and premature contact with the enter key. We excluded these five from our analysis and then calculated WPM and CER from the remaining set of 715 phrases. In addition to dividing the data by typing condition, we also divided it by session to enable us

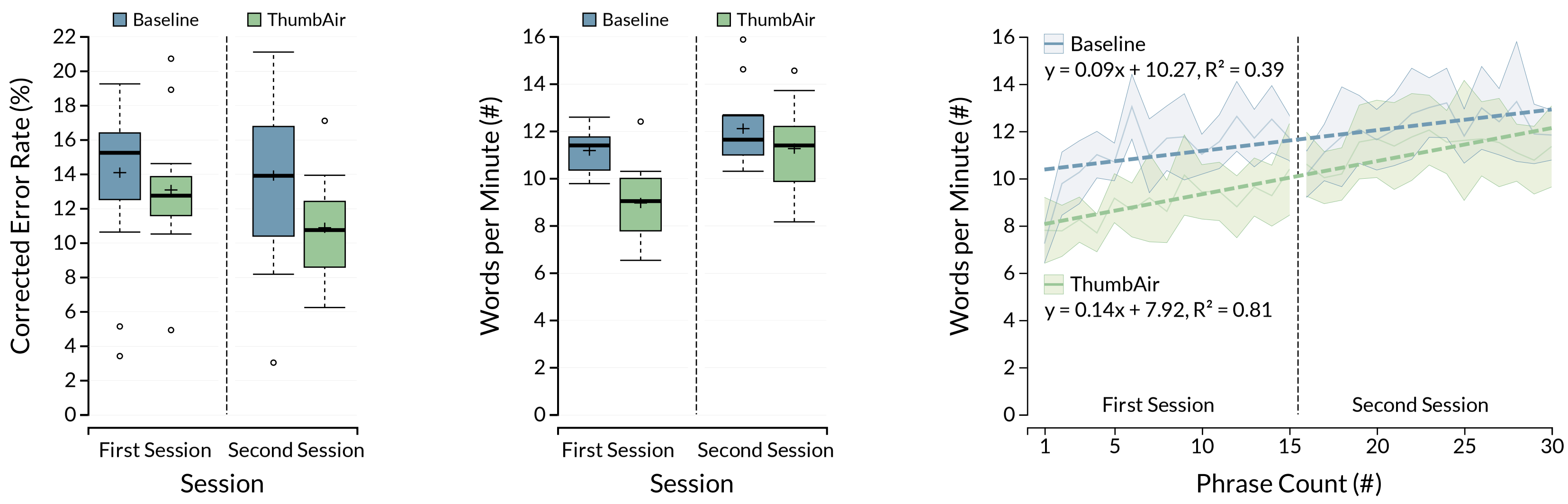


Figure 40: Box plots showing CER (left) and WPM (center) recorded in the comparison study. Right figure shows how WPM varies as more phrases are typed (shaded area represents 95% CI) and includes linear regression lines (which are dotted), equations and fit. Note that high variability for each per-phrase mean is expected and that interpretation of this plot should focus on the regression lines and fit—these reflect the observed trends over the whole duration of the study.

to examine short term learning rates. This data is shown in Figure 40 (center, left) and we analyzed it with a pair of two-way repeated-measures ANOVAs. For WPM, we recorded a full suite of significant differences: the interaction (F(1, 11)=10.545, p=0.008, η_G^2=0.050) and main effects of condition (F(1, 11)=9.377, p=0.011, η_G^2=0.206) and session (F(1, 11)=70.960, p<0.001, η_G^2=0.224). The interaction effect indicates that performance changes between the two sequential sessions differed between the two typing conditions. To explore these trends, we ran paired t-tests on the data from each session individually. The results show that while Baseline leads to a significantly higher WPM in the first session (p=0.001), there is no significant difference in second session data (p=0.167). This suggests that the main effect of condition is primarily due to the improved performance recorded for Baseline in the first session and that the main effect of session mainly reflects the performance improvement observed for ThumbAir in the second session. To explore this change in more detail, we plotted WPM data by phrase entered (see Figure 40 right) and fit linear regression lines: ThumbAir shows a steeper slope (0.1402 vs 0.0877 for Baseline) and good fit (r^2 = 0.81). While per-phrase means show, as expected, high variability, the trends over the full set remain relatively clear and consistent. We conclude that participants initially struggled due to unfamiliarity with ThumbAir's layout. However, they were able to rapidly pick it up: after typing 15 phrases, performance was not significantly different to that attained using a more standard QWERTY design. This strong performance likely reflects our emphasis on maintaining close QWERTY similarity during the design of ThumbAir's keyboard layout. We also note that basic performance levels recorded are representative of those in the literature: our mean Baseline WPM of 12.09 WPM is 20% higher than that recorded for the keyboard design and implementation it is based on [45]. This boost that may be due to a wide variety of factors, most likely to improvements in finger tracking technology. In terms of CER, no significant differences were recorded in either interaction or main effect. In addition, we note these CER figures are typical of those reported in closely related studies of HMD based bare hand text entry (7.6% to 13.3%) [45, 267], but somewhat elevated when compared more mature platforms such as smartphones (approximately 5%) [373]. Finally, MSD error rates were low in both ThumbAir (Mean=0.34%, SD=0.30%) and Baseline (Mean=0.15%, SD=0.27%). This indicates that participants corrected the vast majority of typing errors that occurred in the study.

Subjective feedback data for Borg CR10 and NASA TLX are shown in Table 10. We analyzed this data with Bonferroni corrected paired t-tests, ultimately uncovering significant differences in two measures. Participants reported higher levels of perceived exertion at the shoulder (*p-value* = 0.049) in the Baseline condition. Specifically, they reported that the Baseline condition led to moderate levels of perceived shoulder exertion versus very weak levels during use of ThumbAir. In addition, they reported increased physical demand in the Baseline condition (*p-value* = 0.045). These data suggest that the

Table 10: Mean subjective data recorded in the baseline study for NASA TLX and Borg CR10. Figures in brackets show SD.

	Nasa TLX						Borg CR10				
	Mental	**Physical**	**Temporal**	**Perf.**	**Effort**	**Frust.**	**Shoulder**	**Arm**	**Forearm**	**Wrist**	**Hand**
Baseline	2.83 (2.1)	6.17 (3.9)	4.92 (4.6)	6.5 (4.4)	7.25 (5.3)	4.58 (5.8)	2.79 (2.2)	1.71 (1.9)	1.5 (2)	1.46 (1.9)	1.5 (2.2)
ThumbAir	3.33 (2.3)	2.58 (2.6)	4.17 (4.2)	6.42 (5.2)	7.42 (5.5)	4.42 (4.5)	1.08 (0.8)	0.83 (0.7)	0.75 (0.7)	0.83 (0.9)	2.04 (1.6)

Baseline condition suffered from "gorilla arm" [75, 348] but that the more relaxed posture of ThumbAir, with the hands lowered and the arms close to the body, largely avoided this. This result is in line with prior work in which such poses have been suggested as targets for reducing exertion levels during input tasks [75]. In addition, we analyzed the social acceptance questionnaire by the simple expedient of summing the number of locations and groups participants selected. This creates counts for each condition of the number of locations in which participants would perform input, and the number of groups they would perform it in front of. We then tested these counts using paired t-tests, recording significant differences for both locations (p=0.002) and groups (p=0.001). Specifically, while participants indicated they would use the Baseline condition in a mean of 2.33 locations (SD 1.23)), they indicated they would use ThumbAir in a mean of 4.5 locations (SD 1.78). For groups this data was 4.08 (SD 1.08) for Baseline and 5.42 (SD 0.79) for ThumbAir. These results indicate that participants felt it would be socially acceptable to use ThumbAir in a greater number of real world contexts. Qualitative recorded at the end of the study shed light on this issue. With ThumbAir six participants noted the input actions resembled current practices, such as use of a smartphone. This familiarity increased their sense the input would be acceptable in social situations. Additionally five described the small inputs motions as discreet or hard to observe, factors that were also viewed to make them more acceptable. These findings are in line with prior work [217]. In contrast for Baseline, five participants noted the input would be "strange", "weird" or "odd", and four further suggested this problem is due to the fact the required finger motion is "uncommon" or otherwise unintelligible to bystanders. More positively, two noted that they felt that it would be more socially acceptable if they provided an explanation for their behavior, or readily envisaged a future when this type of input becomes socially acceptable due to its prevalence.

6.11 Discussion and Conclusion

This paper describes the process of designing an in-air thumb typing keyboard for the inside-out hand tracker available on a standalone off-the-shelf HMD. It targets a casual use-scenario in which an AR or VR HMD user lacks external input devices but wishes to type on their headset to meet various needs, such as sporadic messaging [374], note taking [375], authentication [376] or other personal information management tasks [377]. It takes account of factors such as comfort [372], workload [269] and, reflecting the fact that input may take place in a range of settings, the social acceptability [349] of the input by focusing on small and discreet input actions [217]. Our design process was data-driven and empirical. First, we sampled performance over a broad range of target locations to understand viable regions in which to locate keys and how users would move towards them. Based on the results, we designed an arrangement of four targets for each hand in ready reach of thumb when it is rapidly articulated with gross rotations of the metacarpal joint. These movements are comfortable, quick and relatively large scale. As such, they are easy to track (comfortably within the HMD's reported finger tracking accuracy [39]) and also easy to execute in the absence of haptic feedback (from the user's point of view). We then used these targets in a follow up study examining bigram performance—tapping targets in sequential pairs. We used the data from this study to refine target locations and design a novel target selection system (a classifier based on finger intersections with a virtual surface) that achieves a high level of accuracy (98.2%). In addition we used this data in a computational process to inform the design of a novel keyboard layout. The layout we ultimately selected emphasized a high level of QWERTY similarity and low levels of ambiguity with respect to entered words. In a third study, we evaluated the target locations, selection system and keyboard layout in word repetition and phrase typing tasks, ultimately achieving final WPMs of 27.1 and 13.73, respectively. We believe this level of performance, achieved on an on-the-shelf HMD hand tracker, represents a meaningful achievement. Both word repetition and phrase entry speeds approached or exceeded those previously attained in closely related prior work that relies on dedicated worn hardware [180, 260, 262] or external optical tracking systems [87, 191, 267]. We ground this data with a final study that compares performance with ThumbAir against a baseline in-air typing design in a phrase-entry task. In addition to objective measures we collect qualitative measures of perceived exertion, workload and social acceptability. The results indicate that while objective performance between the two designs is broadly similar, ThumbAir offers advantages in terms of the qualitative metrics: it induces significantly lower levels of perceived exertion and physical demand and is rated as acceptable for use in a wider range of situations that the baseline. These are key qualities that match our goal of creating a typing system to support a range of casual HMD use scenarios. In sum, our work demonstrates that rapid, accurate, comfortable and socially acceptable in-air thumb typing performance can be achieved using currently available consumer hardware.

Despite these positive conclusions, there are a number of limitations to this work and, similarly, a wide range of directions for future study. Perhaps most foundationally, while the metacarpal flexes we study represent the dominant thumb motion performed by our users, it is simple and arguably fails to take advantage of the high dexterity, and degrees-of-freedom, available to the thumb [354]. Future work might further investigate the potential for detailed in-air thumb motions involving lateral movements, flexes of inter-phalangeal joint, or more complex compound actions such as ”pokes”. It would also be important to extend the studies reported in this paper with work that examines a range of more diverse settings. We used a seated pose to avoid fatigue during our first three prolonged studies and a standing pose in our shorter fourth study. A valuable complement to this lab-based work would be an exploration of in-air typing performance while walking [262]). These more mobile scenarios are commonplace in HMD applications and particularly relevant for AR: our system needs be assessed in them. The performance we observe while seated and standing may not fully generalize to more dynamic poses and settings. It would also be interesting to explore the impact of customization on performance. While we note our studies showed no variations that could be explained by hand size, and we used a fixed virtual hand size throughout our work, mismatches between a users' body and its representation in AR/VR are known to exert a wide range of effects [378]. It would be interesting to evaluate whether using scaling or other types of customization to minimize mismatches between a user's real and virtual hands can boost performance in the types of text entry tasks studied here. In addition, while the studies in the paper target various aspects of typing performance, we did not conduct a prolonged, multi-session study [191, 337]. Data about in-air thumb-typing from such a study would closely complement the results we report here.

Finally, we also note that while this paper documents the design and evaluation of a keyboard layout, a full typing system is a more sophisticated artifact. Immediate future work will need to redesign or integrate a wide range of features into our system to truly enable users to type. For example, as our keyboard is ambiguous, a interface for word selection needs to be defined. As the proposed layout shows a high Top 2% [161] score of 98.1%, a simple system featuring two word choices is likely viable. These limited selections could therefore be activated by simple hand gestures, such as flexing all fingers (or making a fist) with the left or right hands. However, it is also worth noting that word selection and prediction systems do not universally result in improved text entry speeds [379]. Although a two choice design may be practical for ThumbAir, it would also need be highly accurate in order to allow users to achieve greater WPMs. The cost of errors in word prediction systems are high. Alternatively, an effective design might integrate advanced feedback, such as display of confidence in its predictions [379] which may allow users to better recover from errors. Finally, we note a realistic keyboard will also require mode switching to support entry of numbers or symbols. This might be achieved by a palm open gesture with one hand which could be overlaid with a numeric keypad or other symbol entry system [380]. By continuing to develop our system to achieve full text entry functionality we hope to enable users of current AR and VR HMDs to use the hand tracking built into their existing headsets to type comfortably, discreetly, rapidly and accurately.

6.12 Appendix.

Targets and Key Locations

In the range of motion study, we chose 18 target locations above each hand based on users' ability to select them rapidly and accurately with their thumbs. We used these locations to derive a paraboloid surface and an arrangement of four key locations, which we use in bigram study. We present the locations of these targets here. Table 11 shows coordinates for the 18 selected targets. Table 12 shows the coordinates of the refined key locations, which we use in text entry study. The refined key locations were derived with the mean target centroid position by calculating all ellipses from the collision maps for each target.

Table 11: Coordinates of the 18 targets on the left hand used to derive the paraboloid surface and four key locations used in bigram study. Units are in mm and all coordinates are relative to the center of the palm (as returned by the Oculus SDK). Locations were mirrored for the right hand. The table also depicts the normalized scores of the three metrics (success rate, touch time, and wrist angle) used to calculate the target performance scores (see Section 6.6).

Target ID	1	2	41	42	43	81	82	9	49	50	89	90	17	57	58	97	98	65
X Coordinate	6.25	18.75	6.25	18.75	31.25	6.25	18.75	6.25	6.25	18.75	6.25	18.75	6.25	6.25	18.75	6.25	18.75	6.25
Y Coordinate	-20	-20	-32.5	-32.5	-32.5	-45	-45	-20	-32.5	-32.5	-45	-45	-20	-32.5	-32.5	-45	-45	-32.5
Z Coordinate	-25	-25	-25	-25	-25	-25	-25	-12.5	-12.5	-12.5	-12.5	-12.5	0	0	0	0	0	12.5
Success rate	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Touch time	0.91	0.80	1.00	0.95	0.74	0.92	0.94	0.81	0.95	0.88	0.88	0.91	0.83	0.90	0.87	0.88	0.87	0.87
Wrist angle	0.93	0.86	1.00	1.00	0.87	0.91	0.94	0.86	0.95	0.93	0.85	0.93	0.78	0.84	0.84	0.80	0.87	0.80

Table 12: Coordinates of the refined key locations used in text entry study for the left hand. Units are in mm and all coordinates are relative to the center of the palm (as returned by the Oculus SDK). Locations were mirrored for the right hand.

Key Name	Radial Close	Radial Far	Ulnar Close	Ulnar Far
X Coordinate	28.89	26.96	15.65	12.97
Y Coordinate	-20.73	-42.29	-20.73	-42.29
Z Coordinate	-28.84	-21.17	-16.78	-10.05

VII Conclusion

7.1 Summary

This thesis aimed to address the constrained interaction experience for the resource-limited wearable devices. Various prior works tackled this problem by developing complex interaction techniques or extending the limited design space to the alternative spaces with additional sensors such as other parts of devices, around-devices, physical hand-held devices, or on-body. To address this problem, this thesis argues that the dexterity of finger movements will create novel interaction techniques with high popularity and familiarity by using the built-in sensors in the wearable devices. The four scenarios for smartwatches and HMDs supported the general claim of this thesis that understanding how users move their fingers during input will enable increases in the expressiveness of the interaction techniques we can create for resource-limited wearable devices.

The first scenario examined the comfort range of angle-based touch input on the small touchscreen of smartwatches. Two empirical studies explored the static and dynamic angle-based touch inputs on the design factors in fingers, angles, and touching poses. The touching poses included viable finger regions for generating angle variation on the smartwatches such as the flat and side of the index finger and pair of the index and middle fingers. The results of two studies based on both objective performance and subjective comfort ratings yielded the comfortable angles combined with the metrics of input time and accuracy for each touching pose. In particular, the classifiers for recognizing the finger regions achieved an accuracy of 95.9% and 95.6% with the raw touch data of contact areas during the static and dynamic touches, respectively. Based on the results of two empirical studies, the data-based design guideline was presented to recommend comfortable angle range, target size, gesture type, distinctive touches, and finger flexion. Finally, the potential interaction techniques using the comfortable angles and touch-region classifiers showed the feasibility of angle-based touch input techniques that can improve the expressiveness of touch input on the small touchscreen of smartwatches.

The second scenario explored the touch profiles generated by the unique input context of smartwatches. This scenario aimed to address the distinctiveness of contact areas of each finger (thumb, index, and middle fingers) will enable finger identification to recognize which of a finger is responsible for a touch. The finger identification can assign different functions to each finger to extend the limited input space. The first empirical study collected raw touch images through tapping and swiping tasks with natural and exaggerated touch conditions. The exaggerated input condition aimed at maximizing the distinctiveness of touch profiles. The finger classifiers were constructed with touch profiles (e.g., angle, major and minor axis length, and eccentricity of the contact area) extracted from the raw touch images. The classifiers with the touch profiles from the exaggerated touch condition achieved an accuracy of up to 98% due to the maximized distinctiveness. Meanwhile, the classifiers under natural touch conditions remained challenging to distinguish the index and middle fingers, yielding an accuracy of up to 93%. Two types of keyboards with two or three letters on each key were built to evaluate these classifiers. The keyboard applications based on the finger identification systems demonstrated the contributions of this thesis to increases in the expressiveness of touch input on the small touchscreen of smartwatches.

Next, to support the general claim of this thesis with various wearable scenarios, the other two scenarios explored the properties of in-air finger movements for HMDs. The third scenario examined the characteristics of in-air finger stroking during unconstrained in-air typing tasks. The fast sequential inputs with prevalent finger stroking skills will enable interaction techniques that can improve the input expressiveness for the small form factor wearable devices. The unconstrained typing tasks exhibited variations in a strategy of in-air finger stroking with different patterns of hand movement: either moving the entire hand and finger together (HAND strategy) or only the finger (FINGER strategy). Then, through the analysis of finger kinematics, the unconstrained in-air finger stroking showed faster velocity (762mm/s), shorter time duration (322ms), and lower amplitude (49mm) than in constrained typing tasks with specific behavior instruction [47]. The correlated movements were also analyzed to understand the unintended movements of other fingers during the intentional movements of a single finger. The overall amplitude ratio of unconstrained in-air finger stroking (46.9-83.5%) was higher than in constrained typing (27.5-60.6%) [47]. The analysis of the time interval between the sequential finger stroking inputs showed an average time overlap of 43ms. The last analysis presented the amplitude, velocity, and distribution of individual in-air keys. The results of various analysis methods showed the complexity of fast sequential finger stroking in the air. Based on the results, the data-based recommendations were suggested for future researchers and developers to design input interfaces with in-air finger stroking.

The last scenario explores the properties of thumb movements in the air to develop an in-air typing system using a built-in hand tracking system in HMDs. To understand the thumb movement for input, the first study examined the viable touch locations by considering the efficient range of in-air thumb touch motion having high accuracy and speed. The locations of key were determined to avoid physical occlusion by fitting a paraboloid surface that crosses the viable touch locations. The second study examined the performance of sequential input for the paired keys. Based on the results, a classifier was constructed to determine the intended target on the paraboloid surface. The visual cues for the final keys were re-arranged on the centroid of each distribution. The results of two studies showed that the number of viable touch locations is limited to four per hand. To improve the input expressiveness in this limitation, a computational design process reviewed the potential keyboard layouts to assign the letters to a total of eight keys. The final two studies evaluated the performance of in-air thumb typing using the ambiguous keyboard with word prediction system, showing final WPMs of 27.1 and 13.73 in word repetition and phrase entry tasks respectively. The subjective ratings on the experience of the designed in-air thumb typing system showed a very weak level of perceived exertion with the relaxed postures that can avoid "gorilla arm effect" [75]. The discreet thumb movements during the in-air thumb typing showed higher acceptability for social situations than in the baseline keyboard. The ambiguous keyboard design with fast, comfortable, and subtle thumb motions in the air showed the feasibility of in-air typing that enables increases in input expressiveness for resource-limited wearable devices.

7.2 Design Considerations on finger input techniques for wearables

This section provides design considerations for future researchers and developers to design interaction techniques based on the properties of finger motions that will improve the expressiveness of dexterous inputs for resource-limited wearable devices. The discussion of this section includes properties of finger movements on various perspectives of design factors: contact area on the surface, finger motions in the air, usability, data collection, and novel keyboard interfaces.

Contact areas on the surface

The raw touch images generated from the reliable capacitive sensors have high resolution and sensitivity that are sufficient to detect finger movements during touch input [94]. The contact area on the surface could provide plentiful information about the touch such as angle, major and minor axis length, and eccentricity. This touch profile is a promising feature to improve the expressiveness of touch input by estimating the finger status or recognizing touching fingers.

The touch profile can infer a contacted finger area on the touchscreen to estimate the touching finger poses or provide multiple input options in the same touch area. For example, Chapter III, using the contact areas by flat and side of the index finger, presented the region-aware classifiers that achieved an accuracy of 95.9% and 95.6% in inferring the contacted finger region during static and dynamic touches respectively. The high accuracy indicates that the classification of finger regions can be reliable across users and various touch tasks. The unique oval shapes of fingers would enhance the distinguishability of finger regions on hard surfaces. Meanwhile, the unique finger input context by the wearable form factor could generate distinctive touch profiles of each finger during touch input on the surfaces of the device.

The distinctiveness of touch profiles can lead to the development of finger identification systems. For example, Chapter IV revealed that each finger can form distinguishable contact areas on the touchscreen of smartwatches. The collected touch profiles of each finger during tapping and swiping tasks were used to develop the finger identification system that showed an accuracy of up to 93% and 98% in natural and exaggerated touch poses respectively. The relatively low accuracy in the natural touch poses suggests that the touch profiles of the index and middle fingers are less distinctive, while the exaggerated touch poses can generate more distinctive touch profiles on the thumb, index, and middle fingers even at the cost of performing them in touch time. In addition, the static finger touch provided more distinctive touch profile than in dynamic finger motion on the surface. The moving fingers with more strong force may generate variation on contact area. In sum, the various scenarios in this thesis showed the potentiality of touch profiles that enable increases in the expressiveness of touch input. I recommend future researchers and developers carefully consider the contact regions in fingers, the contact area on the touchscreen, and the unique touch poses by wearable form factors while designing interfaces based on the touch profiles.

Finger motions in the air

The bare-hand interaction can enable dexterous and familiar interaction techniques like manipulating physical things with hands in real environments. However, the in-air finger inputs would generate more complex finger motions due to the hand pose in specific design spaces and the lack of haptic feedback. To achieve reliable input interfaces, it is fundamentally required to understand various perspectives of finger motions.

Among various in-air gestures, in-air finger stroking has simple, fast, and discrete finger motions that will enable fast sequential inputs such as text entry. The results of Chapter V contributed to the in-depth analysis of the in-air finger stroking. The kinematics of a single finger for in-air stroking showed faster (762mm/s), shorter (322ms), and lower amplitude (49mm), while other fingers simultaneously had correlated movements with an amplitude ratio of 46.9-83.5%. The sequences of in-air finger stroking showed a time overlap of 43ms. The high unintended movements of other fingers could complicate the detection of intended finger input. In addition, the not inconsiderable overlaps in sequential inputs (23.8% of the extension phase of the previous input) could blur the boundaries of individual inputs. On the other hand, the touch input in the air needs to carefully design the arrangement of virtual targets by considering finger motions. For example, Chapter VI revealed an efficient range of thumb touch motions in performance to specify viable locations of virtual targets for thumb touch input. Through the performance of sequential input tasks, the final target locations were selected to lead the accurate and fast thumb motions. The final design space included four targets per hand to avoid physical occlusion. The limited number of targets may be due to the immature hand tracking systems of commercial HMDs that are insufficient to perform subtle bare-hand manipulation in the design space. These findings from the two in-air finger input scenarios in this thesis will be helpful for future researchers and developers to build in-air finger input interfaces with accurate and fast sequential inputs for HMDs or other wearables that have limited input space on devices.

Usability based on properties of finger movements

Interaction techniques in wearable devices frequently used in daily life should carefully consider the usability along with input performance. This thesis explored the three factors that can affect overall usability during input: comfort, perceived exertion, and social acceptance.

The touch input techniques based on angles can have the limited range of comfortable touches by requiring users to make specific postures before or during touch. Chapter III in this thesis addressed the comfort during angle-based touch input on the touchscreen of smartwatches. The side of index finger enabled comfortable input over a large range of angles (between 60° and 180°) due to effortless finger flexion. The comfortable range of angles in touch with the flat region of the index finger was between 90° and 180°. The unique input context of smartwatches limited the comfortable angles (between 180° and 270°) on angle-based touches using the pair of index and middle fingers, requiring twisting of the wrist when making touches beyond a certain angle. In spite of the limited angles, the pair touches recorded lower touch time than other side and flat touches. The high familiarity of pair touches with index and middle fingers may lead this results. The comfort ratings were moderately correlated with preparation times ($r=0.546$, $p<0.001$) due to the difficulty of making the touching poses for uncomfortable angles. In the comfort range of angle-based touches, the minimum angular width of the target would be 20° and 15° on static and dynamic touches, respectively. Based on the findings, I recommend that future researchers and developers should design interaction techniques by combining and contrasting the performance with the usability that will increase the reliability of interaction techniques.

In addition, Chapter VI provided evidence that the design space of in-air finger input can impact the perceived exertion and workload. During the in-air thumb typing proposed in this thesis, the relaxed input pose with the hands lowered, the arms close to the body, and small finger motions may result in weak perceived shoulder exertion that avoided "gorilla arm effect" [75]. In terms of social acceptability, the results of Chapter VI showed that discreet and subtle thumb motion in the air was preferred in various social situations than large index poking gestures. The empirical findings on the perceived exertion, workload, and social acceptability will be helpful for future researchers and developers to design interaction interfaces for wearable devices widely used in daily life and public spaces.

Novel interaction techniques

Finger region-aware system with touch profiles: Angle-based touch techniques with finger region-aware systems can generate two options in rotation gesture input with the side and flat of fingers. For example, Chapter III demonstrated the usefulness of angle-based touches and region-aware techniques on smartwatches by suggesting three demo applications: the message, map, and photo apps. An interview study was conducted to evaluate their comfort, efficiency, and learnability. All participants remarked on the value of the region-aware input technique that increased the range of input available on the small touchscreen of smartwatches. Participants also mentioned high learnability due to their consistency and predictability. The comments on the example apps clearly showed the feasibility of the integration of two techniques that enabled increases in the expressiveness of touch input on resource-limited wearable devices. In addition, regardless of the device form factor, the region-aware interaction techniques with angles could be applied to the surface of other wearable devices by considering usability on comfort, since the contact regions on hard surfaces could be independent of the input contexts of wearables.

Virtual keyboards with touch profiles: Finger identification systems are a promising approach to broaden the input range available for enabling high bandwidth tasks on resource-limited wearable devices. To access the high bandwidth tasks with text entry, Chapter IV presented smartwatch virtual keyboards using finger identification systems based on distinctive touch profiles of each finger in natural and exaggerated touch poses. Each keyboard type needed 12 and 16 keys for touch typing by assigning three (Tri-Type) or two (Di-Type) letters to each key, respectively. Phrase typing tasks yielded a mean of 7.53 and 8.08 WPM for Tri-Type and Di-Type keyboards respectively. The typing speed was somewhat limited due to too short exposure to be familiar with the novel interaction technique but still showed reasonable performance, compared to the prior work (10.8 WPM [68]) using finger identification system with additional finger-mounted wearables. The minor figures of up to 1.7% on the wrong-key touch errors indicated that the key size of 6.75mm by 5.75mm or 7.85mm with 0.6mm spacing was sufficient to perform the complex tasks with both natural and exaggerated touch poses. An interesting finding was that typing with two fingers showed better performance in speed and wrong-finger error (8.08 WPM and 8.3%) than with three fingers (7.53 WPM and 10%). It may indicate that the small variation on fingers could lead to short execution time and high accuracy on finger identification systems. The example keyboards showed the feasibility of finger identification systems that will enable high bandwidth tasks requiring the wide input range available on resource-limited wearable devices.

In-air thumb keyboard: The hand tracking systems can enable HMDs to adopt prevalent finger stroking input for typing tasks. This thesis sought to develop in-air typing systems by exploring finger motions in the air. Chapter V investigated the properties of in-air finger motions during unconstrained typing. However, the complex finger motions in the air remained challenging to accurately recognize the intended finger stroking with a low accuracy of up to 86.9%. Meanwhile, Chapter VI introduced a novel in-air typing system with prevalent thumb touch skills to virtual keys. However, the viable location of in-air thumb touch input was limited to eight for both hands. To address this limitation, an ambiguous keyboard layout with a word prediction system [262] was designed by mapping the letters to the eight keys. The evaluation studies for the in-air thumb typing system yielded 27.1 WPM in the word repetition task to mimic the expert performance and 13.73 WPM in the traditional phrase typing task to assess the novice performance. In particular, during the word repetition task, the substantial improvements in typing speed after the single trial (12.30 WPM to 21.5 WPM) indicated that the in-air thumb typing system was easy and efficient in selecting keys. In addition, the low error rate of 2.8% in the final repetition indicated that the experts will be able to accomplish fast and accurate typing. Finally, the in-air thumb typing with a relaxed posture and discreet finger motions led to weak perceived exertion, low workload, and high social acceptability, compared to the prior in-air typing method [45].

Data Collection

In this thesis, the empirically collected data from rigorous study designs can contribute to providing a ground-truth for future researchers and developers to build novel interaction techniques for wearable devices. Chapter III provided specific angles that can enable comfortable angle-based touch input on smartwatches. The angles on various finger regions and poses can be a guideline to design static or dynamic touch input techniques based on rotation. Chapter V collected a total of 25,932 manually labeled finger motions to each in-air key to explore the various characteristics of unconstrained in-air typing. Based on the massive data, various aspects of individual in-air keys were summarized for the invisible in-air keyboard with Qwerty layout, including amplitude, velocity, direction, and three-dimensional fingertip distribution. The geometric features of each in-air key led to a conceptualization of a potential keyboard layout in the air. The details of individual in-air keys and potential keyboard layout will be a ground-truth to building in-air typing systems using finger motions. Chapter VI presented the individual in-air touch performance (e.g., touch success rate, touch time, and global wrist angle changes during touch) to a total of 80 virtual targets in the space above palm (a volume of 6cm by 4.75cm by 4.75cm). Based on the performance metrics, all viable locations for thumb touch input were described. This data will be helpful to develop in-air thumb touch techniques for virtual objects.

7.3 Limitation and Future Work

The interaction techniques proposed in this thesis introduced a variety of promising approaches that can enable increases in the expressiveness of finger input on smartwatches and HMDs. However, the limitations in the scope of the current studies leave future works that will enhance the finger interaction techniques. This section discusses the limitations and immediate future works to extend the current works of this thesis, such as sensing technologies of touchscreen and hand tracking system, wearable types, properties of finger motions, physical occlusion, detection and recognition of in-air finger motions, usability, and mobile environments.

Sensing technologies

In this thesis, the quality of sensing technologies in commercial wearable devices would have affected the overall development process of finger interaction techniques. For example, Chapter IV collected contact areas on a small touchscreen with a total of 64 capacitive cells in a 33mm x 33mm area. The relatively large size of a single cell (approximately 4mm by 4mm) remained challenging to detect the fine details of the contact area with the tip, pad, or side of fingers. The estimated touch profiles by fitting an ellipse with a few cells may generate more overlaps in classification between the index and middle fingers of natural touch poses. Further investigation using a touchscreen with more high resolution is the next step to reveal novel features in contact areas during natural touch inputs. Another future avenue for this issue is to employ image recognition algorithms such as Artificial Neural Networks (ANNs) for the raw or pre-processed touch images. Indeed, future researchers and developers should consider the computational cost of classification models on the limited resources of wearable devices. In addition, the sensing frequency of the touchscreen was limited to around 33Hz which supported a mean of 6.6 frames (a mean touch time of approximately 200ms) for each screen touch with natural poses. Since results showed a higher classification accuracy (2.8% more) with the data of short touches, the more high frequency may lead to more accurate classification rates by capturing fine details of the short finger movements.

The quality of hand tracking systems would have affected the results of the observation study and interface design. For example, the frequency of the hand tracking system in Chapter V was 60Hz which supported the detection of the individual in-air finger stroking with approximately 8.6 frames (143.3ms) and 11.2 frames (187.2ms) for flexion and extension phases, respectively. Despite the in-depth analysis for in-air finger stroking, the observation studies may not have captured the entire finger motions due to the time gap of 16.7ms between two frames (11.7% of the flexion phase). Immediate future work is the validation studies of the suggested feature sets with advanced optical hand tracking systems with high frequency. In Chapter VI, the relatively low sensitivity (9.6° joint angle accuracy [39]) of built-in hand tracking system in commercial HMD constrained the detection of subtle in-air thumb movements. It impacted the interface design of the in-air thumb typing system by allowing only four keys per hand in the efficient range of thumb touch input. I believe that the advancement of hand tracking technologies in near future will enable the in-air touch input techniques to have a high input resolution in this design space.

Types of wearable devices

This thesis mainly focused on exploring the properties of finger input motions suitable for the form factors of popular wearable devices: smartwatches and VR HMDs. Specifically, Chapter III and IV explored the properties of touch input motions on the touchscreen of smartwatches, while Chapter V and VI explored the properties of in-air finger motions for the HMDs. However, the different input contexts of other wearable devices may generate novel finger motions due to their form factors. For example, the location of wearable devices can produce their unique touch input poses on the surface, or the touch input for wearable devices near face basically require an eyes-free interaction that needs to rely on the proprioceptive sense [85]. In addition, the sensor locations for the input of wearable devices may variate the finger input motions. Especially, the bare-hand input motions in the air can be designed differently depending on the location and detection range of hand tracking system built in wearables or additional devices. For example, Chapter V observed downward finger motions by using a Leap motion device placed below hands, while Chapter VI examined the thumb motions with the specific palms-up hand postures due to the inside-out HMD-based tracking system. Therefore, a next step is the investigation of the special properties of finger input motions on various form-factors of other wearable devices, such as armbands, clothes, belts, necklace, or earring.

The attributes of the surfaces on wearable devices can also variate the properties of finger input motions. In terms of the on-surface input, this thesis focused on the properties of touch motions on the rigid and flat surface of the touchscreen, which can extract the touch profiles of fully contacted fingers. However, touch input on soft or patterned surfaces such as cloth-type wearables may generate different finger motions while performing static and dynamic inputs depending on the surface deformation and friction. Also, the surface flatness by the inherent curves or edges in wearable form factor can affect the finger input by constraining the continuous motions or comfortable input ranges. Immediate future work is the investigation of the unrevealed finger motions on various surface types with different hardness, texture, and flatness.

Properties of fingers

The dexterity of fingers with the distinct anatomy, high sensitivity, and high degrees of freedom can enable various finger motions during input on the surface and in the air. Within this space, the scope of this thesis includes a range of specific finger motions for two types of wearable devices. Future works will need to extend the current studies to other fingers, regions, motions, and poses. For example, the finger identification and region-aware classifiers focused on the contact areas from the flat part or the side of fingers. Extending this research scope to other regions (e.g., fingertip and knuckle) or status of fingers (e.g., joint angles of fingers) is a promising next step to increase the range of input available more for resource-limited wearable devices. In addition, the empirical studies for identifying the comfort range of angle-based touch interaction focused on yaw angles of fingers on the touchscreen of smartwatches. Many prior works noted the feasibility of angle-based finger input with various axes including pitch [94, 101, 284] and roll [290] on the surface of mobile devices. Therefore, immediate future works will need to explore the comfort range of touch input on pitch and roll or combination of multiple orientations.

The in-air finger input techniques in the current studies also adopted specific finger motions. For example, the in-air thumb typing system used the metacarpal flexes of thumb to select keys. It led to a total of eight keys above two palms. The dexterous thumb can also move in lateral direction, which will be able to extend the design space of in-air thumb touch input to more large space other than the above-palm. Using multiple fingers towards virtual targets above palms could also enable increases in the expressiveness of the in-air touch input in the limited size of design space around hands. In sum, immediate future work is to explore a wide range of in-air touch motions with various fingers and dimensions to viable locations around hands.

Another important issue of finger input techniques is the individual difference in viable finger motions. The hand size and flexibility can affect the performance and usability of a static interface or recognizer. For example, Chapter IV provided evidence that the per-user recognizer showed the better performance than of the static recognizer. However, Chapter VI presented only static interface with the fixed sized of virtual hand, since the actual hand size had no significant effect on the performance of thumb touch motion. In future work, the effects of interface or recognizer customization for anthropometry need to be explored in the perspectives of performance and usability.

Physical occlusion

The immature hand tracking systems are still insufficient to perform subtle touch input to the small and dense virtual targets in the air. It can cause an issue of physical occlusion during in-air touch input to the virtual targets. For example, Chapter VI addressed this issue by recognizing the intended target selection from in-air thumb motions. The classifiers were constructed based on the aggregated distribution during thumb touch motions on a paraboloid surface across the virtual targets. While the promising performance of classifiers (98.2%) led to accurate in-air touch input by avoiding physical occlusion, only four touch locations were viable per hand. An ambiguous keyboard layout was required to include all letters with a total of eight keys typing tasks. Despite the quick thumb touch movements, the ambiguous layout confused novice users during typing. Therefore, future work is to increase the in-air key resolution while avoiding physical occlusion in dense environments. One approach is to investigate more distinct features from the in-air touch motions so that the key classifier can be improved to increase distinguishable touch locations. Another approach is the development of thumb interaction techniques to virtual targets within a viable volume presented in this study. An important consideration in this approach is the selection method of virtual keys by passing through unintended keys located in the viable volume.

Detection and Recognition of in-air finger stroking

The in-air typing with ten fingers [47, 306] is a promising finger input technique on HMDs with built-in hand tracking systems. However, many commercial HMDs still adopt hand-held controllers for typing due to the lack of robust detection algorithm for the sequences of fast in-air finger stroking. To achieve it, this thesis (Chapter V) provided the properties of finger motions during unconstrained in-air typing. Among various properties of finger motions, the relatively clear palm movements (23mm) and the high velocity (764ms/s) of the intended finger can be significant features to enable the detection of in-air finger stroking. Examining the feasibility of the revealed features is immediate future work to develop an accurate detection system.

The recognition of intended finger stroking is more complex. The static and per-user classifiers based on the feature set extracted from this study showed a limited accuracy of up to 86.9%. The false recognition might occur due to the highly complex finger motions such as highly correlated movements (up to 83.5%) and long overlaps of sequential inputs (23.8% of the extension phase). Further investigation needs to reveal more distinct features that can lower the complexity of the fast sequential finger motions. Interestingly, this observation study revealed in-air finger typing strategies that can variate the finger motions in the air. While the brief analysis in this thesis ascertained two strategies with different palm movements, it left room for several other strategies to exist. Understanding various in-air finger movements is fundamentally required to build the reliability of finger stroking detection and finger recognition. A logical next step is more intensive analysis to explore the potential strategies with advanced optical hand tracking system.

Muscle fatigue, comfort, and social acceptance

The issues in usability can affect the performance of interaction techniques. This thesis explored various perspectives of usability including comfort, muscle fatigue, and social acceptance, while further investigation will need to validate the current results. The high muscle fatigue during continuous input can disturb the reliability of interaction techniques. For example, the in-air thumb typing system in the comparison study of Chapter VI provided evidence that a relaxed hand posture showed low levels of perceived shoulder exertion during in-air gesture input rather than of the baseline. However, other in-air interaction techniques in this thesis need to understand muscle fatigue. Especially, the in-air finger stroking for typing (V) may cause not negligible burden on the forearm due to the fast and continuous input in the air. The short-term lab study with a break after each block limited truly measuring muscle fatigue. Future work is the investigation of muscle fatigue on the actual usage scenarios of the in-air typing systems including fast and continuous finger stroking.

In terms of comfort, the studies in Chapter III rated the comfort during static and dynamic angle-based touch inputs on the touchscreen of smartwatches. The results led to the specific range of comfortable angle-based touches with high performance, while the small sample size with up to 18 participants carefully concluded the results of the subjective feedback. Future work is a validation of the current conclusion with a large set of samples. Also, the studies for developing the in-air thumb typing system (VI) estimated the comfort with the changes of wrist angle during the input. A validation study is the next step to access the truly perceived comfort.

Social acceptability is an essential factor especially to achieve a daily use of novel interaction techniques in public places. For example, the in-air thumb typing system in VI showed high social acceptability due to the discreet and prevalent finger movements near the torso [86]. However, the in-air finger stroking gestures showed relatively large finger motions with an average travel distance of up to 86.22mm due to the lack of haptic feedback. The distribution of in-air keys formed a larger size of keyboard layout with width and height of approximately 250mm and 100mm respectively than the normal size of a physical keyboard layout (approximately 180mm x 50mm). Future works include the investigation of its social acceptability in public places or the development of subtle in-air finger typing. Mid-air haptic feedback to the fingertip can lead to reduced finger motions by providing strong key confirmation [98–100].

Mobile environments

The applications of wearable devices (e.g., smartwatches and AR HMDs) are common and useful in mobile scenarios such as walking [262] or running [256]. Nonetheless, this thesis conducted lab-based studies using sitting or standing postures and fixed settings to avoid any confound factors such as fatigue and distraction. The study results may not generalize to the performance of mobile scenarios with dynamic input postures and settings. Future works will need to explore the impact of mobile scenarios for the suggested interaction techniques.

7.4 Conclusion

This thesis explored how users move their fingers during input to develop novel interaction techniques for various interaction scenarios on wearable devices. The four scenarios on the surface and in the air supported the general claim of this thesis that understanding how users move their fingers during input will enable increases in the expressiveness of the interaction techniques we can create for resource-limited wearable devices. The first study investigated the comfort range of angle-based touch input on the touchscreen of smartwatches. The empirical studies revealed that the unique input context, in which the touching hand approaches the small touchscreen with a limited range of angles, generated the specific comfort ranges of touch input on the variations in the angles, finger regions, and touch poses. In addition, the novel interaction technique of region-aware rotation was proposed to increase the expressiveness of touch input with angles. The second study revealed that the unique input context of smartwatches generated distinctive touch profiles of fingers on the touchscreen. The finger identification systems with the touch profiles led to the development of two virtual keyboards by distinguishing two and three fingers in the natural and exaggerated touch poses respectively. The results of the first and second scenarios showed the feasibility of touch input motions that increased the expressiveness of touch input on the small touchscreen of smartwatches. In addition to the finger motion on the surface, the finger input motions in the air were examined to support the general claim of this thesis with various scenarios. The third scenario explored finger stroking input during unconstrained in-air typing that requires fast sequential input. The observation study revealed various perspectives of finger motions while conducting the sequences of in-air finger stroking, such as strategies, kinematics, correlated movements, inter-keystroke relationship, and in-air keys' distribution. The last scenario investigated the efficient range of in-air thumb touch motions to the virtual targets above the palm to develop the in-air thumb typing system for HMDs. The prevalent thumb motions enabled fast in-air typing on the ambiguous keyboard layout with a total of eight virtual keys. The empirical studies in the third and fourth scenarios showed the feasibility of in-air finger motions that enabled the fast sequential input for the complex typing tasks on HMDs.

A series of empirical studies for various finger input motions presented contributions to the development of finger-based interaction techniques on resource-limited wearable devices, through design considerations, novel interaction techniques, empirically collected data, and discussions. First, this thesis provides a range of practical knowledge for finger motions during input on the surface and in the air. For example, the unique input context to the small and fixed touchscreen of smartwatches generates distinctive touch profiles for different finger-regions or fingers (Chapter III and IV). The unconstrained finger stroking in the air complicates the detection and recognition of fast sequential input (Chapter V). In-air thumb touch techniques have a limited number of keys within the efficient range of thumb touch motion (Chapter VI). Second, this thesis has contribution to the usability of finger input motions. For example, the unique input context of smartwatches generates the limited comfort range of static and dynamic touch inputs with angles (Chapter III). Third, the novel interaction techniques based on the properties of finger input motions contribute to increases in the expressiveness of input on resource-limited wearable devices by widening the range of input available. For example, the finger region-aware systems (Chapter III) and virtual keyboards (Chapter IV) using the touch profiles provide two or three available input options on each target without any additional sensors or wearables. In-air thumb typing system (Chapter VI) based on the distribution of thumb motions also provides three or four input options on each key to enable fast sequential input with a built-in hand tracking system in commercial HMD. Lastly, the collected data from empirical studies is also a contribution to providing a ground-truth for developing novel finger interaction techniques. For example, this thesis provides the specific comfort ranges of angle-based touch input (Chapter III), the in-depth analysis for in-air finger stroking (Chapter V), and the viable locations of in-air thumb touch above palm (Chapter VI). In sum, this thesis provides clear guidelines for future researchers and developers to develop novel interaction techniques based on finger motions for resource-limited wearable devices.

Through the series of empirical studies, I argue that objective and subjective results, novel interaction techniques, and discussions over various finger input scenarios in this thesis support the general claim that understanding how users move their fingers during input will enable increases in the expressiveness of the interaction techniques we can create for resource-limited wearable devices. The additional discussions on the contributions and design considerations will be helpful for future researchers and developers to implement this type of finger-based interaction system. Furthermore, this thesis seeks to suggest the scope of future research works that will achieve efficient, convenient, and comfortable interaction systems on various types of wearable devices.

References

[1] "Wearable," Cambridge Dictionary, 2022. [Online]. Available: https://dictionary.cambridge.org/dictionary/english/wearable

[2] R. S. Bisht, S. Jain, and N. Tewari, "Study of wearable iot devices in 2021: Analysis & future prospects," in *2021 2nd International Conference on Intelligent Engineering and Management (ICIEM)*. IEEE, 2021, pp. 577–581.

[3] F. Kerber, C. Hirtz, S. Gehring, M. Löchtefeld, and A. Krüger, "Managing smartwatch notifications through filtering and ambient illumination," in *Proceedings of the 18th international conference on human-computer interaction with mobile devices and services adjunct*, 2016, pp. 918–923.

[4] S. C. Lee and T. Starner, "Buzzwear: alert perception in wearable tactile displays on the wrist," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2010, pp. 433–442.

[5] H. Ali and H. Li, "Evaluating a smartwatch notification system in a simulated nursing home," *International journal of older people nursing*, vol. 14, no. 3, p. e12241, 2019.

[6] W. C. Giang, L. Hoekstra-Atwood, and B. Donmez, "Driver engagement in notifications: A comparison of visual-manual interaction between smartwatches and smartphones," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, no. 1. Sage Publications Sage CA: Los Angeles, CA, 2014, pp. 2161–2165.

[7] I. Oakley, D. Lee, M. R. Islam, and A. Esteves, "Beats: Tapping gestures for smart watches," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15. New York, NY, USA: ACM, 2015, pp. 1237–1246. [Online]. Available: http://doi.acm.org/10.1145/2702123.2702226

[8] C. E. King and M. Sarrafzadeh, "A survey of smartwatches in remote health monitoring," *Journal of healthcare informatics research*, vol. 2, no. 1, pp. 1–24, 2018.

[9] Z. Bezovski, "The future of the mobile payment as electronic payment system," *European Journal of Business and Management*, vol. 8, no. 8, pp. 127–132, 2016.

[10] M. M. Luna, F. A. A. de Melo Nunes, H. A. D. do Nascimento, J. Siqueira, E. F. de Souza, T. H. Nascimento, R. M. da Costa *et al.*, "Text entry on smartwatches: A systematic review of literature," in *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, vol. 2. IEEE, 2018, pp. 272–277.

[11] D. Perlman, A. Samost, A. G. Domel, B. Mehler, J. Dobres, and B. Reimer, "The relative impact of smartwatch and smartphone use while driving on workload, attention, and driving performance," *Applied ergonomics*, vol. 75, pp. 8–16, 2019.

[12] M. Kheirkhahan, S. Nair, A. Davoudi, P. Rashidi, A. A. Wanigatunga, D. B. Corbett, T. Mendoza, T. M. Manini, and S. Ranka, "A smartwatch-based framework for real-time and online assessment and mobility monitoring," *Journal of biomedical informatics*, vol. 89, pp. 29–40, 2019.

[13] T. Fokkema, T. Kooiman, W. P. Krijnen, C. P. Van der Schans, and M. De Groot, "Reliability and validity of ten consumer activity trackers depend on walking speed," *Med Sci Sports Exerc*, vol. 49, no. 4, pp. 793–800, 2017.

[14] E. Jovanov, "Preliminary analysis of the use of smartwatches for longitudinal health monitoring," in *2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2015, pp. 865–868.

[15] L. Chang, J. Lu, J. Wang, X. Chen, D. Fang, Z. Tang, P. Nurmi, and Z. Wang, "Sleepguard: Capturing rich sleep information using smartwatch sensing data," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, pp. 1–34, 2018.

[16] X. Sun, L. Qiu, Y. Wu, Y. Tang, and G. Cao, "Sleepmonitor: Monitoring respiratory rate and body position during sleep using smartwatch," *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, vol. 1, no. 3, pp. 1–22, 2017.

[17] D. Phan, L. Y. Siong, P. N. Pathirana, and A. Seneviratne, "Smartwatch: Performance evaluation for long-term heart rate monitoring," in *2015 International symposium on bioelectronics and bioinformatics (ISBB)*. IEEE, 2015, pp. 144–147.

[18] A. N. Koshy, J. K. Sajeev, N. Nerlekar, A. J. Brown, K. Rajakariar, M. Zureik, M. C. Wong, L. Roberts, M. Street, J. Cooke *et al.*, "Smart watches for heart rate assessment in atrial arrhythmias," *International journal of cardiology*, vol. 266, pp. 124–127, 2018.

[19] H. Y. Lee, D.-J. Lee, J. Seo, S.-H. Ihm, K.-i. Kim, E. J. Cho, H. C. Kim, J. Shin, S. Park, I.-S. Sohn *et al.*, "Smartphone/smartwatch-based cuffless blood pressure measurement: a position paper from the korean society of hypertension," *Clinical Hypertension*, vol. 27, no. 1, pp. 1–8, 2021.

[20] B. Bent, P. J. Cho, M. Henriquez, A. Wittmann, C. Thacker, M. Feinglos, M. J. Crowley, and J. P. Dunn, "Engineering digital biomarkers of interstitial glucose from noninvasive smartwatches," *NPJ digital medicine*, vol. 4, no. 1, pp. 1–11, 2021.

[21] S. Y. Lee and K. Lee, "Factors that influence an individual's intention to adopt a wearable healthcare device: The case of a wearable fitness tracker," *Technological Forecasting and Social Change*, vol. 129, pp. 154–163, 2018.

[22] K.-J. Brickwood, A. D. Williams, G. Watson, and J. O'Brien, "Older adults' experiences of using a wearable activity tracker with health professional feedback over a 12-month randomised controlled trial," *Digital health*, vol. 6, p. 2055207620921678, 2020.

[23] B. H. Thomas, "A survey of visual, mixed, and augmented reality gaming," *Computers in Entertainment (CIE)*, vol. 10, no. 1, pp. 1–33, 2012.

[24] M. Herz and P. A. Rauschnabel, "Understanding the diffusion of virtual reality glasses: The role of media, fashion and technology," *Technological Forecasting and Social Change*, vol. 138, pp. 228–242, 2019.

[25] T. Hartmann, "Entertainment in virtual reality and beyond: The influence of embodiment, co-location, and cognitive distancing on users' entertainment experience," *The Oxford handbook of entertainment theory*, 2021.

[26] L. Li, F. Yu, D. Shi, J. Shi, Z. Tian, J. Yang, X. Wang, and Q. Jiang, "Application of virtual reality technology in clinical medicine," *American journal of translational research*, vol. 9, no. 9, p. 3867, 2017.

[27] L. Freina and M. Ott, "A literature review on immersive virtual reality in education: state of the art and perspectives," in *The international scientific conference elearning and software for education*, vol. 1, no. 133, 2015, pp. 10–1007.

[28] S. Barteit, L. Lanfermann, T. Bärnighausen, F. Neuhann, C. Beiersmann *et al.*, "Augmented, mixed, and virtual reality-based head-mounted devices for medical education: Systematic review," *JMIR serious games*, vol. 9, no. 3, p. e29080, 2021.

[29] Z. Gao and A. Lecuyer, "A vr simulator for training and prototyping of telemanipulation of nanotubes," in *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, ser. VRST'08. ACM, 2008, pp. 101–104.

[30] S. Philippe, A. D. Souchet, P. Lameras, P. Petridis, J. Caporal, G. Coldeboeuf, and H. Duzan, "Multimodal teaching, learning and training in virtual reality: a review and case study," *Virtual Reality & Intelligent Hardware*, vol. 2, no. 5, pp. 421–442, 2020.

[31] D. Mourtzis, V. Zogopoulos, and F. Xanthi, "Augmented reality application to support the assembly of highly customized products and to adapt to production re-scheduling," *The International Journal of Advanced Manufacturing Technology*, vol. 105, no. 9, pp. 3899–3910, 2019.

[32] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE TRANSACTIONS on Information and Systems*, vol. 77, no. 12, pp. 1321–1329, 1994.

[33] D. Cheng, Q. Wang, Y. Liu, H. Chen, D. Ni, X. Wang, C. Yao, Q. Hou, W. Hou, G. Luo *et al.*, "Design and manufacture ar head-mounted displays: A review and outlook," *Light: Advanced Manufacturing*, vol. 2, no. 3, pp. 350–369, 2021.

[34] G. Evans, J. Miller, M. I. Pena, A. MacAllister, and E. Winer, "Evaluating the microsoft hololens through an augmented reality assembly application," in *Degraded environments: sensing, processing, and display 2017*, vol. 10197. SPIE, 2017, pp. 282–297.

[35] P. A. Rauschnabel, "Virtually enhancing the real world with holograms: An exploration of expected gratifications of using augmented reality smart glasses," *Psychology & Marketing*, vol. 35, no. 8, pp. 557–572, 2018.

[36] P. Parekh, S. Patel, N. Patel, and M. Shah, "Systematic review and meta-analysis of augmented reality in medicine, retail, and games," *Visual computing for industry, biomedicine, and art*, vol. 3, no. 1, pp. 1–20, 2020.

[37] L.-H. Lee and P. Hui, "Interaction methods for smart glasses: A survey," *IEEE access*, vol. 6, pp. 28 712–28 732, 2018.

[38] V. Angelov, E. Petkov, G. Shipkovenski, and T. Kalushkov, "Modern virtual reality headsets," in *2020 International congress on human-computer interaction, optimization and robotic applications (HORA)*. IEEE, 2020, pp. 1–5.

[39] D. Abdlkarim, M. Di Luca, P. Aves, S.-H. Yeo, R. C. Miall, P. Holland, and J. M. Galea, "A methodological framework to assess the accuracy of virtual reality hand-tracking

systems: A case study with the oculus quest 2," *bioRxiv*, 2022. [Online]. Available: https://www.biorxiv.org/content/early/2022/02/20/2022.02.18.481001

[40] V. Clay, P. König, and S. Koenig, "Eye tracking in virtual reality," *Journal of eye movement research*, vol. 12, no. 1, 2019.

[41] O. Quest, "Vr headsets, games, and equipment," 2021. [Online]. Available: https://www.oculus.com/

[42] A. Zenner and A. Krüger, "Drag: on: A virtual reality controller providing haptic feedback based on drag and weight shift," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.

[43] A. Forsberg, K. Herndon, and R. Zeleznik, "Aperture based selection for immersive virtual environments," in *Proceedings of the 9th annual ACM symposium on User interface software and technology*, 1996, pp. 95–96.

[44] T. Grossman and R. Balakrishnan, "The design and evaluation of selection techniques for 3d volumetric displays," in *Proceedings of the 19th annual ACM symposium on User interface software and technology*, 2006, pp. 3–12.

[45] M. Speicher, A. M. Feit, P. Ziegler, and A. Krüger, "Selection-based text entry in virtual reality," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–13. [Online]. Available: https://doi.org/10.1145/3173574.3174221

[46] R. Arora, R. H. Kazi, D. M. Kaufman, W. Li, and K. Singh, "Magicalhands: Mid-air hand gestures for animating in vr," in *Proceedings of the 32nd annual ACM symposium on user interface software and technology*, 2019, pp. 463–477.

[47] X. Yi, C. Yu, M. Zhang, S. Gao, K. Sun, and Y. Shi, "Atk: Enabling ten-finger freehand typing in air based on 3d hand tracking data," in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, ser. UIST '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 539–548. [Online]. Available: https://doi.org/10.1145/2807442.2807504

[48] C. Bullion and H. Gurocak, "Haptic glove with mr brakes for distributed finger force feedback," *Presence*, vol. 18, no. 6, pp. 421–433, 2009.

[49] F. Müller, J. McManus, S. Günther, M. Schmitz, M. Mühlhäuser, and M. Funk, "Mind the tap: Assessing foot-taps for interacting with head-mounted displays," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–13.

[50] A. Rahimi, H. Patel, H. Ajmal, and S. Haghani, "The design and implementation of a vr gun controller with haptic feedback," in *2019 ieee international conference on consumer electronics (icce)*. IEEE, 2019, pp. 1–2.

[51] R. A. Ruddle, E. Volkova, and H. H. Bülthoff, "Learning to walk in virtual reality," *ACM Transactions on Applied Perception (TAP)*, vol. 10, no. 2, pp. 1–17, 2013.

[52] X. Chen, T. Grossman, and G. Fitzmaurice, "Swipeboard: a text entry technique for ultra-small interfaces that supports novice to expert transitions," in *Proceedings of the 27th annual ACM symposium on User interface software and technology*, 2014, pp. 615–620.

[53] T. Grossman, X. A. Chen, and G. Fitzmaurice, "Typing on glasses: Adapting text entry to smart eyewear," in *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 2015, pp. 144–152.

[54] M. R. Islam, D. Lee, L. S. Jahan, and I. Oakley, "Glasspass: Tapping gestures to unlock smart glasses," in *Proceedings of the 9th Augmented Human International Conference*, 2018, pp. 1–8.

[55] S. Azenkot and S. Zhai, "Touch behavior with different postures on soft smartphone keyboards," in *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services*, 2012, pp. 251–260.

[56] W. Guo and J. Wang, "Smartrsvp: Facilitating attentive speed reading on small screen wearable devices," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2017, pp. 1640–1647.

[57] K. A. Siek, Y. Rogers, and K. H. Connelly, "Fat finger worries: How older and younger users physically interact with pdas," in *Proceedings of the 2005 IFIP TC13 International Conference on Human-Computer Interaction*, ser. INTERACT'05. Berlin, Heidelberg: Springer-Verlag, 2005, pp. 267–280. [Online]. Available: https://doi.org/10.1007/11555261_24

[58] I. Oakley and D. Lee, "Interaction on the edge: offset sensing for small devices," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2014, pp. 169–178.

[59] S. T. Perrault, E. Lecolinet, J. Eagan, and Y. Guiard, “Watchit: simple gestures and eyes-free interaction for wristwatches and bracelets,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 1451–1460.

[60] D. Ahlström, K. Hasan, E. Lank, and R. Liang, “Tiltcrown: Extending input on a smartwatch with a tiltable digital crown,” in *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia*, 2018, pp. 359–366.

[61] J. Kim, J. He, K. Lyons, and T. Starner, “The gesture watch: A wireless contact-free gesture based wrist interface,” in *2007 11th IEEE International Symposium on Wearable Computers*. IEEE, 2007, pp. 15–22.

[62] H. Wen, J. Ramos Rojas, and A. K. Dey, “Serendipity: Finger gesture recognition using an off-the-shelf smartwatch,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’16. New York, NY, USA: ACM, 2016, pp. 3847–3851. [Online]. Available: http://doi.acm.org/10.1145/2858036.2858466

[63] C. Harrison, D. Tan, and D. Morris, “Skinput: appropriating the body as an input surface,” in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2010, pp. 453–462.

[64] A. Guo and T. Paek, “Exploring tilt for no-touch, wrist-only interactions on smartwatches,” in *Proceedings of the 18th international conference on human-computer interaction with mobile devices and services*, 2016, pp. 17–28.

[65] H.-S. Yeo, J. Lee, A. Bianchi, and A. Quigley, “Watchmi: Applications of watch movement input on unmodified smartwatches,” in *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, ser. MobileHCI ’16. New York, NY, USA: ACM, 2016, pp. 594–598. [Online]. Available: http://doi.acm.org/10.1145/2957265.2961825

[66] I. Oakley, C. Lindahl, K. Le, D. Lee, and M. R. Islam, “The flat finger: Exploring area touches on smartwatches,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’16. New York, NY, USA: ACM, 2016, pp. 4238–4249. [Online]. Available: http://doi.acm.org/10.1145/2858036.2858179

[67] X. Yi, C. Yu, W. Xu, X. Bi, and Y. Shi, “Compass: Rotational keyboard on non-touch smartwatches,” in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 705–715.

[68] A. Gupta and R. Balakrishnan, "Dualkey: Miniature screen text entry via finger identification," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: ACM, 2016, pp. 59–70. [Online]. Available: http://doi.acm.org/10.1145/2858036.2858052

[69] M. D. Dunlop, A. Komninos, and N. Durga, "Towards high quality text entry on smartwatches," in *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, 2014, pp. 2365–2370.

[70] S. Oney, C. Harrison, A. Ogan, and J. Wiese, "Zoomboard: a diminutive qwerty soft keyboard using iterative zooming for ultra-small devices," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 2799–2802.

[71] K. Kato and H. Miyashita, "Creating a mobile head-mounted display with proprietary controllers for interactive virtual reality content," in *Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, 2015, pp. 35–36.

[72] C. Yu, K. Sun, M. Zhong, X. Li, P. Zhao, and Y. Shi, "One-dimensional handwriting: Inputting letters and words on smart glasses," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 71–82.

[73] S. Ahn and G. Lee, "Gaze-assisted typing for smart glasses," in *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 857–869. [Online]. Available: https://doi.org/10.1145/3332165.3347883

[74] M. Serrano, B. M. Ens, and P. P. Irani, "Exploring the use of hand-to-face input for interacting with head-worn displays," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014, pp. 3181–3190.

[75] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani, "Consumed endurance: A metric to quantify arm fatigue of mid-air interactions," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 1063–1072. [Online]. Available: https://doi.org/10.1145/2556288.2557130

[76] G. Wesche and H.-P. Seidel, "Freedrawer: a free-form sketching system on the responsive workbench," in *Proceedings of the ACM symposium on Virtual reality software and technology*, 2001, pp. 167–174.

[77] J. S. Pierce, A. S. Forsberg, M. J. Conway, S. Hong, R. C. Zeleznik, and M. R. Mine, "Image plane interaction techniques in 3d immersive environments," in *Proceedings of the 1997 symposium on Interactive 3D graphics*, 1997, pp. 39–ff.

[78] G. Shoemaker, L. Findlater, J. Q. Dawson, and K. S. Booth, "Mid-air text input techniques for very large wall displays," in *Proceedings of Graphics Interface 2009*, ser. GI '09. CAN: Canadian Information Processing Society, 2009, p. 231–238.

[79] H. Jiang and D. Weng, "Hipad: Text entry for head-mounted displays using circular touchpad," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, IEEE. USA: IEEE Computer Society, 2020, pp. 692–703.

[80] C. Boletsis and S. Kongsvik, "Controller-based text-input techniques for virtual reality: an empirical comparison," *International Journal of Virtual Reality (IJVR)*, vol. 19, no. 3, 2019.

[81] A. O. S. Feiner, "The flexible pointer: An interaction technique for selection in augmented and virtual reality," in *Proc. UIST*, vol. 3, 2003, pp. 81–82.

[82] W. Xu, H.-N. Liang, A. He, and Z. Wang, "Pointing and selection methods for text entry in augmented reality head mounted displays," in *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, IEEE. USA: IEEE Computer Society, 2019, pp. 279–288.

[83] A. Esteves, E. Velloso, A. Bulling, and H. Gellersen, "Orbits: Gaze interaction for smart watches using smooth pursuit eye movements," in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, ser. UIST '15. New York, NY, USA: ACM, 2015, pp. 457–466. [Online]. Available: http://doi.acm.org/10.1145/2807442.2807499

[84] A. Esteves, D. Verweij, L. Suraiya, R. Islam, Y. Lee, and I. Oakley, "Smoothmoves: Smooth pursuits head movements for augmented reality," in *Proceedings of the 30th annual acm symposium on user interface software and technology*, 2017, pp. 167–178.

[85] S. G. Gustafson, B. Rabe, and P. M. Baudisch, "Understanding palm-based imaginary interfaces: the role of visual and tactile cues when browsing," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 889–898.

[86] Y.-C. Tung, C.-Y. Hsu, H.-Y. Wang, S. Chyou, J.-W. Lin, P.-J. Wu, A. Valstar, and M. Y. Chen, "User-defined game input for smart glasses in public space," in *Proceedings of the 33rd Annual ACM conference on human factors in computing systems*, 2015, pp. 3327–3336.

[87] J. Dudley, H. Benko, D. Wigdor, and P. O. Kristensson, "Performance envelopes of virtual keyboard text input strategies in virtual reality," in *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. Washington, DC, USA: IEEE Computer Society, 2019, pp. 289–300.

[88] J. J. Dudley, K. Vertanen, and P. O. Kristensson, "Fast and precise touch-based text entry for head-mounted augmented reality with variable occlusion," *ACM Trans. Comput.-Hum. Interact.*, vol. 25, no. 6, Dec. 2018. [Online]. Available: https://doi.org/10.1145/3232163

[89] A. Rahman and A. Al-Jumaily, "Design and development of a bilateral therapeutic hand device for stroke rehabilitation," *International Journal of Advanced Robotic Systems*, vol. 10, no. 12, p. 405, 2013.

[90] H. Z. Tan, M. A. Srinivasan, C. M. Reed, and N. I. Durlach, "Discrimination and identification of finger joint-angle position using active motion," *ACM Transactions on Applied Perception (TAP)*, vol. 4, no. 2, pp. 10–es, 2007.

[91] C. Zhang, J. Yang, C. Southern, T. E. Starner, and G. D. Abowd, "Watchout: extending interactions on a smartwatch with inertial sensing," in *Proceedings of the 2016 ACM International Symposium on Wearable Computers*, 2016, pp. 136–143.

[92] C. Harrison, J. Schwarz, and S. E. Hudson, "Tapsense: enhancing finger interaction on touch surfaces," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*, 2011, pp. 627–636.

[93] J. Kim and I. Oakley, "Sonarid: Using sonar to identify fingers on a smartwatch," in *CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–10.

[94] R. Xiao, J. Schwarz, and C. Harrison, "Estimating 3d finger angle on commodity touchscreens," in *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*, ser. ITS '15. New York, NY, USA: ACM, 2015, pp. 47–50. [Online]. Available: http://doi.acm.org/10.1145/2817721.2817737

[95] H. V. Le, S. Mayer, and N. Henze, "Infinitouch: Finger-aware interaction on fully touch sensitive smartphones," in *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, 2018, pp. 779–792.

[96] C. Holz and P. Baudisch, "The generalized perceived input point model and how to double touch accuracy by extracting fingerprints," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2010, pp. 581–590.

[97] A. Markussen, M. R. Jakobsen, and K. Hornbæk, "Vulture: A mid-air word-gesture keyboard," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '14. ACM, 2014, pp. 1073–1082.

[98] S. Brewster, F. Chohan, and L. Brown, "Tactile feedback for mobile interactions," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '07. ACM, 2007, pp. 159–162.

[99] S. Lee and S. Zhai, "The performance of touch screen soft buttons," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '09. ACM, 2009, pp. 309–318.

[100] E. Hoggan, S. Brewster, and J. Johnston, "Investigating the effectiveness of tactile feedback for mobile touchscreens," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '08. ACM, 2008, pp. 1573–1582.

[101] S. Mayer, P. Gad, K. Wolf, P. W. Woźniak, and N. Henze, "Understanding the ergonomic constraints in designing for touch surfaces," in *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '17. New York, NY, USA: ACM, 2017, pp. 33:1–33:9. [Online]. Available: http://doi.acm.org/10.1145/3098279.3098537

[102] Z. Yang, C. Chen, Y. Lin, D. Wang, H. Li, and W. Xu, "Effect of spatial enhancement technology on input through the keyboard in virtual reality environment," *Applied ergonomics*, vol. 78, pp. 164–175, 2019.

[103] K. Sun, W. Wang, A. X. Liu, and H. Dai, "Depth aware finger tapping on virtual displays," in *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 283–295. [Online]. Available: https://doi.org/10.1145/3210240.3210315

[104] N. Marquardt, J. Kiemer, and S. Greenberg, "What caused that touch? expressive interaction with a surface through fiduciary-tagged gloves," in *ACM International Conference on Interactive Tabletops and Surfaces*, 2010, pp. 139–142.

[105] A. Sugiura and Y. Koseki, "A user interface using fingerprint recognition: holding commands and data objects on fingers," in *Proceedings of the 11th annual ACM symposium on User interface software and technology*, 1998, pp. 71–79.

[106] J. Wang and J. Canny, "Fingersense: augmenting expressiveness to physical pushing button by fingertip identification," in *CHI'04 Extended Abstracts on Human Factors in Computing Systems*, 2004, pp. 1267–1270.

[107] Y.-F. Shao, M. Chang-Ogimoto, R. Pointner, Y.-C. Lin, C.-T. Wu, and M. Chen, "Swipekey: a swipe-based keyboard design for smartwatches," in *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 2016, pp. 60–71.

[108] F. Poirier and M. Belatar, "Uniwatch: A soft keyboard for text entry on smartwatches using 3 keys," in *International Conference on Human-Computer Interaction*. Springer, 2016, pp. 341–349.

[109] A. Mottelson, C. Larsen, M. Lyderik, P. Strohmeier, and J. Knibbe, "Invisiboard: maximizing display and input space with a full screen text entry method for smartwatches," in *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 2016, pp. 53–59.

[110] D.-Y. Huang, M.-C. Tsai, Y.-C. Tung, M.-L. Tsai, Y.-T. Yeh, L. Chan, Y.-P. Hung, and M. Y. Chen, "Touchsense: expanding touchscreen input vocabulary using different areas of users' finger pads," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014, pp. 189–192.

[111] L. A. Leiva, A. Sahami, A. Catala, N. Henze, and A. Schmidt, "Text entry on tiny qwerty soft keyboards," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15. New York, NY, USA: ACM, 2015, pp. 669–678. [Online]. Available: http://doi.acm.org/10.1145/2702123.2702388

[112] S. Boring, D. Ledo, X. A. Chen, N. Marquardt, A. Tang, and S. Greenberg, "The fat thumb: Using the thumb's contact size for single-handed mobile interaction," in *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services Companion*, ser. MobileHCI '12. New York, NY, USA: ACM, 2012, pp. 207–208. [Online]. Available: http://doi.acm.org/10.1145/2371664.2371711

[113] F. Wang, X. Cao, X. Ren, and P. Irani, "Detecting and leveraging finger orientation for interaction with direct-touch surfaces," in *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '09. New York, NY, USA: ACM, 2009, pp. 23–32. [Online]. Available: http://doi.acm.org/10.1145/1622176.1622182

[114] S. Rogers, J. Williamson, C. Stewart, and R. Murray-Smith, “Anglepose: Robust, precise capacitive touch tracking via 3d orientation estimation,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’11. New York, NY, USA: ACM, 2011, pp. 2575–2584. [Online]. Available: http://doi.acm.org/10.1145/1978942.1979318

[115] J. Vogelsang, F. Kiss, and S. Mayer, “A design space for user interface elements using finger orientation input,” in *Mensch und Computer 2021*, 2021, pp. 1–10.

[116] J. Rekimoto and C. Schwesig, “Presenseii: bi-directional touch and pressure sensing interactions with tactile feedback,” in *CHI’06 extended abstracts on Human factors in computing systems*, 2006, pp. 1253–1258.

[117] S. Heo and G. Lee, “Force gestures: augmenting touch screen gestures with normal and tangential forces,” in *Proceedings of the 24th annual ACM symposium on User interface software and technology*, 2011, pp. 621–626.

[118] H. Xia, T. Grossman, and G. Fitzmaurice, “Nanostylus: Enhancing input on ultra-small displays with a finger-mounted stylus,” in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, 2015, pp. 447–456.

[119] C.-H. Su, L. Chan, C.-T. Weng, R.-H. Liang, K.-Y. Cheng, and B.-Y. Chen, “Naildisplay: bringing an always available visual display to fingertips,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 1461–1464.

[120] G. Blasko and S. Feiner, “An interaction system for watch computers using tactile guidance and bidirectional segmented strokes,” in *Eighth International Symposium on Wearable Computers*, vol. 1. IEEE, 2004, pp. 120–123.

[121] Y. Kubo, B. Shizuki, and J. Tanaka, “B2b-swipe: Swipe gesture for rectangular smartwatches from a bezel to a bezel,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’16. New York, NY, USA: ACM, 2016, pp. 3852–3856. [Online]. Available: http://doi.acm.org/10.1145/2858036.2858216

[122] A. Neshati, B. Rey, A. S. Mohommed Faleel, S. Bardot, C. Latulipe, and P. Irani, “Bezelglide: Interacting with graphs on smartwatches with minimal screen occlusion,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–13.

[123] F. Mo and J. Zhou, “Adapting smartwatch interfaces to hand gestures during movements: offset models and the c-shaped pattern of tapping,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 7, pp. 8099–8117, 2021.

[124] D. Dobbelstein, G. Haas, and E. Rukzio, "The effects of mobility, encumbrance, and (non-)dominant hand on interaction with smartwatches," in *Proceedings of the 2017 ACM International Symposium on Wearable Computers*, ser. ISWC '17. New York, NY, USA: ACM, 2017, pp. 90–93. [Online]. Available: http://doi.acm.org/10.1145/3123021.3123033

[125] C. J. Turner, B. S. Chaparro, and J. He, "Typing on a smartwatch while mobile: a comparison of input methods," *Human factors*, vol. 63, no. 6, pp. 974–986, 2021.

[126] S. Ahn, S. Heo, and G. Lee, "Typing on a smartwatch for smart glasses," in *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, 2017, pp. 201–209.

[127] G. Singh, W. Delamare, and P. Irani, "D-swime: A design space for smartwatch interaction techniques supporting mobility and encumbrance," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: ACM, 2018, pp. 634:1–634:13. [Online]. Available: http://doi.acm.org/10.1145/3173574.3174208

[128] H.-S. Yeo, J. Lee, A. Bianchi, and A. Quigley, "Sidetap & slingshot gestures on unmodified smartwatches," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 2016, pp. 189–190.

[129] R. Darbar, P. K. Sen, and D. Samanta, "Presstact: Side pressure-based input for smartwatch interaction," in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2016, pp. 2431–2438.

[130] P. C. Wong, K. Zhu, X.-D. Yang, and H. Fu, "Exploring eyes-free bezel-initiated swipe on round smartwatches," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–11.

[131] M. Funk, A. Sahami, N. Henze, and A. Schmidt, "Using a touch-sensitive wristband for text entry on smart watches," in *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, 2014, pp. 2305–2310.

[132] L. Saviot, F. Brudy, and S. Houben, "Wristband. io: expanding input and output spaces of a smartwatch," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2017, pp. 2025–2033.

[133] K. Lyons, D. Nguyen, D. Ashbrook, and S. White, "Facet: a multi-segment wrist worn system," in *Proceedings of the 25th annual ACM symposium on User interface software and technology*, 2012, pp. 123–130.

[134] K. Klamka, T. Horak, and R. Dachselt, “Watch+ strap: Extending smartwatches with interactive strapdisplays,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–15.

[135] Y. Ahn, S. Hwang, H. Yoon, J. Gim, and J.-h. Ryu, “Bandsense: pressure-sensitive multi-touch interaction on a wristband,” in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, 2015, pp. 251–254.

[136] E. Brulé, G. Bailly, M. Serrano, M. Teyssier, and S. Huron, “Investigating the design space of smartwatches combining physical rotary inputs,” in *Proceedings of the 29th Conference on l’Interaction Homme-Machine*, 2017, pp. 13–20.

[137] C. Harrison and S. E. Hudson, “Abracadabra: wireless, high-precision, and unpowered finger input for very small mobile devices,” in *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, 2009, pp. 121–124.

[138] A. Withana, R. Peiris, N. Samarasekara, and S. Nanayakkara, “zsense: Enabling shallow depth gesture recognition for greater input expressivity on smart wearables,” in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 3661–3670.

[139] S. S. Arefin Shimon, C. Lutton, Z. Xu, S. Morrison-Smith, C. Boucher, and J. Ruiz, “Exploring non-touchscreen gestures for smartwatches,” in *Proceedings of the 2016 chi conference on human factors in computing systems*, 2016, pp. 3822–3833.

[140] J. Han, S. Ahn, K. Park, and G. Lee, “Designing touch gestures using the space around the smartwatch as continuous input space,” in *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, 2017, pp. 210–219.

[141] J. McIntosh, P. Strohmeier, J. Knibbe, S. Boring, and K. Hornbæk, “Magnetips: Combining fingertip tracking and haptic feedback for around-device interaction,” in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.

[142] J. Wagner, M. Nancel, S. G. Gustafson, S. Huot, and W. E. Mackay, “Body-centric design space for multi-surface interaction,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 1299–1308.

[143] M. Weigel, V. Mehta, and J. Steimle, “More than touch: understanding how people use skin as an input surface for mobile computing,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014, pp. 179–188.

[144] X. Han, H. Seki, Y. Kamiya, and M. Hikizu, "Wearable handwriting input device using magnetic field," in *SICE Annual Conference 2007*. IEEE, 2007, pp. 365–368.

[145] C. Harrison, H. Benko, and A. D. Wilson, "Omnitouch: wearable multitouch interaction everywhere," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*, 2011, pp. 441–450.

[146] S.-Y. Lin, C.-H. Su, K.-Y. Cheng, R.-H. Liang, T.-H. Kuo, and B.-Y. Chen, "Pub-point upon body: exploring eyes-free interaction and methods on an arm," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*, 2011, pp. 481–488.

[147] G. Laput, R. Xiao, X. A. Chen, S. E. Hudson, and C. Harrison, "Skin buttons: Cheap, small, low-powered and clickable fixed-icon laser projectors," in *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '14. New York, NY, USA: ACM, 2014, pp. 389–394. [Online]. Available: http://doi.acm.org/10.1145/2642918.2647356

[148] M. Ogata, Y. Sugiura, Y. Makino, M. Inami, and M. Imai, "Senskin: adapting skin as a soft interface," in *Proceedings of the 26th annual ACM symposium on User interface software and technology*, 2013, pp. 539–544.

[149] S.-C. Lim, J. Shin, S.-C. Kim, and J. Park, "Expansion of smartwatch touch interface from touchscreen to around device interface using infrared line image sensors," *Sensors*, vol. 15, no. 7, pp. 16 642–16 653, 2015.

[150] C. Zhang, A. Bedri, G. Reyes, B. Bercik, O. T. Inan, T. E. Starner, and G. D. Abowd, "Tapskin: Recognizing on-skin input for smartwatches," in *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*, 2016, pp. 13–22.

[151] M. Zhang, Q. Dai, P. Yang, J. Xiong, C. Tian, and C. Xiang, "idial: Enabling a virtual dial plate on the hand back for around-device interaction," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, pp. 1–20, 2018.

[152] M. Chen, P. Yang, and P. Li, "You can write numbers accurately on your hand with smart acoustic sensing," in *International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*. Springer, 2017, pp. 207–217.

[153] S. Sridhar, A. Markussen, A. Oulasvirta, C. Theobalt, and S. Boring, "Watchsense: On-and above-skin input sensing through a wearable depth sensor," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 3891–3902.

[154] Y. Zhang, J. Zhou, G. Laput, and C. Harrison, "Skintrack: Using the body as an electrical waveguide for continuous finger tracking on the skin," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 1491–1503.

[155] J. Zhou, Y. Zhang, G. Laput, and C. Harrison, "Aurasense: enabling expressive around-smartwatch interactions with electric field sensing," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 2016, pp. 81–86.

[156] R. Xiao, T. Cao, N. Guo, J. Zhuo, Y. Zhang, and C. Harrison, "Lumiwatch: On-arm projected graphics and touch input," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–11.

[157] I. Oakley and S. O'Modhrain, "Tilt to scroll: Evaluating a motion based vibrotactile mobile interface," in *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics Conference*. IEEE, 2005, pp. 40–49.

[158] A. Crossan, J. Williamson, S. Brewster, and R. Murray-Smith, "Wrist rotation for interaction in mobile contexts," in *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, 2008, pp. 435–438.

[159] M. Rahman, S. Gustafson, P. Irani, and S. Subramanian, "Tilt techniques: investigating the dexterity of wrist-based input," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2009, pp. 1943–1952.

[160] J. Gong, Z. Xu, Q. Guo, T. Seyed, X. A. Chen, X. Bi, and X.-D. Yang, "Wristext: One-handed text entry on smartwatch using wrist gestures," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–14. [Online]. Available: https://doi.org/10.1145/3173574.3173755

[161] J. Gong, X.-D. Yang, and P. Irani, "Wristwhirl: One-handed continuous smartwatch input using wrist gestures," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, ser. UIST '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 861–872. [Online]. Available: https://doi.org/10.1145/2984511.2984563

[162] K. Sun, Y. Wang, C. Yu, Y. Yan, H. Wen, and Y. Shi, "Float: One-handed and touch-free target selection on smartwatches," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: ACM, 2017, pp. 692–704. [Online]. Available: http://doi.acm.org/10.1145/3025453.3026027

[163] F. Kerber, P. Lessel, and A. Krüger, "Same-side hand interactions with arm-placed devices using emg," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '15. New York, NY, USA: ACM, 2015, pp. 1367–1372. [Online]. Available: http://doi.acm.org/10.1145/2702613.2732895

[164] H. Kurosawa, D. Sakamoto, and T. Ono, "Myotilt: A target selection method for smartwatches using the tilting operation and electromyography," in *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 2018, pp. 1–11.

[165] P. Strohmeier and J. McIntosh, "Novel input and output opportunities using an implanted magnet," in *Proceedings of the Augmented Humans International Conference*, 2020, pp. 1–5.

[166] V. Becker, L. Fessler, and G. Sörös, "Gestear: combining audio and motion sensing for gesture recognition on smartwatches," in *Proceedings of the 23rd International Symposium on Wearable Computers*, 2019, pp. 10–19.

[167] Y. Iravantchi, M. Goel, and C. Harrison, "Beamband: Hand gesture sensing with ultrasonic beamforming," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–10.

[168] D. Kim, O. Hilliges, S. Izadi, A. D. Butler, J. Chen, I. Oikonomidis, and P. Olivier, "Digits: Freehand 3d interactions anywhere using a wrist-worn gloveless sensor," in *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '12. New York, NY, USA: ACM, 2012, pp. 167–176. [Online]. Available: http://doi.acm.org/10.1145/2380116.2380139

[169] A. Dementyev and J. A. Paradiso, "Wristflex: low-power gesture input with wrist-worn pressure sensors," in *Proceedings of the 27th annual ACM symposium on User interface software and technology*, 2014, pp. 161–166.

[170] S. Aoyama, B. Shizuki, and J. Tanaka, "Thumbslide: An interaction technique for smartwatches using a thumb slide movement," in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2016, pp. 2403–2409.

[171] T. Han, K. Hasan, K. Nakamura, R. Gomez, and P. Irani, "Soundcraft: Enabling spatial interactions on smartwatches using hand generated acoustics," in *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, 2017, pp. 579–591.

[172] Y. Zhang and C. Harrison, "Tomo: Wearable, low-cost electrical impedance tomography for hand gesture recognition," in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, 2015, pp. 167–173.

[173] J. McIntosh, A. Marzo, and M. Fraser, "Sensir: Detecting hand gestures with a wearable bracelet using infrared transmission and reflection," in *Proceedings of the 30th annual ACM symposium on user interface software and technology*, 2017, pp. 593–597.

[174] J. McIntosh, A. Marzo, M. Fraser, and C. Phillips, "Echoflex: Hand gesture recognition using ultrasound imaging," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 1923–1934.

[175] Y. Zhang, T. Gu, C. Luo, V. Kostakos, and A. Seneviratne, "Findroidhr: Smartwatch gesture input with optical heartrate monitor," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, pp. 1–42, 2018.

[176] W. Kienzle and K. Hinckley, "Lightring: always-available 2d input on any surface," in *Proceedings of the 27th annual ACM symposium on User interface software and technology*, 2014, pp. 157–160.

[177] M. Weigel and J. Steimle, "Deformwear: Deformation input on tiny wearable devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 2, pp. 1–23, 2017.

[178] Q. Zhang, D. Wang, R. Zhao, Y. Yu, and J. Jing, "Write, attend and spell: Streaming end-to-end free-style handwriting recognition using smartwatches," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 3, pp. 1–25, 2021.

[179] C. Loclair, S. Gustafson, and P. Baudisch, "Pinchwatch: a wearable device for one-handed microinteractions," in *Proc. MobileHCI*, vol. 10. Citeseer, 2010.

[180] D.-Y. Huang, L. Chan, S. Yang, F. Wang, R.-H. Liang, D.-N. Yang, Y.-P. Hung, and B.-Y. Chen, "Digitspace: Designing thumb-to-fingers touch interfaces for one-handed and eyes-free interactions," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 1526–1537. [Online]. Available: https://doi.org/10.1145/2858036.2858483

[181] C. Zhang, A. Waghmare, P. Kundra, Y. Pu, S. Gilliland, T. Ploetz, T. E. Starner, O. T. Inan, and G. D. Abowd, "Fingersound: Recognizing unistroke thumb gestures using a ring," *Proceedings of*

the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 1, no. 3, pp. 1–19, 2017.

[182] C. Zhang, Q. Xue, A. Waghmare, R. Meng, S. Jain, Y. Han, X. Li, K. Cunefare, T. Ploetz, T. Starner *et al.*, "Fingerping: Recognizing fine-grained hand poses using active acoustic on-body sensing," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–10.

[183] P. C. Wong, K. Zhu, and H. Fu, "Fingert9: Leveraging thumb-to-finger interaction for same-side-hand text entry on smartwatches," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–10. [Online]. Available: https://doi.org/10.1145/3173574.3173752

[184] J. D. Hincapié-Ramos, K. Ozacar, P. P. Irani, and Y. Kitamura, "Gyrowand: Imu-based raycasting for augmented reality head-mounted displays," in *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*, 2015, pp. 89–98.

[185] J. Hartmann and D. Vogel, "An evaluation of mobile phone pointing in spatial augmented reality," in *Extended abstracts of the 2018 CHI conference on human factors in computing systems*, 2018, pp. 1–6.

[186] P. Mohr, M. Tatzgern, T. Langlotz, A. Lang, D. Schmalstieg, and D. Kalkofen, "Trackcap: Enabling smartphones for 3d interaction on mobile head-mounted displays," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–11.

[187] Y.-W. Chow, "3d spatial interaction with the wii remote for head-mounted display virtual reality," *World Academy of Science, Engineering and Technology 50*, 2009.

[188] E. Jones, J. Alexander, A. Andreou, P. Irani, and S. Subramanian, "Gestext: Accelerometer-based gestural text-entry systems," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '10. New York, NY, USA: Association for Computing Machinery, 2010, p. 2173–2182. [Online]. Available: https://doi.org/10.1145/1753326.1753655

[189] B. Jackson, L. B. Caraco, and Z. M. Spilka, "Arc-type and tilt-type: Pen-based immersive text input for room-scale vr," in *Symposium on Spatial User Interaction*, 2020, pp. 1–10.

[190] Z. Zhang, M. Sun, B. Gao, and L. Wang, "2-thumbs typing: A novel bimanual text entry method in virtual reality environments," in *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 2021, pp. 530–531.

[191] J. Son, S. Ahn, S. Kim, and G. Lee, "Improving two-thumb touchpad typing in virtual reality," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1–6. [Online]. Available: https://doi.org/10.1145/3290607.3312926

[192] Y. R. Kim and G. J. Kim, "Hovr-type: Smartphone as a typing interface in vr using hovering," in *Proceedings of the ACM Conference on Virtual Reality Software and Technology*, ser. VRST '16. ACM, 2016, pp. 333–334.

[193] L. Zhang, H. Bai, M. Billinghurst, and W. He, "Is this my phone? operating a physical smartphone in virtual reality," in *SIGGRAPH Asia 2020 XR*, 2020, pp. 1–2.

[194] F. Matulic, A. Ganeshan, H. Fujiwara, and D. Vogel, "Phonetroller: Visual representations of fingers for precise touch input with mobile phones in vr," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3411764.3445583

[195] F. Zhu and T. Grossman, "Bishare: Exploring bidirectional interactions between smartphones and head-mounted augmented reality," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–14. [Online]. Available: https://doi.org/10.1145/3313831.3376233

[196] R. Darbar, A. Prouzeau, J. Odicio-Vilchez, T. Lainé, and M. Hachet, "Exploring smartphone-enabled text selection in ar-hmd," in *Graphics Interface 2021*, 2021.

[197] L. H. Lee, Y. Zhu, Y.-P. Yau, P. Hui, and S. Pirttikangas, "Press-n-paste: Copy-and-paste operations with pressure-sensitive caret navigation for miniaturized surface in mobile augmented reality," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. EICS, pp. 1–29, 2021.

[198] D. Yu, K. Fan, H. Zhang, D. Monteiro, W. Xu, and H.-N. Liang, "Pizzatext: Text entry for virtual reality systems using dual thumbsticks," *IEEE transactions on visualization and computer graphics*, vol. 24, no. 11, pp. 2927–2935, 2018.

[199] Q. Zhou, G. Fitzmaurice, and F. Anderson, "In-depth mouse: Integrating desktop mouse into virtual reality," in *CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–17.

[200] C. Ware and K. Lowther, "Selection using a one-eyed cursor in a fish tank vr environment," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 4, no. 4, pp. 309–322, 1997.

[201] M. A. Nacenta, S. Sallam, B. Champoux, S. Subramanian, and C. Gutwin, "Perspective cursor: perspective-based interaction for multi-display environments," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 2006, pp. 289–298.

[202] M. Nancel, E. Pietriga, O. Chapuis, and M. Beaudouin-Lafon, "Mid-air pointing on ultra-walls," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 22, no. 5, pp. 1–62, 2015.

[203] I. Choi, E. Ofek, H. Benko, M. Sinclair, and C. Holz, "Claw: A multifunctional handheld haptic controller for grasping, touching, and triggering in virtual reality," in *Proceedings of the 2018 CHI conference on human factors in computing systems*, 2018, pp. 1–13.

[204] J. Lee, M. Sinclair, M. Gonzalez-Franco, E. Ofek, and C. Holz, "Torc: A virtual reality controller for in-hand high-dexterity finger interaction," in *Proceedings of the 2019 CHI conference on human factors in computing systems*, 2019, pp. 1–13.

[205] H. Benko, C. Holz, M. Sinclair, and E. Ofek, "Normaltouch and texturetouch: High-fidelity 3d haptic shape rendering on handheld virtual reality controllers," in *Proceedings of the 29th annual symposium on user interface software and technology*, 2016, pp. 717–728.

[206] R. Takada, T. Isomoto, W. Yamada, H. Manabe, and B. Shizuki, "Extensionclip: Touch point transfer device linking both sides of a smartphone for mobile vr environments," in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–6.

[207] J. Gugenheimer, D. Dobbelstein, C. Winkler, G. Haas, and E. Rukzio, "Facetouch: Enabling touch interaction in display fixed uis for mobile virtual reality," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 2016, pp. 49–60.

[208] C. Gao, X. Zhang, and S. Banerjee, "Conductive inkjet printed passive 2d trackpad for vr interaction," in *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, 2018, pp. 83–98.

[209] A. Ohlei, T. Winkler, D. Wessel, and M. Herczeg, "Evaluation of direct manipulation methods in augmented reality environments using google glass." in *ISMAR Adjunct*, 2018, pp. 266–269.

[210] J. Lee, B. Kim, B. Suh, and E. Koh, "Exploring the front touch interface for virtual reality headsets," in *Proceedings of the 2016 chi conference extended abstracts on human factors in computing systems*, 2016, pp. 2585–2591.

[211] C.-Y. Wang, M.-C. Hsiu, P.-T. Chiu, C.-H. Chang, L. Chan, B.-Y. Chen, and M. Y. Chen, "Palmgesture: Using palms as gesture interfaces for eyes-free input," in *Proceedings of the 17th*

International Conference on Human-Computer Interaction with Mobile Devices and Services, 2015, pp. 217–226.

[212] C.-Y. Wang, W.-C. Chu, P.-T. Chiu, M.-C. Hsiu, Y.-H. Chiang, and M. Y. Chen, "Palmtype: Using palms as keyboards for smart glasses," in *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 2015, pp. 153–160.

[213] T. Azai, S. Ogawa, M. Otsuki, F. Shibata, and A. Kimura, "Selection and manipulation methods for a menu widget on the human forearm," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 357–360. [Online]. Available: https://doi.org/10.1145/3027063.3052959

[214] T. Azai, S. Ushiro, J. Li, M. Otsuki, F. Shibata, and A. Kimura, "Tap-tap menu: body touching for virtual interactive menus," in *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, 2018, pp. 1–2.

[215] B. Yi, X. Cao, M. Fjeld, and S. Zhao, "Exploring user motivations for eyes-free interaction on mobile devices," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012, pp. 2789–2792.

[216] M. Mahmoud and P. Robinson, "Interpreting hand-over-face gestures," in *International Conference on Affective Computing and Intelligent Interaction*. Springer, 2011, pp. 248–255.

[217] D. Lee, Y. Lee, Y. Shin, and I. Oakley, "Designing socially acceptable hand-to-face input," in *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 711–723. [Online]. Available: https://doi.org/10.1145/3242587.3242642

[218] K. Yamashita, T. Kikuchi, K. Masai, M. Sugimoto, B. H. Thomas, and Y. Sugiura, "Cheekinput: turning your cheek into an input surface by embedded optical sensors on a head-mounted display," in *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, 2017, pp. 1–8.

[219] Y. Weng, C. Yu, Y. Shi, Y. Zhao, Y. Yan, and Y. Shi, "Facesight: Enabling hand-to-face gesture interaction on ar glasses with a downward-facing camera vision," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–14.

[220] R. Lissermann, J. Huber, A. Hadjakos, and M. Mühlhäuser, "Earput: Augmenting behind-the-ear devices for ear-based interaction," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, 2013, pp. 1323–1328.

[221] D. Dobbelstein, P. Hock, and E. Rukzio, "Belt: An unobtrusive touch input device for head-worn displays," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 2135–2138.

[222] J. Henderson, J. Ceha, and E. Lank, "Stat: Subtle typing around the thigh for head-mounted displays," in *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*, 2020, pp. 1–11.

[223] J. Gugenheimer, "Nomadic virtual reality: Exploring new interaction concepts for mobile virtual reality head-mounted displays," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 2016, pp. 9–12.

[224] S. Yi, Z. Qin, E. Novak, Y. Yin, and Q. Li, "Glassgesture: Exploring head gesture interface of smart glasses," in *IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications*. IEEE, 2016, pp. 1–9.

[225] Y. Yan, C. Yu, X. Yi, and Y. Shi, "Headgesture: Hands-free input approach leveraging head movements for hmd devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 4, pp. 1–23, 2018.

[226] C. Yu, Y. Gu, Z. Yang, X. Yi, H. Luo, and Y. Shi, "Tap, dwell or gesture? exploring head-based text entry techniques for hmds," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 4479–4488. [Online]. Available: https://doi.org/10.1145/3025453.3025964

[227] K. Dobosz, D. Popanda, and A. Sawko, "Head-based text entry methods for motor-impaired people," in *International Conference on Man–Machine Interactions*. Springer, 2019, pp. 3–11.

[228] W. Xu, H.-N. Liang, Y. Zhao, T. Zhang, D. Yu, and D. Monteiro, "Ringtext: Dwell-free and hands-free text entry for mobile head-mounted displays using head motions," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 5, pp. 1991–2001, 2019.

[229] T. Hutchinson, K. White, W. Martin, K. Reichert, and L. Frey, "Human-computer interaction using eye-gaze input," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, no. 6, pp. 1527–1534, 1989.

[230] C. H. Morimoto and M. R. Mimica, “Eye gaze tracking techniques for interactive applications,” *Computer vision and image understanding*, vol. 98, no. 1, pp. 4–24, 2005.

[231] R. J. Jacob, “What you look at is what you get: eye movement-based interaction techniques,” in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1990, pp. 11–18.

[232] L. E. Sibert and R. J. Jacob, “Evaluation of eye gaze interaction,” in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 2000, pp. 281–288.

[233] H. M. Park, S. H. Lee, and J. S. Choi, “Wearable augmented reality system using gaze interaction,” in *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. IEEE, 2008, pp. 175–176.

[234] P. Majaranta, U.-K. Ahola, and O. Špakov, “Fast gaze typing with an adjustable dwell time,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2009, pp. 357–360.

[235] T. Toyama, D. Sonntag, A. Dengel, T. Matsuda, M. Iwamura, and K. Kise, “A mixed reality head-mounted text translation system using eye gaze input,” in *Proceedings of the 19th international conference on Intelligent User Interfaces*, 2014, pp. 329–334.

[236] A. Hyrskykari, H. Istance, and S. Vickers, “Gaze gestures or dwell-based interaction?” in *Proceedings of the Symposium on Eye Tracking Research and Applications*, 2012, pp. 229–232.

[237] W. Delamare, T. Han, and P. Irani, “Designing a gaze gesture guiding system,” in *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 2017, pp. 1–13.

[238] T. Hirzle, J. Gugenheimer, F. Geiselhart, A. Bulling, and E. Rukzio, “A design space for gaze interaction on head-mounted displays,” in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.

[239] W. Chen, X. Cui, J. Zheng, J. Zhang, S. Chen, and Y. Yao, “Gaze gestures and their applications in human-computer interaction with a head-mounted display,” *arXiv preprint arXiv:1910.07428*, 2019.

[240] M. Vidal, A. Bulling, and H. Gellersen, “Pursuits: spontaneous interaction with displays based on smooth pursuit eye movement and moving targets,” in *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, 2013, pp. 439–448.

[241] M. Khamis, C. Oechsner, F. Alt, and A. Bulling, "Vrpursuits: interaction in virtual reality using smooth pursuit eye movements," in *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, 2018, pp. 1–8.

[242] L. Sidenmark, C. Clarke, X. Zhang, J. Phu, and H. Gellersen, "Outline pursuits: Gaze-assisted selection of occluded objects in virtual reality," in *Proceedings of the 2020 chi conference on human factors in computing systems*, 2020, pp. 1–13.

[243] R. Aigner, D. Wigdor, H. Benko, M. Haller, D. Lindbauer, A. Ion, S. Zhao, and J. Koh, "Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for hci," *Microsoft Research TechReport MSR-TR-2012-111*, vol. 2, p. 30, 2012.

[244] S. Sridhar, A. M. Feit, C. Theobalt, and A. Oulasvirta, "Investigating the dexterity of multi-finger input for mid-air text entry," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15. New York, NY, USA: ACM, 2015, pp. 3643–3652. [Online]. Available: http://doi.acm.org/10.1145/2702123.2702136

[245] S. Lin, H. F. Cheng, W. Li, Z. Huang, P. Hui, and C. Peylo, "Ubii: Physical world interaction through augmented reality," *IEEE Transactions on Mobile Computing*, vol. 16, no. 3, pp. 872–885, 2016.

[246] L. H. Lee, T. Braud, K. Y. Lam, Y. P. Yau, and P. Hui, "From seen to unseen: Designing keyboard-less interfaces for text entry on the constrained screen real estate of augmented reality headsets," *Pervasive and Mobile Computing*, vol. 64, p. 101148, 2020.

[247] E. Brasier, O. Chapuis, N. Ferey, J. Vezien, and C. Appert, "Arpads: Mid-air indirect input for augmented reality," in *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 2020, pp. 332–343.

[248] H. Zhang, Y. Yin, L. Xie, and S. Lu, "Airtyping: a mid-air typing scheme based on leap motion," in *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*, 2020, pp. 168–171.

[249] C. R. Foy, J. J. Dudley, A. Gupta, H. Benko, and P. O. Kristensson, "Understanding, detecting and mitigating the effects of coactivations in ten-finger mid-air typing in virtual reality," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–11.

[250] M. Ogata, Y. Sugiura, H. Osawa, and M. Imai, "iring: intelligent ring using infrared reflection," in *Proceedings of the 25th annual ACM symposium on User interface software and technology*, 2012, pp. 131–136.

[251] F. S. Parizi, E. Whitmire, and S. Patel, "Auraring: Precise electromagnetic finger tracking," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 4, pp. 1–28, 2019.

[252] K.-Y. Chen, K. Lyons, S. White, and S. Patel, "utrack: 3d input using two magnetic sensors," in *Proceedings of the 26th annual ACM symposium on User interface software and technology*, 2013, pp. 237–244.

[253] L. Chan, R.-H. Liang, M.-C. Tsai, K.-Y. Cheng, C.-H. Su, M. Y. Chen, W.-H. Cheng, and B.-Y. Chen, "Fingerpad: private and subtle interaction using fingertips," in *Proceedings of the 26th annual ACM symposium on User interface software and technology*, 2013, pp. 255–260.

[254] Z. Xu, P. C. Wong, J. Gong, T.-Y. Wu, A. S. Nittala, X. Bi, J. Steimle, H. Fu, K. Zhu, and X.-D. Yang, "Tiptext: Eyes-free text entry on a fingertip keyboard," in *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, 2019, pp. 883–899.

[255] J. Kim, W. Delamare, and P. Irani, "Thumbtext: Text entry for wearable devices using a miniature ring," in *Graphics Interface*, 2018.

[256] R. Boldu, A. Dancu, D. J. Matthies, P. G. Cascón, S. Ransir, and S. Nanayakkara, "Thumb-in-motion: Evaluating thumb-to-ring microgestures for athletic activity," in *Proceedings of the Symposium on Spatial User Interaction*, ser. SUI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 150–157. [Online]. Available: https://doi.org/10.1145/3267782.3267796

[257] F. Kuester, M. Chen, M. E. Phair, and C. Mehring, "Towards keyboard independent touch typing in vr," in *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '05. New York, NY, USA: Association for Computing Machinery, 2005, p. 86–95. [Online]. Available: https://doi.org/10.1145/1101616.1101635

[258] A. Peshock, J. Duvall, and L. E. Dunne, "Argot: A wearable one-handed keyboard glove," in *Proceedings of the 2014 ACM International Symposium on Wearable Computers: Adjunct Program*, ser. ISWC '14 Adjunct. New York, NY, USA: Association for Computing Machinery, 2014, p. 87–92. [Online]. Available: https://doi.org/10.1145/2641248.2641266

[259] L. H. Lee, K. Y. Lam, T. Li, T. Braud, X. Su, and P. Hui, "Quadmetric optimized thumb-to-finger interaction for force assisted one-handed text entry on mobile headsets," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 3, no. 3, sep 2019. [Online]. Available: https://doi.org/10.1145/3351252

[260] E. Whitmire, M. Jain, D. Jain, G. Nelson, R. Karkar, S. Patel, and M. Goel, "Digitouch: Reconfigurable thumb-to-finger input and text entry on head-mounted displays," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 3, sep 2017. [Online]. Available: https://doi.org/10.1145/3130978

[261] H.-R. Tsai, T.-Y. Wu, D.-Y. Huang, M.-C. Hsiu, J.-C. Hsiao, Y.-P. Hung, M. Y. Chen, and B.-Y. Chen, "Segtouch: Enhancing touch input while providing touch gestures on screens using thumb-to-index-finger gestures," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 2164–2171. [Online]. Available: https://doi.org/10.1145/3027063.3053109

[262] D. Lee, J. Kim, and I. Oakley, "Fingertext: Exploring and optimizing performance for wearable, mobile and one-handed typing," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3411764.3445106

[263] L. Chan, Y.-L. Chen, C.-H. Hsieh, R.-H. Liang, and B.-Y. Chen, "Cyclopsring: Enabling whole-hand and context-aware interactions through a fisheye ring," in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, 2015, pp. 549–556.

[264] F. Guimbretière and C. Nguyen, "Bimanual marking menu for near surface interactions," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2012, pp. 825–828.

[265] M. Prätorius, D. Valkov, U. Burgbacher, and K. Hinrichs, "Digitap: An eyes-free vr/ar symbolic input device," in *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 9–18. [Online]. Available: https://doi.org/10.1145/2671015.2671029

[266] M. Soliman, F. Mueller, L. Hegemann, J. S. Roo, C. Theobalt, and J. Steimle, "Fingerinput: Capturing expressive single-hand thumb-to-finger microgestures," in *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, ser. ISS '18. New York,

NY, USA: Association for Computing Machinery, 2018, p. 177–187. [Online]. Available: https://doi.org/10.1145/3279778.3279799

[267] J. Fashimpaur, K. Kin, and M. Longest, "Pinchtype: Text entry for virtual and augmented reality using comfortable thumb to fingertip pinches," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–7. [Online]. Available: https://doi.org/10.1145/3334480.3382888

[268] B. Lafreniere, C. Gutwin, A. Cockburn, and T. Grossman, "Faster command selection on touchscreen watches," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: ACM, 2016, pp. 4663–4674. [Online]. Available: http://doi.acm.org/10.1145/2858036.2858166

[269] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Human Mental Workload*, ser. Advances in Psychology, P. A. Hancock and N. Meshkati, Eds. North-Holland, 1988, vol. 52, pp. 139–183. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166411508623869

[270] G. A. Borg, "Psychophysical bases of perceived exertion," *Med sci sports exerc*, vol. 14, no. 5, pp. 377–381, 1982.

[271] A. Colley and J. Häkkilä, "Exploring finger specific touch screen interaction for mobile phone user interfaces," in *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design*, ser. OzCHI '14. New York, NY, USA: ACM, 2014, pp. 539–548. [Online]. Available: http://doi.acm.org/10.1145/2686612.2686699

[272] R. Xiao, G. Laput, and C. Harrison, "Expanding the input expressivity of smartwatches with mechanical pan, twist, tilt and click," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '14. New York, NY, USA: ACM, 2014, pp. 193–196. [Online]. Available: http://doi.acm.org/10.1145/2556288.2557017

[273] H. V. Le, S. Mayer, P. Bader, and N. Henze, "Fingers' range and comfortable area for one-handed smartphone interaction beyond the touchscreen," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: ACM, 2018, pp. 31:1–31:12. [Online]. Available: http://doi.acm.org/10.1145/3173574.3173605

[274] E. Hoggan, J. Williamson, A. Oulasvirta, M. Nacenta, P. O. Kristensson, and A. Lehtiö, "Multi-touch rotation gestures: Performance and ergonomics," in *Proceedings of the SIGCHI*

Conference on Human Factors in Computing Systems, ser. CHI '13. New York, NY, USA: ACM, 2013, pp. 3047–3050. [Online]. Available: http://doi.acm.org/10.1145/2470654.2481423

[275] Q. Nguyen and M. Kipp, "Orientation matters: Efficiency of translation-rotation multitouch tasks," in *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '14. New York, NY, USA: ACM, 2014, pp. 2013–2016. [Online]. Available: http://doi.acm.org/10.1145/2556288.2557399

[276] R. W. Soukoreff and I. S. MacKenzie, "Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in hci," *Int. J. Hum.-Comput. Stud.*, vol. 61, no. 6, pp. 751–789, Dec. 2004. [Online]. Available: http://dx.doi.org/10.1016/j.ijhcs.2004.09.001

[277] J. Zhao, R. William Soukoreff, and R. Balakrishnan, "Exploring and modeling unimanual object manipulation on multi-touch displays," *Int. J. Hum.-Comput. Stud.*, vol. 78, no. C, pp. 68–80, Jun. 2015. [Online]. Available: http://dx.doi.org/10.1016/j.ijhcs.2015.02.011

[278] S. Voelker, K. I. Øvergård, C. Wacharamanotham, and J. Borchers, "Knobology revisited: A comparison of user performance between tangible and virtual rotary knobs," in *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*, ser. ITS '15. New York, NY, USA: ACM, 2015, pp. 35–38. [Online]. Available: http://doi.acm.org/10.1145/2817721.2817725

[279] S. Mayer, H. V. Le, and N. Henze, "Designing finger orientation input for mobile touchscreens," in *Proceedings of the 2018 ACM International Conference on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '18. New York, NY, USA: ACM, 2018, p. 1. [Online]. Available: http://doi.acm.org/10.1145/3229434.3229444

[280] M. Bachynskyi, G. Palmas, A. Oulasvirta, J. Steimle, and T. Weinkauf, "Performance and ergonomics of touch surfaces: A comparative study using biomechanical simulation," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15. New York, NY, USA: ACM, 2015, pp. 1817–1826. [Online]. Available: http://doi.acm.org/10.1145/2702123.2702607

[281] C. Holz, S. Buthpitiya, and M. Knaust, "Bodyprint: Biometric user identification on mobile devices using the capacitive touchscreen to scan body parts," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15. New York, NY, USA: ACM, 2015, pp. 3011–3014. [Online]. Available: http://doi.acm.org/10.1145/2702123.2702518

[282] H. V. Le, S. Mayer, P. Bader, and N. Henze, "A smartphone prototype for touch interaction on the whole device surface," in *Proceedings of the 19th International*

Conference on Human-Computer Interaction with Mobile Devices and Services, ser. MobileHCI '17. New York, NY, USA: ACM, 2017, pp. 100:1–100:8. [Online]. Available: http://doi.acm.org.library.unist.ac.kr:8080/10.1145/3098279.3122143

[283] H. Gil, D. Lee, S. Im, and I. Oakley, "Tritap: Identifying finger touches on smartwatches," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 3879–3890. [Online]. Available: https://doi.org/10.1145/3025453.3025561

[284] S. Mayer, H. V. Le, and N. Henze, "Estimating the finger orientation on capacitive touchscreens using convolutional neural networks," in *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, ser. ISS '17. New York, NY, USA: ACM, 2017, pp. 220–229. [Online]. Available: http://doi.acm.org/10.1145/3132272.3134130

[285] G. Norman, "Likert scales, levels of measurement and the "laws" of statistics," *Advances in health sciences education*, vol. 15, no. 5, pp. 625–632, 2010.

[286] J. D. Evans, *Straightforward statistics for the behavioral sciences*. Brooks/Cole, 1996.

[287] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: An update," *SIGKDD Explor. Newsl.*, vol. 11, no. 1, pp. 10–18, Nov. 2009. [Online]. Available: http://doi.acm.org/10.1145/1656274.1656278

[288] M. S. Hancock, S. Carpendale, F. D. Vernier, D. Wigdor, and C. Shen, "Rotation and translation mechanisms for tabletop interaction," in *Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, ser. TABLETOP '06. Washington, DC, USA: IEEE Computer Society, 2006, pp. 79–88. [Online]. Available: https://doi.org/10.1109/TABLETOP.2006.26

[289] X. A. Chen, T. Grossman, D. J. Wigdor, and G. Fitzmaurice, "Duet: Exploring joint interactions on a smart phone and a smart watch," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '14. New York, NY, USA: ACM, 2014, pp. 159–168. [Online]. Available: http://doi.acm.org/10.1145/2556288.2556955

[290] A. Roudaut, E. Lecolinet, and Y. Guiard, "Microrolls: Expanding touch-screen input vocabulary by distinguishing rolls vs. slides of the thumb," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '09. New York, NY, USA: ACM, 2009, pp. 927–936. [Online]. Available: http://doi.acm.org/10.1145/1518701.1518843

[291] C. Stewart, M. Rohs, S. Kratz, and G. Essl, "Characteristics of pressure-based input for mobile devices," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2010, pp. 801–810.

[292] Q. Roy, Y. Guiard, G. Bailly, É. Lecolinet, and O. Rioul, "Glass+ skin: An empirical evaluation of the added value of finger identification to basic single-touch interaction on touch screens," in *IFIP Conference on Human-Computer Interaction*. Springer, 2015, pp. 55–71.

[293] A. Goguey, M. Nancel, G. Casiez, and D. Vogel, "The performance and preference of different fingers and chords for pointing, dragging, and object transformation," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 4250–4261.

[294] E. Ghomi, S. Huot, O. Bau, M. Beaudouin-Lafon, and W. E. Mackay, "Arpège: learning multi-touch chord gestures vocabularies," in *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces*, 2013, pp. 209–218.

[295] C. Holz and P. Baudisch, "Fiberio: a touchscreen that senses fingerprints," in *Proceedings of the 26th annual ACM symposium on User interface software and technology*, 2013, pp. 41–50.

[296] M. Weigel, T. Lu, G. Bailly, A. Oulasvirta, C. Majidi, and J. Steimle, "Iskin: flexible, stretchable and visually customizable on-body touch sensors for mobile computing," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 2991–3000.

[297] T. S. Saponas, D. S. Tan, D. Morris, J. Turner, and J. A. Landay, "Making muscle-computer interfaces more practical," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2010, pp. 851–854.

[298] P. Ewerling, A. Kulik, and B. Froehlich, "Finger and hand detection for multi-touch interfaces based on maximally stable extremal regions," in *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*, 2012, pp. 173–182.

[299] I. S. MacKenzie and R. W. Soukoreff, "Phrase sets for evaluating text entry techniques," in *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '03. New York, NY, USA: Association for Computing Machinery, 2003, p. 754–755. [Online]. Available: https://doi.org/10.1145/765891.765971

[300] M. Yasumoto and T. Teraoka, "Shadow shooter," in *Proceedings of the 2015 Virtual Reality International Conference*, ser. VRIC'15. ACM, 2015, p. Article No. 17.

[301] Y. Chen, X. Su, F. Tian, J. Huang, X. L. Zhang, G. Dai, and H. Wang, "Pactolus: A method for mid-air gesture segmentation within emg," in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA'16. ACM, 2016, pp. 1760–1765.

[302] I. Choi and S. Follmer, "Wolverine: A wearable haptic interface for grasping in vr," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, ser. UIST'16. ACM, 2016, pp. 117–119.

[303] J. Walker, B. Li, K. Vertanen, and S. Kuhl, "Efficient typing on a visually occluded physical keyboard," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 5457–5461. [Online]. Available: https://doi.org/10.1145/3025453.3025783

[304] P. Knierim, V. Schwind, A. M. Feit, F. Nieuwenhuizen, and N. Henze, "Physical keyboards in virtual reality: Analysis of typing performance and effects of avatar hands," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 2018, p. 345.

[305] M. Lee and W. Woo, "Arkb: 3d vision-based augmented reality keyboard," in *Proceedings of the 13th International Conference on Artificial Reality and Telexistence*, ser. ICAT '03, 2003, pp. 54–57.

[306] H.-S. Yeo, X.-S. Phang, T. Ha, W. Woo, and A. Quigley, "Titan: Exploring midair text entry using freehand input," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 3041–3049. [Online]. Available: https://doi.org/10.1145/3027063.3053228

[307] A. Oulasvirta, S. Kim, and B. Lee, "Neuromechanics of a button press," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 2018, p. 508.

[308] T. A. Salthouse, "Perceptual, cognitive, and motoric aspects of transcription typing," *Psychological Bulletin*, vol. 99, no. 3, pp. 303–319, 1986.

[309] M. McGill, D. Boland, R. Murray-Smith, and S. Brewster, "A dose of reality: Overcoming usability challenges in vr head-mounted displays," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 2143–2152. [Online]. Available: https://doi.org/10.1145/2702123.2702382

[310] D. Buschek, O. Schoenleben, and A. Oulasvirta, "Improving accuracy in back-of-device multitouch typing: a clustering-based approach to keyboard updating," in *Proceedings of the 19th international conference on Intelligent User Interfaces*. ACM, 2014, pp. 57–66.

[311] S. Sarcar, J. P. Jokinen, A. Oulasvirta, Z. Wang, C. Silpasuwanchai, and X. Ren, "Ability-based optimization of touchscreen interactions," *IEEE Pervasive Computing*, vol. 17, no. 1, pp. 15–26, 2018.

[312] L. Findlater, J. O. Wobbrock, and D. Wigdor, "Typing on flat glass: Examining ten-finger expert typing patterns on touch surfaces," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 2453–2462. [Online]. Available: https://doi.org/10.1145/1978942.1979301

[313] L. Findlater and J. Wobbrock, "Personalized input: Improving ten-finger touchscreen typing through automatic adaptation," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '12. New York, NY, USA: Association for Computing Machinery, 2012, p. 815–824. [Online]. Available: https://doi.org/10.1145/2207676.2208520

[314] J. R. Kim and H. Z. Tan, "A study of touch typing performance with keyclick feedback," in *2014 IEEE Haptics Symposium (HAPTICS)*. USA: IEEE Computer Society, 2014, pp. 227–233.

[315] ——, "Haptic feedback intensity affects touch typing performance on a flat keyboard," in *Proceedings of EuroHaptics 2014*, ser. EuroHaptics 2014. Springer, 2014, pp. 369–375.

[316] ——, "Effect of information content in sensory feedback on typing performance using a flat keyboard," in *Proceedings of the IEEE World Haptics Conference*, ser. WHC2015. IEEE, 2015, pp. 228–234.

[317] J. Goodman, G. Venolia, K. Steury, and C. Parker, "Language modeling for soft keyboards," in *Proceedings of the ACM Intelligent User Interfaces*, ser. IUI '02. ACM, 2002, pp. 194–195.

[318] J. O. Wobbrock and B. A. Myers, "Analyzing the input stream for character-level errors in unconstrained text entry evaluations," *ACM TOCHI*, vol. 13, no. 4, pp. 458–489, 2014.

[319] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of the leap motion controller," *Sensors*, vol. 13, no. 5, pp. 6380–6393, 2013.

[320] S. Zhu, T. Luo, X. Bi, and S. Zhai, "Typing on an invisible keyboard," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI

’18. New York, NY, USA: ACM, 2018, pp. 439:1–439:13. [Online]. Available: http://doi.acm.org/10.1145/3173574.3174013

[321] A. M. Feit, D. Weir, and A. Oulasvirta, “How we type: Movement strategies and performance in everyday typing,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’16. New York, NY, USA: ACM, 2016, pp. 4262–4273. [Online]. Available: http://doi.acm.org/10.1145/2858036.2858233

[322] P. C. Price, R. Jhangiani, I.-C. A. Chiang *et al.*, *Research methods in psychology*. BCCampus, 2015.

[323] Letterfrequency, “Letterfrequency: Digraph frequency,” 2018, http://letterfrequency.org/.

[324] T. Carger, S. A. Seah, B. Long, B. Drinkwater, and S. Subramanian, “Ultrahaptics: Multi-point mid-air haptic feedback for touch surfaces,” in *Proceedings of the ACM Symposium on User Interface Software and Technology*, ser. UIST ’13. ACM, 2013, pp. 505–514.

[325] I. Hwang, H. Son, and J. R. Kim, “Airpiano: Enhancing music playing experience in virtual reality with mid-air haptic feedback,” in *Proceedings of IEEE World Haptics Conference*, ser. WHC ’17. IEEE, 2017, pp. 213–218.

[326] B. McCaul and A. Sutherland, “Predictive text entry in immersive environments,” in *IEEE Virtual Reality 2004*. IEEE, 2004, pp. 241–242.

[327] P. M. Fitts and J. R. Peterson, “Information capacity of discrete motor responses.” *Journal of experimental psychology*, vol. 67, no. 2, p. 103, 1964.

[328] C. Cruz-Neira, M. Fernández, and C. Portalés, “Virtual reality and games,” 2018.

[329] G. I. Lee and M. R. Lee, “Can a virtual reality surgical simulation training provide a self-driven and mentor-free skills learning? investigation of the practical influence of the performance metrics from the virtual reality robotic surgery simulator on the skill learning and associated cognitive workloads,” *Surgical endoscopy*, vol. 32, no. 1, pp. 62–72, 2018.

[330] M. A. Livingston, L. J. Rosenblum, D. G. Brown, G. S. Schmidt, S. J. Julier, Y. Baillot, J. E. Swan, Z. Ai, and P. Maassel, “Military applications of augmented reality,” pp. 671–706, 2011. [Online]. Available: https://doi.org/10.1007/978-1-4614-0064-6_31

[331] F. Manuri and A. Sanna, “A survey on applications of augmented reality,” *ACSIJ Advances in Computer Science: an International Journal*, vol. 5, no. 1, pp. 18–27, 2016.

[332] C. Pollalis, A. Gilvin, L. Westendorf, L. Futami, B. Virgilio, D. Hsiao, and O. Shaer, "Artlens: Enhancing museum visitors' engagement with african art," in *Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems*, ser. DIS '18 Companion. New York, NY, USA: Association for Computing Machinery, 2018, p. 195–200. [Online]. Available: https://doi.org/10.1145/3197391.3205435

[333] A. Burova, J. Mäkelä, J. Hakulinen, T. Keskinen, H. Heinonen, S. Siltanen, and M. Turunen, "Utilizing vr and gaze tracking to develop ar solutions for industrial maintenance," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–13. [Online]. Available: https://doi.org/10.1145/3313831.3376405

[334] S. Rajaram and M. Nebeling, "Paper trail: An immersive authoring system for augmented reality instructional experiences," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, ser. CHI '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3491102.3517486

[335] R. Guo, Y. Yang, J. Kuang, X. Bin, D. Jain, S. Goodman, L. Findlater, and J. Froehlich, "Holosound: Combining speech and sound identification for deaf or hard of hearing users on a head-mounted display," in *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, ser. ASSETS '20. New York, NY, USA: Association for Computing Machinery, 2020. [Online]. Available: https://doi.org/10.1145/3373625.3418031

[336] R. Darbar, J. Odicio-Vilchez, T. Lainé, A. Prouzeau, and M. Hachet, "Text selection in ar-hmd using a smartphone as an input device," in *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, IEEE. USA: IEEE Computer Society, 2021, pp. 526–527.

[337] A. Oulasvirta, A. Reichel, W. Li, Y. Zhang, M. Bachynskyi, K. Vertanen, and P. O. Kristensson, "Improving two-thumb text entry on touchscreen devices," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '13. New York, NY, USA: Association for Computing Machinery, 2013, p. 2765–2774. [Online]. Available: https://doi.org/10.1145/2470654.2481383

[338] M. Goel, L. Findlater, and J. Wobbrock, "Walktype: Using accelerometer data to accomodate situational impairments in mobile touch screen text entry," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '12. New York, NY,

USA: Association for Computing Machinery, 2012, p. 2687–2696. [Online]. Available: https://doi.org/10.1145/2207676.2208662

[339] E. Clarkson, J. Clawson, K. Lyons, and T. Starner, "An empirical study of typing rates on mini-qwerty keyboards," in *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '05. New York, NY, USA: Association for Computing Machinery, 2005, p. 1288–1291. [Online]. Available: https://doi.org/10.1145/1056808.1056898

[340] I. S. MacKenzie and R. W. Soukoreff, "Text entry for mobile computing: Models and methods, theory and practice," *Human–Computer Interaction*, vol. 17, no. 2-3, pp. 147–198, 2002.

[341] K. Palin, A. M. Feit, S. Kim, P. O. Kristensson, and A. Oulasvirta, "How do people type on mobile devices? observations from a study with 37,000 volunteers," in *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '19. New York, NY, USA: Association for Computing Machinery, 2019. [Online]. Available: https://doi.org/10.1145/3338286.3340120

[342] X. Bi, B. A. Smith, and S. Zhai, "Quasi-qwerty soft keyboard optimization," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '10. New York, NY, USA: Association for Computing Machinery, 2010, p. 283–286. [Online]. Available: https://doi.org/10.1145/1753326.1753367

[343] D.-M. Pham and W. Stuerzlinger, "Hawkey: Efficient and versatile text entry for virtual reality," in *25th ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '19. New York, NY, USA: Association for Computing Machinery, 2019. [Online]. Available: https://doi.org/10.1145/3359996.3364265

[344] J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson, "Text entry in immersive head-mounted display-based virtual reality using standard keyboards," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. USA: IEEE Computer Society, 2018, pp. 159–166.

[345] M. McGill, S. Brewster, D. P. De Sa Medeiros, S. Bovet, M. Gutierrez, and A. Kehoe, "Creating and augmenting keyboards for extended reality with the k eyboard a ugmentation t oolkit," *ACM Transactions on Computer-Human Interaction*, vol. 29, no. 2, pp. 1–39, 2022.

[346] M. R. Zhang, S. Zhai, and J. O. Wobbrock, "Typeanywhere: A qwerty-based text entry solution for ubiquitous computing," in *Proceedings of the 2022 CHI Conference on Human Factors in*

Computing Systems, ser. CHI '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3491102.3517686

[347] A. Gupta, C. Ji, H.-S. Yeo, A. Quigley, and D. Vogel, "Rotoswype: Word-gesture typing using a ring," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1–12. [Online]. Available: https://doi.org/10.1145/3290605.3300244

[348] S. Jang, W. Stuerzlinger, S. Ambike, and K. Ramani, "Modeling cumulative arm fatigue in mid-air interaction based on perceived exertion and kinetics of arm motion," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 3328–3339. [Online]. Available: https://doi.org/10.1145/3025453.3025523

[349] D. Ahlström, K. Hasan, and P. Irani, "Are you comfortable doing that? acceptance studies of around-device gestures in and for public settings," in *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services*, ser. MobileHCI '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 193–202. [Online]. Available: https://doi.org/10.1145/2628363.2628381

[350] H. Gil, Y. Shin, H. Son, I. Hwang, I. Oakley, and J. R. Kim, "Characterizing in-air eyes-free typing movements in vr," in *26th ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '20. New York, NY, USA: Association for Computing Machinery, 2020. [Online]. Available: https://doi.org/10.1145/3385956.3418963

[351] H. Jiang, D. Weng, Z. Zhang, and F. Chen, "Hifinger: One-handed text entry technique for virtual environments based on touches between fingers," *Sensors*, vol. 19, no. 14, 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/14/3063

[352] W. Hutama, H. Harashima, H. Ishikawa, and H. Manabe, "Hmk: Head-mounted-keyboard for text input in virtual or augmented reality," in *The Adjunct Publication of the 34th Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 115–117. [Online]. Available: https://doi.org/10.1145/3474349.3480195

[353] Z. Xu, W. Chen, D. Zhao, J. Luo, T.-Y. Wu, J. Gong, S. Yin, J. Zhai, and X.-D. Yang, "Bitiptext: Bimanual eyes-free text entry on a fingertip keyboard," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New

York, NY, USA: Association for Computing Machinery, 2020, p. 1–13. [Online]. Available: https://doi.org/10.1145/3313831.3376306

[354] C. E. Lang and M. H. Schieber, "Human finger independence: limitations due to passive mechanical coupling versus active neuromuscular control," *Journal of neurophysiology*, vol. 92, no. 5, pp. 2802–2810, 2004.

[355] B. I. Binder-Markey and W. M. Murray, "Incorporating the length-dependent passive-force generating muscle properties of the extrinsic finger muscles into a wrist and finger biomechanical musculoskeletal model," *Journal of biomechanics*, vol. 61, pp. 250–257, 2017.

[356] A. Ng, S. A. Brewster, and J. H. Williamson, "Investigating the effects of encumbrance on one- and two- handed interactions with mobile devices," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 1981–1990. [Online]. Available: https://doi.org/10.1145/2556288.2557312

[357] D. Lee, S. Lee, and I. Oakley, "Nailz: Sensing hand input with touch sensitive nails," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–13. [Online]. Available: https://doi.org/10.1145/3313831.3376778

[358] G. W. Lesher, B. J. Moulton, and D. J. Higginbotham, "Optimal character arrangements for ambiguous keyboards," *IEEE Transactions on Rehabilitation Engineering*, vol. 6, no. 4, pp. 415–423, 1998.

[359] A. Komninos and M. Dunlop, "Using a smart-watch as a input device for text," *IEEE Pervasive Computing*, vol. 13, no. 4, pp. 50–58, 2014.

[360] A. M. Feit, "Assignment Problems for Optimizing Text Input," Doctoral thesis, School of Electrical Engineering, 2018. [Online]. Available: http://urn.fi/URN:ISBN:978-952-60-8016-1

[361] K. Go and Y. Endo, "Catkey: Customizable and adaptable touchscreen keyboard with bubble cursor-like visual feedback," in *Human-Computer Interaction – INTERACT 2007*, C. Baranauskas, P. Palanque, J. Abascal, and S. D. J. Barbosa, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 493–496.

[362] P. Norvig, "Natural language corpus data," *Beautiful data*, pp. 219–242, 2009.

[363] T. Brants, "Web 1t 5-gram version 1," *http://www. ldc. upenn. edu/Catalog/CatalogEntry. jsp? catalogId= LDC2006T13*, 2006.

[364] D. R. O'Day and R. A. Calix, "Text message corpus: Applying natural language processing to mobile device forensics," in *2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, IEEE. USA: IEEE Computer Society, 2013, pp. 1–6.

[365] K. Vertanen and P. O. Kristensson, "Mining, analyzing, and modeling text written on mobile devices," *Natural Language Engineering*, vol. 27, no. 1, pp. 1–33, 2021.

[366] ——, "A versatile dataset for text entry evaluations based on genuine mobile emails," in *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 295–298. [Online]. Available: https://doi.org/10.1145/2037373.2037418

[367] X. Bi and S. Zhai, "Ijqwerty: What difference does one key change make? gesture typing keyboard optimization bounded by one key position change from qwerty," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 49–58. [Online]. Available: https://doi.org/10.1145/2858036.2858421

[368] S. Zhai and P. O. Kristensson, "Interlaced qwerty: Accommodating ease of visual search and input flexibility in shape writing," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '08. New York, NY, USA: Association for Computing Machinery, 2008, p. 593–596. [Online]. Available: https://doi.org/10.1145/1357054.1357149

[369] S. MacKenzie, "A note on calculating text entry speed," 2021. [Online]. Available: https://www.yorku.ca/mack/RN-TextEntrySpeed.html

[370] R. W. Soukoreff and I. S. MacKenzie, "Metrics for text entry research: An evaluation of msd and kspc, and a new unified error metric," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '03. New York, NY, USA: Association for Computing Machinery, 2003, p. 113–120. [Online]. Available: https://doi.org/10.1145/642611.642632

[371] I. S. MacKenzie and R. W. Soukoreff, "A character-level error analysis technique for evaluating text entry methods," in *Proceedings of the Second Nordic Conference on Human-Computer Interaction*, ser. NordiCHI '02. New York, NY, USA: Association for Computing Machinery, 2002, p. 243–246. [Online]. Available: https://doi.org/10.1145/572020.572056

[372] G. Borg, G. Ljunggren, and R. Ceci, "The increase of perceived exertion, aches and pain in the legs, heart rate and blood lactate during exercise on a bicycle ergometer," *European journal of applied physiology and occupational physiology*, vol. 54, no. 4, pp. 343–349, 1985.

[373] N. Banovic, V. Rao, A. Saravanan, A. K. Dey, and J. Mankoff, "Quantifying aversion to costly typing errors in expert mobile text entry," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 4229–4241. [Online]. Available: https://doi.org/10.1145/3025453.3025695

[374] C.-Y. Hsieh, Y.-S. Chiang, H.-Y. Chiu, and Y.-J. Chang, "Bridging the virtual and real worlds: A preliminary study of messaging notifications in virtual reality," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–14. [Online]. Available: https://doi.org/10.1145/3313831.3376228

[375] Y.-T. Chen, C.-H. Hsu, C.-H. Chung, Y.-S. Wang, and S. V. Babu, "Ivrnote: design, creation and evaluation of an interactive note-taking interface for study and reflection in vr learning environments," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, IEEE. USA: IEEE Computer Society, 2019, pp. 172–180.

[376] C. George, M. Khamis, E. von Zezschwitz, M. Burger, H. Schmidt, F. Alt, and H. Hussmann, "Seamless and secure vr: Adapting and evaluating established authentication systems for virtual reality." NDSS, 2017.

[377] J. Liebers, P. Horn, C. Burschik, U. Gruenefeld, and S. Schneegass, "Using gaze behavior and head orientation for implicit identification in virtual reality," in *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3489849.3489880

[378] M. Pyasik, G. Tieri, and L. Pia, "Visual appearance of the virtual hand affects embodiment in the virtual hand illusion," *Scientific reports*, vol. 10, no. 1, pp. 1–11, 2020.

[379] N. Banovic, T. Sethapakdi, Y. Hari, A. K. Dey, and J. Mankoff, "The limits of expert text entry speed on mobile keyboards with autocorrect," in *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '19. New York, NY, USA: Association for Computing Machinery, 2019. [Online]. Available: https://doi.org/10.1145/3338286.3340126

[380] T. Ogitani, Y. Arahori, Y. Shinyama, and K. Gondow, "Space saving text input method for head mounted display with virtual 12-key keyboard," in *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, IEEE. USA: IEEE Computer Society, 2018, pp. 342–349.

Acknowledgements

First and foremost, I would like to express the deepest appreciation to my supervisor, Prof. Ian Oakley, for his patience, motivation, and immense knowledge. His guidance and advice carried me through the stages of my Ph.D. He has always given support for the in-depth discussions about various research problems that I am interested in. Besides my advisor, I would like to thank the committee members: Prof. Kyungho Lee, Prof. Jaeyeon Lee, Prof. Seongkook Heo, and Prof. Jin Ryong Kim, for their insightful comments and suggestions. In particular, I would like to express my sincere gratitude to Prof. Jin Ryong Kim for leading me working on diverse exciting research projects.

I am grateful to members of Interactions lab: DoYoung, Youngeun, Hongmin, Yonghwan, Eunyong, Jiwan, Youryang, Hyunmi, Mintra, Rasel, Liza, Marian, and Brandon, for their academic support and friendship. They have been helping and encouraging me throughout the years. They have given me a lifetime unforgettable memory at UNIST. I am also so thankful to all my colleagues especially Min-joong, Hyungki, Hyeongmin, Seobin, Eunji, and Soohwan. I hope that our fruitful cooperation in the research field will continue. I would like to extend my sincere thanks to my friends: Sangcheol, Dongha, Ilho, and Moongyu, for encouraging me at challenging moments during this journey.

Words cannot express my gratitude to the participants who contributed to the research during my academic journey. They earnestly followed the study instructions and inspired me with their brilliant ideas.

I am extremely grateful to my parents, brother, and grandparents for their support, patience, and love throughout my life. Last but not least, it is my privilege to thank my awesome wife, Dami. She stood by my side from the very beginning of this long Ph.D. journey. I could not have undertaken this journey without her constant encouragement, strong emotional support, and impeccably delicious meals. She is an important inspiration for me. Her endless love always motivates me to remain focus towards achieving various milestones in life.