Eyes-free Gestural Interaction with Smartphone Using Haptic Feedback

Milad Jamalzadeh

A thesis submitted for the degree of Doctor of Philosophy in Information Technology

Thesis director:
Prof. Laurent Grisoni, Université de Lille, France

Supervisor:
Dr. Yosra Rekik, Université Polytechnique Hauts-de-France, France

These researches were conducted at MINT Group at CRIStAL laboratory of Lille University

JURY Members	:		
Reviewers	:	Prof. Laurence Nigay	Université Grenoble Alpes, France.
		Prof. Jean Vanderdonckt	Université catholique de Louvain, Belgique.
Examinators	:	Dr Anke Brock	ENAC, Toulouse, France.
		Dr Vincent Levesque	École de technologie supérieure, Montréal, Canada.
President	:	Prof. Romain Rouvoy	Université de Lille, France.

Date: October 3, 2024

This project has received funding from the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No 860114.

Interaction Gestuelle Sans Regarder avec le Smartphone en Utilisant le Retour Haptique

Milad Jamalzadeh

Une thèse présentée pour obtenir le titre de Docteur en Informatique

Directeur de thèse :
Prof. Laurent Grisoni, Université de Lille, France

Encadrante de thèse :
Dr. Yosra Rekik, Université Polytechnique Hauts-de-France, France

Ces recherches ont été menées au sein du groupe MINT au laboratoire CRIStAL de l'Université de Lille

MEMBRES DU JURY	:		
RAPPORTEURS	:	PROF. Laurence Nigay	Université Grenoble Alpes, France.
		PROF. Jean Vanderdonckt	Université catholique de Louvain, Belgique.
EXAMINATEURS	:	DR Anke Brock	ENAC, Toulouse, France.
		DR Vincent Levesque	École de technologie supérieure, Montréal, Canada.
PRÉSIDENT	:	PROF. Romain Rouvoy	Université de Lille, France.

Date: 3 octobre 2024

Ce projet a reçu un financement du programme de recherche et d'innovation Horizon 2020 de l'Union européenne au titre de l'accord de subvention n° 860114.

Abstract

Smartphones have become integral to modern life and play a crucial role in various situations, even when users cannot fully focus on the device. In scenarios where visual attention is divided, such as driving, the ability to draw gestures with one hand on a touchscreen without looking at the phone is highly beneficial. These gestures allow users to perform actions like answering calls or sending predefined messages effortlessly. In addition, haptic feedback can partially substitute for visual feedback in visually impaired situations. Haptic feedback can inform users of successful inputs or provide information without requiring visual attention. In this context, this thesis explores the enhancement of one-handed and eyes-free gestural interaction with smartphones through the haptic channel. The research is structured into multiple studies, each addressing different aspects of eyes-free gesture-based interaction with touchscreens.

First, we investigate the production of eyes-free gestures in various situational impairment scenarios. We show differences between how users produce gestures on smartphones in an eyes-free context compared to gestures made when they have direct sight of the phone. This is followed by two studies investigating how environmental factors, such as user movement speed, phone location, and multitasking, affect the production of eyes-free gestures. In the presence of an attention saturating task, users finish producing gestures in a shorter time. However, when a user is involved in a physically demanding task, their finger moves faster in a smaller area.

We then carried out three studies with the aim of improving gestural input on touchscreens using haptic feedback in visually impaired scenarios. We first examine the efficiency of haptic feedback for gestural input on a smartphone placed inside a pocket. Then, we compare localized feedback from the smartphone vibration motor against distal feedback from a smartwatch. The findings suggest that providing the user with continuous haptic feedback improves user confidence in gesture input, with localized feedback preferred. Then, we focused on improving the targeting task on a virtual knob in the presence of an attention-saturating task. Our results showed that the use of background haptic texture on the perimeter of the virtual knob plus haptic texture on the detents increases the performance of users in targeting task on the knob.

Then, we focus on using haptic feedback to improve the output part of the interaction. We present *Hap2Gest*, an interaction concept for smartphones using eyes-free gestures and haptic feedback. Users draw a gesture to invoke a command and receive information through haptic feedback on a subsequent gesture. An elicitation study showed a preference for gestures with clear corners. Circular gestures were also popular in some scenarios, so we tested haptic feedback on free-hand circular gestures. First, we assessed users' perception of the length of eyes-free circular gestures, finding that they perceive the arc length as shorter than it is. Next, we tested users' ability to retrieve numerical data on circular gestures using *Hap2Gest* method. The results showed significant variation in the shapes of freely produced arcs, leading to inaccurate estimates of the arc's location. Thus, we recommend using gestures with clear corners for the *Hap2Gest* interaction technique.

This thesis contributes to the field of human-computer interaction by providing a comprehensive analysis of haptic feedback in eyes-free gesture-based interactions. The findings offer valuable insights for designing effective interaction techniques and improving the usability of smartphones in visually impaired scenarios.

Keywords: Eyes-Free Interaction, Situational impairment scenarios, Input/Output, Gesture-Based Interaction, Haptic Feedback, Smartphone Interaction, One-handed interaction.

Résumé

Les smartphones sont devenus essentiels à la vie moderne et jouent un rôle crucial dans diverses situations, même lorsque les utilisateurs ne peuvent pas se concentrer pleinement sur l'appareil. Dans des scénarios où l'attention visuelle est divisée, comme la conduite, la capacité de dessiner des gestes avec une main sur un écran tactile sans regarder le téléphone est très bénéfique. Ces gestes permettent aux utilisateurs d'effectuer des actions telles que répondre aux appels ou envoyer des messages prédéfinis sans effort. De plus, le retour haptique peut partiellement remplacer le retour visuel dans des situations de déficience visuelle. Le retour haptique peut informer les utilisateurs des entrées réussies ou fournir des informations sans nécessiter d'attention visuelle. Dans ce contexte, cette thèse explore l'amélioration de l'interaction gestuelle à une main et sans regarder avec les smartphones grâce au canal haptique. La recherche est structurée en plusieurs études, chacune abordant différents aspects de l'interaction gestuelle sans regarder avec les écrans tactiles.

Tout d'abord, nous étudions la production de gestes sans regarder dans divers scénarios de déficience situationnelle. Nous montrons les différences entre la manière dont les utilisateurs produisent des gestes sur les smartphones dans un contexte sans regarder par rapport aux gestes réalisés lorsqu'ils ont une vue directe sur le téléphone. Cela est suivi de deux études examinant comment les facteurs environnementaux, tels que la vitesse de déplacement de l'utilisateur, la localisation du téléphone et le multitâche, affectent la production de gestes sans regarder. En présence d'une tâche saturant l'attention, les utilisateurs produisent les gestes en un temps plus court. Cependant, lorsqu'un utilisateur est impliqué dans une tâche physiquement exigeante, son doigt se déplace plus rapidement dans une zone plus petite.

Ensuite, nous avons réalisé trois études visant à améliorer l'entrée gestuelle sur les écrans tactiles en utilisant le retour haptique dans des scénarios de déficience visuelle. Nous examinons d'abord l'apport du retour haptique pour la production gestuelle sur un smartphone placé dans une poche. Ensuite, nous comparons le retour localisé du moteur de vibration du smartphone au retour distant d'une montre connectée. Les résultats suggèrent que fournir à l'utilisateur un retour haptique continu améliore sa confiance pour produire les gestes, avec une préférence pour le retour localisé. Ensuite, nous nous concentrons sur l'amélioration de la tâche de ciblage dans un bouton virtuel en présence d'une tâche saturant l'attention. Nos résultats ont montré que l'utilisation d'une texture haptique de fond sur le périmètre du bouton virtuel ainsi que d'une texture haptique sur les crans augmente les performances des utilisateurs dans la tâche de ciblage sur le bouton.

Ensuite, nous nous concentrons sur l'utilisation du retour haptique pour améliorer la partie sortie de l'interaction. Nous présentons *Hap2Gest*, un concept d'interaction pour smartphones utilisant des gestes sans regarder et le retour haptique. Les utilisateurs dessinent un geste pour invoquer une commande et reçoivent des informations via le retour haptique sur un geste ultérieur. Une étude d'élicitation a montré une préférence pour les gestes avec des coins nets. Les gestes circulaires étaient également populaires dans certains scénarios, nous avons donc testé le retour haptique sur des gestes circulaires dessinés à main levée. Tout d'abord, nous avons évalué la perception des utilisateurs de la longueur des gestes circulaires sans regarder, constatant qu'ils perçoivent la longueur de l'arc comme plus courte qu'elle ne l'est. Ensuite, nous avons testé la capacité des utilisateurs à récupérer des données numériques sur des gestes circulaires en utilisant la méthode *Hap2Gest*. Les résultats ont montré une variation significative dans les formes des arcs produits librement, conduisant à des estimations inexactes de la position de l'arc. Ainsi, nous recommandons d'utiliser des gestes avec des coins nets pour la technique d'interaction *Hap2Gest*.

Cette thèse contribue au domaine de l'interaction humain-machine en fournissant une analyse complète du retour haptique dans les interactions gestuelles sans regarder. Les résultats offrent des informations précieuses pour concevoir des techniques d'interaction efficaces et améliorer l'utilisabilité des smartphones dans des scénarios de déficience visuelle.

Mots-clés : Interaction sans regarder l'écran, Scénarios d'incapacité situationnelle, entrées/sorties, Interaction basée sur les gestes, Retour haptique, Interaction avec *smartphone*, Interaction à une main,

Acknowledgements

First, I would like to thank my supervisor, Prof. Yosra Rekik, for spending an immense amount of time to guide me step by step during my Ph.D. studies and being present at all times for help. Then I want to thank my thesis director, Prof. Laurent Grisoni, for his valuable insights and helping me to find the track of my Ph.D.

My Ph.D. was funded by the European Union Horizon 2020 research and innovation program under the Marie Skodowska-Curie grant agreement No 860114. I am grateful for the opportunities given to me by this project, the strong network I was able to create in this project, and the opportunities to travel to different countries for networking and research. I would like to thank the researchers who hosted me during my secondment and the help they provided me for the research that I was conducting there. Prof. Radu-Daniel Vatavu, Mihail Terenti, and Alexandru Dancu in Suceava, Romania; Prof. Gualtiero Volpe in Genoa, Italy; and Prof. André Mouraux and Iqra Shahzad in Louvain, Belgium.

I express my gratitude to the jury members, Professor Laurence Nigay, Professor. Jean Vanderdonckt, Dr. Anke Brock, Dr. Vincent Levesque and Professor Romain Rouvoy, for dedicating their valuable time to review my thesis and for their valuable comments and insights.

In the end, I would like to give special thanks to Prof. Frederic Giraud and Geremie Postdam for managing and organizing the Multitouch project and the help they provided during my Ph.D.

To my loving family.

Contents

List of Figures

List of Tables

Introduction

In recent years, touch-based mobile devices have become ubiquitous in our daily lives. Touch input, including tapping and gestures, is currently the main interaction mechanism used in smartphones [220]. A common method of using a smartphone involves initially reaching for it and then interacting with it with one hand while viewing its screen. Touchscreen displays can simultaneously receive commands from the user (input) and present visual data to the users (output) [76]. Visual feedback stands out as the most prevalent mechanism for presenting output information to users in smartphone interaction [105]. In sum, touchscreens are the main means of interaction with smartphones by detecting and locating movements of the fingers and providing the output information through the visual channel.

The small size of smartphones, in addition to the richness of the proposed applications, has made them sufficiently portable to be used in various aspects of life. This includes scenarios where users are on the go [73], driving [12], working [74], having a conversation [184], or even when the phone is not directly in front of them, such as when it is in their pocket [190] or bag [134]. These scenarios can divert the user's visual focus from the smartphone and cause visual impairment in the interaction with the smartphone. Visual impairment refers to a functional limitation of one or both eyes or the visual system, which can present in several forms such as diminished visual sharpness, decreased contrast sensitivity, loss of visual field, or visual distortions [227]. Visual impairments can also occur in specific scenarios in which an individual with normal vision temporarily loses the ability to see the touchscreen interface directly. Such circumstances may involve glare, smoke, or while operating a vehicle [160]. Consequently, it has become important for many designers to support eyes-free interaction for visually impaired scenarios.

Several eyes-free interaction techniques have been developed to enable smartphone input in visually impaired scenarios. These interaction techniques have been implemented through different sensory channels. For example, voice commands use the auditory channel [61]. On the other hand, touch input is an interesting solution for eyes-free input. Stroke gestures can be produced eyes-free on touchscreens [165]. Stroke gestures can be viewed from two different perspectives in the field of human-computer interaction. From a computer perspective, capacitive touch sensors, which are the dominant types of touchscreen today, measure the electrical charges on the touchscreen. These measurements are used to detect the points of contact between the skin and the touchscreen, and to track the fingers on the touchscreen [147]. However, from user's perspective stroke gestures are result of two sensory perceptions. First, tactile perception detects the contact between the finger and the touchscreen. Second, proprioception tracks the movement of the finger on the touchscreen. These two perceptions are part of our haptic perception of the surroundings [142]. In this thesis, we will focus on user's perspective in the interaction with smartphone and consider stroke gestures as an input method on smartphones through haptic channel. In a similar way, we consider the movements of a smartphone in a 3D space by the user, recorded by the built-in accelerometer of the smartphone, as the haptic input for the smartphone.

The haptic channel has also been employed as an output modality to deliver information to the user. In particular, haptic has been used to reduce the need for other forms of feedback. For example, haptic feedback appears to be effective in replacing visual and audio feedback and is usually rapidly noticed during visually impaired scenarios, such as when the user is focused on another primary task [195, 35, 41, 47, 175, 73]. However, the technologies for generating tactile feedback are still quite limited and not as advanced as display technologies. On numerous Android devices, developers have access to only a limited set of high-level functions to control the vibration motor. These functions simply turn the motor on and off, individually or in a repetitive pattern, providing only basic and limited manipulation of the parameters of the typical vibration signal such as amplitude and frequency. Recently, Android has begun to provide developers with greater control over the amplitude of the motor, but this feature is limited to newer models and requires compatible motors [214]. Designing an interaction concept that can be implemented with a simple vibration motor will make that design accessible to more people.

In this context, instead of isolating tactile feedback and stroke gestures from each other, we will consider both as part of the haptic channel. In this thesis, we will study the use of the haptic channel in visually impaired scenarios for interaction with the smartphone. We will extend the limited bandwidth of the tactile channel in smartphones, due to the limited capabilities of vibrotactile actuators, by incorporating the proprioception understanding of finger positions. Combining these two will increase the vocabulary of interaction, compared to using each of them solely. This will enable users to have richer interaction with smartphones in visually impaired scenarios. We conduct most of our studies on smartphones with simple vibration motors so that the results of our research can be implemented on almost all smartphones.

Objective of the thesis

The objective of this thesis is to better understand how in visually impaired scenarios the haptic channel (vibrotactile feedback along with stroke gestures) can be used to improve interaction with touch devices, in particular smartphones. Studies have shown the benefits of employing haptic as a medium for both input and output (I/O) operations on touchscreens [31, 161]. Despite these efforts, almost all current interaction techniques and I/O operations on commercial touchscreen devices rely exclusively on visual feedback [160]. Our goal is to improve the role of haptic-based interaction with smartphones in visually impaired scenarios.

To achieve this objective, our research is structured into three parts. First, we are interested in understanding how different contextual factors influence the production of eyes-free gestures on a smartphone. We want to establish a body of new scientific knowledge on mental models of eyes-free gesture production in different visually impaired scenarios to propose design recommendations to improve interactions in these contexts.

In the second part of our studies, we use the haptic channel as an input modality for eyes-free interaction with the smartphone. Haptic feedback is essential for improving the usability of eyes-free interactions with smartphones. By providing tactile sensations or vibrations in response to user actions, it offers a non-visual confirmation that is especially beneficial in scenarios where visual attention is limited, such as when the phone is tucked away in a pocket or when users are occupied with other tasks. In this context, our aim is to understand which type of tactile feedback improves the eyes-free input gestures on a touch device.

Similarly, we want to explore how we can utilize the eyes-free gestures for transferring output data from a smartphone to users, using tactile feedback. Our goal is to develop an interaction technique that combines tactile perception with proprioception in smartphone interaction. It is necessary to know the user preferences for this interaction technique and measure the user performance in this new concept. This interaction concept with the results of previous studies will allow full haptic channel interaction with smartphones, input and output, in visually impaired scenarios.

Contributions

Our research has produced the following contributions:

1. The study of how eyes-free gestures are produced on smartphones compared to when visual feedback is available, and how various situational impairment scenarios including user movement speed, phone location, and the presence of attention-demanding tasks, impact the production of eyes-free gestures and possible mental model behind them.

2. Understanding user preferences for vibration pattern for eyes-free gestural input in two visually impaired scenarios (smartphone in the pocket and presence of an attention saturating task).

3. Investigating the preference of user's to receive haptic feedback locally or distally for in-pocket interaction.

4. We proposed a new interaction concept, called *Hap2Gest*, that allows input and output interaction with the smartphone using only a haptic channel without the need for additional hardware.

5. Understanding user preferences for gesture shape and vibration patterns in *Hap2Gest* method and showing the challenges of using circular eyes-free gestures in this method.

6. Suggesting to consider gestures as haptic input for touchscreens and combining their tactile perception and proprioception to maximize the bandwidth of information retrieval through haptic channel.

Thesis structure

This thesis is organized as follows:

Chapter 1 provides a comprehensive review of the literature on smartphone interaction in situational impairment scenarios using different sensory modalities. The chapter begins by exploring the sensory modalities involved in smartphone interaction, including vision, haptic, auditory, and multimodal approaches. It then delves into the concept of Situationally-Induced Impairments and Disabilities (SIID), highlighting how various contextual factors can hinder mobile device usability. Then it covers gestural interaction on touch devices. Subsequent sections review relevant studies on the use of haptic channels for interaction with touch devices in these impaired contexts, with a specific focus on gesture-based and haptic-based interactions by dividing the interaction into two parts: input and output.

Chapter 2 of this thesis examines the production and characteristics of eyes-free gestures in various contexts, comparing them to gestures made when touchscreen is visible to user. The chapter first explores how gestures differ when visual feedback is absent and their effectiveness in scenarios with limited visual attention. Then it investigates how the speed of user movement and phone location affect the features and performance of eyes-free gestures. Finally, the chapter delves into the impact of attention-saturating tasks on the production of eyes-free gestures, considering factors such as hand dominance and multitasking. This chapter aims to deepen the understanding of eyes-free gesture interactions with mobile devices.

Publications

- Rekik, Y., Guettaf, A., **Jamalzadeh, M.**, & Grisoni, L. (2024, June). *Comparing Eyes-free Gestures to Gestures Produced in the Presence or Absence of Visual Feedback on Mobile Device.* In *Proceedings of the 2024 International Conference on Advanced Visual Interfaces* (pp. 1-5).

- **Jamalzadeh, M.**, Rekik, Y., Grisoni, L., Vatavu, R. D., Volpe, G., & Dancu, A. (2023, August). *Effects of Moving Speed and Phone Location on Eyes-Free Gesture Input with Mobile Devices.* In *IFIP Conference on Human-Computer Interaction* (pp. 469-478). Cham: Springer Nature Switzerland.

- **Jamalzadeh, M.**, Rekik, Y., & Grisoni, L. (2023, November). *The Effect of Attention Saturating Task on Eyes-Free Gesture Production on Mobile Devices.* In *Companion Proceedings of the 2023 Conference on Interactive Surfaces and Spaces* (pp. 27-31).

Chapter 3 explores the use of haptic feedback in enhancing eyes-free gesture input on touch devices. The first study examines how haptic feedback improves gestures performed on pocket fabric, enhancing user confidence in the eyes-free gesture input. The second study investigates user preferences for the location of haptic feedback, either locally in the pocket or distally on a smartwatch. The third study, unlike previous studies that focused on physical limitations for interaction with a smartphone, focuses on improving input on a touch device when visual attention is needed simultaneously elsewhere. We will investigate which type of haptic feedback will improve the performance of participants in targeting task while simultaneously being occupied with a primary attention saturating task.

Publications

- *Draw & Feel: Enhancing In-Pocket Interaction with Mobile Devices Using Haptic Feedback.* This paper is currently in preparation for submission.
- Rekik, Y., Guettaf, A., Rupin, M., **Jamalzadeh, M.**, & Grisoni, L. (2024, June). *Enhancing Touch Circular Knob with Haptic Feedback when Performing Another Saturating Attention Primary Task.* In *Proceedings of the 2024 International Conference on Advanced Visual Interfaces* (pp. 1-9).

Chapter 4, on the other hand, focuses on the output part of the interaction. The goal of this chapter is to propose an interaction concept that enriches the data that can be transferred on touchscreens to the user with a simple smartphone vibration motor. The first study introduces this interaction concept called *Hap2Gest* that combines gestures and vibration patterns and allows interaction with the smartphone through the haptic channel. The next two studies examine the use of circular gestures in the *Hap2Gest* method, with the aim of assessing its feasibility and effectiveness. The second study investigates the consistency and accuracy of users drawing arc or circular gestures. The third study builds on these data by implementing circular gestures in the *Hap2Gest* method, where users draw gestures and retrieve information about the gesture using vibration cues.

My Title

- **Jamalzadeh, M.**, Rekik, Y., Dancu, A., & Grisoni, L. (2023, August). *Hap2Gest: An Eyes-Free Interaction Concept with Smartphones Using Gestures and Haptic Feedback.* In *IFIP Conference on Human-Computer Interaction* (pp. 479-500). Cham: Springer Nature Switzerland.
- **Jamalzadeh, M.**, Rekik, Y., Shahzad, I., & Grisoni, L. (2024, September) *Eyes-free Circular Gestures on Smartphones.* In Adjunct Proceedings of the 26th International Conference on Mobile Human-Computer Interaction (pp. 1-5).

Finally, we conclude this thesis with a conclusion and discussion of the results of our work for eyes-free interaction with smartphones, before exploring various directions for future research.

Chapter 1

Literature Overview

Human computer interaction is conducted using sensory channels. The sensory modalities which can be used in interaction depend on both sides of the interaction: computational device and human. On the computer side, the technologies available on the device limit the extent to which each sensory channel can be used in the interaction. This is varied for each device. For example, smartphones usually have sensors such as an accelerometer that is not available on a desktop computer. In contrast, on the human side, as long as they do not experience any kind of sensory disability such as blindness, they have similar sensory means to interact with computers [138]. Although people can have similar sensory means for the interaction with a device, the extent to which they can leverage each sensory modality depends on their environmental factors. Their sensory channels can be impaired by various environmental factors [128]. For example, the ability to perceive auditory signals deteriorates greatly in a noisy environment. Thus, it is important in the human-computer interaction to consider sensory impairments which can be caused by different environmental factors. Having knowledge of these impairments can help designers choose the correct sensory modality for the interaction.

In this chapter, first, we will give an overview of different sensory modalities available for the interaction with smartphones. Then we will explain different types of visual situational impairments that a smartphone user can experience and how gestures and haptic feedback can be used to address these impairments. We will cover how the haptic channel has been used as input and output for interaction with touch devices. We have categorized related work in this field into input and output, from the perspective of user sensory modalities. For example, we consider stroke gestures as a haptic modality for interaction with the smartphone. The production of stroke gestures from the user perspective is based on the detection of finger and touchscreen contact (tactile perception) and spatial understanding of how finger moves on the touchscreen (proprioception). Moreover, when haptic perception is used to provide information to the smartphone about the intention and context of the user, we consider it as the input part of the interaction. For example, when the gesture produced by the user is recognized to invoke a command or when the accelerometer sensor data is used to detect whether the user is sitting still or walking. However, if haptic perception is used to transfer information from the smartphone to users, we consider it to be the output part of the interaction. For example, a user explores a haptic surface to feel the simulated texture of a fabric.

1.1 Sensory Modalities in Smartphone Interaction

In everyday interactions, people use all their senses and various facial and bodily expressions. For example, a simple experience, such as having coffee with a friend, involves multiple sensory inputs and outputs, including smell, taste, vision, touch, and sound [154]. However, today our engagement in technology is primarily governed by user interfaces that are highly dependent on vision and audio, with tactile feedback playing a comparatively minor role [154]. The user interface is a crucial element of a computer. Creating suitable user interfaces for computers requires understanding how users interact with computers. This is a branch of computer science called human-computer interaction (HCI). HCI emphasizes users and evaluates how various hardware and software components of a computer enhance the input/output interaction between the user and the system [250].

Many electronic devices that we use in our daily lives provide inputs and outputs that need to be processed by our senses. The computing literature often distinguishes sharply between input and output. However, almost all human-computer interactions require both input and output to be effective. Input and output are merged using interaction techniques that allow the user to perform a low-level task. For example, in a traditional graphical user interface, users scroll through a document by clicking or dragging the mouse (input) within a scrollbar displayed on the screen (output) [86].

Although computers are arguably the most sophisticated and versatile devices with which we interact, smartphones are the ones with which we interact the most. Smartphones began gaining popularity in developed nations in the late 2000s and in developing countries in the early 2010s. They possess computational capabilities, connectivity, and multimedia features similar to those of computers but are more compact and portable. For example, today, a significant number of people access the Internet through smartphones.

Smartphones integrate a variety of sensors and hardware into a single device to enhance the user experience and utility [158]. For example, touchscreens are used to detect the location of the fingers for command invocation (input) and present visual data to the user (output) via a graphical user interface [250]. In general, the input of a smartphone consists of the information sensed about the physical environment. This includes detecting touch gestures using touchscreen, detecting phone movement using acceleremetor and gyroscope. The output of a smartphone can include any emission or modification of the physical environment, such as changing the image displayed on the screen, playing a sound using speakers, or tactile feedback using vibrotactile actuators [86, 206].

Despite the myriad of senses available for interacting with smartphones, the fields of human-computer interaction (HCI) have historically focused predominantly on vision and audition [154]. This thesis explores research that extends beyond audiovisual interfaces, emphasizing the role of the haptic channel.

1.1.1 Vision

The range of input and output methods has expanded considerably over time. Vision serves as one of the modalities that can be utilized to perceive outputs and to issue commands to the system (input) [131]. Face-based input, using the front camera of the phone, has been demonstrated to be an effective means of interaction with smartphones. The user's face can be easily detected using the front-facing camera of the smartphone, while the user holds the phone. This makes the face an effective channel for interaction with the device [129]. Face-based input has been used for autoscreen rotation [43], authentication [50], expanding the vocabulary of mobile interaction [258], and camera control [44]. In addition, user gaze information can be used for natural scrolling [135]. Eye movement and blinking can be used for mobile browsing and text entry [196]. Recently, facial gestures have appeared in some standard systems, such as smiling in the Huawei smartphone camera for shutter release. Taking advantage of the face-engaged input channel alone provides a number of new interaction possibilities.

Among the various feedback mechanisms used for mobile devices, visual feedback stands out as the most prevalent output. For example, during gesture performance, though some gestures can be executed in a way that is less visually demanding [148], they still rely on the user's vision to convey information about progress or input outcomes [14]. Visual feedback was found to be an effective method for enhancing user learning when dealing with new applications or systems [14]. Nevertheless, Niels et al. [85] could not confirm the hypothesis that indicating the user's touched position improves learnability. However, this approach did reduce error rates in typing, as demonstrated by Kristensson et al. [112], who showed that providing a visual preview of the presently recognized command during the user's stroke articulation decreased error rates with shorter gestures in certain situations. Complete dependence on visual feedback, however, might be reconsidered due to the existence of alternative interaction methods that offer output information through other sensory channels [51], particularly in situations where visual attention is restricted.

1.1.2 Audition

Voice serves as an auditory communication channel that transmits sound information from a speaker to multiple listeners. Voice input is used in a wide range of tasks, including text entry, communication, and sending voice commands. However, there are two main challenges with voice input. First, users worry about the privacy risks of disclosing their personal information while speaking. Second, they suffer the inconvenience of repeatedly speaking the wake-up word or pressing a button during multiple rounds of voice input [251]. Furthermore, it will be much less efficient when used in a noisy environment [58].

There are difficulties when using speech output as opposed to graphical interfaces. One can view a lot of information at once, especially in a graphical interface. When only the auditory channel is used as the output, it is a slow medium of interaction and requires users to exert more cognitive effort. Moreover, in visual interfaces it is possible to give a command concurrently, view the output, and modify it, but in auditory interfaces concurrent input and output are less convenient and more challenging [143].

Voice has become an integral part of the user experience when interacting with computers, particularly those employing artificial intelligence (AI) [197]. Recent progress in speech technology has led to the widespread adoption of voice assistants such as Amazon's Alexa, Apple's Siri, and Google's Google Assistant. One of the names that are widely used for voice-based technologies is "conversational agents" [59]. Conversational agents (CAs) are systems capable of engaging in dialogue with users [116]. This definition includes both input and output in the interaction.

1.1.3 Haptic

The sense of touch is fundamental in most of our day-to-day interactions. With the success of mobile devices, touch-based interaction has become the dominant method of interacting with computing systems. The shift towards entirely touch-based interfaces is expected to intensify. Most of our day-to-day interactions depend solely on the location of touches [66]. On touch screens, the user collaborates with the framework by contacting objects or performing different touch motions [92] such as flicks and swipes [231]. Since most human–computer interaction is mediated through touch, the tactile sense should be placed to play a richer role in such interactions than it currently does [236].

Beyond touches and swipes, stroke gesture input enables users with even more flexibility to perform tasks efficiently [253] and with low cognitive effort [186]. Gesturing has emerged as an important interaction paradigm for entering text and commands on computers. Stroke gestures are frequently employed on mobile devices where keyboard shortcuts are either unavailable or require significantly more cognitive effort, allowing users to complete various tasks efficiently and with confidence. For example, stroke gestures are used as shortcuts to invoke commands in menus [121], to enter text quickly [111] or call app functions directly [10]. Another way to leverage the haptic feedback in smartphones for input in addition to the touchscreen is to use the built-in accelerometer [205, 72]. The recent generation of smartphones includes MEMS-based accelerometer sensors by default [72]. The accelerometer sensor measures constant (gravity), time-varying (vibrations), and quasi-static (tilt) acceleration forces, which affect the device on the three axes (x, y, and z) in meters per second squared (m/s2) [137].

High-resolution haptic feedback systems for mobile phones have recently become widespread, enabling the transfer of detailed information via the haptic channel. For example, the TouchSense technology created by Immersion has been recently deployed in LG smartphones[1]. In a similar manner, Apple devices use the Taptic Engine within their Force Touch feature[2]. At first, the vibrotactile stimuli in standard mobile phones conveyed basic information like alerts. Nevertheless, modern operating systems such as iOS enable the association of vibration sequences (either predefined or customized) to contacts in the address book or to built-in services. Similar functionality can be achieved on Android-based smartphones using applications such as Good Vibrations[3]. This application allows the configuration of vibration patterns not only for different

[1] https://www.businesswire.com/news/home/20170809005178/en/Immersion-LG-Electronics-Expand-Business-Relationship-Include

[2] https://www.imore.com/science-behind-taptics-and-force-touch

[3] https://wbouvy.com/goodvibrations

applications but also for specific contacts [69]. Vibrotactile feedback enhances privacy since the vibrations are felt exclusively by the user. Additionally, these stimuli enhance the user experience by providing an extra communication channel beyond the traditional audio and visual channels. In certain situations, such as noisy settings like concerts, they may even serve as a substitute for audio [90].

1.1.4 Multimodal

Mobile users are usually engaged in multiple activities simultaneously, which naturally leads to multimodal interfaces. The term "multimodal interfaces" refers to interactive systems that make use of human skills to communicate through speech, gesture, touch, facial expression, and other modalities, bringing more advanced pattern detection and classification techniques to human-computer interaction [224]. Multimodal interfaces are especially appropriate for mobile devices due to their small screens and the environment in which they are used (*e.g.*, on buses or trains, or while driving), making single-mode visual interaction inadequate [2]. In mobile usage scenarios, the user's attention is often split, with interactions occurring in brief sessions as users juggle between the device and their surroundings. This indicates that integrating tactile feedback with visual feedback is likely to be the most effective strategy for these scenarios [2].

Multimodal interaction has proved to be a promising way for developing more accessible applications [174]. When using the multimodal interfaces, users select the preferred modality to execute the desired function, such as playing music by tapping their smartphone or using voice commands. While each modality has unique advantages and disadvantages, users can benefit from both by selecting the modality that is most appropriate for their situation. Touch input provides a familiar and immediate operation, while voice input enables eyes-free operation. Depending on their task, purpose, and context, users switch between modalities to achieve their goals.

In general, the design of successful and usable multimodal systems remains a challenge for human–computer interaction (HCI) researchers, as user's modality selection or preference can vary widely across different contexts and usage scenarios [37]. When using a multimodal interface, users tend to select the most efficient mode according to the situation, and this is an essential element of interface design [110]. Budiu [4] defined the interaction cost as the total physical and mental effort necessary for interaction, and users tend to minimize it by switching modalities for maximizing the effectiveness of their actions. In the field of HCI, various measures have been developed to measure the efforts of interaction. One of the most popular measures is the Nasa-TLX, which measures task workload in seven dimensions [79]. Each dimension is measured by a single questionnaire.

[4] `https://www.nngroup.com/articles/interaction-cost-definition/`

1.2 Situationally-Induced Impairment in Mobile Interaction

The mobile interaction introduces a new paradigm in which the guidelines for effective design in a stable desktop environment might not be suitable for mobile contexts [240]. In the mobile context, environmental factors are more versatile, as the user can use a smartphone in much wider scenarios than a desktop computer. For instance, in contrast to desktop computers, which are generally utilized in a consistent and controlled setting (for example, where the user is seated without significant exposure to harsh lighting or weather elements), smartphones can be operated in various settings, such as inside, outside, in noisy or quiet areas, crowded spaces, and among others. The user's interaction with the mobile device in these settings can be impaired by various contextual factors [92]. Impaired abilities to interact with the device due to contextual factors are known as Situationally-Induced Impairments and Disabilities (SIID) or situational impairments [198]. The term situationally-induced impairments (also known as situational impairments) was first introduced by Sears [198]. In this section, we discuss the effect of the environmental context on the interaction with portable devices and how it can impair the interaction. Especially we will cover these situational impairment scenarios for touch devices, which is the focus of this thesis. Furthermore, we will explain the potentials of gestural input and haptic feedback to improve the interaction with touch devices in situational impairment scenarios.

1.2.1 Contextual factors

The contextual factors that can cause SIID can be classified by five dimensions [1]: physical context, temporal context, task context, social context and technical context. In the following, we will give the definition of each factor and the ones we will cover in this thesis.

Physical context refers to the observable characteristics of the situation where human-mobile computing interaction occurs. This includes the spatial setting, functional area and space, detected environmental properties, movements and mobility, as well as the artifacts in presence [95]. For instance the interaction can take place in different locations. In this thesis, we run all our experiments in indoor environment (*i.e.*, laboratory and gym). Another observable characteristic is mobility. We cover these conditions: sitting on a chair, standing and walking and jogging on a treadmill. There are other mobility conditions which we didn't cover but has been studies earlier such as walking on a straight path [152].

Temporal context refers to how a user engages with a mobile device over time in various aspects, including duration, different times of the day to years, the circumstances preceding and following usage, time-related actions, and synchronization [95]. In the second study in the next chapter, we study the effect of different movement speeds of a user on the production of eyes-free gestures.

Task context of use for mobile interaction describes in particular the *multitasking* and possible *interruptions* that are related to the execution of the task Task context pertains to how the overall situation demands one's attention [95]. Engaging in *multitasking* is viewed as an effective method for dividing attention or cognitive resources when interacting with small screen and wearable devices. It is recognized as a significant contributor to SIIDs [1]. In this thesis, we used walking and presence of attention saturating task in some of our studies.

Social Context refers to the individuals who are present, their traits and functions, the interpersonal dynamics, and the surrounding cultural environment that impact the user's engagement with a mobile device [95]. Majority of studies in this SIID field is conducted individually in separate sessions [1]. We followed the same protocol in our experiments and didn't cover the interpersonal relation effect in our studies.

Technical Context refers to the relationship between various relevant systems and services, such as devices, applications, and networks, their interoperability, access to informational artifacts, and mixed reality, in relation to the user's interaction with the mobile computer [95]. In this thesis in most of experiments the main system users were interacting with was a smartphone. However, laptop, smartwatch, keyboards, and haptic surfaces were also present in different experiments.

1.2.2 Examples of situational impairment scenarios

There are numerous possible scenarios of situational impairment. Ambient temperature [68], ambient noise [191], encumbrance [150], user movement [63], and ambient light [216] are examples of potential situational impairments. Covering all the scenarios which can cause situational impairment and earlier studies that have been conducted for each of them requires a long review which is outside the scope of our literature review. Interested readers can refer to the work of Akpinar [1] for a comprehensive overview. Here, instead, we will give an overview of the situational impairments we used in the following chapters and earlier works that have covered these scenarios. We will cover scenarios which will cause visual impairment for interaction with a smartphone such as user's mobility condition, and presence of an attention-saturating task.

The interaction with touch surfaces under different mobility conditions is well studied in the literature, including taking note when sitting [49], entering text when walking [45], wearable touch surfaces when standing [52], and searching the web on a treadmill [80]. Some studies have investigated the effects of user movement on interaction with mobile devices. Marshall and Tenant [133], for example, note four challenges for humans trying to interact while on the go: (1) cognitive load (limited attention resources); (2) physical constraints (nonmobile activities may place constraints on physical resources); (3) terrain (external environment affects how a user will interact); and (4) other people (movement activities often involve a social element). The effect of the size of the key on the text entry with the stylus on the PDA while walking is examined [136]; they reported that the error rate decreases and the speed of the content passage increases for the bigger keys. Walking has been found to have an adverse effect on performance during mobile interaction. The effects of walking on reading and selection tasks are investigated and revealed that the exhibition decreased and the intellectual burden expanded [194]. The researchers found that task completion time, error rate and workload measures (*e.g.*, mental and physical demands) differed significantly while performing tapping tasks under the conditions of being seated, walking on a treadmill, and walking in a freestyle with obstacles [125]. Walking has also been found to cause deterioration in text legibility [144], as well as reading comprehension and cognitive performance, measured by a word search task [13].

With the increasing sophistication of mobile devices, engaging in multiple activities at once (for example, browsing the Internet or sending a text message while walking) becomes more appealing to accommodate the busy lives of users. Consequently, a greater number of users may experience situational impairments than in the past [193].

In this thesis, we focus on mobile device *situational visual impairment* - visual impairment arising from user context - since few studies have explicitly explored the use of a smartphone in an eyes-free configuration [216]. There are various scenarios in which an eyes-free interaction is required. It can be argued that the main use of eyes-free techniques is to create accessible interfaces to assist people with visual impairments [160].

1.2.3 Interaction with touch devices in situational impairments scenarios

Previous research has identified that interactions with mobile devices constitute a significant change, which requires the development of new interaction rules that may not align well with those created for the traditional desktop setting [240]. For example, in situational impairment scenarios, people often opt to operate their mobile devices without looking at the screen, using one hand, and keeping their head upright. This behavior is influenced by various environmental, social, device-related and personal factors (*e.g.*, harsh lighting, participation in a meeting, small screen size, personal motivation) [103]. Furthermore, designing interaction techniques for contexts of situational impairment can result in improved solutions for people with disabilities (permanent impairment). Similarly, investigating the needs of disabled people can inspire innovations that effectively address situational impairments [20].

For example, one difference between interaction with smartphones and computers is that in mobile HCI the way that the phone is grasped should be taken into account. The one-handed touch interaction is a commonly used hand grip for smartphones while walking [151], standing [55], working or talking with someone [185], being encumbered [30], being under some distractions [148], or when physical and visual attention is hosted by the other hand [26]. The one-handed interaction on mobile devices is widely documented and reviewed in the literature. Eardley et al. [55] findings indicate that one-handed interaction implies more phone movements than asymmetric bimanual interaction, in particular with lying body posture. In these situations, the thumb of the hand that grips the device is typically the sole finger used for the touch input [25]. However, due to the biomechanical limitations of the thumb, only a partial area of the touch screen is easily accessible by the thumb [101].

Another aspect to consider is which sensory modalities should be used in the interaction. There is some information about the modalities that are the best for situational impairment scenarios. Under typical workload conditions, the combination of visual and audio feedback is more effective for single tasks. In contrast, the pairing of tactile and visual feedback proves to be most beneficial for multitasking and during periods of high workload [29]. Tactile feedback can offer considerable potential to mobile users with situationally induced impairments and disabilities. Vibrations provided by mobile devices can be beneficial in certain challenging environments where other signal forms may fail (*e.g.*, in noisy areas where sound cues from the phone would typically be obscured).

Another key benefit of touchscreen computing is that it allows users to interact with highly versatile and dynamic devices while performing other tasks concurrently. A common instance is handling an in-vehicle display (such as interacting with menus, buttons, and scroll bars) while driving. Although entertainment systems in vehicles can be effectively managed using mechanical and manual controls, these physical options are rapidly being replaced by touchscreen interfaces to capitalize on their enhanced adaptability and flexibility [209].

Another difference between touch interfaces and desktop computers that should be considered in the interaction design is how input commands are given. Desktop computers usually use keyboards alongside the mouse or touchpad for input. On the other hand, mouse and keyboards are replaced with touch sensors in touch interfaces. This has made gesture inputs widely available on modern interfaces.

1.2.4 Gestural interaction in visually impaired scenarios

The gesture interaction provides a control interface that reduces the need for visual attention to interact with a device [96]. Numerous stroke gestures have been introduced in scientific studies for visually impaired interactions. These include both free-form gestures (*i.e.*, operands and mnemonic gestures [185, 26]) and mark-based gestures (*i.e.*, on-axis rectilinear strokes [165, 179, 26]). Some researchers allow users to initiate their gestures from any point on the screen [39, 26], whereas others require users to perform bezel gestures, starting at the device's edge [179, 247, 26], since bezel gestures can enhance eyes-free interaction by providing tactile feedback from the device's borders [247]. Different gesture-based interaction techniques have been proposed. For example, Kubo et al. [113] introduced the B2B-Swipe for an eye-free gesture from a bezel to a bezel on rectangular touchscreens, in particular for smartwatches. Negulescu et al. [149] studied the cognitive demands of an eyes-free tap, swipe, or move on a smartphone in distracted scenarios. Tinwala et al. [218] introduced an eyes-free text entry technique on touchscreens using graffiti strokes. Anthony et al. [4] explored the effectiveness of visual feedback for stroke gestures performed on a mobile touchscreen. Their research concluded that gestures performed with or without visual feedback vary in a way that makes them difficult to understand, and this is true across different age groups. However, no significant differences were found in gesture recognition accuracy regardless of visual feedback presence, although users generally favored gestures accompanied by visual feedback.

1.2.5 Role of haptic channel for interaction with smartphones in situational impairment scenarios

Recent advances in technology have significantly increased the implementation of haptic feedback in human-computer interaction. The haptic channel is a fundamental component of our sensory experience in daily life, including in situational impairment scenarios. In these scenarios, we can look at the haptic channel from two angles. First, how situational impairments affect the perception of haptic feedback. For example, Guettaf et al. [73] showed that walking deteriorates tactile texture recognition. In another study [74], they also showed that the attention-saturating task and the cognitively demanding task cause deterioration in the recognition of haptic textures on a haptic surface. Chen et al. [42] investigated how performing a cognitively demanding task, specifically typing text, influenced the ability to recognize spatio-temporal vibrotactile patterns. Their research demonstrated a significant impact of the primary task on the recognition rate. Second, the use of the haptic channel to provide distinguishable input and output has the potential to improve interaction with smartphones [20].

Smartphones are often manipulated by highly sensitive skin regions, hands, with tactile discrimination capabilities that reach the nanoscale [203]. Haptic actuators were missing on desktop computers. However, smartphones are now widely used in daily life and usually are equipped with some type of haptic actuators. So, it can be beneficial for users to use the vibrotactile actuators available in their hand to a greater extent. Progress in both hardware (touch input sensitivity and presentation of rich tactile textures using haptic surfaces) and software (gesture recognition) aspects of haptic technology have been key contributors to the success and popularity of the haptic feedback method and the interaction approach [227, 160].

In the next section, we will provide an overview of gestural interaction with smartphones. It is followed by a section covering the role of the haptic channel in the eyes-free interaction with smartphones as input and output.

1.3 Touch Gestural Interaction with Smartphones

A stroke gesture consists of one or more touch strokes [176]. A touch stroke is the movement trajectories of a pointer, such as a finger, stylus, etc.; which contact points begin with pointer-down and end with pointer-up from the screen [253]. The number of possible gestures could be as large as more than 1000 [242]. The most common single-touch gestures are clicking (or pressing, tapping), swiping (or flicking), and dragging (or scrolling, panning, sliding). In addition to the type of touchscreen gestures, it is necessary to pay attention to the directions of the gestures [94]. In this section, we will explain which features are used to record and recognize stroke gestures on the touchscreen. Then, we will explain how elicitation studies are used to choose gestures by focusing on end-user preferences. Finally, we will review studies that have captured gestures using other means than capacitive touchscreens.

1.3.1 Gesture features

The effectiveness and ease of use of touch gestures are determined by their specific attributes and the recognition methods used. These methods consider a variety of gesture features that are constantly evolving due to advancements in screen technology and algorithms. Most of the studies and technologies related to gestures incorporate a blend of absolute geometric and kinematic characteristics as descriptors [5, 176], such as gesture length, gesture height, gesture width, gesture area, gesture duration and gesture speed. In this thesis, we consider geometric and kinematic features in order to characterize the produced gestures.

Understanding how people articulate stroke gestures is important for developing robust recognizers and designing a good set of gestures [92]. Therefore, researchers have proposed many techniques to analyze gestures, including examining consistency among users [5], differences between the user population in gesture articulation [202], and the impact of input devices on gesture performance [222]. In addition, researchers have employed a variety of measures to characterize users' performance with stroke gesture input (*e.g.*, GECKo [5] and GHoST [230]).

Most of these techniques used a combination of geometric and kinematic features to analyze gestures; for example, Kane et al. broke down the gestures delivered by blind individuals and found that visually impaired individuals want to use corners and edges on the screen [98]. They additionally revealed that the visually impaired individuals produce longer, more extensive, and slower gestures with larger size varieties. Tu et al. examined differences and similarities between pen and finger gestures using kinematic and geometric features [222]; they found that finger gestures were faster and larger than pen gestures, but finger and pen gestures were similar in axial symmetry, proportional shape distance, and articulation time. Shaw and Anthony analyzed the children's touch gestures and highlighted features that can be used to create a set of gestures and new gesture recognizers that are more customized for children [202].

Most of the research on the characteristics that make touch-based gestures usable has been conducted in the presence of visual feedback. Nevertheless, numerous similarities in actions and preferences can be observed between sighted and eyes-free gestures [32]. Situational impairments can also influence the effectiveness of touch-screen devices, especially when on the move. Bradgon et al. found that uni-stroke gestures initiated from the edge of the display were more dependable in these scenarios compared to those performed at other screen locations [27].

1.3.2 Stroke Gesture recognition

Stroke gestures recognition involves categorizing movement patterns made with a pointing device or fingers, which are utilized as shortcuts to specify scope and commands [253]. 2D stroke gestures are commonly recorded by tracing them on a touch-sensitive surface and representing them as sequential 2D (x, y) points [211]. Optionally, stroke gestures include a third dimension, timestamp (t_i). Thus, the goal of a 2D stroke gesture recognizer is to categorize an unidentified candidate gesture by matching it with pre-stored template gestures, which were collected from different gesture classes during the training phase. This process involves computing a (dis-) similarity measure between the candidate gesture and the template classes [130].

In particular, the popular $-family recognizers [244, 6, 9, 229] permit to recognize both single- and multi-stroke gestures. The $-family recognition methods come with certain restrictions. Specifically, $1 and Protractor are limited to recognizing unistroke gestures [245, 122]. $1 recognizer involves four key steps: resampling, rotation, scaling, and translation. Initially, the trajectory points are resampled into an isometric set of points, allowing gestures with various movement speeds to be comparable. For rotation invariance, the trajectory is then adjusted so that the angle between its centroid and its initial point is $0°$. Subsequently, the trajectory is scaled non-uniformly to fit within a reference square and translated such that its centroid is at the origin (0,0). The similarity between two gestures is ultimately determined by computing the average Euclidean distance between corresponding points [33]. The Protractor employs a preprocessing technique that is similar to the $1 recognizer. However, it varies in how it manages orientation sensitivity and scaling, maintaining the gesture's aspect ratio. Additionally, instead of using the Euclidean distance, it utilizes angular cosine to measure the similarity between gestures [122].

On the other hand, $N and $N-Protractor incorporate support for multistroke gestures [7, 8]. The $N approaches achieve this by considering multistrokes as unistrokes obtained by connecting individual strokes "in the air" [7],which enables them to use the same matching algorithm as $1. However, since the order and direction of the stroke may differ between the users drawing the same symbol, $N must generate all possible permutations of a given multi-stroke [7, 8], causing an explosion in both memory and execution. More sophisticated recognizers are also available, such as HMM [200], DTW [145], and statistical classifiers such as Rubine's popular classifier [188].

1.3.3 Gesture elicitation studies

Approaches that incorporate users into the design process have become more popular following the research of Wobbrock et al. [243]. They developed a collection of gestures defined by users for touchscreens by demonstrating the result of an action and prompting participants to create a gesture that would lead to that outcome. After gathering all participants' gesture designs, the researchers measured the level of consensus among the proposed gestures using an agreement score. The gestures with the highest agreement rates were deemed the most suitable for each action. Since that time, the elicitation method has been extensively employed to develop gestures for free-hand TV control [140] and smart glasses [223], unmanned aerial vehicles [166], handheld objects [201], deformable displays [219], and assistance for blind individuals [98, 181].

Over the years, the original method has been refined with an updated formula for computing agreement [57]. *Agreement* is a fundamental element of end-user elicitation, which signifies when actions or gestures suggested by the study participants are, for the most part, identical or significantly similar [3]. *Agreement* is essential for categorizing analogous actions or symbols and for computing a score that measures the degree of agreement among participants in their suggestions [234]. Agreement has been assessed through multiple methods in the context of end-user elicitation. The *Agreement score* A, introduced by Wobbrock et al., is probably the most frequently used measure of agreement in end-user elicitation studies [234]. The *Agreement score* A was also implemented in the first hand-gesture elicitation research conducted by Wobbrock et al. in 2009 [242]. It is calculate using the following formula:

$$A(r) = \sum_{P_i \subseteq P} (|\frac{P_i}{P_t}|)^2 \tag{1.1}$$

where P is the set of all proposals for referent r, $|P|$ the size of the set, and P_i subsets of identical proposals from P.

An inconvenience of the A measure is that it never reaches 0, even when all participants are in disagreement with each other. The minimum value attainable by A is $\frac{1}{N}$, which depends on the number of proposals put forward by the participants [234]. So, instead, Vatavu and Wobbrock [233]. proposed the Agreement rate (AR) measure which has a lower bound of zero and upper bound of zero. Here is their formula:

$$AR(r) = \frac{|P|}{|P|-1} \sum_{P_i \subseteq P} (|\frac{P_i}{P_t}|)^2 - \frac{1}{|P|-1} \tag{1.2}$$

where P is the set of all proposals for referent r, $|P|$ is the size of the set, and P_i are the subsets of identical proposals from P. In this thesis, we used *Agreement rate (AR)* in our elicitation studies.

1.3.4 Gesture interaction beyond smartphone touchscreens

The main way to interact with a smartphone is through the front-facing touchscreen and a few physical buttons located on the sides. However, there are several situations where touching the screen or buttons is not practical, such as when the phone is in a pocket, lying on a table while the user's hands are dirty or wet, when touching the screen obstructs the view, or when one hand is holding the phone and the other is busy with another task [254]. Here we present different alternatives that have been suggested to extend the gestural interaction beyond direct touch with the touchscreen.

Back of the device interaction: Interacting with the back of the device (BoD) has become a popular method for improving the user interface of the mobile device [115]. The use of the back surface of a smartphone for one-handed interaction (BoD input) has shown potential to mitigate well-established issues related to thumb-based touch input, particularly reachability [91]. Thanks to advances in recent smartphone technologies, particularly those featuring rear displays (such as Samsung's foldable phones), detecting input from the back of the device is becoming more feasible [48]. Gestures executed on the back of the device can be identified by the integrated camera [246] or the sensors present on the mobile device [71].

Vibration and acoustic sensing surfaces: Several gesture sensing schemes based on different types of surfaces have been proposed which instead of capacitive sensing uses vibration or accoustic signals. Vibration-based [126, 127] methods facilitate authentication through tapping locations or gestures on solid surfaces with considerable precision. TapSense [78] leverages tapping sounds to identify touch movements on touch screens. Both SoundWrite [255] and Ipanel [40] detect the acoustic signals produced by a finger sliding on tables for sensing. Nevertheless, these sound-based recognition techniques are highly susceptible to interference from ambient noise.

In-air gestures: Various systems that use sound for gesture recognition have been introduced to identify gestures made in the air [187, 207]. Soundwave [75], Multiwave [167], and AudioGest [187] use the Doppler effect to recognize pre-defined gestures. However, the Doppler effect only gives coarse-grained movement speeds. Therefore, these approaches can only identify a limited number of gestures that exhibit unique speed characteristics [208].

Secondary wearable devices: Significant advances in sensor technology have led to the development of compact and powerful wearable input devices. One notable example is the smartwatch, which has transitioned into mainstream culture. Previous studies have explored wearable input systems in various forms, including smartwatches, rings [104], necklaces [256], and wristbands [108]. These wearable devices offer a convenient alternative to physically accessing the phone for certain tasks, effectively eliminating the need to reach in a manual way.

Textile sensors: Sensing techniques applied to textiles encompass various methods such as object recognition, NFC, and capacitive sensing. Capacitive sensing, despite its challenges for fabrics and wearables, offers the potential for touch input, hand gesture recognition, and posture detection. Notable projects in this domain include Tasca [249], which introduces a fabric-based sensor capable of detecting touch and pressure within pockets, as well as identifying metallic, non-metallic and tagged objects. Another project, Jacquard [171], employs a novel conductive yarn and multi-touch woven panels to transform regular clothing into interactive devices. Previous research has shown that the front of the thigh represents an optimal location to place touchpad-like interfaces, including textile sensors, when standing, sitting, or kneeling [215].

In-pocket interaction: It is also possible to extend the interaction with the touchscreen by allowing touch gestures through clothes and fabrics or indirectly interact with the smartphone when it is inside the pocket. The domain of on-leg and in-pocket interaction is an active area of research, encompassing concepts such as interactive pockets (*e.g.*, integrated into a pair of pants). These interactive pockets provide a comfortable, private, and readily accessible means of touch-based interactions with other computing devices. Numerous sensing techniques have been investigated to facilitate pocket interaction. There are three primary approaches to achieving in-pocket interaction: using secondary wearable devices, employing textile sensors, or leveraging smartphones' sensors. In-pocket techniques such as Tap [182] and Whack [93] take advantage of the inertial measurement unit (IMU) of smartphones to capture quick gestural commands. STAT [84] uses a smartphone screen on the thigh to enable text input based on tap and gesture-based words on a head display. Pocket interaction can also be used to interact with other electronic devices such as ambient displays and augmented reality glasses. Smart pockets [226] use pocket-based gestures, such as placing hands in specific pockets, as input for large ambient displays. PocketThumb [52] integrates a touch interface into a pocket, allowing control of wearable devices such as augmented reality glasses. These advances contribute to the development of eyes-free interfaces.

Touch-enhanced motion techniques: A wide range of previous work combined touch input with the built-in accelerometer of mobile devices. Hinckley et al. [87] introduced the terminology of touch-enhanced motion techniques that combines touch input and implicit changes in the accelerometer. For example, a touch coupled with a subsequent tilt detected by the accelerometer can facilitate one-handed zooming. Similarly, holding an item on the touchscreen and then shaking the device can provide a quick method to delete files. Similar gestures were explored, especially for interaction with wall displays using a mobile phone. Hassan et al. [81] introduced the Chucking gesture in which users tap and hold an icon on the touchscreen, followed by a toss measured by the accelerometer to transfer the file to the wall display. To transfer items between public displays using a mobile phone, Boring et al. [24] proposed a similar gesture in which users hold an object on the touchscreen and move mobile devices between displays. The researchers also used the built-in accelerometer to improve text entry on mobile devices. This includes the use of device orientation to resolve ambiguity on a T9 keyboard [239] and the improvement of one-handed gestural text input on large mobile devices [252].

In the next section, we explain how haptic feedback has been used to improve the interaction with touch devices in visually impaired scenarios.

1.4 Eyes-free Interaction with Smartphone Through Haptic Channel

Direct touch interaction on mobile phones revolves around screens that compete for visual attention. Even in the presence of visual feedback, smartphone screens are significantly smaller than those of computers, resulting in intentionally restricted mobile applications compared to their desktop counterparts. Current touchscreen gadgets extensively use single-finger swipe to navigate viewports. This method relies heavily on graphical widgets for various operations. For instance, users need to alternate between swiping fingers (to adjust the viewport) and manipulating small graphical handles (to set the selection range). The presence of widgets on touchscreens not only limits the space available for the main content, but also gives rise to usability challenges. For example, accurately targeting and manipulating objects with fingers can be laborious on screens of all sizes [155]. Therefore, it should come as no surprise that an extensive amount of research in HCI is dedicated to improving interaction with touch devices.

The issue of interacting with small touchscreens becomes even more challenging in visually impaired scenarios. In the absence of visual feedback users need to interact with smartphone eyes-free. There is some understanding of how eyes-free interaction can be enhanced with the aid of vibrotactile feedback to support interactions [159]. Vibrotactile feedback is widely used in smartphones to provide alerts and messages and to enhance the user experience through an additional feedback channel [70]. Vibrotactile signals can convey non-visual information using various vibration patterns by altering the frequency, intensity, and duration of the vibrations [28, 35]. However, touchscreens do not offer any distinct haptic cues to the finger(s) other than the sensation of touching a smooth, featureless glass surface. To address this lack of inherent haptic cues and to facilitate interaction with on-screen control elements, touchscreen-based haptic interactions must depend on external haptic feedback (e.g., vibration, friction, or electrostatic signals) and kinesthetic hand movements during screen exploration [212].

1.4.1 Tactile feedback devices

Researchers have introduced various novel methods to provide haptic feedback on touchscreens [15, 118, 259, 235, 178]. The touch screen, in combination with haptic actuators, can be considered a tactile interface. Tactile interfaces can be categorized into three types depending on the implementation methods.

First, modulating friction between the finger and the touch screen [106] through haptic surfaces [15, 118, 235, 178]. Haptic surfaces produce haptic feedback directly on the user's fingertip by altering the friction between the skin and the glass surface [17]. Two primary technologies have been developed on haptic surfaces: electrovibration technologies [15], which increase friction between the finger and the surface, and ultrasonic technologies, which decrease friction using the squeeze film effect [118, 235].

Second, haptic feedback can be provided on touchscreens using intermediary components. Roudaut et al. [184] developed two prototypes capable of moving the finger on the screen to replicate a specific gesture, such as a letter or symbol, without requiring visual attention.

Third, delivering direct vibration below 500 Hz between the finger and the touch screen. For example, Zhao et al. [259] generated the perception of a moving tactile sensation on a tablet by integrating multiple vibrators into the device. Vibrotactile actuators remain the most straightforward and prevalent means of implementing haptic feedback on these smartphones. They offer benefits such as straightforward and cost-effective technology and simple control. However, the disadvantage is that they have limited expressiveness [173].

1.4.2 Using haptic feedback to improve touch input

To avoid the extensive use of limited screen space or the extensive use of menus, a promising solution to increase interaction bandwidth is to enrich the input vocabulary [66]. Increasing the input bandwidth of touch devices, if it is independent of vision, can enhance interaction for visually impaired users. Many works studied new dimensions for increasing the input vocabulary in touch devices, such as orientation [64], the force exerted during a contact [65], or with which finger [67] or part of the hand [56] interacts. These methods expand touch devices' input vocabulary beyond 2D touch points. On the other hand, in some studies, instead of trying to increase the input bandwidth, they focused on improving the quality of input by providing feedback to user. Haptic feedback has been used to improve the touch screen input. It allows the development of user interfaces that require minimal user attention [164]. Hoggan et al. [89]examined text entry accuracy using fingers on a mobile device with and without haptic feedback. Their study revealed that haptic feedback integration led to improved accuracy. Snibbe et al. [204] introduced multiple rendering strategies to provide physical feedback to continuous interactions. Liao et al. [124] improved the effectiveness of multiple dwell selection using active haptic ticks.

Especially many studies have used haptic feedback for targeting tasks on touchscreens. Targeting is a crucial activity in human-computer interaction and has been the subject of numerous studies. Levesque et al. [119] studied the effect of a variable friction interface on targeting tasks. They found that variable friction improves performance and can have a positive impact on enjoyment, engagement, and the sense of realism experienced by users of touch interfaces. Similarly, Casiez et al. [34] examined the impact of ultrasonic friction modulation on an indirect target acquisition task using a haptic trackpad. Their findings indicate an enhancement of approximately 9% in targeting performance when there are no distractors, with comparable performance in the presence of distractors. They suggested that performance improvement was caused by tactile information feedback and that the feedback is more effective when friction is increased on targets using a simple step function. Zhang et al. [257] complement these previous studies and quantify how haptic feedback with electrovibration can be used to improve user performance. In particular, four haptic feedback designs—Line Center, Line Leading Edge, Line Background, and Fill—were evaluated to determine where to place the haptic feedback on the target. They observed a 7.5% enhancement in performance when the whole target was filled with a tactile texture. Kalantari et al. [97] extended this work to ultrasonic devices. Bernard et al. [18] investigated the effect of gradual tactile feedback on visual cues for a sliding task. They showed that after training with visual and auditory feedback, users can perform the task eyes-free.

Eyes-free input to an interface is also useful when the user is simultaneously involved in another task. For example, in environments requiring visual attention, such as driving, the lack of haptic feedback could significantly hinder safety, making this critical feedback [168]. Different studies [189, 47, 77, 19] have been conducted to determine the effect of tactile feedback on user interaction when performing a secondary task and a primary task that saturates attention. Several related studies have reported that haptic feedback improves response time and performance [120, 38]. Beruscha et al. [19] compared three feedback conditions (visual, visual and haptic, and haptic) when selecting target buttons on a touchscreen during simulated driving. Their findings indicated that haptic feedback reduces gaze away time from the road and perceived workload. When visual feedback is added, gazes on the touchscreen are more frequent. Cockburn et al. [47] investigated users' performances when interacting with a touchscreen covered with a static stencil overlay while driving in a 2D emulator. Their findings indicate that with stencil, the targeting task is faster and demands less visual attention to the touchscreen device. Harrison et al. [77] investigated the relevance of dynamic buttons displays based on pneumatic actuation when the user is performing simultaneously an attention saturating primary task. They used the same attention-saturating dual task framework as in [26] in which the task consisted of keeping a moving circle centered on a fixed crosshair. Their findings showed that pneumatic displays perform as well as physical buttons with fewer glances towards the surface when performing the primary task.

1.4.3 Using haptic feedback to enrich output information on touchscreens

Interactive systems integrate hardware and software components to allow users to execute various categories of tasks. Since the advent of interactive computing systems, vision and audition have served as the main channels for user feedback. Extensive research has been conducted by the human-computer interaction community on leveraging these feedback modalities for interaction techniques. Contrarily, interactive systems have yet to fully harness the potential of the sense of touch, even though it offers numerous benefits across diverse fields such as well-being [82], immersion [180], and assistive technologies [23]. Haptic feedback has the potential to substitute visual and auditory information.

There have been many attempts to convey speech and / or language through the haptic channel for situational impairment scenarios. A recent review provides a comprehensive summary [100] of hands-free devices designed to transmit speech and language through tactile means, distinguishing actual language units and words by encoding speech components such as phonemes, letters, and Morse code. A notable study addressing the potential data throughput of the skin, specifically the bandwidth of vibrotactile communication, is that of Novich et al. [153]. They revealed that vibrotactile patterns encoded in both spatial and temporal dimensions via a haptic vest outperform those encoded solely in spatial dimensions. Zhao et al. [260] demonstrated that users could memorize

20 haptic words after 30 minutes of training using six actuators on the forearm, while Tan et al. [210] achieved the acquisition of 500 words following a 10-day training period. At the application layer, WhatsHap was created by Marino et al. [132], facilitating conversation between two users by translating speech into tactile sensations on their arms.

Many studies have broadened the application of tactile feedback for the interpretation of graphical information on touchscreens. This includes in particular the representation of maps [170], the tracking and understanding of lines and graphs [213, 62], and the illustration of graphical shapes [212]. Further research has shown that it is feasible to pan and zoom in large-scale vibrotactile maps non-visually, even when they extend beyond the display of the device [162, 163].

However, the bandwidth of information that can be provided by these studies is limited by the haptic technology used. An other body of studies has focused on increasing the output bandwidth of tactile using in current technologies. For example, ActiVibe changes the duration and pattern of vibration in an actuator with fixed vibration frequency [36]. On the other hand, one can leverage the variability in the touch input as a means to increase the output bandwidth. Rekik et al. [175] improved the haptic feedback bandwidth on a touch interface by enabling users to dynamically choose among various tactile feedback channels through changes in the speed of the finger. Similarly to speed, the orientation of the device has been leveraged to create various haptic channels [21]. Roudaut et al. [184] instead used a different approach. They installed a force-feedback system on the touchscreen that can move the touchscreen beneath the user finger. However, this method relies on external hardware that is not available on commercial smartphones. Even if an approach does not need any external hardware, the question arises as to which type of haptic feedback should be used for each scenario. When designing the output of an interface, it is important to look beyond the hardware and emphasize the ability and preferences of users for the available interaction means.

1.4.4 Ability-Based Design for Haptic Interaction

The properties of haptic feedback (such as duration and interval of the stimulus, presentation location, amplitude, and frequency) can be modified to generate unique experiences that can be incorporated into an interface. However, these signals are often chosen at random by interface designers without considering the meanings that users might attribute to tactile feedback [172]. In addition to user preferences for input gestures, numerous recent studies have employed the elicitation methodology to investigate user preferences for haptic feedback. Haptics are typically used as feedback mechanisms or notifications. Lawrence et al. [225] investigated user preferences for mid-air haptic feedback linked to gestures to interact with an augmented reality menu. Kim et al. [107] conducted an elicitation study on haptic feedback patterns for social touch using the SwarmHaptics haptic display. Wei et al. [238] carried out a comparable experiment on social touch mediated on touchscreens.

Although this information on the use of haptic feedback looks promising and intriguing, most of it is based on hardware that is not available on commercial mobile devices. It is preferable for solutions to use built-in mobile device sensors to avoid the imposition of additional equipment on users [192]. Recognizing the abilities of people whose senses are impaired is integral to the ability-based design approach (that is, consideration of what users can do instead of what they cannot do, forcing systems to adapt to users rather than the other way around) [241].

For this reason, in this thesis, we mainly focus on introducing interaction techniques that use widely available sensors in commercial smartphones that allow for input using the haptic channel (touch sensor, accelerometer, and gyroscope) in visually impaired scenarios. There are ways to enrich the input/output vocabulary without any additional hardware. For example, eyes-free gestures on touchscreens can be a solution. They can reduce the cognitive load [27], memorize as motor skills, help users concentrate on their activities [114], typically do not require dedicated screen space [27], and can be operated with one hand. Another widely available sensor in smartphones is the accelerometer. For example, accelerometer sensor data can be used to detect walking and adapt the keyboard to reduce typing errors [63]. Automatic detection of SIIDs during mobile interaction opens the way to relevant interface adaptations; thus allowing interactions with mobile devices to be more appropriate to the user's context [217].

1.4.5 Rethinking stroke gestures and phone movements as input for smartphone through haptic channel

Haptic perception arises when the mechanoreceptors present in the skin, muscles, tendons, and joints are stimulated by the manual exploration of an object in space. The brain integrates the information about the parts of the body touching the object with the information about the position of the parts of the body in space to form the haptic perception [60]. In the context of interaction with the smartphone, haptic perception does not form only when the vibrotactile actuator in the smartphone vibrates but also forms when users are giving a command to the phone. Any kind of input to the smartphone that is done through hands will form a haptic perception for the user, such as tapping or stroking a finger on the touchscreen or moving smartphone in 3D space.

When the user makes a gesture on the touchscreen using the fingers, the smartphone detects it using the capacitive touch sensors [147]. However, from the user's perspective, stroke gestures are the result of two sensory perceptions. First, tactile perception detects the contact between the finger and the touchscreen. Second, proprioception tracks the movement of the finger on the touchscreen. These two perceptions are part of our haptic perception of the surroundings [142]. Similarly, when the user moves the smartphone in 3D space using his/her hand to invoke a command such as rotating the screen, the smartphone senses these movements using the accelerometer and gyroscope, and a haptic perception forms in the user brain about this movement even if the user cannot see the phone.

From user's perspective, we consider stroke gestures and phone movements as haptic modalities for interaction with the smartphone. When the haptic perception is formed as a result of the user wanting to provide information to the smartphone about the intention and his/her context, we consider it as input part of the interaction. On the other hand, when haptic perception is formed as a result of the transfer of information from the smartphone to users, we consider it as the output part of the interaction. For example, a user might explore a haptic surface to feel the simulated texture of a fabric.

1.5 Conclusion

In this chapter, we provide an overview of the sensory channels available for interacting with a smartphone. We provide some examples of previous research that has used each of these sensory channels. We explain the advantages and limitations of each sensory channel. Then, we covered the environmental scenarios that can cause impairment in the interaction with smartphones. We explained how these situational impairments have been shown to affect the interaction with touch devices in previous studies. Then, we explained the benefits of using gestures and haptic feedback in such scenarios, especially visually impaired scenarios. We covered how gestures are analyzed and recognized in smartphones. Then, we explained how in this thesis we look at the human-computer interaction from user perception perspective. We then covered previous studies which have used haptic channels to interact with smartphones. We explained how from user's perspective stroke gestures, recorded by touchscreen, and phone movement, recorded by accelerometer and gyroscope, are part of user's haptic channel, which are used for invoking command on smartphone (input). Similarly, tactile feedback felt from vibrotactile actuators or haptic surfaces by the user is part of the haptic channel but is used to retrieve information (output). We also explained how some studies, called elicitation studies, investigated the preferences of users for gestures and haptic feedback. Most of the studies we covered were conducted in visually impaired scenarios. In the following chapters, we present our studies and our contribution to the literature. We divide our studies into three parts. First, we investigate the effect of different situational impairment scenarios on the production of eyes-free gestures. Then, we present how to use haptic feedback to improve gestural input on touchscreen. Finally, we introduce an interaction concept that allows complete eye-free gestural interaction with the touchscreen and information retrieval through the haptic channel.

Chapter 2

Production of Eyes-free Gestures in Different Scenarios

Eyes-free gestures on smartphones are crucial as they improve user accessibility and safety, allowing efficient device interaction without the need to visually engage with the screen. This is especially beneficial in situations causing visual impairment, either pathological (*i.e.*, individuals with visual impairments) or situational nature (*e.g.*, walking, multitasking, encumbrance, screen glare, screen occlusion, etc.) [227]. Most eyes-free gestures are designed to be performed with one hand [183], making them convenient and practical for users who need to operate their devices while their other hand is occupied. This chapter delves into the production of single-handed, eyes-free gestures caused by different environmental factors for sighted users. In particular, this chapter explores three distinct studies, each focusing on a specific question related to eyes-free gestures production with smartphones.

The primary aim of the first study is to determine in what ways gestures made eyes-free differ from those made when the phone is visible. In particular, as some existing techniques (*e.g.*, drawing applications) provide the users a visual trace of the produced gesture, while others do not, our study aims to investigate how eyes-free gestures differ from those produced in the presence or absence of a visual trace when the user can see the mobile device?

In the second study, we will explore how real-life contextual factors can affect the production of eyes-free gestures. For example when the user is walking or when the smartphone is in a closed small location such as a bag. We investigate whether such contextual conditions will affect the production of eyes-free gestures. In particular, the goal of this study is to explore how environmental factors, such as how movement speed of the user and location of the phone, influence the characteristics of eyes-free gesture input.

Another important use case for eyes-free gestures is when a user is multitasking. Eye-free gestures are more common when the other task is the primary task, such as when performing a saturation attention task, such as driving. Involving in another primary task can potentially change the way gestures are produced. The third study aims to investigate how an attention-saturating task and the hand used for eyes-free gesture input the production of gesture and user experience?

This chapter is organized into three sections, each dedicated to a distinct study of how eyes-free gestures are produced under various conditions. In each section, we begin by introducing the objective of the study, followed by a detailed description of the experiment design. Next, we present the results of the study, which are discussed in depth. Finally, each section concludes with a summary that summarizes the key findings and insights gained from the study.

2.1 A Comparative Study of Eyes-Free and Visual-Based Mobile Gestures

There is a difference in the dynamic behind the production of gestures when visual feedback is available versus when it is missing during smartphone interactions. When visual feedback is present, users can see and adjust their gestures in real time. In contrast, in the absence of visual feedback, users rely on their proprioception and muscle memory, which can lead to variations in the shape and execution of gestures [130]. These variations could potentially affect the accuracy of gesture recognition, as the geometry of gestures directly influences the accuracy with which the device recognizes them. In this section, we will study how the production of eyes-free gestures differs from those made when the user can see the touchscreen (with visual trace of gestures or without visual trace of gestures).

2.1.1 Experiment

In this study, we compare eyes-free gestures with those produced when one can see the phone, either in the presence of visual feedback (*i.e.*, the gesture trace visible to the user) or in the absence of visual feedback (*i.e.*, the gesture trace is not visible to the user), in one-handed mobile device interaction.

2.1.1.1 Participants and Apparatus

(a) (b)

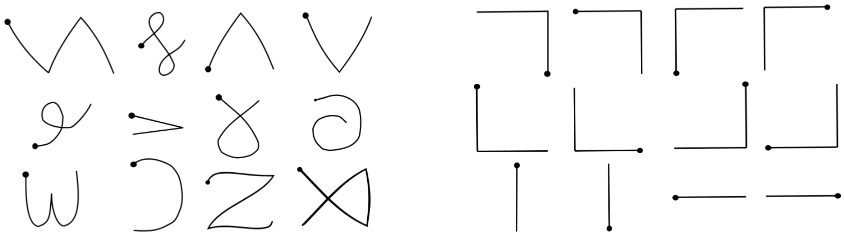
Figure 2.1: The experimental setup for the three feedback conditions. (a) With and without visual trace conditions, and (b) Eyes-free condition.

21 right-handed volunteers (8 women and 13 men) participated in the experiment. They were between 23 and 36 years old ($mean = 29.33, SD = 4.09$).

We used a Huawei Y7 Pro running an Android operating system with an 8.1.0 version. The dimensions of the phone were 6.23"× 3.02"×0.32". We developed our main application using JavaScript on the React Native framework. The application was designed to support the three visual conditions and permit us to log all input events generated by our participants in an embedded database. For the *eyes-free* condition, in order to display the required gesture to draw, we implemented another application in JavaScript, which ran on a Dell laptop machine with a 13-inch LCD display screen with a desktop resolution of 1920×1080 pixels. Both applications communicated through the TCP protocol. The laptop application was only responsible for displaying gestures, and all computing was still done on the phone side. We added a small piece of cardboard to the top of the screen to allow users to perform bezel gestures from the top of the screen without engaging the Android status bar. The dimensions of the cardboard were $1cm$ in height and $0.2cm$ in thickness, covering $0.35cm$ of the screen. The user's hands were videotaped using two Sony HDR cameras as shown in Figures 2.1.

2.1.1.2 Gesture Set

We consider the same gesture sets used in [26]: mark-based gestures and free-form gestures. Each set is composed of 12 unistroke gestures carefully selected by the authors [26] and are frequently used in the literature (*e.g.*, [139, 88, 123, 244, 11]. The free-form gestures were composed of operands, rationally invariant, and mnemonic gestures and were simple to perform, did not include any on-axis marks, were distinct enough, and could be drawn imprecisely [26] (see Figure 2.2a). Similarly to marking menus, mark-based gestures consist of on-axis rectilinear mark segments like those in the primary compass directions (N, S, E, w or up, down, left, right) and off-axis marks of 45° angles (NE, SE, SW, NW), which can form a compound path, , *e.g.*, "down" followed by "left" [26] (see Figure 2.2b). These two sets were shown to be producible without looking at the phone [26]. For example, when performing those gestural interactions while being distracted with another primary task, the user's gaze toward the phone is only 3.5% of the time.



(a) (b)

Figure 2.2: Gesture sets used in the experiment. (a) Freeform gestures, and (b) Mark-based gestures.

2.1.1.3 Design

The experiment used a 3 × 2 × 2 within-subject design with these factors: visual *feedback*, gestures *set* and gesture *beginning*. Visual *feedback* covers three conditions: (1) *with visual trace*, which means that participants can see the phone and have a visual trace of the articulated gesture; (2) *without visual trace* feedback, which means participants can see the phone but not the trace of the articulated gesture, and (3) an *eyes-free*, where participants cannot see the phone and nor the gesture trace. *Gesture set* and *gesture beginning* are similar to the work of Bragdon et al. [26]. *Gesture set* covers two conditions: *freeform* gestures and *mark-based* gestures. Each set is composed of 12 unistroke gestures (see Figure 2.2). Gesture *beginning* covers two conditions: (1) *free*, where gesture can start from anywhere on the screen, and (2) *bezel*, where gesture should start from one of the four edges of the screen.

2.1.1.4 Task & Procedure

Participants held the phone with their dominant hand and used the thumb of their dominant hand to draw gestures on the screen (single-handed grip), while sitting in front of a desk. For the *with/without visual trace* conditions, participants interacted only with the phone. A preview of the gesture to be drawn was shown at the top left of the screen (Figure 2.3). After drawing gestures, participants could click on the "next" button, which was placed on the bottom right of the screen, to draw the next gesture. To be sure that the locations of the preview of the current gesture and the "next" button will not interfere with gesture articulation, we followed Eardley et al.'s [54] work and placed them in a non-functional area of the thumb for single-handed grips. For the *eyes-free* condition, participants dominant hand was put under the desk to ensure that the participants could not see the smartphone. A preview of the gesture to be drawn was shown on the laptop in front of them.

In the experiment phase, the three visual *feedback* conditions were counter-balanced among our 21 participants, and within each visual *feedback*, the two

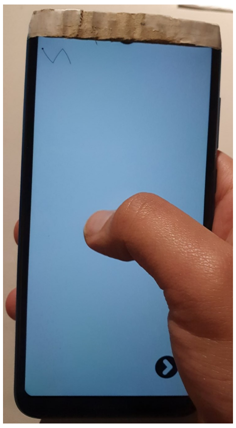


Figure 2.3: A preview of gesture was showing on the corner of the screen for with/without visual trace conditions.

gestures *beginnings* were counterbalanced. Within each gesture *beginning*, the two gesture *sets* were also counterbalanced. For each set of gestures, the twelve types of gestures and their five repetitions were randomly presented, resulting in a total of 720 gestures (= 3 *feedback* × 2 *beginning* × 2 *set* × 12 types of gestures × 5 repetitions) drawn per participant. The experiment took one hour on average to complete.

2.1.2 Results

Our results include gesture features, gesture recognition, and mobile directional movements. All analyses used a multi-way ANOVA. Tukey tests were used for post hoc analysis when significant effects were found. Only significant effects and interactions are reported. We also analyzed qualitative observations.

2.1.2.1 Gesture Features

We selected six geometric features: (a) gesture length, (b) gesture height, (c) gesture width, (d) gesture area, (e) gesture duration, and (f) gesture speed. These features have been used in the literature on gesture recognition and analysis literature [4, 188, 22, 232, 177] in order to characterize how users produce strokes.

- **Gesture length** is the cumulative distance between the first touch event recorded for the gesture and the last. There were significant main effects of *feedback* ($F_{2,40} = 53.49, p < .0001$), *set* ($F_{1,20} = 360.15, p < .0001$) and *beginning* ($F_{1,20} = 18.00$, $p = .0004$) on *gesture length* with significant *set*× *beginning* ($F_{1,20} = 14.20$, $p = .0012$) interaction. Post hoc tests revealed significant differences among the three feedback conditions. Gestures made with visual trace ($mean = 7.66cm, SD = 0.11cm$) were the shortest, followed by those made without visual trace ($mean = 8.68cm, SD = 0.11cm$), and then eyes-free gestures ($mean = 10.21cm, SD = 0.12cm$) ($p < .05$). Furthermore, freeform gestures (bezel: $mean = 11.89cm, SD =$

$.14cm$, free: $mean = 10.58cm, SD = .11cm$) were longer than mark-based gestures (bezel: $mean = 6.77cm, SD = .08cm$, free: $mean = 6.16cm, SD = .07cm$) ($p < .05$). Freeform condition produced also, significantly longer gestures when they started from the edge than when they started on a free space ($p < .05$).

- **Gesture height** is the height of the smallest bounding box that contains the gesture ($max_{\mathbf{y}}$ - $min_{\mathbf{y}}$). There were significant main effects of *feedback* ($F_{2,40} = 38.10$, $p < .0001$), *set* ($F_{1,20} = 8.78$, $p = .0077$) and *beginning* ($F_{1,20} = 26.75$, $p < .0001$) on *gesture height*. Post hoc comparisons revealed that eyes-free gestures ($mean = 44.98cm, SD = .51cm$) are significantly higher than gestures produced in the absence of visual trace ($mean = 3.87cm, SD = .04cm$) or in the presence of visual trace ($mean = 3.41cm, SD = .04cm$) ($p < .05$)). Freeform gestures ($mean = 4.03cm, SD = .03cm$) were significantly higher than mark-based gestures ($mean = 3.82cm, SD = .04cm$) ($p < .05$). Bezel gestures ($mean = 4.22cm, SD = .04cm$) were significantly higher than free gestures ($mean = 3.63cm, SD = .03cm$).

- **Gesture width** is the width of the smallest bounding box that contains the gesture ($max_{\mathbf{x}}$ - $min_{\mathbf{x}}$). There were significant main effects of *feedback* ($F_{2,40} = 67.42$, $p < .0001$), *set* ($F_{1,20} = 205.08$, $p < .0001$) and *beginning* ($F_{1,20} = 30.61$, $p < .0001$) on *gesture width* with significant *feedback*$\times$*set* ($F_{2,40} = 13.37$, $p < .0001$), *set*$\times$ *beginning* ($F_{1,20} = 89.36$, $p < .0001$) and *feedback*$\times$*set*$\times$*beginning* ($F_{2,40} = 6.63$, $p = .0032$) interactions. Post hoc comparison showed that mark-based gestures were significantly narrower than freeForm gestures for all feedback and beginning conditions ($p < .05$). For mark-based gestures, eyes-free gestures were significantly wider than both gestures produced in the presence and absence of visual trace ($p < .05$). For freeForm gestures, for each feedback condition, bezel gestures were articulated with significantly larger width than those started on a free space ($p < .05$). When starting from the edge, in the presence of visual trace, articulated gestures were significantly narrower than those in the absence of visual trace or eyes-free ($p < .05$) with a significant difference between no-visual trace and eyes-free conditions($p < .05$). For freeForm gestures, when starting on a free position, eyes-free gestures were produced with significantly larger width than gestures produced in the presence or absence of visual trace ($p < .05$) with no significant differences between the presence and the absence of visual trace.

- **Gesture area** is the surface area of the smallest bounding box containing the gesture (height $\times$ width). There were significant main effects of *feedback* ($F_{2,40} = 52.90$, $p < .0001$), *set* ($F_{1,20} = 95.67$, $p < .0001$) and *beginning* ($F_{1,20} = 29.45$, $p < .0001$) on *gesture area* with significant *feedback*$\times$*set* ($F_{2,40} = 14.65$, $p < .0001$), *set*$\times$ *beginning* ($F_{1,20} = 30.25$, $p < .0001$) and *feedback*$\times$*set*$\times$*gesture beginning* ($F_{2,40} = 9.53$, $p = .0004$) interactions. Post hoc comparison showed that mark gestures were articulated with significantly smaller area than freeform gestures ($p < .05$). For mark gestures, eyes-free gestures were articulated with significantly larger area than gestures produced under the remaining feedback conditions ($p < .05$). For freeForm gestures, when staring the gesture from the edge, eyes-free gestures were articulated with significantly bigger area than gestures produced in the remainder feedback conditions with no visual trace being significantly bigger than those produced in the presence of visual trace ($p < .05$). In contrast, when the gesture starts freely, eyes-free gestures were produced with significantly bigger surface than gestures produced in the remaining feedback conditions ($p < .05$) with no significant difference between visual and no visual conditions.

- **Gesture duration** is the time elapsed while drawing the gesture, that is, the time of the last touch event registered for the gesture minus the time of the first touch event. There were significant main effects of *feedback* ($F_{2,40} = 8.0$, $p = .0011$), *set* ($F_{1,20} = 71.95$, $p < .0001$) and *gesture beginning* ($F_{1,20} = 5.31$, $p = .0319$) on *gesture duration* with significant *feedback*× *set* ($F_{2,40} = 6.63$, $p = .0032$) and *set*× *beginning* ($F_{1,20} = 4.91$, $p = .0383$) interactions. Post hoc comparison showed that for each feedback and beginning condition, freeForm gestures implies significantly more time to be drawn than mark gestures ($p < .05$). For freeForm gestures, the gesture duration in eyes-free condition ($mean = 1558ms, SD = 34ms$) was significantly bigger than no-visual trace ($mean = 1317ms, SD = 27ms$) with no significant difference with the visual one ($mean = 1431ms, SD = 30ms$). For that gesture set, starting the gesture from the edge ($mean = 1500ms, SD = 25ms$) tended to take significantly more time than when being free ($mean = 1371ms, SD = 24ms$) ($p < .05$). Interestingly, for mark gestures there were no significant differences between the different feedback and beginning conditions ($p > .05$).
- **Gesture speed** is the average speed registered over all the touch events belonging to a gesture (length/duration). There were significant main effects of *feedback* ($F_{2,40} = 21$, $p < .0001$) and *set* ($F_{1,20} = 15.79$, $p = .0007$) on *gesture speed*. Post hoc comparison showed that gestures produced in the presence of visual trace ($mean = 7.89cm/s, SD = .15$) were significantly slower than both gestures produced in the absence of visual trace ($mean = 9.62cm/s, SD = .15$) or eyes-free ($mean = 9.58cm/s, SD = .14$)($p < .05$). Freeform gestures ($mean = 9.03cm/s, SD = .08$) were significantly faster than mark gestures ($mean = 8.16cm/s, SD = .07$).

2.1.2.2 Gesture Recognition

As we found how gestures are produced eyes-free, in the presence, or absence of visual trace, it is important to study the impact of these differences on gesture recognition. $1 recognizer [244] was found inadequate for 1-D gestures that are presented in our mark set, So, we considered $N-protractor [9] recognizer for both sets. However, $N generates for each multistroke gesture all unistroke permutations [6]. In order to be independent of gesture direction and since we are concerned with stroke direction especially for mark-based gestures where two different gestures may share the same stroke shape but with different directions, *e.g.*, "up" and "down" gestures, we removed this feature. We then performed user-dependent training (which is appropriate for gesture recognition on personnel devices such as mobile phones [6, 9, 229]), in which recognition rates were individually calculated for each participant with the same methodology as in [244, 6, 9, 229]. Note that we conducted separate tests for each set of gestures generated in different feedback and beginning conditions. We found significant main effects of the feedback ($F_{2,40} = 7.03$, $p = .0024$) and *beginning* ($F_{1,20} = 14.69$, $p = .001$) on gesture recognition. The post hoc comparison showed that gestures produced eyes-free ($mean = 95.53\%, SD = .14\%$) were significantly less accurate than both gestures produced in the presence ($mean = 96.88\%, SD = .12\%$) or absence of visual trace ($mean = 97.27\%, SD = .11\%$) ($p < .05$)) with no significant difference between gestures produced in the presence and absence of visual trace ($p > .05$)). Free gestures ($mean = 97.68\%, SD = .11\%$) were significantly more accurate than bezel gestures ($mean = 95.44\%, SD = .11\%$) ($p < .05$).

2.1.2.3 Mobile Directional Movement

We considered the same dependent variables as in [54] to characterize the movement of the phone: Alpha (z axis), Beta (x axis), and Gamma (y axis) using the built-in accelerometer and gyroscope (see Figure 2.4) For each of the directional axes, using the built-in accelerometer and gyroscope we captured the total deviation made around this axis, calculated as the difference between the largest and the smallest value. Data were captured during gesture articulation.

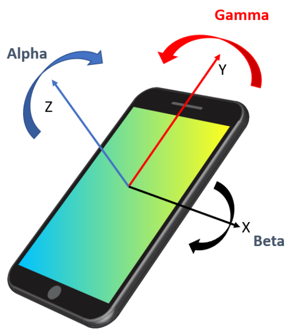


Figure 2.4: Axes and orientations of the smartphone.

- **Alpha deviation – deviation around z axis** There were significant main effects of *feedback* ($F_{2,40} = 3.31, p = .0464$), *set* ($F_{1,20} = 21.11, p = .0002$) and *beginning* ($F_{1,20} = 18.81$, $p = .0003$) on *alpha deviation* with significant *feedback*$\times$ *beginning* ($F_{2,40} = 5.24$, $p = .0094$) interaction. Post hoc comparisons showed that producing freeform gestures ($mean = 7.54°, SD = .21$) caused more alpha deviation than when producing mark gestures ($mean = 7.54\circ, SD = .21)(p < .05$). For eyes-free condition, starting the gesture from a free space ($mean = 6.04°, SD = .17$) caused less alpha deviation than when starting the gesture from the edge ($mean = 9.70°, SD = .65$). When starting the gesture on a free space, gestures produced eyes-free caused less alpha deviation than gestures produced in the absence of visual trace ($mean = 8.78°, SD = .19)(p < .05$).

- **Beta deviation – deviation around x axis** There were significant main effects of *feedback* ($F_{2,40} = 7.61$, $p = .0016$) and *beginning* ($F_{1,20} = 32.52$, $p < .0001$) on *beta deviation* with significant *feedback*$\times$ *beginning* ($F_{2,40} = 9.17$,

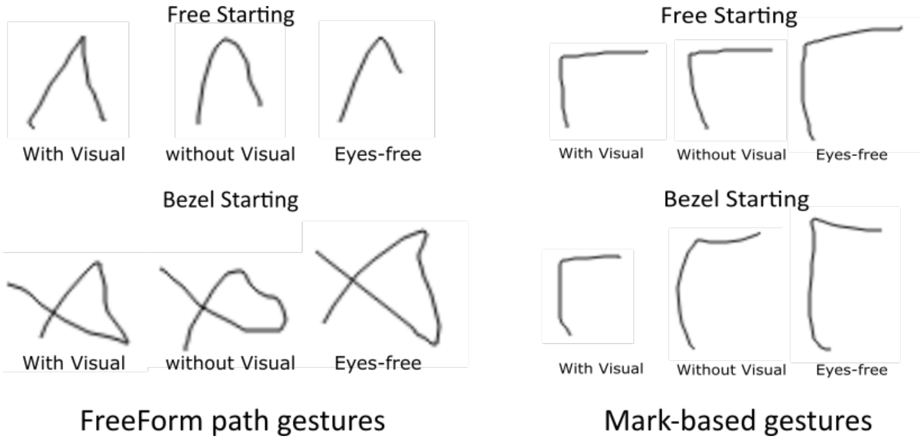


Figure 2.5: Examples of different gestures articulations produced in the three visual feedback conditions.

$p = .0005$) interactions. Post hoc comparisons showed that for each feedback condition, gestures started from the edge caused more beta deviation than those which started on a free space (visual: (1) bezel: $mean = 7.15°, SD = .21$; (2) free: $mean = 5.11°, SD = .13$, no-visual trace: bezel: $mean = 8.58\circ, SD = .23$; (2) free: $mean = 5.72°, SD = .15$, eyes-free: bezel: $mean = 9.38°, SD = .25$; (2) free: $mean = 5.67°, sd.14)(p < .05)$. For bezel gestures, the presence of visual trace caused less beta deviation than other conditions ($p < .05$).

- **Gamma deviation – deviation around y axis** There were significant main effects of *feedback* ($F_{2,40} = 23.21$, $p < .0001$), *set* ($F_{1,20} = 35.58$, $p < .0001$) and *beginning* ($F_{1,20} = 88.21$, $p < .0001$) on *gamma deviation*. Post hoc comparisons showed that, eyes-free gestures ($mean = 10.46°, SD = .19$) produced significantly more gamma deviation than both gestures produced on the presence ($mean = 7.43°, SD = .15$) or the absence ($mean = 8.22°, SD = .16$) of visual trace($p < .05$). Freeform gestures ($mean = 9.67°, SD = .14$) and bezel gestures ($mean = 10.81°, SD = .16$) produced significantly more gamma deviation than mark gestures ($mean = 7.74°, SD = .14$) and free gestures ($mean = 6.60°, SD = .10$) ($p < .05$).

2.1.2.4 Observations

The participants felt confident in their gesture production under the three feedback conditions. However, they noticed deformations when provided with visual feedback, especially if they had initially performed without it. Figure 2.5 illustrates gesture articulations under eyes-free and visual feedback conditions (with gesture trace, and without gesture trace). Six participants found drawing freeform gestures easier and more comfortable than mark-based gestures. They perceived mark-based gestures as initially simple but found it challenging to draw them perfectly when provided with visual trace. Additionally, they felt that freeform gestures appeared better executed than mark gestures in the presence of visual trace. However, and interestingly, participants noted that visual feedback allowed for greater accuracy, enabling them to correct their actions. This observation aligns with the findings of Anthony et al. [4], indicating a preference for visual guidance in gesture drawing. Finally, most of the participants found that the bezel gestures are more challenging and require more effort compared to free gestures. Some participants reported feeling "*more comfortable*" initiating bezel gestures from the right or left edges rather than the upper or lower edges, finding them easier to reach. These findings are consistent with the research by Karlson et al. [102].

2.1.3 Discussion

Our findings indicate that eyes-free gestures were produced with significantly longer length and height, caused more phone movement around the y axis, and were faster than gestures produced in the presence or absence of visual trace. These results are consistent across different gesture sets and gesture beginnings. In contrast, gestures made in the presence of *visual trace* were significantly shorter in length, height, and width (only when starting from the edge) and slower than those made *without visual trace* or in *eyes-free* conditions. In terms of gesture recognition, our findings indicate that eyes-free gestures were significantly less accurate than the two visual conditions, with no significant differences between gestures produced in the presence or absence of visual trace. Consequently, for eyes-free gestures, which are typically longer, larger, and faster, we recommend employing recognizers that rely on geometric and kinematic gesture descriptors with caution, such as those proposed by Rubine[188](p. 335).

When producing gestures in the *eyeless* condition, bezel gestures were less accurate and implied more phone movements around all axes than free gestures. The same results were also observed for the two other visual conditions, except for the directional movements around the z axis, and were consistent across the different set conditions. These findings suggest that it is preferable to use freely initiated gestures rather than bezel gestures for better accuracy and reduced phone movement. However, if bezel gestures are necessary, we recommend to opt for left and right edges, which aligns with the findings of [199].

Mark-based gestures took considerably less space to be drawn in an *eyes-free* interaction compared to *freeForm* gestures: they were shorter in height and length. Consequently, they generated less directional movements to be performed than the *freeForm* set around both the y and z axes. The same results were also observed for the two other visual conditions, and were consistent across the different set conditions. These results suggest that gestures composed of at most two mark segments took considerably less space for articulation than freeform gestures, making them more suitable for small areas.

2.1.4 Summary

In this study, we present the results of an investigation concerning *eyes-free* stroke gesture articulation for a single-handed mobile interaction. Our findings indicated that gestures articulated in an *eyes-free* interaction were geometrically different from gestures generated in the presence or absence of visual trace, implied more rotational movement around the y axis, and were less well recognized compared to gestures made under other feedback conditions. In addition, we captured several observations concerning users' hand movements and articulation behavior when generating gestures under the different feedback conditions. For example, designers are recommended to use free gestures rather than bezel gestures and to use the right and left edges if using bezel gestures is required.

Given the distinct nature of gesture articulation under eyes-free conditions, it is important to recognize that these results may not be generalizable to all eyes-free gestures, especially when considering different situational impairment scenarios. This raises open questions about how various situational factors might further influence eyes-free gesture input. In the following sections, we will focus on three specific scenarios: user movement speed, phone location, and the presence of a primary attention-saturating task. By examining these scenarios, we aim to understand how these particular scenarios impact the production of eyes-free gestures, thus paving the way for a more comprehensive understanding of eyes-free gesture production.

2.2 Effects of Movement Speed and Phone Position on Eyes-Free Gesture Input with Mobile Devices

In continuation of our exploration into eyes-free gesture articulation, this study investigates the effect of user's moving speed and phone location on the articulation characteristics of gesture input. When moving, visual focus on the environment is more necessary, and looking at the smartphone in this scenario can be potentially dangerous and cause accidents [157]. That is why having an eyes-free interaction with smartphone can be beneficial in this scenario. Another scenario of eyes-free interaction with the smartphone is when one does not want to remove the phone from the bag due to lack of privacy or the danger of robbery in a crowded space [156]. In this study, our aim is to understand how these two scenarios influence eyes-free gesture production.

2.2.1 Experiment

We conducted an experiment to evaluate the effect of the user's movement speed and phone location on eye-free gesture articulation on a mobile device.

2.2.1.1 Participants and apparatus

Since the average speeds of walking and jogging between the sexes are different and the same speed can cause different mental and physical loads for the different sexes [46], we conducted experiments with only 12 male participants to avoid increasing the number of independent variables in the analysis. The ages of the participants were between 20 and 34 years ($mean = 26.8$, $SD = 4.3$). All participants were right-handed, without any known mobility impairment, and had been using smartphones for several years.

We collect stroke gestures using our custom software on a Samsung Galaxy S7 smartphone running Android 6.0.1. The smartphone was attached to a 1.5-meter cellphone lanyard to ensure that it did not fall. The smartphone's screen size is 5.65"×2.78" with a display resolution of 1440×2560 pixels and a pixel

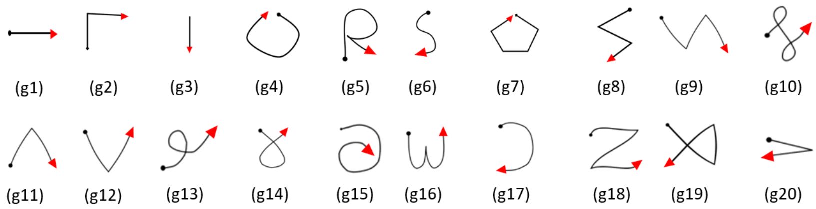


Figure 2.6: Gesture set used in the experiment.

density of 227 pixels per cm. The smartphone screen was mirrored on a Samsung Galaxy Tab 7 tablet in front of the participants while they were standing, walking, or jogging on a FreeMotion Reflex T11.8 treadmill. The dimensions of the shoulder bag used in the experiments were $20cm \times 20cm \times 5cm$.

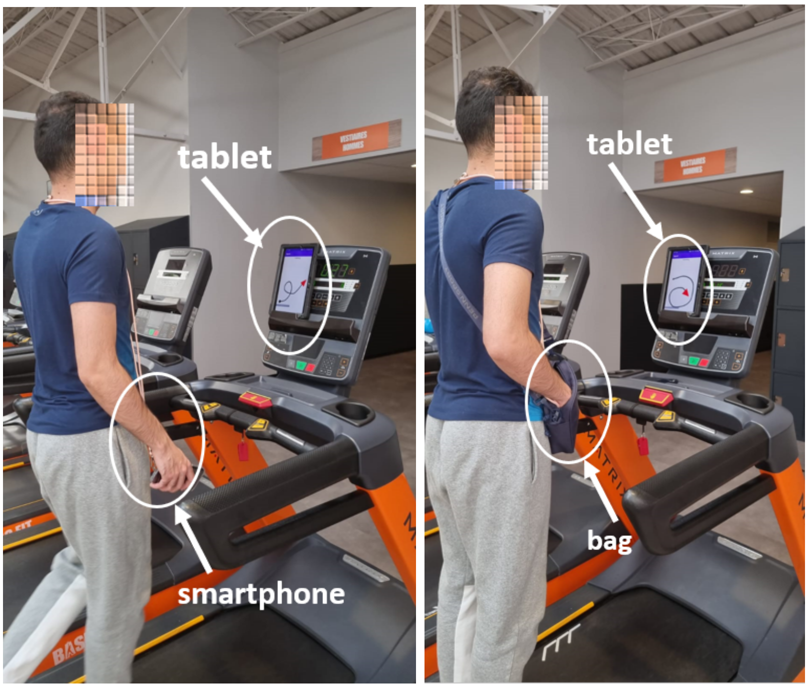
2.2.1.2 Gesture set

Based on the results and the guideline of the previous study, it is better to use free-from gestures for eyes-free interaction. So, for this study and the following ones, we used a set of 20 gestures which includes primarily freeform gestures (unlike theprevious study which included 50 percent freeform gestures). In the new gesture set, we keep the most popular mark-based gestures, such as swiping in the vertical and horizontal directions. These gestures were selected from previous works (*e.g.*, [27, 11, 244]) and were composed of operands, letters, mark segments, rationally invariant, and mnemonic gestures (see Figure 2.6).

2.2.1.3 Design

The experiment used a 3 × 2 within-subject design with two factors: *moving speed* and *smartphone location*. We followed [99] and chose to control the movement speed during the experiment. As in [99], the rationale for fixing the walking speed is that, first, we assume that users will not be able to slow down or stop walking to use their mobile device and, second, by doing so, the effects of impaired walking are maximized, as users cannot slow down if the task becomes difficult. *Moving speed* covers three conditions: (1) standing at 0 km/h, (2) walking at 4.6 km/h, and (3) jogging at 8 km/h. The moving speed values are defined through preliminary experiments with two participants. We chose speeds that are different and fast enough so that they have the potential to cause an effect while still being comfortable for participants to hold and use their smartphones. Participants in the preliminary experiment reported high frustration and difficulty in conducting experiments at high speeds, such as 10 km/h. Therefore, we did not consider running conditions. *Phone location* describes how the phone is held by the dominant hand and covers two conditions: (1) *free* where the phone is held freely alongside the body (see Figure 2.7a) and (2) *bag* where the phone is held inside a shoulder bag (see Figure 2.7b).

2.2.1.4 Task & Procedure



(a) (b)

Figure 2.7: The two configurations in which participants held the phone in the experiments. Figure (a) shows how participants held the smartphone freely in their hand. Figure (b) shows how participants held smartphone in a shoulder bag.

During the experiment, participants stand on a treadmill in a gym. The participants were then asked to hold the phone with their dominant hand and use the thumb of their dominant hand to draw gestures on the screen without looking at the phone. In the free condition, participants were asked to hold the phone alongside their body (see Figure 2.7a). In bag condition, participants were asked to hold the phone inside a shoulder bag (see Figure 2.7b). A preview of the gesture they had to draw was shown on a tablet placed in front of them on the treadmill. Since all gestures in this experiment were single strokes, as soon as they lifted their finger from the touchscreen, the next gesture appeared. In case of a false entry, participants always had the option to use the back button of the phone to return to previous gestures.

In the experiment phase, two conditions for the location of the phone were randomly presented to the participants. For each phone location, the three moving-speed conditions were also randomly presented. Each participant in total performed 600 gestures (=2 *phone locations* $\times$ 3 *moving speeds* $\times$ 20 gesture types $\times$ 5 repetitions). For each moving speed and phone location condition, the gestures were presented to the participant in a random order. The experiment took 30 minutes on average to complete.

2.2.2 Results

Our results include gesture features and smartphone directional movements. We also analyzed the qualitative observations. All analyzes used a two-way ANOVA. Tukey post-hoc tests were used when significant effects were found. Only significant effects and interactions are reported.

2.2.2.1 Gesture Features

We used the same gesture features as previous study to analyse how shapes of gestures changed by moving speed and phone location.

- **Gesture length.** We found significant main effects of *speed moving* ($F_{2,22} = 27.232$, $p < .0001$) and *phone location* ($F_{1,11} = 10.659$, $p = .008$) on *gesture length*. Jogging ($mean = 10.63cm, SD = 4.53cm$) led participants to enter gestures with the longest lengths, then walking ($mean = 10.20cm, SD = 4.63cm$) and standing ($mean = 9.67cm, SD = 4.55cm$). The post hoc test confirms differences between all pairs ($p < .05$). Post-hoc tests show that the jogging ($mean = 10.63cm, SD = 4.53cm$) led participants entering gestures with significantly longer lengths than both walking ($mean = 10.20cm, SD = 4.63cm$) and standing ($mean = 9.67cm, SD = 4.55cm$), and walking led to gestures with significantly longer lengths than standing. Holding phone freely ($mean = 10.52cm, SD = 4.53cm$) led also to gestures with significantly longer lengths than when holding it in a bag ($mean = 9.85cm, SD = 4.45cm$)($p < .05$).

- **Gesture height.** We found significant main effects of *moving speed* ($F_{2,22} = 6.444$, $p = .006$) and *phone location* ($F_{1,11} = 23.963$, $p = .0005$) on *gesture height*. Post hoc tests show that standing ($mean = 3.70cm, SD = 1.77cm$) caused participants to produce gestures with significantly smaller heights than walking ($mean = 3.83cm, SD = 1.74cm$) and jogging ($mean = 3.87cm, SD = 1.64cm$))($p < .05$). Holding phone freely ($mean = 3.97cm, SD = 1.64cm$) led to gestures with significantly higher heights than holding the phone in a bag ($mean = 3.65cm, SD = 1.68cm$)($p < .05$).

- **Gesture width.** We found significant main effects of *moving speed* ($F_{2,22} = 18.615$, $p < .0001$) on *gesture width*. Larger widths than walking ($mean = 3.15cm, SD = 1.16cm$) and standing ($mean = 3.06cm, SD = 1.20cm$)($p < .05$).

- **Gesture area.** We found significant main effects of *moving speed* ($F_{2,22} = 7.415$, $p = .004$) and *phone location* ($F_{1,11} = 11.823$, $p = .006$) on *gesture area*. Post-hoc tests show that jogging ($mean = 13.32cm^2, SD = 8.86cm^2$) led to gestures with significantly larger area than both walking ($mean = 12.99cm^2$, $SD = 9.52cm^2$) and standing ($mean = 12.33cm^2, SD = 9.58cm^2$)($p < .05$). Holding phone freely ($mean = 13.76cm^2, SD = 8.86cm^2$) produced gestures with significantly larger area than when the smartphone was in a bag ($mean = 12.05cm^2, SD = 8.89cm^2$)($p < .05$).

- **Gesture duration.** We found no significant main effects on *gesture duration* nor interaction between moving and grasping ($p > 0.057$).

- **Gesture speed.** We found significant main effects of *moving speed* ($F_{2,22} = 15.966$, $p < .0001$) and *phone location* ($F_{1,11} = 13.596$, $p = 0.004$) on *gesture speed*. Post hoc tests show that the jogging ($mean = 12.40cm/s, SD = 3.99cm/s$) determined the participant to produce significantly faster gestures than walking ($mean = 10.84cm/s, SD = 3.52cm/s$) and standing ($mean = 10.81cm/s, SD = 3.55cm/s$)($p < .05$). Holding phone freely ($mean = 12.25cm/s, SD = 3.99cm/s$) led to significantly faster gestures than holding the phone in the bag ($mean = 10.45cm/s, SD = 3.63cm/s$)($p < .05$).

2.2.2.2 Mobile Directional Movement

We considered the same dependent variables than previous study to characterise the phone's movement: Alpha (z-axis), Beta (x-axis) and Gamma (y-axis)

Alpha deviation – deviation around the z-axis There were significant main effects of *moving speed* ($F_{2,22} = 107.146$, $p < .0001$) and *phone location* ($F_{1,11} = 7.038$, $p = 0.023$) on *alpha*. Post hoc tests show that during jogging ($mean = 14.34°, SD = 9.14°$) participants held a smartphone with a significantly higher deviation around the z-axis than both during walking ($mean = 9.19°, SD = 7.06°$) and standing ($mean = 4.91°, SD = 5.20°$), and walking had a higher deviation around the z-axis compared to standing. We also found that the deviation of the phone around the z-axis while holding it freely ($mean = 11.47°, SD = 9.14°$) is significantly larger than when it is held in a bag ($mean = 7.83°, SD = 7.39°$)($p < .05$).

Beta deviation – deviation around the x-axis. There was a significant main effect of *moving speed* ($F_{2,22} = 68.322$, $p < .0001$) on *beta*. Post hoc tests show that during jogging ($mean = 22.75°, SD = 11.59°$) participants held a smartphone with a significantly higher deviation around the x-axis than during walking ($mean = 16.32°, SD = 9.90°$) and standing ($mean = 6.43°, SD = 6.18°$), and walking had a larger deviation around the x-axis compared to standing ($p < .05$). We did not find any significant difference between the deviation of the phone around the x-axis while holding it freely ($mean = 16.09°, SD = 11.59°$) and when it is held in a bag ($mean = 15.06°, SD = 11.95°$).

Gamma deviation – deviation around the y-axis. There were significant main effects of *moving speed* ($F_{2,22} = 93.213$, $p < .0001$) on *gamma* with *moving speed* × *phone location* ($F_{2,22} = 8.538$, $p < .002$) interaction. Post hoc tests show that, when standing (respectively, walking), holding the smartphone freely ($mean = 5.60°, SD = 4.36°$) (respectively, $mean = 9.39°, SD = 6.65°$) implies a larger deviation around the y axis than when holding the phone in a bag ($mean = 2.65°, SD = 3.07°$) (respectively, $mean = 6.442°, SD = 4.728°$)($p < .05$). However, when the user runs, the gamma deviation is significantly larger when the smartphone is placed in a bag ($mean = 12.10°, SD = 7.52°$) compared to when it is held freely ($mean = 10.56°, SD = 3.07°$)($p < .05$).

Qualitative findings During the experiments, some participants reported that when they jogged they felt like they wanted to draw gestures faster. Some participants found that some gestures were more complex and needed more focus to draw, which can be challenging to draw in a real-time scenario, where they need to also remember the gesture shape. In particular, our participants found letter-shaped gestures with curves and corners more complex to draw than the remaining gesture shapes. No participant reported the task of imitating the gesture that was shown to them mentally difficult. Jogging at 8 km/h was physically difficult for some participants and they had to take breaks between.

2.2.3 Discussion

Our key finding is that an increase in the speed of movement results in longer, larger and faster stroke gestures. Conversely, when the phone is held in a bag, the gestures produced are slower, shorter, and smaller than when the phone is held freely. We also found no significant interaction ($p > .05$), suggesting that these results are consistent across different phone locations and moving speeds. Consequently, for walking and jogging, as well as for holding the phone in a bag, recognizers that rely on geometric and kinematic gesture descriptors, such as those described by [188] (p. 335), should be used with caution.

Interestingly, our results showed that gesture production time was not affected by moving speed or phone location. Specifically, moving faster or holding the phone freely led to faster gestures without requiring less time to draw the gesture, but instead resulted in longer gestures, which caused the gesture entry speed to increase. This result is consistent with findings from motor control theory, which demonstrates a dependency between writing speed and path length [237]. This finding suggests that long gestures can serve as convenient shortcuts for different user-moving speeds, as people compensate for the extra gesture length with increased gesture speed.

With regard to smartphone orientation, a similar trend was observed. Generally, higher speed of movement resulted in greater deviation in the alpha and beta angles. These findings can be explained by the increased movement of the phone as a result of the user's body movement. Consequently, designers should account for this additional movement of the phone when considering motion gestures in their designs for walking contexts [149].

However, an opposite effect was observed when the phone is held in a bag compared to when it is held freely, particularly for the alpha deviation and for gamma deviation when standing or walking. These findings can be explained by the reduced space available for manipulating the phone in a shoulder bag, which results in a more stable phone position.

Our findings also indicate that gesture shapes, such as letters with complex geometries (i.e., with a mixture of curves and lines), were more difficult to draw. Therefore, we recommend that designers avoid using such gesture shapes for moving contexts. If complex shapes are necessary, they should be designed to be easy to articulate to facilitate learning and memorization.

Finally, after showing how moving speed and phone location effect the production of eyes-free gestures we recommend designers to leverage build-in sensors of the smartphone to detect the context of use and adapt the recognizers accordingly. For example, accelerometer can be used to detect jogging and walking while ambient light sensor can be used to detect if the smartphone is used in a closed and small location. If the ambient light sensor is not detecting any significant light and yet the phone is detecting touch gestures it's a strong indication that the interaction is eyes-free.

2.2.4 Summary

In this study, we investigated the effects of moving speed and phone location on the articulation characteristics of gesture input. Our key findings reveal that faster moving speeds result in longer, larger, and faster stroke gestures, whereas holding the phone in a bag produces slower, shorter, and smaller gestures compared to holding the phone freely. In particular, gesture production time was not impacted by moving speed or phone location. Instead, faster gestures were produced without reducing the drawing time. We also observed that higher moving speeds caused more deviation in the phone's orientation angles, emphasizing the need for designers to consider additional phone movement when creating motion gestures for mobile contexts. In contrast, holding the phone in a bag resulted in less phone movement, particularly for certain orientation angles.

In the next study, we examine another factor that can affect the production of eyes-free gestures and the user experience: multitasking.

2.3 Investigating the effect of Attention-Saturating Tasks on Eyes-Free Gesture Execution on Mobile Devices

One common case for eyes-free interaction with smartphones is when visual attention is needed for another task that is more important and the primary task [103]. Especially for smartphones since they are portable and their usage has increased significantly, it is common to carry and use smartphones while being engaged in another task. Involving in two tasks can hinder the execution of the primary task. In this context, touch gestures can be executed with minimal visual attention, making them more susceptible to distractions [26]. In this study, we focus on the effect of an attention demanding primary task on eyes-free gestural interaction.

2.3.1 Experiment

We conducted an experiment to evaluate the effect of an attention-saturating task on eyes-free gesture articulation on a mobile device.

2.3.1.1 Participants and apparatus

13 right-handed participants (7 males and 6 females) volunteered to participate in our experiment. The ages of the participants ranged from 20 to 31 years ($mean = 24.3$, $SD = 3.5$). All participants had used smartphones for several years. We collected stroke gestures using our custom software on a Samsung Galaxy S7 smartphone running Android 6.0.1. The smartphone screen size is 5.65"×2.78" with a display resolution of 1440×2560 pixels and a pixel density of 227 pixels per cm.

2.3.1.2 Gesture set

In this study, we used the same set of gestures as in the previous study.

2.3.1.3 Design

The experiment used a 2×2 within-subject design with two factors: *activity* and *used hand*. *Activity* presents if the participant is doing the attention saturating task, *i.e.*, *centering the ball (CB)* or just waiting for the notification that we name here the *control condition (Ctrl)*. The *used hand* presents the hand used for the secondary task (*i.e.*, the used hand to hold the phone and draw the gesture without eyes) and covered two conditions: right hand (RH) and left hand (LH).

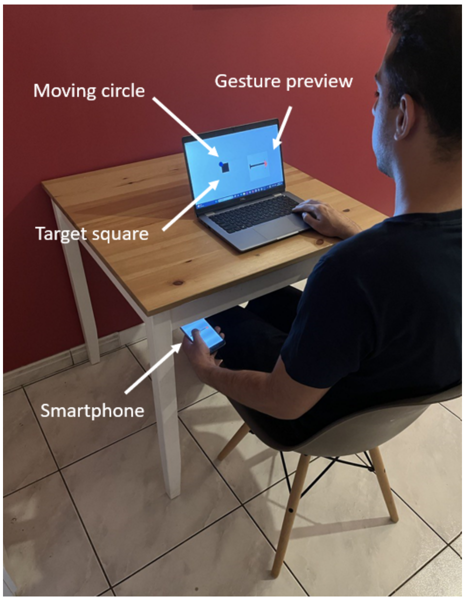


Figure 2.8: The setup used for the experiment. The smartphone was held under the table to maintain eyes-free condition.

2.3.1.4 Task & Procedure

After participants comfortably sat at a desk with laptops, they completed a demographic questionnaire. Next, we explained the experiment and its requirements, including primary and secondary tasks. The participants then engaged in the experiment: performing a primary task on laptop and drawing eye-free gestures on a smartphone (secondary task) upon laptop notifications. (see Figure 2.8). In the control condition, participants only had to wait for the notifications displayed on the laptop screen to perform the eyes-free gesture. The rationale for adding the control condition is to better understand the effect of the attention-saturating task on eyes-free gesture production.

For the primary task that saturates the attention, we followed [74] and used the same task that saturates the attention. Our attention-saturating task featured a circle that randomly moved according to a two-dimensional Perlin noise function. Participants were asked to keep the circle centered over a cross-hairs displayed in the center of the square as the best they could, and this by contracting its movement using the arrow keys of the laptop keyboard. The participants were told that keeping the circle centered on the crosshair was the most important task. However, we did not measure their performance on the primary task.

In the right-hand scenario, participants held the phone with their right hand, using their right thumb to draw gestures on the screen without looking. Similarly, in the left-hand scenario, they used their left hand. In both cases, the free hand was used to interact with the laptop during attention-saturating tasks. All gestures were single strokes; a new gesture appeared as soon as the finger left the screen. For corrections, participants could use the back button of the phone.

In the experiment phase, the two activity conditions were randomly presented to the participants. For each activity condition, the two hand conditions were also randomly presented. Each participant in total performed 400 gestures (=2 *activities* × 2 *hand* × 20 gestures × 5 repetitions). For each activity condition and hand condition, the gestures were presented to the participant in random order. The experiment took 20 minutes on average to complete.

2.3.2 Results

Our results include gesture features, smartphone directional movements, and questionnaire responses.

2.3.2.1 Gesture Features

We used the same geometric features as in two previous studies to analyze the change in the shape of gestures. Repeated measures ANOVA was used for non-parametric factorial analysis. Tukey tests were used post hoc when significant effects were found. Only significant results are reported below.

Gesture length. We found a significant main effect of *activity* ($F_{1,12} = 41.135$, $p < 0.001$) and *used hand* ($F_{1,12} = 5.323$, $p = 0.040$) on *gesture length*. *Centering the ball condition* ($mean = 6.82cm, SD = 3.22cm$) led participants to enter gestures with shorter lengths than *the control condition* ($mean = 7.44cm, SD = 3.51cm$). The drawing gesture with the left hand ($mean = 7.47cm, SD = 3.50cm$) led the participants to enter gestures with lengths longer than the drawing gesture with the right hand ($mean = 6.79cm, SD = 3.23cm$).

Gesture height. We found a significant main effect of *activity* ($F_{1,12} = 34.327$, $p < 0.001$) on *gesture height*. *Centering the ball condition* ($mean = 2.40cm, SD = 1.02cm$) led participants to enter gestures with a shorter height than *the control condition* ($mean = 2.64cm, SD = 1.19cm$).

Gesture width. We found a significant main effect of *activity* ($F_{1,12} = 17.998$, $p = 0.001$) on *gesture width*. *Centering the ball condition* ($mean = 2.33cm, SD = 0.88cm$) led participants to enter gestures with width smaller than *the control condition* ($mean = 2.49cm, SD = 0.97cm$).

Gesture area. We found a significant main effect of *activity* ($F_{1,12} = 38.801$, $p < 0.001$) on *gesture area*. *Centering the ball condition* ($mean = 7.07cm^2, SD = 4.72cm^2$) led participants to enter gestures with smaller area than *the control condition* ($mean = 7.17cm, SD = 5.87cm$).

Gesture duration. We found a significant main effect of *activity* ($F_{1,12} = 7.422$, $p = 0.019$) on *gesture duration*. *Centering the ball condition* ($mean = 0.96seconds, SD = 0.65seconds$) led participants to enter gestures with duration shorter than *the control condition* ($mean = 1.05seconds, SD = 0.67seconds$).

	RH+Ctrl		***LH+Ctrl***		***RH+CB***		***LH+CB***		**Friedman**
	mean	sd	mean	sd	mean	sd	mean	sd	$\chi^2(3)$
Performance	6.15	1.81	6.23	1.59	6.61	2.02	6.53	1.76	4.18
Temporal demand	2.23	1.42	2.31	1.60	2.54	1.45	2.00	1.68	6.95
Physical demand	3.08	1.66	3.06	1.71	3.15	2.27	4.23	2.71	16.70
Mental demand	3.84	1.95	3.74	2.26	3.38	1.76	3.69	1.97	0.57
Frustration	2.69	1.55	3.54	1.66	3.23	2.17	5.46	2.54	21.70
Effort	4.30	2.25	4.38	2.06	3.92	2.78	5.46	2.10	16.01

Note: Friedman tests are reported at significance levels p=.05 (*) significance levels. The significant tests are highlighted.

Table 2.1: Mean and SD of the Nasa TLX questionnaire responses, rated on a scale of 1 (very low) to 10 (very high).

Gesture speed. We did not find any significant main effect of *activity* or *used hand* on gesture speed.

2.3.2.2 Mobile Directional Movement

We considered the same dependent variables than previous two studies to characterise the phone's movement: Alpha (z-axis), Beta (x-axis) and Gamma (y-axis) Repeated measures ANOVA was used for nonparametric factorial analysis. Tukey tests were used post hoc when significant effects were found. Only significant results are reported below.

Alpha deviation – deviation around the z-axis. We found a significant main effect of *used hand* ($F_{1,12} = 7.675$, $p = 0.017$) on *alpha deviation*. Using the right hand to manipulate the phone caused significantly more deviation in alpha ($mean = 5.14°, SD = 4.71°$) compared to using the left hand ($mean = 3.40°, SD = 4.48°$).

Beta deviation – deviation around the x-axis, and Gamma deviation – deviation around the y-axis. No significant significant main effect of *used hand* or *activity* was found on *beta deviation* ($mean = 5.83°, SD = 4.84°$) or *gamma deviation* ($mean = 2.83°, SD = 2.84°$).

2.3.2.3 Subjective results

Our participants were also asked to rate the task after each condition in terms of performance, temporal demand, physical demand, mental demand, frustration, effort. Table 2.1 shows the responses to the questionnaire. Friedman tests, followed by Conover post hoc analysis revealed that only physical demand, frustration, and effort were significantly higher than three other conditions when the participant was performing two tasks simultaneously and holding the phone in the left hand ($p < 0.05$).

2.3.3 Discussion

Our primary discovery is that the attention-saturating task influenced the geometric features of the eyes-free gestures, but did not affect their kinematic properties. For example, while centering the ball, the gestures exhibited shorter lengths and smaller height, width, and area compared to the control condition. This task also resulted in gestures being produced in a shorter time without compromising speed. It is crucial to note that the speed remained unchanged because both the duration and the length of the gestures were reduced. We did not observe a significant interaction ($p > .05$), indicating that these results are consistent between different hand conditions.

Therefore, when performing an attention-saturating task while entering eyes-free gestures on mobile devices, recognizers that depend on geometric gesture descriptors, such as those described by [188] (p. 335), should be used with caution.

Additionally, our findings reveal that the hand used influenced the length of the gesture, with the left hand producing longer gesture paths than the right hand but resulting in fewer directional movements around the z-axis.

Our results suggest that during attention-saturating tasks, using the left hand for eyes-free gestures is more frustrating and physically demanding than the other three conditions. Hence, right-handed users should prioritize using their right hand for eye-free gestures on mobile devices when simultaneously engaging in attention-saturating activities.

Based on our results the users tend the shorten the interaction with smartphone when their attention is needed somewhere else. So, we recommend designers to think about if they believe that their application will be interesting to users to use while mentally focusing elsewhere to make the interaction time with application shorter. This can be achieved by requiring shorter time for the interaction. If short interaction with the application is not possible, then the interaction should be designed in a way that it can be divided by user to shorter interactions. For instance, designers should avoid having a long mark-based gesture in their applications, which requires several strokes. Instead, this interaction should be able to be conducted by lifting the finger after each stroke, addressing the attention saturating task, and coming back later to smartphone to conduct the following strokes.

2.3.4 Summary

The study explored how attention-saturating tasks affect eyes-free touch gestures on smartphones. Participants performed gestures with their dominant and non-dominant hands while engaged in another task. The results showed that such tasks altered the geometric features of gestures, making them shorter and faster without affecting speed. No significant differences were found between hand conditions, but left-hand use was more frustrating and physically demanding. The study suggests that for right-handed users, it is preferable to prioritize their dominant right hand over their non-dominant left hand for the secondary task of producing eyes-free gestures, aiming to reduce frustration.

2.4 Conclusion

This chapter has explored the production of eyes-free gestures in various contexts through three comprehensive studies, each providing valuable insights into the factors influencing gesture articulation and recognition on smartphones. These three studies showed that each type of situationally induced impairment can cause a different effect on gesture production. For example, eyes-free gestures compared to gestures produced when the phone is visible are produced in larger size and longer duration without change in speed. On the other hand, an increase in the movement speed of the user increases the size of eyes-free gestures while increasing the speed of gesture production, without changing the duration of gesture production. Meanwhile, the presence of an attention-saturating task decreases the size of eyes-free gestures but does not change the speed of gesture production. Instead, it decreases the gesture production duration. It seems that users use different strategies in different contexts for interaction with smartphones. In the presence of an attention-saturating task where the cognitive focus of the user is needed for the primary task, they try to finish producing gestures in a shorter time so that they can focus their cognitive focus again on the primary task. However, when the user is involved in a physical demanding task without significant cognitive load such as jogging, they do not try to shorten the gesture production time, instead just their finger moves faster in a smaller area. Based on the feedback we received from participants, when the legs are moving with a high speed, they tend to move their fingers faster. Moreover, it seems as the manipulation of the touchscreen becomes more difficult due to high movement speed, the users tend to produce gestures in smaller area. This tendency is similar to one-hand manipulation of the smartphone, which results in smaller gestures [16]. Users seem to be more confident of not dropping the phone when they draw smaller gestures compared to larger gestures. Being aware of the environmental factors of the user can help us adjust the recognizer so that it can have higher accuracy. The built-in sensors of the smartphone can provide some information about the context. For example, if the user is running the acceleremeter data can reveal it or the ambient light sensor of the smartphone can show if the smartphone is in a closed space such as a bag or pocket.

Overall, these studies underscore the complexity of eyes-free gestural interactions and the significant influence of various contextual factors. The findings provide a foundation for developing more intuitive and reliable eyes-free interfaces, emphasizing the need for adaptable gesture recognition systems and thoughtful design considerations to accommodate different user conditions and environments. Although the gesture recognition rates for the visually impaired scenarios were less compared to when visual feedback was present, it is still acceptable, and eyes-free gestures can be used as convenient shortcuts. Now that we have a better understanding of how eyes-free gestures are produced in different situational impairment scenarios, in the next two chapters we can focus on introducing eyes-free specialized interaction concepts with smartphones leveraging the haptic channel. In Chapter 3, we will focus on improving the input channel of interaction using haptic feedback and in Chapter 4 we will use tactile feedback for enhancing the output information while interacting with touchscreens.

Chapter 3

Enhancing Gestural Inputs on Touchscreens Using Haptic Feedback

After understanding how eyes-free gestures are produced in different scenarios, it is now time to implement interaction techniques which use these eyes-free gestures. Since in visual impairment scenarios, visual feedback, which is the main channel of interaction on smartphones, is lacking, other sensory modalities can be used for interaction. In this chapter, we focus on the use of the haptic channel in eyes-free gestural interactions to improve command invocation on touchscreens, especially smartphones. Haptic feedback has shown considerable potential to improve eyes-free interactions. By providing non-visual cues, haptic feedback can enhance the user's ability to interact with their device in situations where visual attention is limited or obstructed [159]. This chapter investigates, through three studies, the utilization of haptic feedback to enhance gesture-based input interactions with touch devices in eyes-free configurations.

More specifically, the aim of the first study is to determine whether haptic feedback can improve the production of eyes-free gestures (input), their recognition accuracy, and user confidence for the interaction on a smartphone placed inside a pant pocket without taking it out. We call this interaction scenario *pocket interaction*. This scenario presents unique challenges, as users must interact indirectly with the smartphone's touchscreen through the pocket fabric. To address these challenges, we experimented with various vibration patterns to determine which ones offer the most assurance to the user and accuracy in recognition for indirect touch gestures.

The second study builds on the results of the first study. After finding the most preferred vibration pattern for pocket interaction. In this study,we investigate the preferred location for haptic feedback when producing eyes-free gestures through the pocket fabric. Specifically, we investigate whether users prefer local vibration on the smartphone, in the pocket (as in the previous study), or distal vibration on a smartwatch.

Unlike the last two studies which were using haptic feedback to come over a physical barrier for input on a touchscreen, the last study focuses on a scenario which includes mental challenge for input on a touchscreen. In this study, we used haptic feedback to improve targeting task on a touch device while being engaged in an attention saturating task. We investigate which type of haptic texture on the virtual knob helps users to perform the targeting task better while maintaining their focus on the primary task.

Through these studies, this chapter aims to expand our understanding of how haptic feedback can be harnessed to enhance eye-free interactions with smartphones, by exploring having feedback on input gestures, to increase confidence in gesture input to users and improving the performance of users while multitasking.

This chapter is organized into three sections, each dedicated to a distinct study of using haptic feedback for improving input in eyes-free gesture-based interaction scenarios with smartphone. Each section starts with an introduction to the study's objective, followed by a comprehensive explanation of the experimental design. Afterward, the results are presented and thoroughly discussed. Lastly, each section ends with a summary that highlights the main findings and insights obtained from the study.

3.1 Enhancing Production of Gestures for In-Pocket Interaction with Mobile Devices Using Haptic Feedback

Being able to interact with smartphone in the pocket of pants without taking it out enables comfortable, private, and readily accessible means of interaction. Researchers have suggested methods to interact with a smartphone without taking it out of the pocket, such as using a secondary device (*i.e.*, smartwatches) or fabric-based touch sensors. However, these methods are based on additional hardware, which is not widely available, such as textile sensors. Here, instead, we focus on using built-in sensors available in smartphones and developing an interaction technique which can be available to the general public. More specifically, this study examines the effect and outcomes of enhancing eyes-free gestures in pocket interactions with haptic feedback.

3.1.1 Experiment

We conducted an experiment to determine the effect of haptic feedback on the production of eye-free gestures on mobile devices in the context of in-pocket interaction.

3.1.1.1 Participants and apparatus

12 right-handed participants (7 males and 5 females) volunteered to participate in our experiment. The ages of the participants ranged from 18 to 38 years ($mean = 29.4, SD = 5.2$). All participants had used smartphones for several years.

We collect stroke gestures using our custom software on a Samsung Galaxy S7 smartphone running Android 6.0.1. The smartphone has a screen size of 5.65" × 2.78" with a display resolution of 1440 × 2560 pixels and a pixel density of 227 pixels per cm. The experiment application was written in Java and run on the

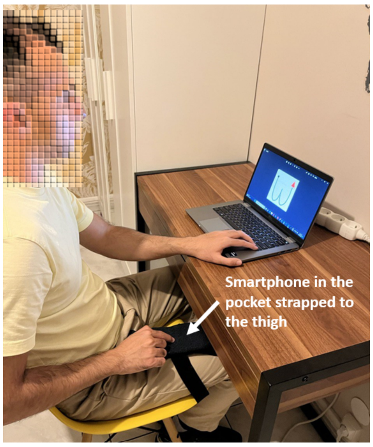


Figure 3.1: The setup used for the experiment. The smartphone was placed in the right front pocket of the pants.

smartphone. Although the smartphone screen was not visible to participants since it was placed in their pocket, we mirrored the smartphone screen to a laptop positioned in front of the participants. This allowed participants to see which gesture they were required to draw on the laptop screen (see Figure 3.1).

3.1.1.2 Gesture Set

We used the same set of gestures as we used in the last two studies in the previous chapter (see Figure 2.6).

3.1.1.3 Design

The experiment utilized a within-subject design with one factor: *feedback*. Feedback encompassed three conditions: *variable vibration* (VV), *continuous vibration* (CV), and *no vibration* (NV).

In the *variable vibration* condition, a 100-ms vibration burst was provided to the user at the beginning of the gesture, followed by a second burst when the user lifted their finger at the end of the gesture. This design offers the user a cue that signals the initiation and completion of the gesture, indicating when the touchscreen no longer detects the touch.

In the *continuous vibration*, the phone continues to vibrate as long as the touch event is detected. In this technique, the phone maintains a consistent vibration intensity throughout the gesture, and once the touch event is no longer detected, the vibration ceases. The rationale behind this haptic feedback is to provide the user with a continuous cue that their finger is detected by a touch sensor, ensuring that the gesture performed is recorded.

We also included a *no vibration* condition, which aligns with most touchscreens available today. In this condition, no haptic feedback is provided when drawing gestures, serving as our baseline.

3.1.1.4 Task & Procedure

During the experiment, the participants were seated in a chair in front of a table, where a laptop was placed. The smartphone was then placed inside the front pocket of their pants on the same side as their dominant hand. The phone was placed in a way that the touchscreen faced outward, while the back of the phone rested on their laps. This setup allowed participants to interact with the touchscreen by pressing on the fabric of their pants, without needing to insert their hand into the pocket.

A preview of the gesture they were instructed to draw was shown on the laptop which was positioned in front of them on the table. Participants were asked to use the index finger of their dominant hand to draw gestures on the fabric of their pants, directly above the touchscreen of the smartphone.

Since all gestures in this experiment consisted of single strokes, the next gesture appeared as soon as the touch event was no longer detected. In the event of a false entry, participants were not allowed to return and modify their gestures. This rule was implemented to facilitate the comparison of participant performance under different conditions.

During the experiment phase, the participants were presented with three vibration conditions in a random order. For each condition, participants were presented 20 gestures and each gesture was repeated five times. The gestures were presented to the participant in random order.

After finishing each of the three feedback conditions, participants completed a NASA-TLX worksheet, plus a 10-point Likert scale questionnaire (strongly disagree to strongly agree) to measure enjoyment while interacting with the mobile device. At the end of the experiment, we asked participants to rank the feedback conditions. The experiment took 20 minutes on average to complete.

3.1.2 Results

Our results included gesture characteristics, gesture recognition rates, smartphone directional movements, and qualitative observations.

3.1.2.1 Gesture Features

We utilized the same six geometric features as in the previous studies from the prior chapter gestures: (a) gesture length, (b) gesture height, (c) gesture width, (d) gesture area, (e) gesture duration, and (f) gesture speed.

All analyzes used Friedman test. Tukey post-hoc tests were used when significant effects were found. Only significant effects and interactions are reported.

Gesture length. Gesture length refers to the cumulative path distance from the first registered touch event to the last. The Friedman test revealed significant differences in gesture duration between the conditions ($\chi^2(2) = 37.44, p < 0.001$). Post hoc analysis showed that the mean length of gestures for CV ($mean = 10.05cm, SD = 3.26cm$) was significantly higher than for both NV ($mean = 9.43cm, SD = 3.81cm$) and VV ($mean = 9.82cm, SD = 3.59cm$).

Gesture height. Gesture height is the height of the bounding box that contains the gesture ($max_{\mathbf{y}}$ - $min_{\mathbf{y}}$). Friedman test revealed significant differences in gesture height between conditions ($\chi^2(2) = 19.02, p < 0.001$). Post hoc analysis showed that the mean length of gestures for CV ($mean = 3.93cm, SD =$

$1.10cm$) was significantly higher than for both NV ($mean = 3.71cm, SD = 1.41cm$) and VV ($mean = 3.67cm, SD = 1.15cm$).

Gesture width. Gesture width is the length of the bounding box that contains the gesture ($max_{\mathbf{x}}$ - $min_{\mathbf{x}}$). Friedman test revealed significant differences in gesture width between conditions ($\chi^2(2) = 50.15, p < 0.001$). Post hoc analysis showed that the mean widths of the gesture of VV ($mean = 2.90cm, SD = 0.66cm$) and NV ($mean = 3.01cm, SD = 0.72cm$) were significantly lower than CV ($mean = 3.25cm, SD = 0.67cm$).

Gesture area. Gesture area is the surface area of the bounding box containing the gesture (height × width). Friedman test revealed significant differences in gesture area between conditions ($\chi^2(2) = 40.21, p < 0.001$). Post hoc analysis showed that the mean gesture area of VV ($mean = 10.94cm^2, SD = 3.41cm^2$) and NV ($mean = 11.16cm^2, SD = 3.37cm^2$) were significantly smaller than CV ($mean = 13.10cm^2, SD = 4.35cm^2$).

Gesture duration. Gesture duration is defined as the time elapsed when entering the gesture, that is, the timestamp of the last touch event registered minus the timestamp of the first touch event. The Friedman test revealed significant differences in gesture duration among the conditions ($\chi^2(2) = 35.66, p < 0.001$). Post hoc analysis showed that the mean duration of gestures of NV ($mean = 0.88s, SD = 1.05s$) was significantly shorter than both CV ($mean = 1.13s, SD = 0.91s$), and VV ($mean = 1.08s, SD = 0.89s$).

Gesture speed. Gesture speed is calculated as the average speed based on all touch events within a gesture (length/duration). The Friedman test revealed significant differences in gesture speed among all conditions ($\chi^2(2) = 67.70, p < 0.001$). Post hoc analysis showed that the mean gesture speed was different between all pairs of conditions. NV ($mean = 13.72cm/s, SD = 4.69cm/s$) had the highest gesture speed, then CV condition ($mean = 11.77cm/s, SD = 3.88cm/s$), and VV had the lowest gesture speed ($mean = 10.77cm/s, SD = 3.51cm/s$).

3.1.2.2 Gesture Recognition

Unlike the studies we conducted in the previous chapter, where there was no physical barrier that can negatively impact touch detection on the touch screen, this interaction technique has a risk of having incomplete gestures which may not be recognizable. The results we derived for the changes in gesture feature is only valuable if the gestures are produced with a good accuracy and are recognizable. So, here it is essential to evaluate the accuracy of gesture recognition in this interaction technique.

As the $1 recognizer [244] was found to be inadequate for 1-D gestures in our mark set, due to its insensitivity to rotation, we opted for the $N-protractor [9] recognizer for our set of gestures. We conducted user-dependent training, which is appropriate for gesture recognition on personal devices such as mobile phones [6, 9, 229]. In this training, the recognition rates were individually calculated for each participant using the same methodology as in previous studies [244, 6, 9, 229]. For each user, we first trained the gesture recognizer on a small set (four candidates for each gesture type were selected) and then tested the recognizer on the remaining samples from that user. This process was repeated 100 times. It is important to note that separate tests were conducted for each set of gestures generated under different feedback conditions.

The Friedman test was conducted to examine differences in the average recognition rate under three conditions. We found no significant differences between the conditions, ($\chi^2(2)(df) = 2.66, p = 0.26$). The NV condition had the lowest average recognition rate $mean = 74.37\%(SD = 19.39\%)$. The recognition rates of CV, and VV were $82.37\%(SD = 13.65\%)$ and $83.67\%(SD = 13.99\%)$, respectively.

3.1.2.3 Mobile Directional Movement

We considered the same dependent variables as previous studies to characterize the phone movement: Alpha (z axis), Beta (x axis), and Gamma (y axis).

Alpha deviation – deviation around z axis The Friedman test revealed significant differences in alpha deviation between conditions ($\chi^2(2) = 111.73, p < 0.001$). Post hoc analysis showed the mean alpha was different between all pairs of conditions. The CV ($mean = 4.36°, SD = 4.11°$) has the highest alpha deviation, followed by the NV condition ($mean = 2.92°, SD = 2.44°$), and the VV had the lowest alpha deviation ($mean = 2.84°, SD = 2.83°$).

Beta deviation – deviation around x axis We found no significant main effect of the vibration condition on beta deviation ($\chi^2(2) = 4.54, p = 0.10$).

Gamma deviation – deviation around y axis The Friedman test revealed significant differences in the alpha deviation between conditions ($\chi^2(2) = 91.02, p < 0.001$). Post hoc analysis showed a significant difference in the mean gamma deviation among all conditions. ($p < 0.02$). CV had significantly higher gamma deviation ($mean = 1.17°, SD = 1.20°$) than VV ($mean = 0.95°, SD = 0.91°$), and NV ($mean = 0.78°, SD = 0.95°$).

3.1.2.4 Subjective Results and Observations

We recall that our participants were asked to rank the three feedback conditions after completing the experiment. In general, CV was ranked first 8 times (66%), second 4 times (33%), and was never ranked third. VV was ranked first 4 times (33%), second 7 times (58%), and third once (8%). NV was never first, second once (8%), and third 11 times (92%). Friedman test revealed significant differences in ranking ($\chi^2(2) = 12.58, p = 0.002$). The pairwise comparison revealed that the mean ranking for all three conditions were significantly different from those of others ($p < 0.02$). NV was the least preferred configuration ($mean = 2.91$), followed by VV ($mean = 1.75$). CV was the most preferred option ($mean = 1.33$).

Our participants were also asked to rate the task after each feedback condition. Overall, questionnaire responses (Table 3.1) show that mean ratings for CV were highest in all seven questions, but only significant for performance ($p = 0.05$) and enjoyment ($p = 0.016$).

The most common comment given by the participants was that in the NV condition they had no clue when the touch was detected by the phone or not until the end of the gesture. They would only know if the next gesture appeared on the computer, indicating that some gesture was detected by the phone. However, they were still unsure at which points the phone started to detect that their finger was touching it. As a result, they were uncertain if the entire gesture was recorded. This lack of feedback in the NV condition made the participants less confident in their performance and gave it a lower score. Figure 3.2 shows how

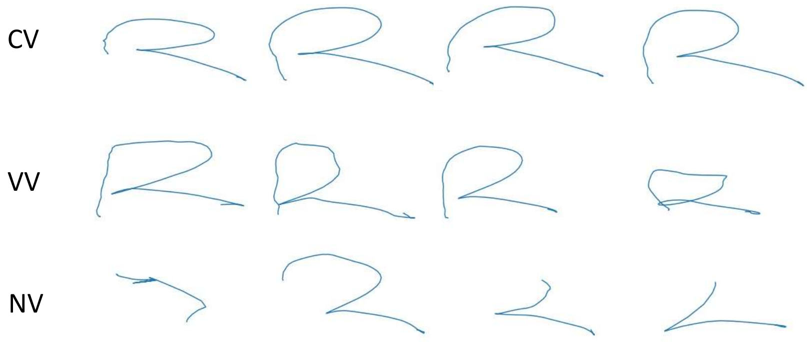


Figure 3.2: Gesture G5 drawn by one participant under different conditions. In the *NV* condition, the gestures are incomplete due to the lack of haptic feedback.

	NV		***VV***		***CV***		**Friedman**
	mean	SD	mean	SD	mean	SD	$\chi^2(2)$
Performance	4.58	2.27	7.50	2.11	8.25	1.29	19.82
Temporal demand	2.17	1.34	2.25	1.42	1.83	1.40	2.07
Physical demand	2.58	1.24	2.42	1.67	2.08	1.38	3.83
Mental demand	4.00	2.63	3.17	2.08	3.00	1.48	0.83
Frustration	3.42	1.93	3.08	1.96	2.75	1.29	0.62
Effort	2.92	1.88	3.08	2.15	2.92	1.88	0.56
Enjoyment	5.00	2.30	6.08	1.68	6.66	1.87	6.54
Ranking	2.91	0.29	1.75	0.62	1.33	0.49	16.17

Note: Friedman tests are reported at significance levels of p=.05 (*). Significant tests are highlighted.

Table 3.1: Mean and SD the of NASA TLX questionnaire and enjoyment responses, rated on a scale of 1 (very low) to 10 (very high).

one of the participants drew gesture g5 in different conditions. It is obvious from the figure that, under the *NV* condition, most gestures were recorded incompletely due to the absence of feedback. Furthermore, the uncertainty in drawing gestures and having to repeat them when they were not logged made the *NV* condition less enjoyable.

For the *CV* and *VV* conditions, participants provided varied feedback. Participants experienced different levels of vibration intensity. Some reported the vibration in the *CV* condition as strong, while others perceived it as weak. The perceived vibration intensity differed among participants, likely due to variations in the contact mechanics between the phone and participants' bodies, such as the thickness of their pants, the tightness of their pants, and the tightness of the pocket. Participants who did not rank *CV* as their first choice often attributed it to the strong intensity of vibrations. We believe that adjusting the intensity of *CV* for each user could make it the most preferred condition.

3.1.3 Discussion

Despite NV implying less phone movement than CV and VV, NV led to significantly faster gesture production, requiring less time to draw gestures and resulting in shorter gesture lengths (except for VV). However, although no significant differences were found in terms of recognition, the average NV recognition rate is smaller than that of CV and VV, achieving a precision rate of 74%, compared to 82% and 84%. This can be explained by the fact that the user's gestures were not completely logged due to participants not noticing that their finger was not detected correctly by the mobile device placed in their pocket. Instead of a fully recorded gesture, it was recorded as smaller incomplete shapes. Although the 82% and 84% mean recognition rate of CV and VV are not ideal, they show pocket interaction as a feasible interaction technique.

When comparing CV with VV, our findings indicate that with the CV condition, the gestures were significantly larger, longer and faster, and produced with higher speed. However, the gesture production time was not affected. In particular, using CV led to faster gestures without requiring less time to draw the gesture, but instead longer gestures, which caused the gesture entering speed to increase. This result was expected due to findings from motor control theory that proved a dependency between writing speed and path length [237]. This finding advocates that long gestures serve as convenient shortcuts for different user's moving speeds, as people compensate for the extra gesture length with increased gesture speed.

Importantly, the general trend we observed when entering gestures is that users are more confident when generating gestures in the presence of haptic feedback. Our participants lacked confidence in NV and ranked it as their least preferred feedback condition. This behavior could be due to the increased fluency in entering gestures when users can feel that the device detects their finger correctly. Especially in the CV condition, the gestures were "nicely" produced, which can be explained by the fact that continuous perceived haptic feedback provides the user with a continuous clue that their gesture is correctly logged, thereby increasing confidence. In contrast, the generation of incomplete gestures was expected in the absence of haptic feedback. NV was also rated as the least accurate and least enjoyable feedback condition. Consequently, we strongly recommend including haptic feedback during eyes-free gesture production on mobile devices for in-pocket interaction contexts.

However, it is important to note that for both CV and VV, gesture recognition rates are still low, only 80. Additionally, both CV and VV involve more phone movement around the y and z axes compared to NV, with CV causing the highest deviations.

3.1.4 Summary

In this study, we explored the impact of haptic feedback on in-pocket smartphone interactions under three conditions: No Vibration (NV), Continuous Vibration (CV), and Variable Vibration (VV). Our findings show that without haptic feedback, gestures were faster but less accurate (74% recognition rate) and users felt less confident. Interestingly, our findings indicate a preference for CV that leads to longer, larger and faster gestures without affecting production time, which we recommend using. However, the gesture recognition rate for CV remains low, around 82 percent (similar to VV). Moreover, continuous vibration caused the highest phone deviations.

The importance of haptic feedback in improving user experience and gesture accuracy for in-pocket smartphone use is evident. Understanding how different feedback types influence input gestures aids in designing more intuitive mobile interfaces. While localized feedback caused no major issues, in the next study, we investigate if distal feedback (*e.g.*, via a smartwatch) could further enhance gesture recognition and reduce phone movement in in-pocket contexts.

3.2 Investigating the Effect of Vibration Location on Input for In-Pocket Interaction

The findings of the previous study suggest that adding local haptic feedback during eyes-free in-pocket gesture input increases confidence. However, it also results in an increase in the movement of the phone, and the recognition rate reaches only 82%, which is better than when haptic feedback is not provided. Although our participants did not express inconvenience with regard to phone movement when the haptic vibration was localized on the phone, we wondered if providing the vibration distally (in a location other than the mobile device where the gesture is produced) could support the production of eyes-free gestures on a pocket mobile device. Henderson et al. [83] demonstrate parity in performance between localized feedback and distal feedback for target acquisition tasks. However, what about gesture production tasks for in-pocket interaction? Could distal haptic feedback improve gesture recognition while minimizing phone movements? In this experiment, we explore the effectiveness of smartwatch-based feedback as an aid to perform gaze-free input gestures on a mobile device in an in-pocket context.

3.2.1 Experiment

We conducted a second study to compare the effectiveness of local vibrotactile feedback (on the smartphone) with distal vibrotactile feedback (a smartwatch) regarding eyes-free gesture production on mobile devices in an in-pocket interaction context. We utilized the same gesture set as in the previous study.

3.2.1.1 Participants and apparatus

24 right-handed participants (8 males and 16 females) volunteered to participate in our experiment. The ages of the participants ranged from 20 to 35 years ($mean = 26.3$, $SD = 4.3$). All participants had used smartphones for several years.

For gesture collection, we used the same custom software and smartphone device as in the previous study. In addition, we introduced a Samsung Galaxy Watch 4 to provide haptic feedback. The smartwatch operated on Wear OS 4, and a custom Wear application was developed to enable vibration patterns on the smartwatch. The application received TCP messages from the smartphone over a local WiFi network to trigger the desired haptic feedback.

3.2.1.2 Gesture Set

We used the same set of gestures as in previous studies (see Figure 2.6).

3.2.1.3 Design

The experiment used a within-subject design with one factor: *vibration location*. *Vibration location* determined the perceived location of the haptic feedback and included two conditions: *local* vibrotactile feedback (*i.e.* on the smartphone) and *distal* vibrotactile feedback (*i.e.*, on a smartwatch).

3.2.1.4 Task & Procedure

In this experiment, 24 participants made eyes-free gestures in their pockets under four conditions. The *local CV* condition was exactly the same as in previous study. However, in the *distance CV* condition, continuous vibration was applied on a smartwatch that was on the left hand side of the participants, as long as the smartphone detected the touch. The participants followed the exact same task and procedure as in the previous experiment. At the end of each condition, participants completed the NASA-TLX questionnaire plus enjoyment. At the end of the experiment, participants ranked the four tested conditions.

3.2.2 Results

Similarly to the previous study, our results include gesture features, gesture recognition, smartphone directional movements, and questionnaire responses.

3.2.2.1 Gesture Features

Gesture length. The Wilcoxon signed-rank test revealed a statistically significant difference between *local CV*'s and *distal Cv*'s gesture length (Z = -7.82, p<0.001), with the *local CV*'s gesture length ($mean = 11.44cm, SD = 4.91cm$) was significantly longer than the length of the *distal CV* gesture ($mean = 10.25cm, SD = 4.1cm$).

Gesture height. The Wilcoxon signed-rank test revealed a statistically significant difference between *local CV*'s and *distal Cv*'s gesture height (Z = -7.26, p<0.001), with the *local CV*'s gesture height ($mean = 4.39cm, SD = 1.92cm$) was significantly higher than *distal CV*'s gesture height ($mean = 3.90cm, SD = 1.77cm$).

Gesture width. The Wilcoxon signed-rank test revealed a statistically significant difference between *local CV*'s and *distal Cv*'s gesture width (Z = -6.71, p<0.001), with the *local CV*'s gesture width ($mean = 3.51cm, SD = 1.19cm$) was significantly larger than *distal CV*'s gesture width ($mean = 3.21cm, SD = 1.02cm$).

Gesture area. The Wilcoxon signed-rank test revealed a statistically significant difference between *local CV*'s and *distal CV*'s gesture area (Z = -7.52, p<0.001), with the *local CV*'s gesture area ($mean = 15.77cm^2, SD = 8.77cm^2$) was significantly bigger than *distal CV*'s gesture area ($mean = 13.09cm^2, SD = 7.83cm^2$).

Gesture duration. The Wilcoxon signed-rank test revealed a statistically significant difference between *local CV*'s and *distal CV*'s gesture production time (Z = -11.63, p<0.001), with the *local CV*'s gesture duration ($mean = 1.19s, SD = 0.57cm^2$) was longer than *distal CV*'s gesture duration ($mean = 1.17s, SD = 0.59cm^2$).

Gesture speed. The Wilcoxon signed-rank test revealed a statistically significant difference between *local CV*'s and *distal CV*'s gesture speed (Z = -5.06, p<0.001), with the *local CV* gesture speed ($mean = 13.05cm/s, SD = 5.01cm/s$), was higher than *distal CV* gesture speed ($mean = 12.20cm/s, SD = 5.65cm/s$).

3.2.2.2 Gesture recognition

The Wilcoxon signed-rank test was conducted to examine the differences in the average recognition rate under two conditions. The Wilcoxon signed-rank test ($Z = -0.22, p = 0.85$) did not reveal any significant differences in the recognition rate between the conditions (local CV: mean=86.67%, SD=9.17% ; distal CV: mean=86.02%, SD=9.82%).

3.2.2.3 Mobile Directional Movement

Alpha deviation – deviation around z axis. The Wilcoxon signed-rank test revealed a statistically significant difference between *local CV*'s and *distal Cv*'s alpha deviation (Z = -3.00, p<0.001), with the *Local CV* had higher alpha deviation (mean = 4.17°, SD = 2.78°), than *Distal CV* alpha deviation (mean = 3.05°, SD = 2.64°).

Beta deviation – deviation around x axis. The Wilcoxon signed-rank test did not find any significant difference between *local CV*'s and *distal Cv*'s beta deviation (Z=-0.39, p=0.69). The mean beta deviation was 1.97° with $SD = 3.35°$.

Gamma deviation – deviation around y axis. The Wilcoxon signed-rank test revealed a statistically significant difference between *local CV*'s and *distal Cv*'s alpha deviation (Z=-6.32, p<0.001), with the *local CV* (mean = 1.25°, SD = 0.94°) had a higher gamma deviation compared to *distal CV* (mean = 0.88°, SD = 1.53°).

3.2.2.4 Qualitative findings

Participants were asked to rank the two feedback conditions after completing the experiment. In general, *local CV* was ranked first 18 times, second 6 times against *distal CV*. The responses to the questionnaire (Table 3.2) show that the mean ratings of enjoyment were significantly higher for the *local CV*.

In some cases, participants reported that there is a delay in the feedback from the smartwatch. For example, sometimes the vibration continues after the user has left the finger on the smartphone. This is due to the wireless nature of communication between the smartphone and the smartwatch. In general, if one uses wireless communication to interact with a remote haptic actuator, the communication will not be as perfect as that of local actuators. Moreover, some participants find the distal vibration to be "*unnatural*" as they are touching one point but receiving feedback elsewhere. On the other hand, the most common

	Local CV		*Distal CV*		Wilcoxon
	mean	SD	mean	SD	W
Performance	7.92	1.83	8.31	1.91	1.46
Temporal demand	1.58	1.23	2.11	1.51	1.33
Physical demand	2.73	1.46	2.44	1.31	-1.4
Mental demand	3.73	1.82.49	4.3	1.95	-1.12
Frustration	3.41	1.96	3.24	2.05	1.2
Effort	3.43	2.44	3.65	2.78	0.78
Enjoyment	6.89	1.53	6.01	1.21	-3.78

Note:Wilcoxon signed-rank tests are reported at p=.05 (*) significance levels. Significant tests are highlighted.

Table 3.2: Mean and SD of NASA TLX questionnaire, enjoyment and ranking responses, rated on a scale of 1 (very low) to 10 (very high).

comment from participants on how preferred distal feedback was that they found the vibration of the smartphone "*unpleasant*". Instead, the vibration of the smartwatch was more "*pleasant*". We received interesting use cases for our interaction concept from the participants. Four participants suggested that this interaction concept can be used to send subtle emergency messages when in danger. For example, one can draw the letter "e" in his pocket to ask for help if he is in danger,

3.2.3 Discussion

Our findings reveal that both phone feedback and watch feedback exhibit similar performance characteristics, particularly with regard to gesture recognition. Consequently, this suggests that smartwatch feedback can serve as a viable alternative to under-finger feedback for providing input confirmation when drawing eyes-free gestures on mobile device for in-pocket context. Furthermore, it is worth noting that watch feedback results in reduced phone movement. This, in turn, argues that smartwatch feedback emerges as a practical alternative to mitigate phone movement while providing interaction location cues.

However, it is important to note that gestures produced during distal haptic feedback (on the watch) exhibit geometric and kinematic differences compared to those produced with local haptic feedback (on the phone). For instance, *distal* gestures were significantly shorter, smaller, and drawn slower than *local* gestures, with a shorter duration of the gesture. A closer look at our results shows that although the length and duration of *distal CV* was significantly less than the *local cv* the differences in mean lengths are greater than the average duration, 12% compared to 2%. This shows that the effect of the location of the distal feedback is stronger on the length of the gesture than on the duration of the gesture. This discrepancy suggests that the location of the haptic feedback delivery influences how users perform gestures on mobile devices. The shorter and smaller gestures observed imply that users may perform gesture input with more caution when relying on feedback from a smartwatch. Similarly, the slower drawing speed of the gestures indicates that users may adapt their motor control and movement patterns when receiving feedback from a distal source. This finding underscores the importance of understanding how users adjust their interaction strategies based on the perceived location of the feedback.

Finally, the majority of participants, 75%, generally preferred the local vibration while they found the interaction more pleasant. The findings imply that participants may have found the local haptic feedback more intuitive, reliable, or satisfying compared to the distal haptic feedback. This could be attributed to factors such as the immediacy and directness of the feedback, as well as the perceived naturalness of the interaction as commented by our participants. It is also notable that some participants reported issues with the distal haptic feedback, such as delays or unnatural sensations. These technical and perceptual challenges associated with distal feedback may have contributed to the preference of the participants for local feedback. Consequently, local phone feedback should be preferred over distal watch feedback whenever possible.

However, despite perceiving distal watch feedback as "*unnatural*", its efficacy remains notable. This observation underscores the complexity of user preferences and the multifaceted nature of interaction design. Interestingly, Vatavu [228] argues for the exploration of "*non-natural interaction design*" as a transformative and creative process. This approach intentionally diverges from users' conventional expectations and experiences of interacting with the physical world. In doing so, it aims to create interactions that are highly usable and effective, despite deviating from traditional norms. In this context, distal watch feedback may serve as an example of such non-natural interaction design, offering unique opportunities for innovative and unconventional user experiences.

3.2.4 Summary

This study further explored the spatial aspects of haptic feedback on pocket interaction with smartphone. It showed that while distal feedback through a smartwatch can be effective, users predominantly prefer haptic cues directly from the smartphone. These findings underscore the importance of haptic feedback in enhancing user confidence and accuracy in eyes-free interactions, especially in constrained environments like pocket-based interactions. Recognizing the value of haptic feedback in such scenarios, our next study focuses on a scenario where the attention of the user to the task is limited due to the presence of another attention-saturating task.

3.3 Enhancing Touch Circular Knob with Haptic Feedback when Performing Another Saturating Attention Primary Task

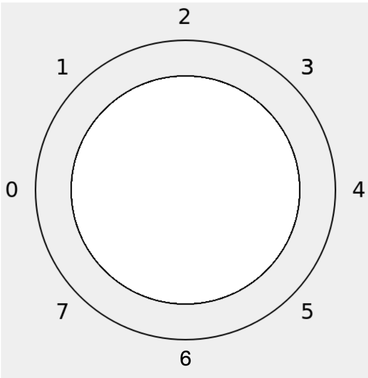
In two previous studies, we showed how haptic feedback can be used to improve the production of gestures and user confidence in gesture input when the smartphone is physically constrained and not directly accessible. In this study, we will explore another context involving visual impairment scenarios: multitasking, and investigate how tactile feedback can enhance eyes-free input interactions. In particular, we are interested in improving user input for targeting tasks on a touchscreen when they are occupied by another attention-saturating primary task. These two tasks make the interaction much more challenging than in previous studies and mayrequire more sophisticated solution. For this reason, in this study, we will use an ultrasonic surface that provide the user with considerably different sensations than those of a simple vibration motor found in current smartphones, delivering tactile textures perceived at the fingertips when moving the finger, in the context of active touch. In this experiment, we investigate the effect of different types of haptic feedback on a virtual knob. Is it also better to have continuous haptic feedback in this scenario, similar to previous studies? If so, since the haptic surface can provide much richer textures, which texture will improve the targeting task better than others.

3.3.1 Haptic feedback designs

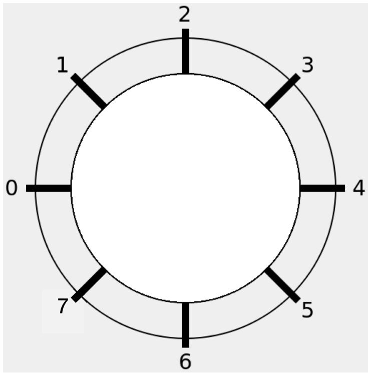
We investigate three different haptic feedback conditions tailored for targeting task that involves navigation along a circular touch path, resembling the operation of a touch-sensitive knob, such as adjusting music volume or managing the air conditioning controls. We also consider a *no feedback* design (see Figure 3.3a), which is equivalent to most touchscreens today and thus serves as our baseline. The three designed haptic feedback are *HapticDetent*, *GradualTexture+HapticDetent*, and *BumpyTexture+HapticDetent*. In *HapticDetent*, each of the detents (points) provides the user with haptic feedback when the user's finger touches them (see Figure 3.3b). This design is inspired by previous work on directional targeting tasks [117, 257], which was shown to improve performance when considering a single task.

The next two proposed designs leverage the first design by adding a continuous haptic feedback, *background texture*, using different tactile textures throughout the entire trajectory. The rationale for adding background texture is to provide the user with a cue that the finger is always in the right trajectory, hopefully allowing the user to reduce visual attention to the targeting task. The rationale for maintaining the *HapticDetent* in these two later designs is to inform users when their finger reaches a detent on the knob. In *GradualTexture+HapticDetent*, the background texture's intensity increases when finger moves in the clockwise direction and decreases in the counterclockwise direction (see Figure 3.3c). This change in background texture intensity allows users to locate their finger on the trajectory. In *BumpyTexture+HapticDetent*, the background texture is higher when the user finger is around a detent on the knob and lower when the user finger is between two successive detents (see Figure 3.3d). The rationale for having an intenser background texture when the user's finger approaches a detent is to prematurely alert the users that their finger is crossing into a new detent area on the knob.

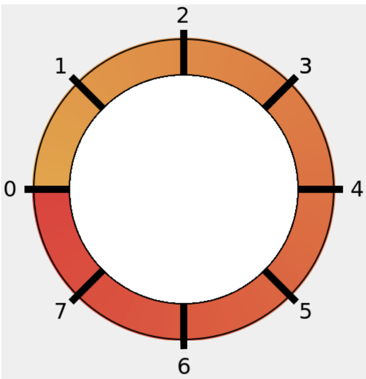
Implementation Tactile feedback was felt only on the perimeter of the circle, and this was determined by computing the distance between the center of the circle and the user's finger (see Figure. 3.3). The haptic feedback signal was rendered according to the position of the user's finger (x,y) on the perimeter of the circle. We calculate the angle formed by the two vectors $\overrightarrow{(x - x_c, y - y_c)}$ and $\overrightarrow{(x_i - x_c, y_c - y_c)}$, where (x_c, y_c) corresponds to the coordinates of the center of the circle and (x_i, y_c) the coordinates of a random point on the y=y_c vector which is parallel to the x axis with a value lower than x_c.



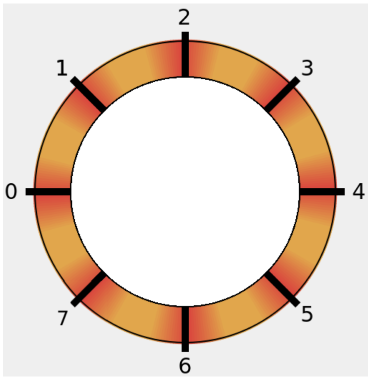
(a) No feedback



(b) HapticDetent



(c) GradualTexture+HapticDetent



(d) BumpyTexture+HapticDetent

Figure 3.3: The different haptic feedback designs. Tactile textures were replaced by visual representation in the figure. In (a), no haptic feedback was provided on the objects or on the trajectory. In (b-d) haptic feedback (through a tactile texture) was perceived when the user finger was on virtual detents. In (c) and (d), an additional background haptic feedback (through the use of different tactile textures) was perceived when moving the finger through the circular trajectory of the knob.

3.3.2 Experiment

We conducted an experiment to compare performance and users preferences among *no feedback*, *HapticDetent*, *GradualTexture+HapticDetent*, and *BumpyTexture+HapticDetent* for target selection on the circular touch knob on the tablet device when performing another primary task that saturates attention.

3.3.2.1 Participants and apparatus

15 participants (3 identified as women and 12 as men) volunteered to participate in our experiment. The ages of the participants were between 24 and 36 years (mean = 27.53 years, sd = 4.61 years). All participants were right-handed.

The primary task was implemented in the JavaScript framework using Node.js runtime and was run on a Dell laptop machine with a 13-inch LCD display screen with a desktop resolution of 1920×1080. The secondary task was implemented in Python 3 using PyQt5, a comprehensive set of Python bindings for Qt v5. The secondary task was then run on the *Xplore Touch* tablet[1] (the commercialized version of the E-VITA device [235]). *Xplore Touch* is a tactile feedback tablet that supports friction modulation by means of ultrasonic vibrations, where the squeeze film effect generates an ultrathin film of air between the finger and the surface when an ultrasound frequency is applied to a display overlay. *Xplore Touch* includes both visual and tactile feedback alongside the tactile display and is equipped with a 7-inch LCD display that includes a capacitive sensor that allows a sampling frequency of 50 Hz, similar to the capabilities of commercial mobile devices. The participants' faces were also videotaped using the integrated laptop camera (RGB, HD fixed focus with a resolution of 1280×720).

3.3.2.2 Design

The experiment used a 4 × 3 × 2 × 2 within-subject design for the factors: *feedback*, *size*, *amplitude* and *direction*. Feedback covers the four haptic feedback conditions: (1) *no feedback*; (2) *hapticObject*; (3) *gradual+hapticObject* and (4) *adaptive+hapticObject*. Size corresponds to the size of the target and objects and covers three conditions: (1) small: 10px (= 0.1494cm), (2) medium: 20px (= 0.2988cm) and (3) large: 30px (=0.4482cm). The amplitude corresponds to the distance between the control line and the target and covers two conditions: (1) short: between objects 3 and 6 (=6.5128cm) and (2) long: between objects 1 and 7 (=13.0256cm). Finally, the direction covers two directions: (1) clockwise and (2) counterclockwise. The task featured a circle with a radius of 160px (=2.39 cm) and a perimeter thickness of 50px(=.75cm).

3.3.2.3 Task & Procedure

The experiment required participants to interact with a primary task on the laptop and to select a target on the haptic circular trajectory on the tablet each time they received a notification. The touchscreen tablet lies flat on the table to the right of the computer.Participants were asked to prioritize the primary task over the secondary one and told that their performances were being measured. To better understand the effect of the primary task on the

[1] `https://www.hap2u.net/haptic-technology`

secondary task, we included a control condition in which participants had to react to the notifications displayed on the laptop screen.

We used the same attention-saturating task as in the previous chapter as the primary task. For the secondary task, participants were asked to perform the target task as quickly and accurately as possible. Participants were then given the exact procedure to follow for each trial:

State 1. A blue control line and a red target line appear on the circular knob on the tablet display. The control line and the target line are the same size. No visual cue was available about the type of haptic feedback.

State 2. Touch the control line and hold it for 0.2 s to free it. Then, after that, the control line is free to move.

State 3. Drag the control line over the target area. After holding for 0.2 s, if the center of mass of the control line was inside the target line, the target turned green to confirm the successful trial and the next trial started. If the finger is raised off the control line during the dragging task, then an error is counted, the target flashes orange, and the trial is repeated. During the dragging task, each time the control line touches an object, the object turns blue. The number of objects placed before the target was two in the short amplitude and five in the long amplitude.

During the experiment phase, the participants started to interact with the primary task. And after a random period of time between 2 and 14 seconds, a notification was shown on the computer screen to indicate to the participants that they could start the secondary task on the tablet surface. The participants were free to choose the appropriate strategy to handle the primary task and the secondary task, while keeping in mind that the primary task is prioritized. Participants have the total freedom to interact with both tasks at the same time or sequentially by switching from one task to another, etc. After ending a trial by dragging the control line to the target line, our software presented the next notification after a random period of time between 2 and 14 seconds.

For the secondary task, the experimental trials were administered as blocks of 14 trials, each block sharing a haptic feedback condition (no feedback, HapticDetent, GradualTexture+HapticDetent, BumpyTexture+HapticDetent), target size (small, medium, large), and movement amplitude (short and long). The first 4 trials (*i.e.*, 2 trials by direction) of each block were discarded to allow for adaptation of the strategy. The initial direction was randomly assigned for each block, then alternated with the inverse direction. The blocks are grouped by feedback condition to allow questionnaire assessment. The four feedback conditions were randomly presented to the participants. Within each feedback condition, the three target *sizes* were counterbalanced. Within each target *size*, the two amplitudes were also counterbalanced. Inside each amplitude condition, the two directions were alternated with seven repetitions for each direction. Participants performed a total of 24 blocks (=4 feedback conditions × 3 target sizes × 2 movement amplitudes), each block is made up of 14 trials (2 directions × 7 repetitions) – a total of 336 trials (=24 blocks × 14 trials) per participant. After finishing the 6-block set of each feedback condition, participants completed a NASA-TLX worksheet, plus a 5-point Likert scale questionnaire (strongly disagree to strongly agree) to measure enjoyment while interacting with the tablet device. At the end of the experiment, we asked participants to rank the feedback conditions and comment on the experiment. The experiment took around 90 minutes to complete. To reduce fatigue, sufficient rest periods were provided between feedback conditions and as required by the participants.

3.3.3 Results

The dependent measures are *accuracy*, *reaction time*, *movement time*, and *total time*. We computed **accuracy** as the ratio of the fraction of correctly performed trials to the total number of trials. *Reaction time* was the interval between the first display of the stimulus and the first touch. It represents the time required for participants to process the stimulus, switch to the secondary task, and decide which object to select. *Movement time* is the interval between the first touch and the selection of the target and represents the time required to physically navigate between objects to select the target. *Total time* was the sum of reaction and movement times. All analyses for dependent measures used a multiway ANOVA. Tukey tests were used post hoc when significant effects were found. Only significant effects and interactions are reported.

3.3.3.1 Accuracy

There was a significant main effect of *size* ($F_{2,28} = 3.33$, $p = .05$) on *accuracy*. Post hoc comparisons using the Bonferroni correction showed that the large size (mean=98%, SD=1.05%) is significantly more accurate than the small size (mean= 95.58% , SD= 1.47%) with no difference from the medium size (mean=96.5%, SD=1.29%). There were no significant interactions ($p > .19$) between the effect of size and feedback, or amplitude, or direction.

3.3.3.2 Reaction Time

There was a significant main effect of *size* ($F_{2,28} = 20.67$, $p < .0001$) on *reaction time* with *amplitude* × *direction* ($F_{1,14} = 4.99$, $p = .0422$) interaction. Large size is the fastest (mean= 2165 ms, SD= 165 ms), followed by medium (mean= 2438 ms, SD=191 ms) and small (mean= 2852 ms , SD= 242 ms). Post hoc tests confirm differences between all pairs ($p < .05$). We also found that, for the clockwise direction, the short amplitude (mean= 2325 ms, SD= 208 ms) is significantly faster than the long amplitude (mean= 2663 ms, SD= 257 ms) ($p < .05$).

3.3.3.3 Movement Time

There were significant main effects of *feedback* ($F_{3,42} = 3.52$, $p = .0228$), *size* ($F_{2,28} = 38.60$, $p < .0001$), *amplitude* ($F_{1,14} = 82.92$, $p < .0001$) and *direction* ($F_{1,14} = 12.80$, $p = .0030$) on *movement time*. Post hoc tests showed that *HapticDetent* (mean 2712 ms, sd 240 ms) condition is significantly slower than the three reminder feedback conditions (*GradualTexture+HapticDetent* (mean= 2108 ms SD= 171 ms), *BumpyTexture+HapticDetent* (mean= 2059 ms, SD= 134 ms) and *no feedback* (mean= 2248 ms, SD= 145 ms)) ($p < .05$).

Both medium (mean= 2150 ms, SD= 134 ms) and large (mean= 1939 ms, SD= 140 ms) sizes are significantly faster than the small size (mean= 2756 ms, SD= 175 ms) ($p < .05$). Unsurprisingly, the short amplitude (mean 1803 ms, sd 97 ms) is significantly faster than the long amplitude (mean 2760 ms, sd 136 ms) ($p < .05$). And the clockwise direction (mean= 2225 ms, sd 121 ms) is significantly faster than the counterclockwise direction (mean= 2338 ms, sd 134 ms) ($p < .05$).

3.3.3.4 Total Time

There were significant main effects of *feedback* ($F_{3,42} = 3.80$, $p = .0169$), *size* ($F_{2,28} = 82.56$, $p < .0001$) and *amplitude* ($F_{1,14} = 52.51$, $p < .0001$) on *total time* with *feedback*×*size* ($F_{6,84} = 2.71$, $p = .0184$) interaction. Post hoc tests showed that for small target size, the four feedback conditions were significantly different from each other ($p < .05$), with *GradualTexture+HapticDetent* feedback (mean 4615 ms, sd 380 ms) beign the fastest, then *BumpyTexture+HapticDetent* (mean 5469 ms, sd 591 ms), followed by *no feedback* (mean 5675 ms, sd 467 ms), and *HapticDetent* (mean=6677 ms, sd 721 ms). We also found that for the *no feedback* condition, the small size is significantly slower than both medium (mean=4519 ms, SD=397 ms) and large (mean=4165 ms, SD=412 ms) sizes ($p < .05$). Similarly, for *BumpyTexture+HapticDetent*, the small size is significantly slower than both medium (mean=4325 ms, SD=527 ms) and large (mean=3751 ms, SD=341 ms) sizes ($p < .05$). For the *HapticDetent*, the three sizes are significantly different from each other, the large is the fastest (mean 4487 ms, SD=468 ms), then the medium (mean=5101 ms, SD=556 ms), and small ($p < .05$). Furthermore, we found that the long amplitude (mean=5298 ms, SD=216 ms) is significantly slower than the short amplitude (mean=4237 ms, SD=203 ms) ($p < .05$).

3.3.3.5 Subjective results and observations

By using participants comments and responses to questionnaires, video analysis, and our observations during the experiment, we accompanied our quantitative data with considerable qualitative data that capture the mental models of users as they perform the task.

Questionnaires responses

Participants were asked to rank the four feedback conditions after completing the experiment. Overall, the *BumpyTexture+HapticDetent* condition was ranked 5 times (33.33%) first, 7 times (46.66%) second, 2 times (13.33%) third and only one time (6.66%) in the last position. The *GradualTexture+HapticDetent* condition was ranked 5 times as first position (33. 33%), 4 times second (26.66%), 4 times third (26.66%) and 2 times fourth (13.33%). In the third position, we found *HapticDetent* who was ranked first six times (6. 66%), third twice (20%), ninth times third (60%) and second twice (13. 33%). In the fourth position, we found the *no feedback* condition, which was chosen 4 times (26.66%) first, while being selected in the second position only once (6. 66%) and 10 times (66.66%) as the least appreciated feedback.

Our participants were also asked to rate the secondary task after each feedback condition in terms of performance, temporal demand, physical demand, mental demand, frustration, effort, and enjoyment. Table 3.3 shows the overall response of participants to this questionnaire. Friedman tests revealed that only mental demand was significantly different depending on which interaction was used ($\chi^2(3)$=11.88, p<.05).

	NF		*HD*		*GT+HD*		*BT+HD*		**Friedman**
	mean	sd	mean	sd	mean	sd	mean	sd	$\chi^2(3)$
Performance	3.73	.48	3.4	.37	3.93	.30	3.8	.20	5.77
Temporal demand	2.6	.49	2.53	.53	2.73	.52	2.53	.42	1.04
Physical demand	2.06	.58	2.13	.60	1.86	.46	1.8	.34	2.21
Mental demand	2.86	.53	2.33	.65	2.06	.44	2	.38	11.88
Frustration	2.26	.48	1.93	.40	1.66	.24	1.6	.49	6.55
Effort	2.46	.42	2.26	.44	2	.38	2	.38	5.46
Enjoyment	3.4	.59	3.4	.56	3.6	.53	3.66	.62	2.28

Note 1:Friedman tests are reported at p=.05 (*) significance levels. Significant tests are highlighted.

Note 2: We used the following abbreviations for feedback conditions: NF for No Feedback, HD for Haptic Detent, GT+HT for Gradual Texture + Haptic Detent, and BT+HT for Bumpy Texture + Haptic Detent.

Table 3.3: Mean and SD of the questionnaire's responses, rated on a scale of 1 (very low) to 10 (very high).

Strategies used to handle the primary and the secondary tasks

During the experiment, similar to [74], we observed three main strategies in managing the primary and secondary tasks, some strategies more dominant in one feedback design than others.

1. Competitive interaction with exclusive attention to the primary task where users interact with both the primary and secondary tasks at the same time while keeping their gaze attention mostly directed toward the primary task with a nearly eyes-free interaction with the secondary task. This strategy was mostly used in *GradualTexture+HapticDetent* and *BumpyTexture+HapticDetent* feedback designs, in particular after performing the first blocks.

2. Reactive interaction with shared attention to both tasks: users mostly interacted with the secondary task and reacted to the primary task only when the ball moved away from the center. In this condition, users gaze attention was shared almost equally between the primary and secondary tasks, with certain cases where the primary task received more gaze attention than the secondary one. This strategy was mostly used in *HapticDetent* technique in particular for large and medium target sizes. This strategy was also used by *GradualTexture+HapticDetent* and *BumpyTexture+HapticDetent* feedback designs, in particular in the first blocks.

3. Divided interaction with exclusive attention to the secondary task, where users start by making sure to center the ball well before switching to the secondary task, and then stop interacting with the primary task and switch to the secondary to select the target. Most of the gaze attention was directed toward the secondary task, with only a few glances toward the primary one. This strategy was mainly used in *no feedback* technique, and it was also observed frequently for the *HapticDetent* feedback design when selecting small targets.

Preferences for two-handed interaction

Concerning the coordination of hands, and contrary to the findings of Guettaf et al. [74], we observed principally one hand coordination used during the experiment and independently of the feedback design used. For instance, all our participants used principally a two-handed interaction: the dominant hand (right hand), which is close to the tablet device, was used to perform the secondary task, while the non-dominant hand (left hand), which is close to the keyboard, remains on the keyboard arrow keys, *i.e.,*, used for the primary task. When the secondary task was finished (*i.e.*, participants had to only perform the primary task), our participants continued to use their nondominant hand to perform the primary task.

GradualTexture+HapticDetent feedback comments

Two participants commented that they are able to know approximately where their finger is on the circular trajectory when using this haptic feedback because of its intensity. Ten participants were confident with the task and able to synchronize both tasks while their gaze attention was mostly directed towards the laptop screen, as they received continuous haptic feedback about if their finger is in the right trajectory.

BumpyTexture+HapticDetent feedback comments

In addition, for the *GradualTexture+HapticDetent* condition, our participants felt that as the haptic feedback is continuously perceived, the secondary task demands less attention and, consequently, synchronizing the primary and secondary tasks is easier. For example, one participant said: " *I do not need to pay too much attention to the secondary task, so I am more focused on the primary task* ". Another participant commented: " *I rest assured even if I do not look at the tablet as the haptic feedback is continuously perceived* ".

Interestingly, a participant comments: " *The different perceived textures help me to direct my finger to the target* ". In addition, one participant found the haptic feedback very useful, particularly for small targets. He said: " *For a small target, I do not think I can make it without textures, my eyes are quickly tired* ".

HapticDetent feedback comments

For *HapticDetent* feedback, our participants felt that despite the fact that the haptic feedback is helpful to determine approximately where they are, it is not sufficient to limit the visual attention needed for the secondary task. In addition, three participants found this technique helpful in determining where the finger is on the circular trajectory of the knob. Interestingly, 11 participants describe perceived haptic feedback on objects as a " *speed bump* " that allows " *the finger to be braked on arrival at the detent and then slide from the middle or end of the detent size, which takes the finger out of the detent*". For small targets, due to this feeling of sliding, nine participants commented that they had to go over the target and then go back to select it. Thus, those participants felt that this haptic feedback is not sufficient, especially for small targets, and that it makes the task more demanding both temporally and mentally, which is why they ranked it in the lowest position.

3.3.4 Discussion

Our findings indicate that when using *GradualTexture+HapticDetent* and *BumpyTexture+HapticDetent* techniques, participants kept their gaze attention mainly focused on the laptop screen (*ie*, the primary task). Furthermore, these techniques resulted in a faster selection of small targets and were less mentally demanding than *HapticDetent* and *no feedback* designs, without compromising accuracy. These observations indicate that perceiving such haptic feedback when performing a targeting secondary task permits participants not only to increase visual attention on the primary task but also to perform the secondary task faster. Since the accuracy of *GradualTexture+HapticDetent* and *BumpyTexture+HapticDetent* is not decreased, one may conjecture that the accuracy of these two designs may eventually increase with practice. Therefore, we recommend that researchers and designers using *GradualTexture+HapticDetent* and *BumpyTexture+HapticDetent* haptic feedback techniques for target selection on a circular trajectory when performing another saturated attention primary task to create an eye-free dialog between the surface and the user.

Interestingly, the *HapticDetent* design permits users to approximately estimate the position of their finger on the circular trajectory by counting the number of times they enter and leave an object; in particular, for large and medium target sizes. This finding is correlated with the study of Liao et al. [124] who showed that integrating haptic feedback into interaction with touchscreens helps users set a value or select a menu by counting the number of vibration ticks from a vibrotactile surface. Although *HapticDetent* is helpful, this design is still not sufficient to limit visual attention to the secondary task; even if it does not demand long eye glances like when selecting a small target or when using a baseline *without feedback* technique. For example, our participants were observed to keep an eye on the tablet when performing the secondary task to keep their finger within the trajectory. This issue can be addressed by using a more sophisticated gesture recognizer that can track knob rotation even when the finger is not exactly inside the expected trajectory. Furthermore, since the trajectory of the targeting task is circular, the secondary task may be more challenging than when the trajectory of the targeting task is directional.

HapticDetent technique demands the longest movement time, and it is also the slowest technique in terms of total time for a small target size. This finding can be explained by the comments of our participants. For example, our participants described the perceived haptic feedback on objects as a "speed bump" that limits the finger speed when reached, and then slips the finger from the middle or end of the target size, which can provoke target overshoots. Consequently, the additional time needed can be explained by target overshoots, and hence, longer distance traveled to select the target. Unfortunately, we have not counted the number of overshoots and the distance traveled. Consequently, future work will investigate these two measures to better understand the drawbacks of *HapticDetent* in terms of speed.

Furthermore, *HapticDetent* was described to be more difficult to deal with small targets. This suggests that it could be interesting to enlarge the space where haptic feedback is perceived by starting it before entering the object. The *GradualTexture+HapticDetent* and *BumpyTexture+HapticDetent* designs fill this gap by providing the user with additional feedback that can prevent the user from approaching a new notch. Other haptic feedback designs could be additionally evaluated, for example, by placing the haptic feedback before the object or at the leading edge of objects, as in the work of [257]. Future work will study these additional designs.

Similarly to Guettaf et al.[74], we observed different user strategies to handle and perform the primary task and the secondary task. However, in our study, the observed strategies depended mainly on the feedback used. For example, competitive interaction was mostly used with the *GradualTexture+HapticDetent* and *BumpyTexture+HapticDetent* techniques; reactive interaction was mostly used with the *HapticDetent* technique; and divided interaction was mostly used with the *no feedback* technique. This finding may help designers choose the appropriate haptic feedback for the secondary targeting task on the touchscreen based on their preferences for how to manage the primary task and the secondary one.

Previous work [18] has shown that, after learning visual and auditory stimuli, using haptic feedback for sliding allows reducing visual attention. For future investigations, it is therefore worth studying the effect of combining auditory, visual, and haptic feedback on reducing attention with respect to the secondary task.

3.3.5 Summary

We designed three haptic feedback designs for targeting tasks on a circular trajectory on a mobile surface when performing another primary task that saturates the attention. We then, conducted an experiment to evaluate and compare these three designs with a *no feedback* design. Our findings show that *GradualTexture+HapticDetent* and *BumpyTexture+HapticDetent* – which use a continuous haptic texture in the background along with a tactile texture at the detents of the virtual knob – improved total time for small targets and decreased the mental demand and the need for visual attention to the secondary task over both *no feedback* and *HapticDetent*, without compromising accuracy. We hope that this work will advance our knowledge for targeting tasks on touchscreen devices and that the *GradualTexture+HapticDetent* and *BumpyTexture+HapticDetent* designs will prove useful by adding them to the growing toolkit of circular targeting tasks on tactile devices when performing another primary task that saturates the attention, as they are seemingly well-suited to perform such tasks.

3.4 Conclusion

This chapter highlights the role that haptic feedback can play in eyes-free interaction by focusing on gestural input on touch devices, user confidence in giving a command to the smartphone, and user preferences for the pattern of haptic feedback. We showed that although continuous local haptic feedback did not cause a significant increase in gesture recognition rate, it increased the user confidence in the production of eyes-free gestures when there is a barrier between finger and touchscreen, in-pocket interaction. Similarly, when the smartwatch was used to provide haptic feedback on a different hand for in-pocket interaction, we did not observe a significant increase in gesture recognition rate. However, using the smartwatch made gestures smaller and slower. This shows that using haptic feedback for input, even if it does not increase a measurable matrix on the smartphone itself, such as the recognition rate, it can have a significant positive effect on the user experience (*i.e.*, increasing user confidence in the interaction) and change the gesture shape. Although in these two studies we applied haptic feedback locally or distally, having two actuators allows us to use them together and create richer haptic feedback by combining them in different ways. For example, a smartphone can vibrate continuously in gesture, while a smartwatch can vibrate at only specific points on gesture (such as the detents on a virtual knob). In the last study, we used a haptic surface for this purpose. Our results showed that even if some discrete haptic feedback is available in the interaction (*i.e.*, on the detents of a virtual knob) still adding a continuous haptic feedback in the background (*i.e.*, on the whole perimeter of a virtual knob) can improve the targeting task. Our findings show that *GradualTexture+HapticDetent* and *BumpyTexture+HapticDetent* reduced total time for targeting small indents and decreased mental demand and the need for visual attention to secondary task over both *no feedback* and *HapticDetent*, without compromising accuracy. In the last study, although the task was more challenging than in the previous two studies, performing a targeting task in the presence of an attention-saturating task, using a haptic surface with richer tactile feedback compared to the smartphone vibration motor had a significant positive effect on the interaction. In sum, in this chapter, we used both the simple vibration motor of the smartphone and the haptic surface to improve the command to touch devices using gestures. In the next chapter, we will focus on improving the output part of the eyes-free interaction using haptic feedback and gestures.

Chapter 4

Increasing the Bandwidth of Haptic Channel for Information retrieval from Smartphones

Unlike previous chapter which focused on input part of the interaction with the smartphone, this chapter focuses on the output, the second part of the interaction. The output part of the interaction with the smartphone is achieved primarily by using vision. However, in the absence of visual feedback, other sensory channels such as haptic can be used for substitution. The bandwidth of the haptic channel is significantly less than that of the visual channel. This makes it beneficial to develop new interaction techniques that can increase the bandwidth of the haptic channel as the output. For this purpose, in this chapter, we introduce a new interaction concept for eyes-free gestural information retrieval from smartphone using a simple vibrotactile motor of smartphones and a user performing gestures on a touchscreen. Through three studies, we investigates the utilization of haptic feedback as an output modality in the context of eyes-free gesture-based interaction with smartphone.

The first study introduces a new two-stage interaction concept, called *Hap2Gest*, that utilizes haptic feedback on gesture trajectories as an output mechanism. This two-stage interaction concept involves users first drawing a gesture to invoke a command, followed by a second gesture during which they receive vibration feedback along the gesture's path. The vibration patterns provide information related to the command initiated by the first gesture. This interaction combines tactile perception with kinesthetic perception to enrich the output information that can be provided by simple vibration motors and finger movement detection using a touchscreen.

The two next studies focuses on the use of circular gestures in the *Hap2Gest* method. This gesture shape was proposed by many participants for the *Hap2Gest* method. The second study explores the consistency and accuracy of users when drawing eyes-free arcs or circular gestures. When users draw circular gestures or arcs, they mainly rely on their propioceptive perception to locate the finger on the trajectory. If the user is asked to draw a half circle and a circle, is the length of the half circle half of the full circle? Does direction matter (clockwise or counterclockwise)? The answers to these questions are necessary to use circular gestures in the *Hap2gest* method.

In the third study, using the data derived from the previous study, we will implement circular gestures in the *Hap2Gest* method. Specifically, this study aims to assess whether users can accurately determine the location of haptic feedback while performing eyes-free circular arc gestures. The objective is to evaluate the feasibility of using haptic feedback with circular arcs to provide users with eyes-free information output.

By investigating these aspects, this chapter aims to introduce and study interaction techniques which combine tactile feedback with proprioception understanding of finger location to increase the bandwidth of data that can be transferred to user as output.

This chapter is organized into three sections. In the first section, we introduce a new interaction concept for completely eyes-free interaction with smartphones. The next two sections investigate the effectiveness of circular gestures in this method. Each section starts with an introduction to the study's objective, followed by a comprehensive explanation of the experimental design. Afterward, the results are presented and thoroughly discussed. Lastly, each section ends with a summary that highlights the main findings and insights obtained from the study.

4.1 Hap2Gest: A Gesture and Haptic Feedback-Based Eyes-Free Interaction Concept

In a visually impaired scenario where vision is not available, using haptic feedback for providing rich information to users is challenging. This is mainly due to the limited bandwidth of vibrotactile actuators available on smartphones, which is significantly less than the bandwidth that mobile screens can provide using visual feedback [109]. The vibrotactile actuators on smartphones can usually provide simple vibrations to confirm or alert users [69]. Is it possible to increase the output bandwidth of the haptic channel in smartphones using the same vibrotactile actuators? The tactile information that can be provided by these vibrotactile actuators is limited by their hardware and the software access to this hardware. However, in this study, we propose to focus on the user, instead of the smartphone, to increase the bandwidth of haptic communication. The haptic channel is composed of two modalities: tactile and proprioception. While interacting with the smartphone the tactile modality is the vibration we feel from vibrotactile actuators and sensing the texture of touchscreen. On the other hand, proprioception is our perception of the location of our fingers and their move [53]. Although the proprioception data are indirectly used for input on smartphones using an accelerometer which detects the movements of smartphones as a result of hand joints movements, to the best of our knowledge the proprioception data is not used as a part of output data for presenting data to users. In this study, we propose to combine proprioception data on how fingers move on touchscreen with tactile signals from vibrotactile actuators to increase the bandwidth of data that can be transmitted from the smartphone to the user

Figure 4.1: *Hap2Gest* concept and context: (left) eyes-free context of use example, (center) command invocation by drawing the input gesture, and then (right) drawing the output gesture and receiving the haptic feedback that corresponds to the output information through this gesture.

using the haptic channel. In the following, we present this interaction concept, called *Hap2Gest*.

4.1.1 Hap2Gest Concept

Hap2Gest is a new eyes-free interaction concept that uses gestures both as the input modality and as part of the output modality, combined with haptic feedback. First, the user draws an eyes-free gesture on the touchscreen of the smartphone to give a command to the phone (input). The user then draws a second eye-free gesture, the same as or different from the first, and can feel one or multiple vibrations on different parts of the gesture (Figure 4.1). The vibration pattern they feel, the number and location of vibrations through the output gesture, are the vocabulary of this interaction concept for output. The output modality uses a combination of the tactile and kinesthetic senses. The vibration created by the vibration motor of a smartphone creates the tactile sense, and the speed and position of the fingers create the kinesthetic sense. Thus, Hap2Gest would enable a less obtrusive way to retrieve information from user's phone, *i.e.*, without looking at the display or turning on his phone.

Our interaction design is similar to Roudaut et al., [184] as we both use gestures to give commands to the smartphone (input) and retrieve information with gestures and haptic feedback (output), in an eyes-free configuration. However, in their work, the output gesture is created by a force feedback system that moves the finger. Consequently, haptic feedback is used to guide the user to draw the output gesture, while the output gesture constitutes the main output information. For example, if the drawn output gesture is "8", the user will understand that he has received eight new messages. However, in our work, the output gesture is a predefined gesture drawn by the user, and the haptic feedback is created using the vibration motor at some points throughout the output gesture. Thus, the vibration patterns (their number and locations) constitute the main output information.

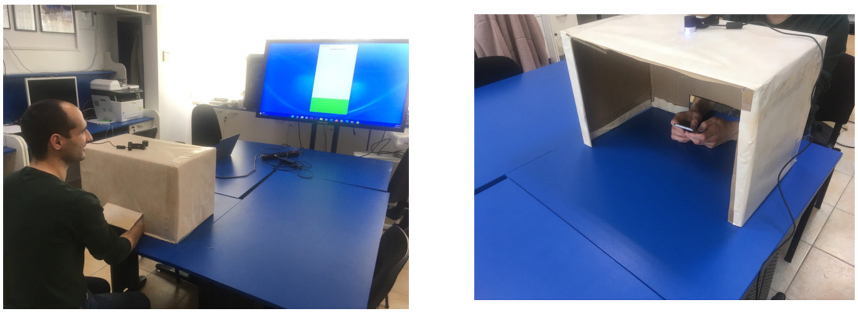


Figure 4.2: The experiment setup. The participant manipulated the smartphone while their hands were inside a box to maintain eyes-free interaction. The user interface of the experiment was displayed on a monitor in front of the participants.

4.1.2 Experiment

In the last 15 years, in particular for input gesture, an impressive body of work has been published on elicitation studies: the design of intuitive gesture commands that are reflective of end-user behavior for controlling all kinds of interactive devices, applications, and systems. In this context, we conducted an elicitation study to determine user preferences for input gestures, output gestures, and the vibration patterns for interacting eyes-free with smartphones for the design of *Hap2Gest*. Unlike earlier elicitation studies, we studied user-defined gestures when haptic feedback is available to provide feedback, using a simple vibration motor, in the absence of visual cues.

4.1.2.1 Overview & Rationale

We conducted an elicitation study to determine users' preferences for the design of *Hap2Gest*: (i) eyes-free input gestures for command invocation on a smartphone in the absence of visual cues, (ii) eyes-free output gestures for receiving the output information, and (iii) vibration patterns to obtain eyes-free output information through the output gesture.

The main premises underlying this research are that (i) a good gesture set needs to be easy to use and remember by the user, and (ii) vibration patterns need to be easy to understand, remember, and recognize by the user. Consequently, to support these main premises, we asked our participants to design gestures that are easier to remember and to come up with vibration patterns that take less time and effort to understand. Additionally, to avoid compromising the system in differentiating between the different commands, we asked participants to try their best to avoid having exactly the same gestures for different commands.

Similarly to previous studies on gesture elicitation [243, 146, 169], we do not want participants to focus on recognizer issues for the defined set of gestures. Consequently, we do not provide participants with recognition feedback during gesture production. We also asked participants to ignore recognition issues by considering the smartphone to be able to understand and recognize any gesture they might wish to perform. In addition, for vibration patterns, since we want to identify user preferences, we do not want participants to focus on tactile rendering issues. Consequently, no haptic feedback was provided to our participants during the task. We also encouraged participants to ignore haptic feedback issues by considering the smartphone to be able to render any vibration pattern they might wish to have.

4.1.2.2 Participants and apparatus

12 participants (three females, and nine males) volunteered to take part in our experiment. Participants' ages were between 20 and 32 years ($mean = 22.9$, $SD = 3.7$). All participants were right-handed and had used smartphones for several years and were familiar with the vibration motors of smartphones.

For the second stage of the study, we used a Samsung Galaxy S6 smartphone running Android 6.0.1. The phone dimensions were 5.65"$\times$ 2.78". The display resolution was 1440$\times$2560 pixels. We developed our application using Java to record the gestures and vibration patterns of the participants. The users' hands were videotaped using a Microsoft LifeCAM Studio webcam. An author observed each session and took detailed notes.

4.1.2.3 Scenarios & Referents

We wanted to create a list of common smartphone commands that users frequently use when visual cues are not available. Moreover, we wanted to cover different types of data that can be provided to users: binary data, categorical data, numerical data, range, or aided with auditory feedback. For this purpose, we considered five interaction scenarios in which users frequently need to execute such commands on the smartphone. Each scenario contains between one and 14 referents. Overall, the experiment included 25 referents. The list of the five interaction scenarios and associated referents is available in Table 4.1.

4.1.2.4 Procedure

First, participants watched a video, available as supplementary material, explaining our proposed interaction concept and the instructions for the experiment. The video served as *priming* [141] since it contained contexts of use and examples so that participants would think more generally about the proposed gestures. The written instructions were then distributed as printed forms, with 25 referents and possible responses for each referent. The participants were then asked to draw the gesture and vibration patterns for each interaction scenario on the paper forms. Participants were asked to pay attention to the following points while giving their answers:

- The participants will draw the gesture with one finger. They are allowed to hold the phone with the same hand or with the other hand.
- Try their best to avoid having the same gesture for different commands.
- Design gestures that are easier to remember.
- Come up with vibration patterns that take less time and effort to understand.

Table 4.1: The five interaction scenarios and the different referents considered in each scenario.

A. Scenario in the presence of auditory feedback	
R1. Accept call	**R2.** Reject call
B. Scenario with a yes/no response	
R3. Is it a call from my favorite contacts?	**R4.** Do I have any missed calls?
R5. Is my phone silent?	**R6.** Do I have new messages?
R7. Do I have new messages from a favorite contact?	**R8.** Do I have new notifications?
R9. Do I have a notification from Instagram?	**R10.** Do I have a notification from Facebook?
R11. Do I have a notification from Twitter?	**R12.** Do I have a notification from WhatsApp?
R13. Accept the call and tell me if it's from a favorite contact.	**R14.** Reject the call and tell me if it's from a favorite contact.
R15. Mute my phone with success feedback.	**R16.** Unmute my phone with success feedback.
C. Scenario with categorical responses	
R17. Which application do I have a notification from (Instagram or Facebook)?	**R18.** Which application do I have a notification from (Instagram, Facebook, or Twitter)?
R19. Which application do I have a notification from (Instagram, Facebook, Twitter, or WhatsApp)?	**R20.** Which application do I have a notification from (Instagram, Facebook, Twitter, WhatsApp, or Telegram)?
R21. How is the weather today? (sunny, cloudy, rainy, or snow)	**R22.** Which day of the week is it? (Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, or Sunday)
R23. Which month is it? (January, February, March, April, May, June, July, August, September, October, November, or December)	
D. Scenario with numerical responses	
R24. How many new notifications do I have? (0, 1, 2, 3, 4, or 5)	
E. Scenario for time range	
R25. At which time today do I have a meeting?	

After finishing their designs on paper, the participants moved on to the next step in the experiment. In this step, participants were asked to enter the solutions they had just designed on paper into a smartphone in an eye-free configuration. For this purpose, the participants held the smartphone in a box to avoid having visual cues. Five participants manipulated the phone with only one hand and the rest manipulated the phone with the dominant hand and held it with the other hand. Then, the papers on which they designed their solutions were given to them, and they were asked to copy their solutions from paper to smartphone one by one. An Android application was developed to capture participants' responses in an eyes-free configuration. The smartphone screen was mirrored on a display in front of the participants to guide them during the experiment and maintain eye-free interaction with the smartphone itself, as shown in Figure 4.2. At the top of the screen, the interaction scenario and the answers were displayed one after another. The participants had to first draw the gesture to give the command and then specify vibration points one by one by drawing the gesture from the beginning to the point where they wished to get the vibration. No visual cues were shown on the display about the touch point or gesture path, to ensure an eye-free condition. At the bottom of the screen, there was a green sliding button, which was used to approve the response by swiping right or to return to the previous steps by swiping left. A paper strap was glued on top of this virtual button so that participants could feel its position without seeing the phone. The average duration of the experiment for each participant was 50 minutes.

4.1.3 Results

Our results include the agreement rate measures, user-defined gestures set, and the user-defined vibration patterns set.

4.1.3.1 Agreement rate measure

To evaluate the degree of consensus among our participants, we used AGATe software (AGreement Analysis Toolkit) software [233] for calculating an agreement rate for each referent. An agreement rate, $AR(r)$, quantifies the magnitude of agreement among the gestures elicited from the participants, where:

$$AR(r) = \frac{|P|}{|P|-1} \sum_{P_i \subseteq P} (|\frac{P_i}{P_t}|)^2 - \frac{1}{|P|-1} \tag{4.1}$$

In Equation 4.1, P is the set of all proposals for referent r, $|P|$ is the size of the set, and P_i is the subsets of identical proposals from P. The range for $AR(r)$ is [0, 1]. In our study, we used the formula above to calculate two agreement rates: (1) the gesture's shape agreement rate and (2) the gesture's speed profile agreement rate.

The gesture shape agreement rate is the same as the agreement rate in previous elicitation studies. In this study, participants propose two gestures for each referent (an input gesture and an output gesture). We assumed that two designs were identical if they had identical input gestures and identical output gestures. After gathering all the proposals of the participants, the authors created a codebook [234] to assess the similarity of the proposals. We considered two gestures identical if they were made from equal number of strokes and the deviation between stroke angles was less than 45 degrees, even if two gestures had differences between their absolute position on the screen, their overall shape size, or strokes' lengths. Consequently, we calculate one gesture agreement rate for each design. However, in 96 percent of the designs, participants proposed the same gesture for input and output. Chance agreement [221, 234] was not considered and corrected, as the user elicitation study was not conducted with a fixed set of nominal categories out of which the participants chose their proposals.

In order to calculate gestures speed agreement rate, we derived the speed profile of gestures by taking the derivative of finger displacement with respect to time. The speed profile of all participants for each referent is available in the supplementary material. Then, we used the same formula as above on the speed profile. The codebook used for evaluating the similarity of speed profiles was

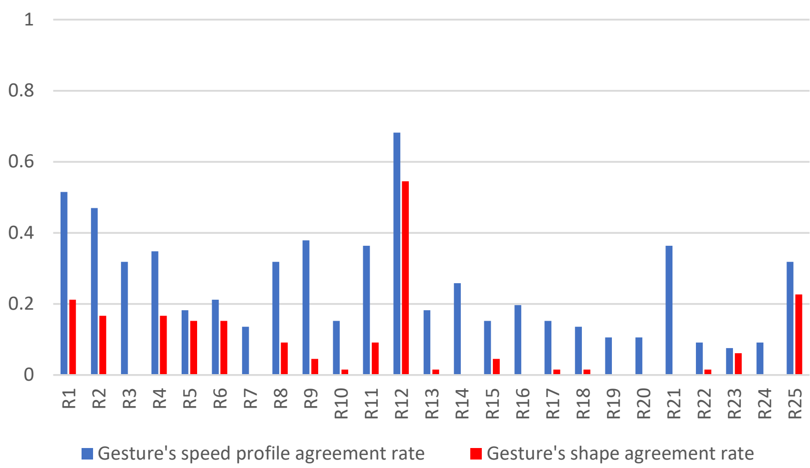


Figure 4.3: The gesture's shape agreement rates and the gesture's speed profile agreement rates are shown for all scenarios.

different from the one used for shape agreement rate. First, the peaks and lows in the speed profile were detected. Then, the speed profile was translated into series of peaks and lows, with the same order as they appear on the time axis. Two speed profiles were considered identical if they had the same number of lows and peaks in same order. To the best of our knowledge, we are the first gesture elicitation study that measures the speed profile agreement rate. The gesture set proposed by participants encouraged us to look at the agreement rate on the speed profile. We observed that the gesture shape proposed by participants for some referents had relatively low agreement, but the agreement rate among the speed profile of the same referent was significantly higher. So, it is more likely to find consensus by using gesture speed profile rather than gesture shape. Figure 4.3 shows the gesture agreement rates and the speed profile agreement rates for all 25 referents.

A paired-samples t-test was performed to compare the gesture agreement rate and the speed profile agreement rate of the 25 referents. There was a significant difference between gesture shape agreement rates ($mean = .081, SD = .122$) and gesture speed profile agreement rates ($mean = 0.252, SD = 0.153$); $t(24) = 8.545, p < 0.00001$. The gesture's shape agreement rate is, in particular, low. Thus, to better understand the cause of this low rate and, in particular, if it depends on the scenario or not, we decided to calculate the gesture shape agreement rate by scenario. We also calculate the speed agreement rate for each scenario in order to determine if the gesture's speed profile agreement rate compensates for the gesture's shape agreement rate in cases where the latter is low.

Finally, the vibration pattern agreement rate VAR is calculated for each referent using the formula below:

$$VAR(r) = \frac{|V|}{|V|-1} \sum_{V_i \subseteq P} (|\frac{V_i}{V_t}|)^2 - \frac{1}{|V|-1} \tag{4.2}$$

where P is the set of all vibration patterns of the proposal of the referent r,

$|P|$ is the size of the set, and P_i are the subsets of identical vibration patterns of P. The range for VAR is [0, 1]. The criteria used to consider two vibration patterns identical for each scenario are explained in the next section.

4.1.3.2 User-defined eyes-free Gestures & vibration patterns sets

In the following, we present, for each scenario, the most used input gestures, output gestures, and vibration patterns, along with the agreement rates for gesture's shapes, gesture's speed profiles, and vibration patterns.

Interaction scenarios in the presence of auditory feedback.

The most common gesture suggested for accepting a call was a straight line drawn from left to right by seven participants, and for rejecting a call, a straight line drawn from right to left by six participants. The mean agreement rates for gesture shape and gesture speed profile are 0.190 and 0.493, respectively.

For accepting or rejecting call scenarios where auditory feedback is available, 75 percent of participants preferred not to have any haptic feedback for either of the referents. The rest of the participants all preferred to have a single vibration at the end of the output gestures, which were identical to the input gestures, for both referents. The vibration agreement rates for both accept and reject call referents were 0.591.

Interaction scenarios with yes/no response.

In this scenario, for eight of the referents the most agreed gestures were gestures with shape of a letter from referent. For instance "N" shaped gesture for "Do I have new notifications?" referent and "M" shaped gesture for "Do I have new messages?" referent. All participants suggested same gesture for input and output for every single referent. For this scenario, the mean agreement rate for gesture shape and gesture speed profile were 0.094 and 0.277, respectively. Figure 4.4 shows the most suggested gestures for yes/no feedback scenario referents. Respectively, Figure 4.5 shows all the gestures suggested by participants for this scenario.

The vibration patterns suggested by participants can be categorized by three parameters. First, the number of locations on gestures where vibration is present. Figure 4.6a shows that users suggested between zero and two vibration points for these interaction scenarios. 96% of users preferred to have a vibration at a single point for yes response, and only four percent preferred to have vibrations at two points on a gesture for yes responses. For no response, most participants, 63%, again preferred to have vibration at one point on the gesture. However, contrary to the yes responses, 37% proposed not having any vibration for the no response.

Second, the vibration patterns can be categorized based on the number of vibrations proposed by the participants. Figure 4.6b shows that the participants suggested between zero and two vibrations for this interaction scenario. Most of the participants, 77% for yes responses and 64% for no responses, suggested one vibration on the gesture for feedback. However, the rest of the participants proposed no vibration for no response and two vibrations for a yes response.

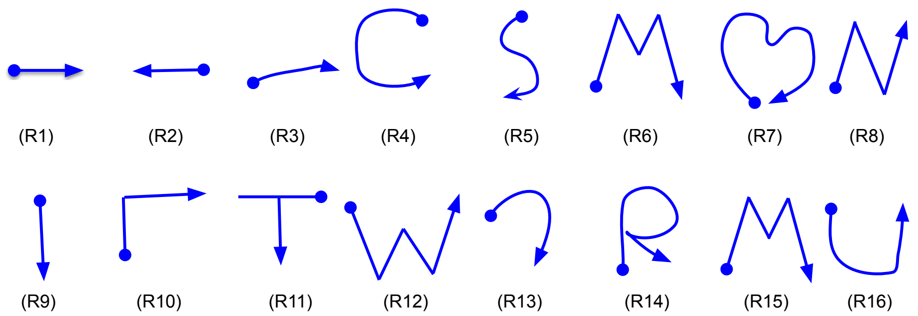


Figure 4.4: The most suggested gestures for referents R1 to R16. The filled circle shows the start point of the gesture. The arrow shows the ending point of the gesture.

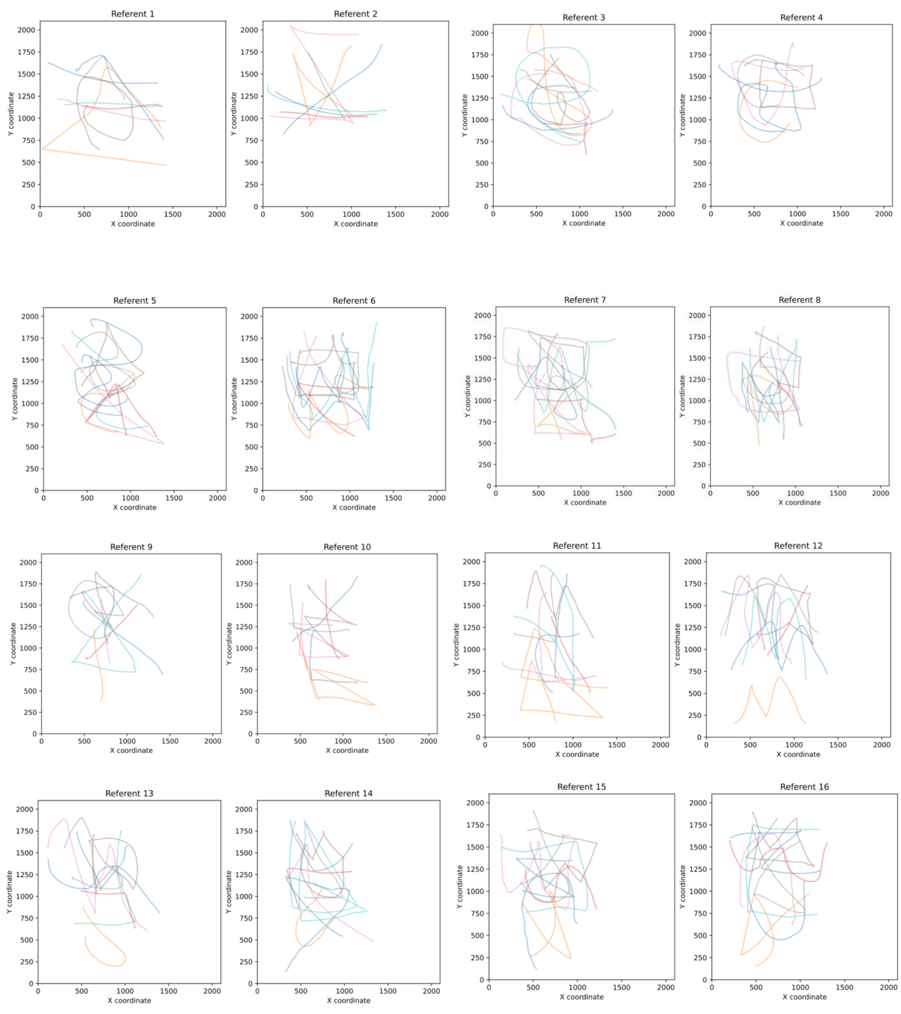
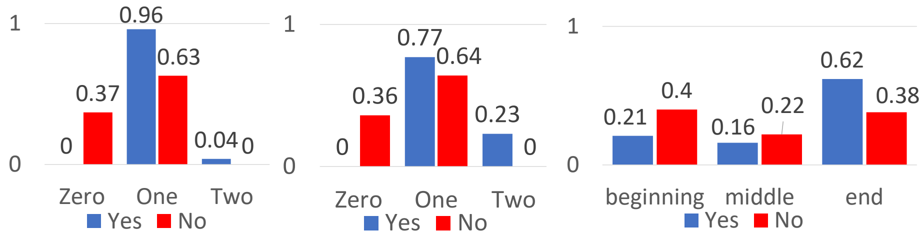


Figure 4.5: The gestures suggested by all participants for R1 to R16 referents. A Different color is dedicated to each participant.



(a) (b) (c)

Figure 4.6: The vibration patterns suggested by users for yes/no response scenarios. Figure (a) shows on how many points participants prefer to have vibration for each response. Figure (b) shows the number of vibrations they prefer to have for each response. Figure (c) shows where on the gesture they preferred to have the vibrations for each response.

Third, the vibration patterns can be grouped based on the position of the vibrations. Figure 4.6c shows the distribution of vibrations in gestures by categorizing them into three groups: at the beginning of the gesture, at the end of the gesture, and between. Our results show that most participants prefer to have a vibration at the end of the gesture for a yes response. However, for no response, vibration is preferred both at the beginning and at the end of gesture.

Using the three criteria stated above to consider the vibration patterns identical, we calculated the vibration agreement rate for all 14 referents. The results indicate that the mean vibration pattern agreement rate for the answer "yes" was ($mean = 0.263, SD = 0.082$) and the mean vibration pattern agreement for the answer "no" was ($mean = 0.242, SD = 0.046$), $t(13) = 3.802, p = 0.001$. Although most of the participants used similar vibration patterns for different referents of this scenario, no participant applied the same vibration pattern to every single referent of this scenario. The most suggested haptic feedback was a vibration at the beginning of the gesture for "no" response and a vibration at the end of the gesture for "yes" response.

Interaction scenario for categorical responses.

This interaction scenario includes seven referents. Four of these referents were analogs, R17 to R20, in the sense that the question asked was the same, but the number of possible answers was different. In all four referents, the question was "What application do I receive notification from" and the possible response was an application from two, three, four, or five applications. The most popular design was suggested by three participants. They differentiated the referents by adding an extra stroke to the end of the previous referent to increase the number of possible answers. Although the gesture shape they used was different, the speed profiles of those gestures were similar. For this set of designs, participants always assigned the vibration to the corners of the gestures. Figure 4.7 shows the most popular design for referents R17 to R20. Two participants used the same gesture for all these analogues referents and assigned different vibration points for each referent. They always assigned the vibration to the same location for the same application in all referents. Two participants used the first letters of the applications for gestures. The next most common suggested gesture shape was drawing the letter from each possible response as both the input and the output gestures and feeling the vibration at the end of the corresponding letter, depending on the response. For example, when the possible answers were “Instagram” or “Facebook”, they drawn “IF” for this referent. When the possible answers were “Instagram”, or “Facebook”, or “Twitter”, or “WhatsApp”, they drew “IFTW”. Figure 4.8 shows all gestures suggested by participants for referents R17 to R20.

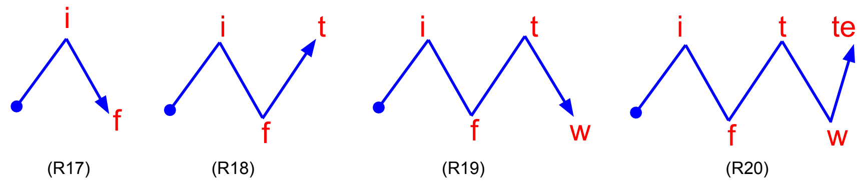


Figure 4.7: The most suggested gestures for referents R17 to R20. The filled circle shows the beginning of the gesture. The arrow shows the end point of the gesture. The red letters show the vibration point for each response. "i" for Instagram, "f" for Facebook, "t" for Twitter, "w" for WhatsApp, and "te" for Telegram.

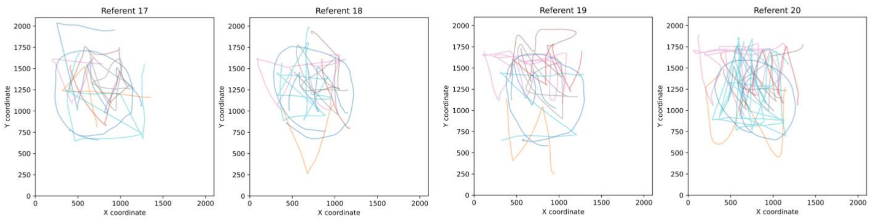


Figure 4.8: The gestures suggested by all participants for referents R17 to R20.

Referents R21, R22, and R23 were unrelated questions with four, seven, and 12 possible responses. For referents R21 and R22 with four and seven possible responses, respectively, 70 percent of participants proposed gestures with clear corners and assigned the vibrations to the corners of the gesture. However, for referent R23, with 12 possible responses, this percentage was 25 percent. Figure 4.9 shows the most suggested referents for these three referents. Respectively, Figure 4.10 shows all gestures suggested by participants for referents R21 to R25.

This interaction scenario included referents with two to twelve possible responses. Our results show that the users prefer to have gestures with clear corners, such as zigzag patterns or geometric shapes such as rectangles or hexagons, and assign the vibration to the corners of these shapes. The gesture shape agreement rate and the gesture speed profile agreement rate were 0.015 and 0.147. The vibration agreement rate for this scenario was 0.056 ± 0.061.

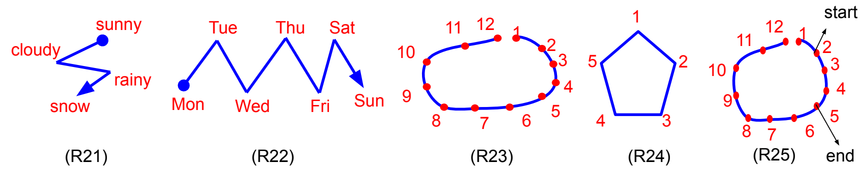


Figure 4.9: The most suggested gestures for referents R21 to R25. (R23) numbers 1 to 12, represent the vibrations for the month from January to December, respectively. (R24) shows that the participants suggest no vibration for the zero feedback and one vibration at the corners of the pentagon. (R25) The design suggested by half of the participants for a time range. They suggested feeling a vibration at the beginning of the time range (i.e., at 2 o'clock in the figure) and one at the end (i.e., at 5 o'clock in the figure) while drawing a circle.

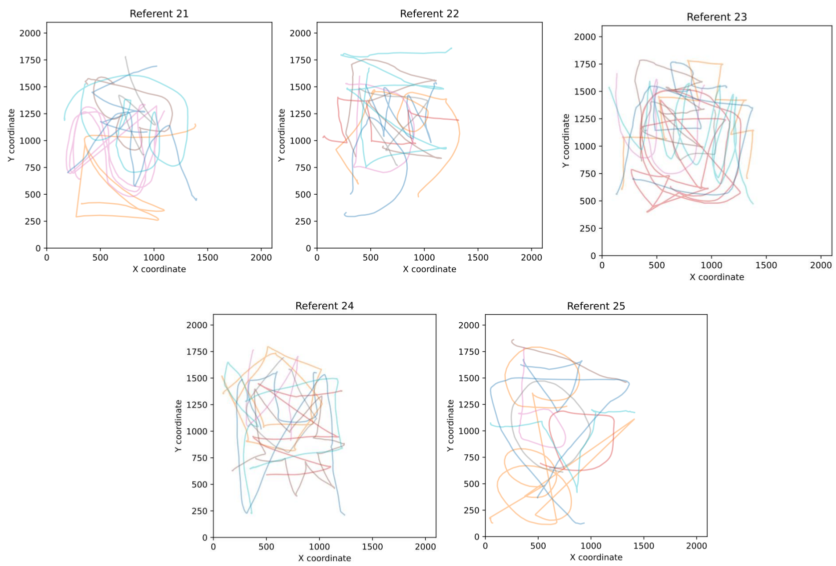


Figure 4.10: All gestures suggested by users for referents R21 to R25.

Interaction scenario for numerical responses from 0 to 5.
For numerical responses, 75% of participants preferred gestures with clear corners, similar to categorical responses. The gesture most suggested for this referent was a pentagon, as shown in Figure 4.9. However, some other designs were also suggested, like having vibration at different places in a straight line (for the "1" response vibration at the beginning of a straight line and for the "5" at the end of the straight line) or varying the number of vibrations (one vibration for the response "1" and five vibrations for the response "5"). In this case, the participant suggested a gesture shaped like the letter "N" as input gesture, and no gesture as output gesture. The gesture shape agreement rate was 0.000 and the gesture speed profile agreement rate was 0.091. The vibration agreement rate for this scenario was 0.106.

Interaction scenario for the time range.

For this scenario, we had one referent that asks in what time range do I have a meeting today. For this referent, unlike other referents, participants did not have access to possible responses when designing their gestures on the paper. We asked them to devise a solution that covers all possible time spans. 50% of the participants suggested a circular shape gesture, which corresponds to a clock, and proposed having one vibration at the beginning of the time range and one at the end. This design was the most popular one (an illustration is shown in Figure 4.9). The other participants suggested a similar idea, but for shapes other than circles, such as triangles. The gesture shape agreement rate and the gesture speed profile agreement rate were 0.439. The vibration agreement rate for this scenario was 0.439.

4.1.4 Discussion

In our study, the overall average agreement rate for the shape of the gesture was very low ($mean = 0.081, SD = 0.122$). The five scenarios in this study can be listed in descending order based on the gesture shape agreement rate as follows: 1) scenario for the time range ($mean = 0.227$), 2) scenario in the presence of auditory feedback ($mean = 0.190$), 3) scenario with yes/no responses ($mean = 0.094$), 4) scenario with categorical responses ($mean = 0.015$), and 5) scenario for numerical responses ($mean = 0.000$). The time range scenario has the highest agreement rate, though based on our participants' feedback, it was the most challenging scenario to design. However, in the end, half of the participants suggested the same design. Their design was based on a simple illustration of a clock's hours on a circle. This shows that the designers should try to use gestures, which are a simple illustration of an object or act related to the scenario. However, a commonly accepted illustration of every scenario is not possible. For instance, there is no common visualization of months or weeks accepted by public. The second scenario with the highest agreement rate was for accept and reject call referents. The gestures suggested by users are very similar to gestures used on many smartphones to accept and reject calls when visual cues are available. So, we recommend using designs that are similar to the gestures that are famously already used on smartphones when possible. For the yes/no response scenario, though the gesture's shape agreement rate was low, the most popular design was to use a letter from the scenario. We believe, this was suggested by many participants because it's easy to remember and the haptic feedback was a simple binary response, which can be easily applied to any gesture with any shape. Among the other scenarios, the categorical and numerical responses scenarios had the lowest rate of agreement for gesture's shape. This may be due to the fact that they were difficult to illustrate with a simple, widely accepted gesture, and the required haptic feedback was more complicated. However, for these two scenarios, the mean agreement rates of the gesture speed profile were higher, 0.147 and 0.091, respectively. This finding suggests that, for such scenarios, the speed of the gesture is more important to users than the shape of the gesture.

Based on our findings, in the following, we discuss the implications of our results for the design of the *Hap2Gest* concept in terms of gesture design, gesture recognition, and haptic design.

4.1.4.1 Gesture design and recognition implications

From our study, we recommend designers and researchers to **use the same gesture for input and output** to increase learnability and memorability. It also simplifies recognition. In our study, although we allowed participants to use different gestures for input and output, our user-designed sets emerged with 96% of the same gesture for input and output. This may be due to the fact that using the same gesture for input and output requires less memory effort for participants. Although participants decided to do this without considering the recognition problem, it is also beneficial to the recognizer. Since the two gestures drawn by users for input and output are quite similar to each other, the recognizer has a reference with which to compare the output gesture. For instance, if a user wants to draw a circle as a gesture, it is a very challenging task to detect on-line where on the circle the touch point is, especially at the initial stages. But when the input gesture is already recorded, the recognizer can easily detect the position on the circle by comparing the incomplete output gesture with the input gesture if there is no significant difference between the input and output gestures.

Although we recommend the designers use the same gesture for input and output in most cases, we don't recommend it when the gesture is too long. Drawing two long gestures can be slow and exhausting. To solve this issue, we recommend excluding the output gesture and providing the information with only vibrotactile messages at the end of the input gesture, or using a shorter input gesture. For example, for referents R22 and R23, where there were 7 and 12 possible responses, respectively, some participants used a short letter as input gesture (*i.e.*, letter "Z" as input gesture and a longer zigzag as output gesture for referent R22).

Our findings show there is a clear relation between the type of gestures and vibration patterns users suggest and the type of output information. First, **stroke gestures with geometric patterns (like rectangle, zigzag, polygon, etc.) should be preferred over alphabetic gestures for commands that can provide the user with many output information**, like in the scenario with categorical responses. What is important in this case is that the number of strokes that make up the output gesture is significant to the participants: **The more strokes the gesture contains, the more corners there are, and the more output information the output gesture can provide the user**. However, when the size of output information is so large, drawing so many strokes can be exhausting. To solve this problem, some participants suggested assigning multiple vibrations to a line or a curve, and not only at corners. In this way, the gesture becomes shorter and easier to draw. The agreement between the shapes of the geometries for such a referent, R23, was very low, but the majority of participants spread the vibrations on straight lines. However, the most popular design was to spread the 12 vibrations over a circle. Though this was the most popular suggestion, it does not show that it would work. Further experiments are required to determine if participants can accurately locate these 12 points on a circle while drawing gestures. If the circular gestures cannot have high accuracy, then the other design suggested by users can be used instead. Some participants, for example, proposed spreading the 12 points across a geometry with four edges with three vibration points on each edge. The corners available in this geometry may result in higher accuracy in locating the vibration points.

In contrast, **alphabetic gestures are interesting to use when considering commands that provide the user with a binary response, such as those in the scenario with a yes or no response**. In this case, the letter that corresponds to the first letter of the name of the application is a good choice for the input and output gestures. For example, to check if the user has received a message, the letter "M" can be used. This finding is correlated with the work of Roudaut et al. [184], where letters are used as output gesture accompanied with haptic feedback to notify the user that he has received a message. Successive letter shapes can also be used, for example, in the case where the user wants to have different notifications from different applications.

For challenging scenarios such as the time range, we recommend using gestures, which are simple illustrations of an object or act related to the scenario. For example, our participants used a circle to illustrate the hours of a clock. However, a commonly accepted illustration of every scenario is not possible. For instance, there is no common visual representation of months or weeks accepted by the public. Finally, **for scenarios where gestures are already famously used on smartphones, such as accepting and rejecting calls, we recommend using these familiar gesture shapes**.

As Wobbrock et al. [243] we advocate gesture reuse to increase learnability and memorability. Our user-designed set emerged with reusable gestures for analogous operations. Interestingly, to exclude ambiguity between different referents, in addition to relying on the target of the gesture as observed by Wobbrock et al. [243], our participants rely on the location of the haptic feedback for the output information. Two participants used the same gesture for all four referents, asking which application they have notifications from. They just added new vibration locations to each additional application while keeping the same gesture for the four referents.

For the categorical and numerical response scenarios, the recognizer should focus more on the speed profile and not on the geometry of the gesture, as there is more agreement between the speed profiles. For example, for the majority categorical referents, users draw a gesture with clear corners, such as a polygon or zigzag, and assign the vibration to the corners of the gesture. Although the mean shape agreement rates of gestures for categorical and numerical scenarios were not high (0.015 and 0.000, respectively), the mean speed profile agreement rates of gestures were significantly higher (0.147 and 0.091, respectively). This shows that the gestures the participants suggested were quite different, but the speed pattern was much more similar. In this case, the speed profile of these gestures is formed from consecutive bumps where, at the corners of the geometry, the corresponding speed is close to zero, at the bottom of bumps in the speed profile, and that is where most users prefer to have the vibrations.

4.1.4.2 Haptic design implications

Designers and researchers should place the output vibration on the corner of the output gesture. For example, in the first scenario, accept and reject call, all the participants who decided to have haptic feedback chose to have it at the end of the gesture. In the second scenario, yes / no responses, 89 percent of participants assigned vibrations at the beginning, end, or in between but at the corners. In the third scenario, categorical responses, the participants' preference was to have gestures with clear corners and assigning vibrations to the corners. However, it was a function of the number of possible responses. The percentage of participants who decided to have shapes with clear corners, such as a zigzag pattern, increased from 50% to 75% when the number of responses increased from 2 to 6. However, for the number of responses greater than 6, this percentage was less, *i.e.*, 33 percent for 12 possible responses. For the numerical response scenario, there was also a high tendency to assign vibrations to corners, 75%. These results show that participants find it easy to feel vibration in corners and tend to focus on the speed pattern of the finger rather than the geometry of the gesture.

The number of successively perceived vibrations could be used to provide users with numerical output information. While this solution was not the most suggested, it is still an interesting way to receive the output information. After drawing the input gesture, the user will simply remain stationary and count the number of distinct perceived vibrations. This method was suggested by one user for R22, R23, and R24 referents. Referent R24's output is intrinsically a number. However, the outputs of the referents R22 and R23 can have an order and can be numerated by the users, *, that is,*, one corresponds to the first month of the year and 12 corresponds to the last month of the year. Consequently, it is interesting to study in future work from both cognitive and precision points of view, for numerical output information, if it is better to perceive different vibrations through the gesture, each of which corresponds to a different number, or to stay stationary after drawing the input gesture while perceiving many successive vibrations (at the same location) such that their number corresponds to the output information.

When auditory feedback is available, haptic feedback can be excluded. For example, for accepting or rejecting call referents, most participants preferred not to have haptic feedback.

However, as our study was carried out without a primary task, this implication can depend on the context of the interaction. For example, in [18], authors found that, after training with visual and auditory feedback, the use of haptic feedback permits users to reduce their attention to the touchscreen.

4.1.5 Summary

In this study, we incorporate an additional dimension, the kinesthetic perception of the finger, into the human-smartphone interaction via the haptic channel. We presented *Hap2Gest*, a two-phase, gesture-based, eyes-free interaction concept that uses the tactile and kinesthetic senses to access information from a smartphone. We conducted the first elicitation study in the literature in which touchscreen gestures are chosen not only based on the task, but also considering the tactile channel for information retrieval. The findings reveal that in over 96 percent of instances, participants prefer using the same gesture for both input and output, which is advantageous for designers, as it simplifies recognition and enables online recognition. Our results also indicate a clear correlation between the types of gestures and the vibration patterns proposed by users and the type of output information. Furthermore, we demonstrated that the agreement rate for the gesture's speed profile is significantly higher than that for the gesture's shape, and it can be utilized by the recognizer when the gesture shape agreement rate is low.

Moreover, generally participants suggested having tactile feedback at the corners of gesture shapes. Since in the corners of gesture the speed profile reaches its local minimum value, designers can focus on the speed profile to locate the point where vibration should be applied. However, users also suggested shapes without clear corners, such as circles, for some scenarios. The lack of clear corners makes the determination of the correct location for vibration more challenging than shapes with clear corners. For this reason, in the next two studies, we will investigate the eyes-free production of gestures without clear corners (arcs and circles) and how accurately they can be used in the *Hap2Gest* method.

4.2 Production of Eyes-free Circular Gestures With Free Trajectory

In this study, we focus on circular gestures. The circular gesture paradigm is not only functional, but also engaging, providing users with a natural and fluid means of interacting with digital content. Moreover, in the previous study, a large number of users suggested the circular gesture to retrieve time information in the *Hap2gest* interaction concept. However, this suggestion was made by participants without actually investigating its feasibility in practice. Here, as a first step, we will study perception of the user about the location of their fingers on eyes-free free-hand circular path. We want to understand when a user is trying to draw a circle partially if their perception of finger location on the circle corresponds to the intended geometrical length.

4.2.1 Design Principles

We introduce three design principles aimed at investigating the use of circular arc gestures with haptic feedback for eyes-free interaction. By exploring different combinations of these principles, designers can create a diverse set of interaction techniques tailored to specific user needs and preferences in eyes-free interaction scenarios.

- **Continuous interaction**: a circular trajectory provides an infinite length for exploration without interruption at constant speed. Consequently, the continuous nature of circular gestures should reduce cognitive load by eliminating the need for users to hold and reposition their finger, especially in an eyes-free interaction context. This seamless flow enhances the user experience, making interactions more intuitive and effortless. For example, a linear trajectory, on the other hand, has a limited length, and the user has to lift the finger and move it to the starting point to continue the interaction. This will result in continues change of exploration speed, especially at the beginning and end of trajectory, which may influence the haptic perception.
- **Reusability**: Gesture reuse involves utilizing the same gesture to perform various tasks by considering gestures primitives, such as gesture dynamics [248]. Using the concept of reuse for circular arc gestures involves

utilizing variations in angles to represent different commands or functions in addition to the movement directions (clockwise and under clockwise). By altering the angle of the circular arc gesture, users can trigger distinct actions or operations, enhancing the versatility and efficiency of interaction. This approach allows for a more nuanced and flexible utilization of circular arc gestures, enabling users to perform a wider range of tasks without the need for additional gestures or complex input mechanisms. Moreover, we extend the principle of reuse to haptic feedback by allowing users to interpret perceived haptic feedback differently based on the gesture performed and where the haptic feedback is perceived along the gesture. For example, in the context of eyes-free interaction, during circular gesture drawing, depending on where the user perceives the haptic feedback, they can retrieve different information corresponding to the drawn circular arc. This adaptive interpretation of haptic feedback enhances the user's understanding of the interaction and optimizes the efficiency of gesture-based tasks.

- **Efficient space usage:** Circular gestures allow for long gestures and interaction on very small gadgets such as smartwatches. Beyond merely drawing individual circular arcs, users can execute successive or repetitive circles and arcs, allowing for extended gestures on small gadgets. This versatility enhances the usability of circular gestures on constrained interfaces, ensuring that users can engage with their devices effectively despite limited space. For instance, users may draw consecutive circles to obtain additional details about retrieved information, discerned through the position of the haptic feedback on the circle, i.e., the drawn circular arc. For example, after the first circle indicates receiving an email, subsequent arcs/circles could provide further information like the email category (personal, professional, etc.), number, etc.

4.2.2 Experiment

The objective of this study is to investigate the accuracy and variability of users' eyes-free drawing of an arc over a circular gestures on a touchscreen mobile device.

4.2.2.1 Participants and apparatus

Twelve right-handed participants (5 females and 7 males) were volunteered for this experiment. The ages of the participants ranged from 23 to 32, ($mean = 25.6, SD = 3.5$). All participants had used smartphones for several years.

The experiment was conducted using a Samsung Galaxy S7 smartphone. An android application was running on the smartphone which was showing arcs and asking participants to draw the arcs in an eyes-free configuration on the smartphone. The smartphone was held by participants under the table to ensure eyes-free configuration. The smartphone screen was casted on a 13-inch Dell laptop which was placed on the table in front of the participants, as shown in Figure 4.11.

Figure 4.11: The setup we used for circular eyes-free gesture evaluation. The user was sitting in front of the computer that was showing instruction. The user was drawing gestures on the smartphone which was held under the table to avoid visual cues.

4.2.2.2 Design

The experiment utilized a within-subject design with two factor: *angle* and *direction*. *Angle* refers to the degree of the arc to be drawn and encompasses 12 conditions: 30°, 60°, 90°, 120°, 150°, 180°, 210°, 240°, 270°, 300°, 330°, and 360° representing a full circle. On the other hand, *direction* pertains to the direction of the drawing movement and comprises two conditions: clockwise and counterclockwise.

4.2.2.3 Task and Procedure

The task required participants to draw an eyes-free arc of various degrees in both clockwise and counterclockwise directions.

The experiment was conducted in a silent room where participants sat comfortably on a chair in front of a desk. They held the smartphone in their right hand and placed it under the table to restrict the view. The participants looked at the monitor screen placed on the desk that displayed the instructions for the task. Participants were provided with a brief training session to familiarize them with the eyes-free drawing setup. Subsequently, they were guided through a series of trials where they drew arcs of different degrees in both clockwise and counterclockwise directions. Each degree of the arc was repeated four times to account for variability.

In the experimental phase, the two *directions* were randomly presented to the participants. Within each *direction*, the 12 *angles* $\times$ 4 repetitions were randomly presented to the participants, resulting in a total of 96 trials per participant (= 2 *directions* $\times$ 12 *angles* $\times$ $\times$ 4 repetitions). After each block of trials, participants took a break.

4.2.3 Results

We collected the touch gestures on the smartphone which included the coordinates of the drawn trajectory and timestamp. We also recorded the direction, degree, and accuracy of each arc drawn by the participants.

The ideal geometric length (idl) of each arc for each user was calculated by average length of 360 degree arcs for each user ($360L$) and nominal degree of the arc (deg) using the following formula:

$$idl = 360L \times \frac{deg}{360} \tag{4.3}$$

A two-way analysis of variance using ARTOol was conducted to examine the effects of direction, and expected angle on idl.

The two-way ANOVA revealed that there was significant main effect of angle ($F(11,264)=79.54$, $p < .001$) on idl. We ran Tukey Post-hoc analysis on the angle factor. This resulted in 66 tests from which 48 pairs of angles had statistically significant means from each other ($p < 0.05$). Here, we will report which arc lengths had the most interference with other angles, failed to reject the hypothesis that their mean is different from other angles, and ($p > 0.05$). The 120° arc's length was not significantly different from that of 4 other arcs (60°, 90°, 150°, and 180°). The 150° arc's length was not significantly different from that of 4 other arcs (90°, 120°, 180°, and 210°). The 180° arc's length was not significantly different from that of 4 other arcs (120°, 150°, 210°, and 240°). The 330° arc's length was not significantly different from that of 4 other arcs (240°, 270°, 300°, and 360°).

4.2.4 Discussion

The key finding is that, with the exception of larger angles, our study reveals that participants tended to overestimate the length of arcs compared to their ideal geometric length. These results are consistent in both directions. This would suggest that when designing interactions based on circular arcs, it would be more natural to position the degrees along the arc with varying distances rather than equal ones. This adjustment can accommodate users' tendency to overshoot, potentially enhancing both accuracy and efficiency when adjusting values along a circular arc.

4.2.5 Summary

In sum, the results of this experiment showed that when users draw an incomplete circle (arc) without looking at the smartphone, their perceived angle and length do not correspond to the actual angle and length of the arc. So, if

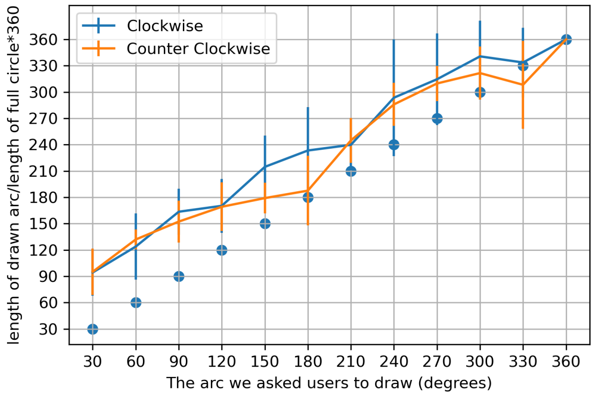


Figure 4.12: The mean lengths of different arcs drawn by user in clockwise and counterclockwise direction, on smartphone.

one wants to design an eyes-free interaction with a smartphone with circular gestures, one should take into account this discrepancy between the perceived angle and the produced angle of arc. In the next study, we will use this new information to increase the precision of the *Hap2Gest* method.

4.3 Using Haptic Feedback on Free Circular Eyes-free Gestures as Output

In this study, we examine the effectiveness of the circular gesture in *Hap2Gest* interaction concept for perceiving the angle of an arc on a mobile device. We will use the results of previous study on eyes-free production of arcs.

4.3.1 Experiment

We conducted an experiment to compare two methods of vibration feedback triggered at calculated arc lengths to assess participants' accuracy in estimating angles of eyes-free drawn arcs on a touchscreen mobile device.

4.3.1.1 Participants and apparatus

We recruited a new sample of 12 right-handed participants (7 males, and 5 females), aiming for a similar demographic as in the previous study to ensure consistency and comparability. Participants had prior experience with touchscreen interactions.

We Utilized the same Samsung Galaxy S7 smartphone used in previous section for this experiment. Participants were asked to draw arcs in both clockwise and counterclockwise directions in an eyes-free configuration. The smartphone was programmed to vibrate when the drawn arc reaches a length calculated based on target arc angle. The gesture length *gl* corresponding to each angle was calculated by using the following formula:

$$gl = 360L \times \frac{deg'}{360} \tag{4.4}$$

where $360L$ corresponds to length of the full circle user just produced to invoke the command before producing each arc, the deg' was the modified target angle of the arc and the number we wanted the user to retrieve using the haptic cues. The modified target angle is the angle we derived from Figure 4.12. We used this modifies angle instead of actual angle since we showed in previous experiment that length of the arcs produced by users is not exactly proportional to the angle of the arc and they tend to overshoot.

4.3.1.2 Design

The experiment utilized a within-subject design with three factor: *arc angle*, *arc direction*, and *reference circle*. We used the same *arc angle* and *arc direction* as we used in previous study. However, *reference circle* refers to the full circle which its length is used to calculate the length of the arcs for each angle and in result the length that they will feel the vibration for each angle at. We considered two *reference circles* in this study: *immediate circle*, and *non-immediate circle*. *Immediate circle* refers to when the reference circle was made by the user exactly before each trial and before starting to draw arcs to feel vibration and make the estimation. The length of each full circle was used as the reference circle for each trial. *Non-immediate circle* refers to when we asked each user to draw 5 circles in each direction. We calculated the length of the reference circle as the average length of the five circles in each direction.

4.3.1.3 Task and Procedure

The participants attended two sessions for two *reference circle* conditions. For each participant, the order of these sessions were random. In the *non-immediate circle* condition, we first asked participant to draw five full circles in clockwise direction and five full circles in counter-clockwise direction. The order of these ten circles for each participant was random. Then, we calculated the average length of the five circles in each direction as the length of full reference circle for each direction. Then, the length of each arc is calculated by multiplying the ratio of arc to full circle found for each angle in the previous experiment and the length of the full reference circle found in this experiment. In each session, after the lengths of each arc calculated by our program, the main experiment starts. In each trial, participants were just instructed to start drawing circular gesture on clockwise or counter-clockwise directions. They had to keep drawing the circular gestures, even if necessary multiple circles, until they felt the vibration. After feeling the vibration they lift their finger from the touchscreen and verbally report to the experimenter what is their estimation of the angle of the arc they drew. The participants were told that the angle can be any number greater than zero. In total we gathered 2880 angle estimations from participants (2 arc directions $\times$ 2 reference circles $\times$ 12 arc angles $\times$ 12 participants $\times$ 5 repetitions).

In the *immediate circle* condition unlike the *non-immediate circle* we didn't have one calculated for one full circle length. Instead before each trial and

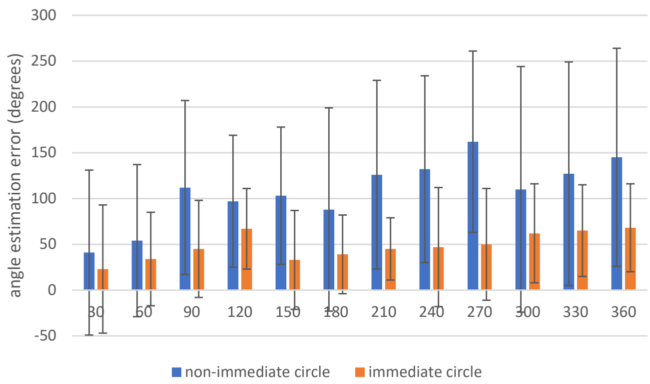


Figure 4.13: The mean and standard deviation of two reference conditions (immediate circle, and non-immediate circle) for each target arc angle.

each angle estimation participants were instructed to draw a full circle in the same direction as the next trial. So, before each trial, the full length circle was calculated independently from other trials. After the full length circle was calculated the arc estimation phase was similar to *non-immediate* condition. Our hypothesis was that when the reference circle is drawn immediately before the arc estimation phase there will be less variance in the arc shapes and the participant's estimation would be more accurate.

4.3.2 Results

A three-way ANOVA was conducted to examine the effects of arc degree, arc direction, and reference circle on angle error. The means and standard deviations for angle error across different levels of *arc degree*, *arc direction*, and *reference circle* are reported below.

Arc degree: There was a significant effect of *arc degree* on *angle estimation error* ($F(11, 2843) = 2017.86, p < 0.001$). The mean and standard deviation (SD) of angle estimation errors are shown in Figure 4.13. In general, a higher *arc degree* caused a higher error in the estimation of angle by the user.

Arc direction: We didn't find a significant effect of arc direction on *angle estimation error* ($F(11, 2843) = 3.49, p = 0.06$). So, we didn't find any evidence that the direction of gesture (clockwise or counterclockwise) affects the accuracy of angle estimation.

Reference circle: There was a significant effect of *reference circle* on *angle estimation error* ($F(11, 2843) = 2017.86, p < 0.001$). The mean angle estimation error of the immediate circle ($48.17°$) was significantly less than the mean angle estimation error of the non-immediate circle ($108.08°$). This shows that asking for the reference circle right before angle estimation will cause less error in angle user's estimation.

We didn't find any interaction between any of the variables.

4.3.3 Discussion

The results indicate that the reference circle used has a significant impact on angle error, with the *immediate circle* method resulting in significantly lower angle errors compared to the *non-immediate circle*.

Arc angle also significantly affects angle error, suggesting that different degrees result in varying levels of error. Generally, the higher the "arc angle" the higher error in estimation was observed. However, the error we observed in angle estimation was relatively high. Although the *immediate circle* caused lower error compared to *non-immediate circle*, it's error is still high and makes it impractical to use.

We recommend two possible solutions for this issue. First, one can use a more sophisticated recognizer, which in real-time can analyze the shape of the gesture made by user and tries to locate user's finger position on the circle. This method is relatively challenging, especially at smaller angles, where there is less time and data available for analysis. The second solution is to fix the radius and position of the circle and force the user to move the finger over the predefined circle. Although fix circle is common on touchscreens, using them in an eyes-free configuration is challenging as the user needs to locate the circle without visual cues. One possible solution to tackle this problem is to provide a haptic texture on the circle and then use a different haptic feedback on the target points.

4.4 Conclusion

This chapter delves into the output part of interaction with smartphones, particularly focusing on the haptic channel as an alternative to the visual channel for eyes-free information retrieval. Given the significantly lower bandwidth of the haptic channel compared to the visual channel, the chapter introduced a new interaction techniques to enhance this bandwidth, leveraging simple vibrotactile motors and user gestures on a touchscreen. In the first study, we introduced *Hap2Gest*, a novel two-stage interaction concept that uses haptic feedback on gesture trajectories as an output mechanism. This approach integrates tactile and kinesthetic perception to enrich the haptic feedback provided by smartphones. The study's findings indicated a high user preference for using the same gesture for both input and output, facilitating simpler recognition and enhancing the design process. Furthermore, the study highlighted the importance of the speed profile of gestures over their shapes in determining the application points for vibration feedback, with a specific user inclination towards tactile feedback at gesture corners. The speed profile of the gesture can help to locate the corners of the gestures. Users tend to reduce the speed of their fingers in corners. This is because users have to change the direction of the finger movement suddenly at corners.

On the other hand, in the next two studies we focused on the feasibility of using circular gestures in *Hap2Gest* method. Using circular gestures in the *Hap2Gest* method is challenging, as there is no clear corner in circles or arcs (incomplete circles). This makes it difficult to locate the right point for the haptic feedback. We tried to address this issue by first examining the consistency and accuracy of users performing eyes-free circular gestures. The investigation revealed a significant discrepancy between users' perceived and actual angles and lengths when drawing arcs without visual feedback. We used this information for applying vibration on arcs in *Hap2Gest*. We used the gesture length to find the correct point for applying vibrations on circular paths. However, our results showed a large error for information retrieval on circular gestures using this method. This shows how challenging it is to use circular gestures for the *Hap2Gest* method. So, we recommend designers to use gestures with clear corners in this interaction technique. If using circular gestures is necessary, we recommend using much more sophisticated gesture recognizer (software) or fixing the circle on the screen or using haptic feedback on the perimeter of the circle and a different texture or intensity on the target point which requires more sophisticated haptic actuators (such as haptic surfaces).

In summary, in this chapter we introduced a new interaction technique, which allows complete eyes-free interaction (input and output) using only haptic channel. We in particular showed how challenging it is to use circular gestures for this method, and we suggest using gestures with clear corners instead.

Chapter 5

Conclusions and Perspectives

In this final chapter, we summarize the key findings from our studies on eyes-free gestures and haptic feedback interactions with smartphones. We also discuss the broader implications of our work for future research and practical applications in mobile interface design.

5.1 Conclusions

This thesis focused on understanding and improving touchscreen interactions, specifically on smartphones, in situational impairment scenarios using haptic channel and eyes-free gestures. Creating designs that focus on situational awareness can result in improved solutions for people with disabilities. Similarly, focusing on the needs of disabled people can inspire innovations that address situational impairments [20]. They both have similar challenges for the interaction through haptic channel. These challenges can be divided into two parts: first, the difficulty in input (invoking commands on smartphones using one hand) without any visual feedback, and second, the difficulty of retrieving information from smartphones using the limited bandwidth of the haptic channel.

First, we delved into differences in gesture production under eyes-free conditions versus visible phone scenarios. Our results revealed significant differences in gesture production and recognition in eyes-free configurations compared to when the smartphone is visible. Then, we studied how some environmental factors can affect the production of eyes-free gestures on mobile devices. We found that a faster speed of user movement resulted in longer, larger, and faster gestures, while holding the phone in a shoulder bag produced slower, shorter, and smaller gestures. Higher movement speeds also caused more deviation in the phone's orientation angles. Finally, we focused on environmental factors caused by multitasking. Our results showed that an attention-saturating task alters the geometric features of gestures, making them shorter and faster. Right-handed users preferred to use their dominant hand for gesture production to reduce frustration. This research demonstrated that situationally induced impairments affect gesture production differently. Environmental factors influence user interaction strategies, and recognizing these factors can help improve gesture recognition accuracy. The built-in smartphone sensors can provide context information, such as accelerometer data indicating running or ambient light sensors showing a closed space.

We explored the potential of haptic feedback in improving eyes-free gestural input with smartphones. Continuous haptic feedback significantly improved user confidence and gesture accuracy during in-pocket interactions compared to no feedback or variable feedback. Although distal feedback through a smartwatch was effective, users predominantly preferred haptic cues directly from the smartphone. This study highlighted the importance of haptic feedback in eyes-free interactions, especially in constrained environments. However, we did not find significant differences in some of our measurements, such as gesture recognition. The large number of similarities we observed shows that it is possible to use haptic feedback for input both distally or locally if depending on the situation. Then, we used haptic feedback to improve gesture input in a completely different way. This time the user directly touched the touchscreen, so there was no doubt by the user about the detection of finger by touchscreen. However, this time the user needed to move his/her finger on a fixed circular path (a virtual knob) and target specific points on the circular path (detents of the virtual knob). For this purpose, we used a two-layer haptic feedback to guide the touch input. First, a background texture was applied on the perimeter of the knob to guide the gesture path. Second, a different texture was applied to the detents of the virtual knob. The presentation of two different tactile textures on the touchscreen required a more sophisticated haptic actuator. So, we used a haptic surface instead of a vibrotactile actuator. We presented three haptic feedback designs for the implementation of the virtual knob. GradualTexture+HapticDetent and BumpyTexture+HapticDetent designs improved performance and reduced mental demand compared to without feedback + HapticDetent alone. This research showed the potential of haptic feedback to improve touchscreen input in visually impaired scenarios. The existence of different haptic actuators provides a large opportunity for designers to use different actuators in their designs according to their needs. It's even possible to use these actuators together to extend the bandwidth of haptic channel if necessary.

Later we focused on leveraging the haptic channel as output in interaction with smartphones. We introduced the Hap2Gest concept, which combines tactile and kinesthetic senses to enhance eyes-free command invocation and information retrieval. Participants preferred using the same gesture for both input and output, simplifying recognition and enabling online recognition for designers. Generally, participants preferred having vibration points on the corners of gestures as it is easier for them to locate while producing the gesture. Gestures that have clear corners, such as zigzag patterns, also make the recognition side easier for the smartphone. Since users need to make a steep change in the direction of finger movement at the corners, the corners are usually easy to spot on gesture profile speed (local minimums). However, participants also suggested some gestures without clear corners, such as circles, where, due to lack of clear corners, the speed profile may not be very useful for the recognition. So, we explored the complexities and possibilities of circular gestures in this method. First, we studied how users produce arcs (incomplete circles) with different angles in eyes-free configuration. Users consistently overestimated the length of arcs in eyes-free circular gestures. Then we studied the ability of users to retrieve information on a circular gesture in the *Hap2Gest* method. Our results showed poor accuracy for information retrieval on circular gestures. The high error rates at higher arc angles highlighted the need for advanced gesture-recognition technologies or fixed circles with haptic textures. This research underscored the critical role of haptic feedback in eyes-free interactions, covering input, output, and user preferences for haptic feedback patterns for output signals and the challenges of using free-hand eyes-free circular gestures in *Hap2Gest* method.

In summary, this thesis explored the effects of different environmental factors on gestural interaction with smartphones and how haptic feedback can enrich eyes-free interactions. We identified user preferences, limitations, and performance outcomes for various interaction techniques. These findings lay the foundation for developing more intuitive and reliable eyes-free interfaces, emphasizing adaptable gesture recognition systems and thoughtful design considerations to accommodate different user conditions and environments.

5.2 Perspectives

Our studies had several common and general limitations. We only covered a few scenarios that can cause situational impairment, and all the experiments were conducted under controlled conditions in a laboratory or gym. Future research could investigate these questions in real-world scenarios such as driving, walking on the street, engaging in daily activities like cooking, or conversing with others. For example, the effect of physical activities, such as jogging, on gesture production is markedly different from that of cognitive tasks. It would also be interesting to study how social factors affect gesture production. For example, how does someone make gestures to check a message on their smartphone without interrupting a conversation? Situational impairments could also affect the perception of haptic signals from the smartphone.

In addition, most of our participants were between 20 and 30 years old. Eye-free interaction with smartphones is likely to be more difficult for older individuals, which warrants further investigation. In addition, in all our experiments, participants used one hand to manipulate the smartphone. Using both hands or the opposite hand could yield different results. Except for Section 2.2, the participants were comfortably seated in a chair during the experiments. In our research, we did not study the effects of mobility or other body postures. The size of the phone may also influence gesture production, an aspect we did not investigate. These are the general limitations of our studies, each of which also had specific limitations.

We used a gesture recognizer only in some studies (Sections 2.1, 3.1, 3.2 and 4.3), and each study used only one type of recognizer. Future studies could explore the effects of different gesture recognizers to determine which works best for eyes-free gestures. More sophisticated recognizers that incorporate additional sensor data, such as accelerometer data to detect user movement, could also be used to adjust the recognizer accordingly. In Section 2.2, we only used male right-handed participants to study the effect of movement speed and phone location on the production of eyes-free gestures. Given the differences in average running speeds between males and females, it would be worthwhile to conduct similar experiments with female or left-handed participants. In addition, other location on the phone, such as an armband during running, could be explored.

In Section 2.3, we used only one type of attention-saturating task and all participants were right-handed. Using other practical and common scenarios, such as driving, for attention saturation could provide more comprehensive insights. In Sections 3.1 and 3.2, the participants used the smartphone in the front pocket of their pants. Different pockets might yield different results. We also used only one type of fabric for the pocket; different fabrics and pocket sizes could influence interaction. For example, in preliminary studies, we observed that loose pockets caused the fabric to move with the finger, negatively affecting the detection of touch by the smartphone.

In the Hap2Gest interaction concept, participants proposed their designs for gesture and vibration patterns without actually trying them for information retrieval. Chapter 4 demonstrated that while some proposals, such as circular gestures, were popular among participants, they were challenging for designers to implement. Similar studies to those in Chapter 4 could be conducted for other gestures and vibration patterns proposed in the Hap2Gest concept. Furthermore, the cognitive load of the Hap2Gest method, particularly in the presence of a primary attention-saturating task, is an important question that needs investigation.

In Section 4.1, we only analyzed the length of the arcs and circles produced by the participants. Future research could use 2D features of the gestures produced for analysis, which is more complex. Similarly, in Section 4.2, our recognizer only used the gesture length to determine where the vibration should be applied. Developing a more sophisticated recognizer that uses 2D information to find the correct vibration point would be beneficial, though challenging, as the recognizer would need to find the correct point in real time. This would be especially difficult for smaller arcs, where there is less time and data for analysis. Finally, in Section 4.3, similar to Section 2.3, we used only one type of attention-saturating task in a controlled laboratory environment.

Overall, the studies presented in this chapter contribute to a deeper understanding of circular gesture interactions in eyes-free contexts. The insights gained highlight the potential for improving user performance and experience through thoughtful design adjustments and the integration of advanced haptic technologies. As we move forward, these findings will serve as a foundation for developing more intuitive and reliable eyes-free interaction techniques, paving the way for their broader application in various real-world scenarios.

Bibliography

[1] Mehmet Elgin Akpınar. *Context-Aware Prediction of User Performance Problems Caused by the Situationally-Induced Impairments and Disabilities*. PhD thesis, Middle East Technical University, 2022.

[2] Akshita, Harini Alagarai Sampath, Bipin Indurkhya, Eunhwa Lee, and Yudong Bae. Towards multimodal affective feedback: Interaction between visual and haptic modalities. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2043–2052, 2015.

[3] Abdullah X Ali, Meredith Ringel Morris, and Jacob O Wobbrock. Crowdsourcing similarity judgments for agreement analysis in end-user elicitation studies. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, pages 177–188, 2018.

[4] Lisa Anthony, Quincy Brown, Jaye Nias, and Berthel Tate. Examining the need for visual feedback during gesture interaction on mobile touchscreen devices for kids. In *Proceedings of the 12th International Conference on Interaction Design and Children*, IDC '13, page 157–164, New York, NY, USA, 2013. Association for Computing Machinery.

[5] Lisa Anthony, Radu-Daniel Vatavu, and Jacob O Wobbrock. Understanding the consistency of users' pen and finger stroke gesture articulation. In *Proceedings of Graphics Interface 2013*, pages 87–94. Canadian Information Processing Society, 2013.

[6] Lisa Anthony and Jacob O. Wobbrock. A lightweight multistroke recognizer for user interface prototypes. In *Proceedings of Graphics Interface 2010*, page 245–252. Canadian Information Processing Society, 2010.

[7] Lisa Anthony and Jacob O Wobbrock. A lightweight multistroke recognizer for user interface prototypes. In *Proceedings of Graphics Interface 2010*, pages 245–252. Canadian Information Processing Society, 2010.

[8] Lisa Anthony and Jacob O Wobbrock. $ n-protractor: a fast and accurate multistroke recognizer. In *Proceedings of Graphics Interface 2012*, pages 117–120. Canadian Information Processing Society, 2012.

[9] Lisa Anthony and Jacob O. Wobbrock. $n-protractor: A fast and accurate multistroke recognizer. In *Proceedings of Graphics Interface 2012*, page 117–120. Canadian Information Processing Society, 2012.

[10] Caroline Appert and Shumin Zhai. Using strokes as command shortcuts: cognitive benefits and toolkit support. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2289–2298, 2009.

[11] Caroline Appert and Shumin Zhai. Using strokes as command shortcuts: Cognitive benefits and toolkit support. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, page 2289–2298, New York, NY, USA, 2009. ACM.

[12] Kenneth Majlund Ba h, Mads Gregers Jæger, Mikael B. Skov, and Nils Gram Thomassen. You can touch, but you can't look: Interacting with in-vehicle systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, page 1139–1148, New York, NY, USA, 2008. Association for Computing Machinery.

[13] Leon Barnard, Ji Soo Yi, Julie A Jacko, and Andrew Sears. Capturing the effects of context on human performance in mobile computing systems. *Personal and Ubiquitous Computing*, 11:81–96, 2007.

[14] Olivier Bau and Wendy E. Mackay. Octopocus: A dynamic guide for learning gesture-based command sets. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, page 37–46. Association for Computing Machinery, 2008.

[15] Olivier Bau, Ivan Poupyrev, Ali Israr, and Chris Harrison. Teslatouch: Electrovibration for touch surfaces. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, page 283–292, New York, NY, USA, 2010. Association for Computing Machinery.

[16] Joanna Bergstrom-Lehtovirta and Antti Oulasvirta. Modeling the functional area of the thumb on mobile touchscreen surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1991–2000, 2014.

[17] Corentin Bernard, Jocelyn Monnoyer, Sølvi Ystad, and Michael Wiertlewski. Eyes-off your fingers: Gradual surface haptic feedback improves eyes-free touchscreen interaction. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–10, 2022.

[18] Corentin Bernard, Jocelyn Monnoyer, Sølvi Ystad, and Michael Wiertlewski. Eyes-off your fingers: Gradual surface haptic feedback improves eyes-free touchscreen interaction. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery.

[19] Frank Beruscha, Wolfgang Krautter, Anja Lahmer, and Markus Pauly. An evaluation of the influence of haptic feedback on gaze behavior during in-car interaction with touch screens. In *2017 IEEE World Haptics Conference (WHC)*, pages 201–206. IEEE, 2017.

[20] Tigmanshu Bhatnagar, Youngjun Cho, Nicolai Marquardt, and Catherine Holloway. Expressive haptics for enhanced usability of mobile interfaces in situations of impairments. *arXiv preprint arXiv:1904.06119*, 2019.

[21] David Birnbaum, Satvir Singh Bhatia, and Stephen D Rank. Orientation adjustable multi-channel haptic device, December 8 2015. US Patent 9,207,764.

[22] Rachel Blagojevic, Samuel Hsiao-Heng Chang, and Beryl Plimmer. The power of automatic feature selection: Rubine on steroids. In *Proceedings of the Seventh Sketch-Based Interfaces and Modeling Symposium*, page 79–86, Goslar, DEU, 2010. Eurographics Association.

[23] BB Blasch, RG Long, and Nora GRIFFINSHIRLEY. Results of a national survey of electronic travel aid use. *Journal of Visual Impairment & Blindness*, 83(9):449–453, 1989.

[24] Sebastian Boring, Dominikus Baur, Andreas Butz, Sean Gustafson, and Patrick Baudisch. Touch projector: mobile interaction through video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2287–2296, 2010.

[25] Sebastian Boring, David Ledo, Xiang 'Anthony' Chen, Nicolai Marquardt, Anthony Tang, and Saul Greenberg. The fat thumb: Using the thumb's contact size for single-handed mobile interaction. In *Proceedings of MobileHCI'12*, page 207–208. ACM, 2012.

[26] Andrew Bragdon, Eugene Nelson, Yang Li, and Ken Hinckley. Experimental analysis of touch-screen gesture designs in mobile environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, page 403–412. Association for Computing Machinery, 2011.

[27] Andrew Bragdon, Eugene Nelson, Yang Li, and Ken Hinckley. Experimental analysis of touch-screen gesture designs in mobile environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 403–412, 2011.

[28] Stephen Brewster and Lorna M. Brown. Tactons: Structured tactile messages for non-visual information display. In *Proceedings of the Fifth Conference on Australasian User Interface - Volume 28*, AUIC '04, page 15–23, AUS, 2004. Australian Computer Society, Inc.

[29] Jennifer L Burke, Matthew S Prewett, Ashley A Gray, Liuquin Yang, Frederick RB Stilson, Michael D Coovert, Linda R Elliot, and Elizabeth Redden. Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. In *Proceedings of the 8th international conference on Multimodal interfaces*, pages 108–117, 2006.

[30] Daniel Buschek, Maximilian Hackenschmied, and Florian Alt. Dynamic ui adaptations for one-handed use of large mobile touchscreen devices. In *IFIP Conference on Human-Computer Interaction*, pages 184–201. Springer, 2017.

[31] Matthew Butler, Leona Holloway, Kim Marriott, and Cagatay Goncu. Understanding the graphical challenges faced by vision-impaired students in australian universities. *Higher Education Research & Development*, 36(1):59–72, 2017.

[32] Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, and Amaury Trujillo. Analyzing visually impaired people's touch gestures on smartphones. *Multimedia Tools and Applications*, 76:5141–5169, 2017.

[33] Alexandre Calado, Paolo Roselli, Vito Errico, Nathan Magrofuoco, Jean Vanderdonckt, and Giovanni Saggio. A geometric model-based approach to hand gesture recognition. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(10):6151–6161, 2022.

[34] Géry Casiez, Nicolas Roussel, Romuald Vanbelleghem, and Frédéric Giraud. Surfpad: riding towards targets on a squeeze film effect. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2491–2500, 2011.

[35] Jessica R. Cauchard, Janette L. Cheng, Thomas Pietrzak, and James A. Landay. Activibe: Design and evaluation of vibrations for progress monitoring. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 3261–3271, New York, NY, USA, 2016. Association for Computing Machinery.

[36] Jessica R Cauchard, Janette L Cheng, Thomas Pietrzak, and James A Landay. Activibe: design and evaluation of vibrations for progress monitoring. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 3261–3271, 2016.

[37] Min Chul Cha and Yong Gu Ji. Context matters: Understanding the effect of usage contexts on users' modality selection in multimodal systems. *International Journal of Human–Computer Interaction*, pages 1–16, 2023.

[38] Landry Delphin Chapwouo Tchakoute and Bob-Antoine J Menelas. Perception of a haptic stimulus presented under the foot under workload. *Sensors*, 20(8):2421, 2020.

[39] Chen Chen, Soon Hau Chua, David Chung, Simon T Perrault, Shengdong Zhao, and Wing Kei. Eyes-free gesture passwords: a comparison of various eyes-free input methods. In *Proceedings of the Second International Symposium of Chinese CHI*, pages 89–92, 2014.

[40] Mingshi Chen, Panlong Yang, Jie Xiong, Maotian Zhang, Youngki Lee, Chaocan Xiang, and Chang Tian. Your table can be an input panel: Acoustic-based device-free interaction recognition. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(1):1–21, 2019.

[41] Qin Chen, Simon T. Perrault, Quentin Roy, and Lonce Wyse. Effect of temporality, physical activity and cognitive load on spatiotemporal vibrotactile pattern recognition. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, AVI '18, New York, NY, USA, 2018. Association for Computing Machinery.

[42] Qin Chen, Simon T Perrault, Quentin Roy, and Lonce Wyse. Effect of temporality, physical activity and cognitive load on spatiotemporal vibrotactile pattern recognition. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, pages 1–9, 2018.

[43] Lung-Pan Cheng, Fang-I Hsiao, Yen-Ting Liu, and Mike Y Chen. irotate: automatic screen rotation based on face orientation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2203–2210, 2012.

[44] Shaowei Chu, Fan Zhang, Naye Ji, Zhefan Jin, and Ruifang Pan. Pan-and-tilt self-portrait system using gesture interface. In *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*, pages 599–605. IEEE, 2017.

[45] James Clawson, Thad Starner, Daniel Kohlsdorf, David P Quigley, and Scott Gilliland. Texting while walking: an evaluation of mini-qwerty text input while on-the-go. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*, pages 339–348, 2014.

[46] J Richard Coast, Jennifer S Blevins, and Brian A Wilson. Do gender differences in running performance disappear with distance? *Canadian Journal of Applied Physiology*, 29(2):139–145, 2004.

[47] Andy Cockburn, Dion Woolley, Kien Tran Pham Thai, Don Clucas, Simon Hoermann, and Carl Gutwin. Reducing the attentional demands of in-vehicle touchscreens with stencil overlays. In *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '18, page 33–42, New York, NY, USA, 2018. Association for Computing Machinery.

[48] Wenzhe Cui, Suwen Zhu, Zhi Li, Zheer Xu, Xing-Dong Yang, IV Ramakrishnan, and Xiaojun Bi. Backswipe: Back-of-device word-gesture interaction on smartphones. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2021.

[49] Liwei Dai, Andrew Sears, and Rich Goldman. Shifting the focus from accuracy to recallability: A study of informal note-taking on mobile information technologies. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 16(1):1–46, 2009.

[50] Alexander De Luca, Alina Hang, Emanuel Von Zezschwitz, and Heinrich Hussmann. I feel like i'm taking selfies all day! towards understanding biometric authentication on smartphones. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, pages 1411–1414, 2015.

[51] Christina Dicke, Katrin Wolf, and Yaroslav Tal. Foogue: Eyes-free interaction for smartphones. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, page 455–458. Association for Computing Machinery, 2010.

[52] David Dobbelstein, Christian Winkler, Gabriel Haas, and Enrico Rukzio. Pocketthumb: A wearable dual-sided touch interface for cursor-based control of smart-eyewear. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(2):1–17, 2017.

[53] R Efe Dogruoz, Natalya S Shelchkova, Drew E Sheets, Charles M Greenspon, and Sliman J Bensmaia. The integration of tactile and proprioceptive signals to achieve haptic object perception. *bioRxiv*, pages 2023–11, 2023.

[54] Rachel Eardley, Anne Roudaut, Steve Gill, and Stephen J. Thompson. *Understanding Grip Shifts: How Form Factors Impact Hand Movements on Mobile Phones*, page 4680–4691. ACM, New York, NY, USA, 2017.

[55] Rachel Eardley, Anne Roudaut, Steve Gill, and Stephen J. Thompson. *Investigating How Smartphone Movement is Affected by Body Posture*, page 1–8. ACM, New York, NY, USA, 2018.

[56] Philipp Ewerling, Alexander Kulik, and Bernd 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*, pages 173–182, 2012.

[57] Leah Findlater, Ben Lee, and Jacob Wobbrock. Beyond qwerty: augmenting touch screen keyboards with multi-touch gestures for non-alphanumeric input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2679–2682, 2012.

[58] Feng Gao, Chaoyang Yu, and Jun Xie. Study on design principles of voice interaction design for smart mobile devices. In *Cross-Cultural Design. Methods, Tools, and Users: 10th International Conference, CCD 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings, Part I 10*, pages 398–411. Springer, 2018.

[59] Radhika Garg, Hua Cui, Spencer Seligson, Bo Zhang, Martin Porcheron, Leigh Clark, Benjamin R Cowan, and Erin Beneteau. The last decade of hci research on children and voice-based conversational agents. In *Proceedings of the 2022 CHI conference on human factors in computing systems*, pages 1–19, 2022.

[60] Edouard Gentaz, Gabriel Baud-Bovy, and Marion Luyat. The haptic perception of spatial orientations. *Experimental brain research*, 187:331–348, 2008.

[61] Debjyoti Ghosh, Can Liu, Shengdong Zhao, and Kotaro Hara. Commanding and re-dictation: Developing eyes-free voice-based interaction for editing dictated text. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 27(4):1–31, 2020.

[62] Nicholas A Giudice, Hari Prasath Palani, Eric Brenner, and Kevin M Kramer. Learning non-visual graphical information using a touch-based vibro-audio interface. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*, pages 103–110, 2012.

[63] Mayank Goel, Leah Findlater, and Jacob 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*, pages 2687–2696, 2012.

[64] Alix Goguey, Géry Casiez, Daniel Vogel, and Carl Gutwin. Characterizing finger pitch and roll orientation during atomic touch actions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2018.

[65] Alix Goguey, Sylvain Malacria, and Carl Gutwin. Improving discoverability and expert performance in force-sensitive text selection for touch devices with mode gauges. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2018.

[66] Alix Goguey and Michael Ortega. Studies and guidelines for two concurrent stroke gestures. *International Journal of Human-Computer Studies*, 170:102942, 2023.

[67] Alix Goguey, Daniel Vogel, Fanny Chevalier, Thomas Pietrzak, Nicolas Roussel, and Géry Casiez. Leveraging finger identification to integrate multi-touch command selection and parameter manipulation. *International Journal of Human-Computer Studies*, 99:21–36, 2017.

[68] Jorge Goncalves, Zhanna Sarsenbayeva, Niels van Berkel, Chu Luo, Simo Hosio, Sirkka Risanen, Hannu Rintamäki, and Vassilis Kostakos. Tapping task performance on smartphones in cold temperature. *Interacting with Computers*, 29(3):355–367, 2017.

[69] FJ González-Cañete, JL López Rodríguez, PM Galdón, and A Díaz-Estrella. Improvements in the learnability of smartphone haptic interfaces for visually impaired users. *PLoS One*, 14(11):e0225053, 2019.

[70] Francisco Javier González-Cañete, José Luís López-Rodríguez, Pedro María Galdón, and Antonio Diaz-Estrella. Improving the screen exploration of smartphones using haptic icons for visually impaired users. *Sensors*, 21(15):5024, 2021.

[71] Emilio Granell and Luis A Leiva. βtap: back-of-device tap input with built-in sensors. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 1–6, 2017.

[72] George Grouios, Efthymios Ziagkas, Andreas Loukovitis, Konstantinos Chatzinikolaou, and Eirini Koidou. Accelerometers in our pocket: Does smartphone accelerometer technology provide accurate data? *Sensors*, 23(1):192, 2022.

[73] Adnane Guettaf, Yosra Rekik, and Laurent Grisoni. Effect of physical challenging activity on tactile texture recognition for mobile surface. *Proc. ACM Hum.-Comput. Interact.*, 4(ISS), November 2020.

[74] Adnane Guettaf, Yosra Rekik, and Laurent Grisoni. Effect of attention saturating and cognitive load on tactile texture recognition for mobile surface. In *Human-Computer Interaction–INTERACT 2021: 18th IFIP TC 13 International Conference, Bari, Italy, August 30–September 3, 2021, Proceedings, Part IV 18*, pages 557–579. Springer, 2021.

[75] Sidhant Gupta, Daniel Morris, Shwetak Patel, and Desney Tan. Soundwave: using the doppler effect to sense gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1911–1914, 2012.

[76] Kyohei Hakka, Toshiya Isomoto, and Buntarou Shizuki. One-handed interaction technique for single-touch gesture input on large smartphones. In *Symposium on Spatial User Interaction*, pages 1–2, 2019.

[77] Chris Harrison and Scott E Hudson. Providing dynamically changeable physical buttons on a visual display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 299–308, 2009.

[78] Chris Harrison, Julia Schwarz, and Scott E Hudson. Tapsense: enhancing finger interaction on touch surfaces. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 627–636, 2011.

[79] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

[80] Morgan Harvey and Matthew Pointon. Searching on the go: the effects of fragmented attention on mobile web search tasks. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 155–164, 2017.

[81] Nabeel Hassan, Md Mahfuzur Rahman, Pourang Irani, and Peter Graham. Chucking: A one-handed document sharing technique. In *Human-Computer Interaction–INTERACT 2009: 12th IFIP TC 13 International Conference, Uppsala, Sweden, August 24-28, 2009, Proceedings, Part II 12*, pages 264–278. Springer, 2009.

[82] Patricia Heidt. Effect of therapeutic touch on anxiety level of hospitalized patients. *Nursing Research*, 30(1):32–37, 1981.

[83] Jay Henderson, Jeff Avery, Laurent Grisoni, and Edward Lank. Leveraging distal vibrotactile feedback for target acquisition. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–11, New York, NY, USA, 2019. Association for Computing Machinery.

[84] Jay Henderson, Jessy Ceha, and Edward Lank. Stat: Subtle typing around the thigh for head-mounted displays. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 1–11, 2020.

[85] Niels Henze, Enrico Rukzio, and Susanne Boll. Observational and experimental investigation of typing behaviour using virtual keyboards for mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, page 2659–2668. Association for Computing Machinery, 2012.

[86] Ken Hinckley, Robert JK Jacob, Colin Ware, Jacob O Wobbrock, and Daniel Wigdor. Input/output devices and interaction techniques., 2014.

[87] Ken Hinckley and Hyunyoung Song. Sensor synaesthesia: touch in motion, and motion in touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 801–810, 2011.

[88] Ken Hinckley, Shengdong Zhao, Raman Sarin, Patrick Baudisch, Edward Cutrell, Michael Shilman, and Desney Tan. Inkseine: In situ search for active note taking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, page 251–260. Association for Computing Machinery, 2007.

[89] Eve Hoggan, Stephen A Brewster, and Jody Johnston. Investigating the effectiveness of tactile feedback for mobile touchscreens. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1573–1582, 2008.

[90] Eve Hoggan, Andrew Crossan, Stephen A Brewster, and Topi Kaaresoja. Audio or tactile feedback: which modality when? In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 2253–2256, 2009.

[91] Christian Holz and Patrick Baudisch. Understanding touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2501–2510, 2011.

[92] Masoumehsadat Hosseini, Hamid Reza Hamidi, and Shokooh Kermanshahani. The impacts of situational visual impairment on usability of touch screens. *Multimedia Tools and Applications*, pages 1–25, 2024.

[93] Scott E Hudson, Chris Harrison, Beverly L Harrison, and Anthony LaMarca. Whack gestures: inexact and inattentive interaction with mobile devices. In *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*, pages 109–112, 2010.

[94] Heejin Jeong and Yili Liu. Effects of touchscreen gesture's type and direction on finger-touch input performance and subjective ratings. *Ergonomics*, 60(11):1528–1539, 2017.

[95] Satu Jumisko-Pyykkö and Teija Vainio. Framing the context of use for mobile hci. *International journal of mobile human computer interaction (IJMHCI)*, 2(4):1–28, 2010.

[96] Raine Kajastila and Tapio Lokki. Eyes-free interaction with free-hand gestures and auditory menus. *International Journal of Human-Computer Studies*, 71(5):627–640, 2013.

[97] Farzan Kalantari, Edward Lank, Yosra Rekik, Laurent Grisoni, and Frédéric Giraud. Determining the haptic feedback position for optimizing the targeting performance on ultrasonic tactile displays. In *2018 IEEE Haptics Symposium (HAPTICS)*, pages 204–209, 2018.

[98] Shaun K. Kane, Jacob O. Wobbrock, and Richard E. Ladner. Usable gestures for blind people: Understanding preference and performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, page 413–422, New York, NY, USA, 2011. Association for Computing Machinery.

[99] Shaun K. Kane, Jacob O. Wobbrock, and Ian E. Smith. Getting off the treadmill: Evaluating walking user interfaces for mobile devices in public spaces. In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services*, page 109–118, New York, NY, USA, 2008. ACM.

[100] Astrid ML Kappers and Myrthe A Plaisier. Hands-free devices for displaying speech and language in the tactile modality—methods and approaches. *IEEE Transactions on Haptics*, 14(3):465–478, 2021.

[101] Amy K Karlson and Benjamin B Bederson. Thumbspace: generalized one-handed input for touchscreen-based mobile devices. In *Proceedings of INTERACT 2007*, pages 324–338. Springer, 2007.

[102] Amy K Karlson, Benjamin B Bederson, and Jose L Contreras-Vidal. Understanding one-handed use of mobile devices. In *Handbook of research on user interface design and evaluation for mobile technology*, pages 86–101. IGI Global, 2008.

[103] Taslim Arefin Khan, Dongwook Yoon, and Joanna McGrenere. Designing an eyes-reduced document skimming app for situational impairments. In *Proceedings of CHI'20*, CHI '20, page 1–14, New York, NY, USA, 2020. Association for Computing Machinery.

[104] Wolf Kienzle, Eric Whitmire, Chris Rittaler, and Hrvoje Benko. Electroring: Subtle pinch and touch detection with a ring. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.

[105] Huhn Kim, Seungyoun Yi, and So-Yeon Yoon. Exploring touch feedback display of virtual keyboards for reduced eye movements. *Displays*, 56:38–48, 2019.

[106] Kiduk Kim, Ji-Hoon Jeong, Jeong-Hyun Cho, Sunghyun Kim, Jeonggoo Kang, Jeha Ryu, and Seong-Whan Lee. Development of a human-display interface with vibrotactile feedback for real-world assistive applications. *Sensors*, 21(2):592, 2021.

[107] Lawrence H. Kim and Sean Follmer. Swarmhaptics: Haptic display with swarm robots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–13, New York, NY, USA, 2019. Association for Computing Machinery.

[108] Konstantin Klamka, Tom Horak, and Raimund Dachselt. Watch+strap: Extending smartwatches with interactive strapdisplays. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–15, New York, NY, USA, 2020. Association for Computing Machinery.

[109] Roberta L Klatzky, Nicholas A Giudice, Christopher R Bennett, and Jack M Loomis. Touch-screen technology for the dynamic display of 2d spatial information without vision: Promise and progress. *Multisensory research*, 27(5-6):359–378, 2014.

[110] Ahmet Baki Kocaballi, Liliana Laranjo, and Enrico Coiera. Understanding and measuring user experience in conversational interfaces. *Interacting with Computers*, 31(2):192–207, 2019.

[111] Per-Ola Kristensson and Shumin Zhai. Shark2: a large vocabulary shorthand writing system for pen-based computers. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 43–52, 2004.

[112] Per Ola Kristensson and Shumin Zhai. Command strokes with and without preview: Using pen gestures on keyboard for command selection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, page 1137–1146. Association for Computing Machinery, 2007.

[113] Yuki Kubo, Buntarou Shizuki, and Jiro 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*, pages 3852–3856, 2016.

[114] Gordon Paul Kurtenbach. *The design and evaluation of marking menus.* Citeseer, 1993.

[115] Huy Viet Le, Sven Mayer, Patrick Bader, and Niels 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*, pages 1–8, 2017.

[116] Oliver Lemon and Olivier Pietquin. *Data-driven methods for adaptive spoken dialogue systems: Computational learning for conversational interfaces*. Springer Science & Business Media, 2012.

[117] Vincent Lévesque, Louise Oram, and Karon MacLean. Exploring the design space of programmable friction for scrolling interactions. In *2012 IEEE Haptics Symposium (HAPTICS)*, pages 23–30, 2012.

[118] Vincent Levesque, Louise Oram, Karon MacLean, Andy Cockburn, Nicholas D. Marchuk, Dan Johnson, J. Edward Colgate, and Michael A. Peshkin. Enhancing physicality in touch interaction with programmable friction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, page 2481–2490, New York, NY, USA, 2011. Association for Computing Machinery.

[119] Vincent Levesque, Louise Oram, Karon MacLean, Andy Cockburn, Nicholas D. Marchuk, Dan Johnson, J. Edward Colgate, and Michael A. Peshkin. Enhancing physicality in touch interaction with programmable friction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2481–2490. ACM, 2011.

[120] Teng Li, Dangxiao Wang, Cong Peng, Chun Yu, and Yuru Zhang. Speed-accuracy tradeoff of fingertip force control with visual/audio/haptic feedback. *International Journal of Human-Computer Studies*, 110:33–44, 2018.

[121] Yang Li. Gesture search: a tool for fast mobile data access. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, pages 87–96, 2010.

[122] Yang Li. Protractor: a fast and accurate gesture recognizer. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 2169–2172, 2010.

[123] Chunyuan Liao, François Guimbretière, Ken Hinckley, and Jim Hollan. Papiercraft: A gesture-based command system for interactive paper. *ACM Trans. Comput.-Hum. Interact.*, 14(4), January 2008.

[124] Yi-Chi Liao, Yen-Chiu Chen, Liwei Chan, and Bing-Yu Chen. Dwell+: Multi-level mode selection using vibrotactile cues. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, page 5–16. Association for Computing Machinery, 2017.

[125] Min Lin, Rich Goldman, Kathleen J Price, Andrew Sears, and Julie Jacko. How do people tap when walking? an empirical investigation of nomadic data entry. *International journal of human-computer studies*, 65(9):759–769, 2007.

[126] Jian Liu, Yingying Chen, Marco Gruteser, and Yan Wang. Vibsense: Sensing touches on ubiquitous surfaces through vibration. In *2017 14th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pages 1–9. IEEE, 2017.

[127] Jian Liu, Chen Wang, Yingying Chen, and Nitesh Saxena. Vibwrite: Towards finger-input authentication on ubiquitous surfaces via physical vibration. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 73–87, 2017.

[128] Xingyu Bruce Liu, Jiahao Nick Li, David Kim, Xiang'Anthony' Chen, and Ruofei Du. Human i/o: Towards a unified approach to detecting situational impairments. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–18, 2024.

[129] Mona Hosseinkhani Loorak, Wei Zhou, Ha Trinh, Jian Zhao, and Wei Li. Hand-over-face input sensing for interaction with smartphones through the built-in camera. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 1–12, 2019.

[130] Nathan Magrofuoco, Paolo Roselli, and Jean Vanderdonckt. Two-dimensional stroke gesture recognition: A survey. *ACM Computing Surveys (CSUR)*, 54(7):1–36, 2021.

[131] Päivi Majaranta, Kari-Jouko Räihä, Aulikki Hyrskykari, and Oleg Špakov. Eye movements and human-computer interaction. *Eye movement research: An introduction to its scientific foundations and applications*, pages 971–1015, 2019.

[132] David Marino, Maurício Fontana de Vargas, Antoine Weill-Duflos, and Jeremy R. Cooperstock. Conversing using whatshap: a phoneme based vibrotactile messaging platform. In *2021 IEEE World Haptics Conference (WHC)*, pages 943–948, 2021.

[133] Joe Marshall and Paul Tennent. Mobile interaction does not exist. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 2069–2078. Association for Computing Machinery, 2013.

[134] Emiliano Miluzzo, Michela Papandrea, Nicholas D Lane, Hong Lu, and Andrew T Campbell. Pocket, bag, hand, etc.-automatically detecting phone context through discovery. *Proc. PhoneSense*, 2010, 2010.

[135] Emiliano Miluzzo, Tianyu Wang, and Andrew T Campbell. Eyephone: activating mobile phones with your eyes. In *Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds*, pages 15–20, 2010.

[136] Sachi Mizobuchi, Mark Chignell, and David Newton. Mobile text entry: relationship between walking speed and text input task difficulty. In *Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*, pages 122–128, 2005.

[137] Zakriya Mohammed, Ibrahim (Abe) M Elfadel, and Mahmoud Rasras. Monolithic multi degree of freedom (mdof) capacitive mems accelerometers. *Micromachines*, 9(11):602, 2018.

[138] Andrew F Monk. *Fundamentals of human-computer interaction*. Academic Press, 2014.

[139] Thomas P. Moran, Patrick Chiu, and William van Melle. Pen-based interaction techniques for organizing material on an electronic whiteboard. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology*, UIST '97, page 45–54, New York, NY, USA, 1997. Association for Computing Machinery.

[140] Meredith Ringel Morris. Web on the wall: Insights from a multimodal interaction elicitation study. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '12, page 95–104, New York, NY, USA, 2012. Association for Computing Machinery.

[141] Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, MC Schraefel, and Jacob O Wobbrock. Reducing legacy bias in gesture elicitation studies. *interactions*, 21(3):40–45, 2014.

[142] Kishore Mukhopadhyay. Proprioception and kinesthesia: The sixth sense organ. *Advances in Health and Exercise*, 1(1):12–17, 2021.

[143] Christine Murad and Cosmin Munteanu. Designing voice interfaces: Back to the (curriculum) basics. In *Proceedings of the 2020 chi conference on human factors in computing systems*, pages 1–12, 2020.

[144] Terhi Mustonen, Maria Olkkonen, and Jukka Hakkinen. Examining mobile phone text legibility while walking. In *CHI'04 extended abstracts on Human factors in computing systems*, pages 1243–1246, 2004.

[145] C. S. Myers and L. R. Rabiner. A comparative study of several dynamic time-warping algorithms for connected-word recognition. *The Bell System Technical Journal*, 60:1389–1409, 1981.

[146] Miguel A. Nacenta, Yemliha Kamber, Yizhou Qiang, and Per Ola Kristensson. Memorability of pre-designed and user-defined gesture sets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, page 1099–1108, New York, NY, USA, 2013. Association for Computing Machinery.

[147] Hyoungsik Nam, Ki-Hyuk Seol, Junhee Lee, Hyeonseong Cho, and Sang Won Jung. Review of capacitive touchscreen technologies: Overview, research trends, and machine learning approaches. *Sensors*, 21(14):4776, 2021.

[148] Matei Negulescu, Jaime Ruiz, Yang Li, and Edward Lank. Tap, swipe, or move: Attentional demands for distracted smartphone input. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, page 173–180. Association for Computing Machinery, 2012.

[149] Matei Negulescu, Jaime Ruiz, Yang Li, and Edward Lank. Tap, swipe, or move: Attentional demands for distracted smartphone input. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '12, page 173–180, New York, NY, USA, 2012. ACM.

[150] Alexander Ng, Stephen A Brewster, and John 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*, pages 1981–1990, 2014.

[151] Alexander Ng, John Williamson, and Stephen Brewster. The effects of encumbrance and mobility on touch-based gesture interactions for mobile phones. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '15, page 536–546, New York, NY, USA, 2015. Association for Computing Machinery.

[152] Hugo Nicolau, Tiago Guerreiro, David Lucas, and Joaquim Jorge. Mobile text-entry and visual demands: reusing and optimizing current solutions. *Universal access in the information society*, 13(3):291–301, 2014.

[153] Scott D Novich and David M Eagleman. Using space and time to encode vibrotactile information: toward an estimate of the skin's achievable throughput. *Experimental brain research*, 233(10):2777–2788, 2015.

[154] Marianna Obrist, Elia Gatti, Emanuela Maggioni, Chi Thanh Vi, and Carlos Velasco. Multisensory experiences in hci. *IEEE MultiMedia*, 24(2):9–13, 2017.

[155] Halla Olafsdottir and Caroline Appert. Multi-touch gestures for discrete and continuous control. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*, pages 177–184, 2014.

[156] Elenice Oliveira, Mangai Natarajan, and Bráulio da Silva. Bus robberies in belo horizonte, brazil: solutions for safe travel. *Crime & Delinquency*, 69(11):2359–2383, 2023.

[157] C Ortiz, S Ortiz-Peregrina, JJ Castro, M Casares-López, and C Salas. Driver distraction by smartphone use (whatsapp) in different age groups. *Accident Analysis & Prevention*, 117:239–249, 2018.

[158] Ekin Ozer, Dongming Feng, and Maria Q Feng. Hybrid motion sensing and experimental modal analysis using collocated smartphone camera and accelerometers. *Measurement Science and Technology*, 28(10):105903, 2017.

[159] Toni Pakkanen. *Supporting Eyes-Free Human–Computer Interaction with Vibrotactile Haptification.* PhD thesis, Tampere University, 2020.

[160] Hari P Palani, Paul DS Fink, and Nicholas A Giudice. Design guidelines for schematizing and rendering haptically perceivable graphical elements on touchscreen devices. *International Journal of Human–Computer Interaction*, 36(15):1393–1414, 2020.

[161] Hari Prasath Palani, Jennifer L Tennison, G Bernard Giudice, and Nicholas A Giudice. Touchscreen-based haptic information access for assisting blind and visually-impaired users: Perceptual parameters and design guidelines. In *Advances in Usability, User Experience and Assistive Technology: Proceedings of the AHFE 2018 International Conferences on Usability & User Experience and Human Factors and Assistive Technology, Held on July 21–25, 2018, in Loews Sapphire Falls Resort at Universal Studios, Orlando, Florida, USA 9*, pages 837–847. Springer, 2019.

[162] HariPrasath Palani and Nicholas A Giudice. Evaluation of non-visual panning operations using touch-screen devices. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*, pages 293–294, 2014.

[163] Hariprasath Palani, Uro Giudice, and Nicholas A Giudice. Evaluation of non-visual zooming operations on touchscreen devices. In *Universal Access in Human-Computer Interaction. Interaction Techniques and Environments: 10th International Conference, UAHCI 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings, Part II 10*, pages 162–174. Springer, 2016.

[164] Jason Pascoe, Nick Ryan, and David Morse. Using while moving: Hci issues in fieldwork environments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(3):417–437, 2000.

[165] Simon T. Perrault, Eric Lecolinet, James Eagan, and Yves Guiard. Watchit: Simple gestures and eyes-free interaction for wristwatches and bracelets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, page 1451–1460, New York, NY, USA, 2013. Association for Computing Machinery.

[166] Ekaterina Peshkova, Martin Hitz, and David Ahlström. Exploring user-defined gestures and voice commands to control an unmanned aerial vehicle. In Ronald Poppe, John-Jules Meyer, Remco Veltkamp, and Mehdi Dastani, editors, *Intelligent Technologies for Interactive Entertainment*, pages 47–62, Cham, 2017. Springer International Publishing.

[167] Corey R Pittman and Joseph J LaViola Jr. Multiwave: Complex hand gesture recognition using the doppler effect. In *Proceedings of the 43rd Graphics Interface Conference*, pages 97–106, 2017.

[168] Matthew J Pitts, Gary Burnett, Lee Skrypchuk, Tom Wellings, Alex Attridge, and Mark A Williams. Visual–haptic feedback interaction in automotive touchscreens. *Displays*, 33(1):7–16, 2012.

[169] Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. User-defined gestures for augmented reality. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, page 955–960, New York, NY, USA, 2013. Association for Computing Machinery.

[170] Benjamin Poppinga, Charlotte Magnusson, Martin Pielot, and Kirsten Rassmus-Gröhn. Touchover map: audio-tactile exploration of interactive maps. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, pages 545–550, 2011.

[171] Ivan Poupyrev, Nan-Wei Gong, Shiho Fukuhara, Mustafa Emre Karagozler, Carsten Schwesig, and Karen E Robinson. Project jacquard: interactive digital textiles at scale. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 4216–4227, 2016.

[172] Huimin Qian, Ravi Kuber, and Andrew Sears. Developing tactile icons to support mobile users with situationally-induced impairments and disabilities. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–2, 2013.

[173] Roope Raisamo, Katri Salminen, Jussi Rantala, Ahmed Farooq, and Mounia Ziat. Interpersonal haptic communication: review and directions for the future. *International journal of human-computer studies*, 166:102881, 2022.

[174] Natacsha Ordones Raposo, Thais Castro, and Alberto Castro. A scheme for multimodal component recommendation. In *Universal Access in*

Human–Computer Interaction. Design and Development Approaches and Methods: 11th International Conference, UAHCI 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017, Proceedings, Part I 11, pages 412–422. Springer, 2017.

[175] Yosra Rekik, Edward Lank, Adnane Guettaf, and Laurent Grisoni. Multi-channel tactile feedback based on user finger speed. *Proceedings of the ACM on Human-Computer Interaction*, 5(ISS), nov 2021.

[176] Yosra Rekik, Radu-Daniel Vatavu, and Laurent Grisoni. Understanding users' perceived difficulty of multi-touch gesture articulation. In *Proceedings of the 16th International Conference on Multimodal Interaction*, pages 232–239, 2014.

[177] Yosra Rekik, Radu-Daniel Vatavu, and Laurent Grisoni. Understanding users' perceived difficulty of multi-touch gesture articulation. In *Proceedings of the 16th International Conference on Multimodal Interaction*, page 232–239, New York, NY, USA, 2014. ACM.

[178] Yosra Rekik, Eric Vezzoli, Laurent Grisoni, and Frédéric Giraud. Localized haptic texture: A rendering technique based on taxels for high density tactile feedback. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 5006–5015, New York, NY, USA, 2017. ACM.

[179] Bradley Rey, Kening Zhu, Simon Tangi Perrault, Sandra Bardot, Ali Neshati, and Pourang Irani. Understanding and adapting bezel-to-bezel interactions for circular smartwatches in mobile and encumbered scenarios. *Proc. ACM Hum.-Comput. Interact.*, 6(MHCI), sep 2022.

[180] Gabriel Robles-De-La-Torre. The importance of the sense of touch in virtual and real environments. *Ieee Multimedia*, 13(3):24–30, 2006.

[181] Marco Romano, Andrea Bellucci, and Ignacio Aedo. Understanding touch and motion gestures for blind people on mobile devices. In Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler, editors, *Human-Computer Interaction – INTERACT 2015*, pages 38–46, Cham, 2015. Springer International Publishing.

[182] Sami Ronkainen, Jonna Häkkilä, Saana Kaleva, Ashley Colley, and Jukka Linjama. Tap input as an embedded interaction method for mobile devices. In *Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 263–270, 2007.

[183] Juan Rosso, Céline Coutrix, Matt Jones, and Laurence Nigay. Simulating an extendable tangible slider for eyes-free one-handed interaction on mobile devices. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, pages 1–9, 2018.

[184] Anne Roudaut, Andreas Rau, Christoph Sterz, Max Plauth, Pedro Lopes, and Patrick Baudisch. Gesture output: Eyes-free output using a force feedback touch surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, page 2547–2556, New York, NY, USA, 2013. Association for Computing Machinery.

[185] Anne Roudaut, Andreas Rau, Christoph Sterz, Max Plauth, Pedro Lopes, and Patrick Baudisch. *Gesture Output: Eyes-Free Output Using a Force Feedback Touch Surface*, page 2547–2556. ACM, New York, NY, USA, 2013.

[186] Quentin Roy, Sylvain Malacria, Yves Guiard, Eric Lecolinet, and James Eagan. Augmented letters: mnemonic gesture-based shortcuts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2325–2328, 2013.

[187] Wenjie Ruan, Quan Z Sheng, Lei Yang, Tao Gu, Peipei Xu, and Longfei Shangguan. Audiogest: Enabling fine-grained hand gesture detection by decoding echo signal. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*, pages 474–485, 2016.

[188] Dean Rubine. Specifying gestures by example. In *Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '91, page 329–337, New York, NY, USA, 1991. ACM.

[189] Annie Rydström, Camilla Grane, and Peter Bengtsson. Driver behaviour during haptic and visual secondary tasks. In *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '09, page 121–127. Association for Computing Machinery, 2009.

[190] T Scott Saponas, Chris Harrison, and Hrvoje Benko. Pockettouch: through-fabric capacitive touch input. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 303–308, 2011.

[191] Zhanna Sarsenbayeva, Niels van Berkel, Eduardo Velloso, Vassilis Kostakos, and Jorge Goncalves. Effect of distinct ambient noise types on mobile interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(2):1–23, 2018.

[192] Zhanna Sarsenbayeva, Niels Van Berkel, Aku Visuri, Sirkka Rissanen, Hannu Rintamaki, Vassilis Kostakos, and Jorge Goncalves. Sensing cold-induced situational impairments in mobile interaction using battery temperature. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):1–9, 2017.

[193] Sidas Saulynas and Ravi Kuber. Understanding and supporting individuals experiencing severely constraining situational impairments. *Universal Access in the Information Society*, 19(4):919–933, 2020.

[194] Bastian Schildbach and Enrico Rukzio. Investigating selection and reading performance on a mobile phone while walking. In *Proceedings of MobileHCI '10*, page 93–102, New York, NY, USA, 2010. ACM.

[195] JJ Scott and Robert Gray. A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving. *Human factors*, 50(2):264–275, 2008.

[196] I Scott MacKenzie and Behrooz Ashtiani. Blinkwrite: efficient text entry using eye blinks. *Universal Access in the Information Society*, 10:69–80, 2011.

[197] Katie Seaborn, Norihisa P Miyake, Peter Pennefather, and Mihoko Otake-Matsuura. Voice in human–agent interaction: A survey. *ACM Computing Surveys (CSUR)*, 54(4):1–43, 2021.

[198] Andrew Sears, Min Lin, Julie Jacko, and Yan Xiao. When computers fade: Pervasive computing and situationally-induced impairments and disabilities. In *HCI international*, pages 1298–1302. Association for Computing Machinery, 2003.

[199] Marcos Serrano, Eric Lecolinet, and Yves Guiard. Bezel-tap gestures: quick activation of commands from sleep mode on tablets. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 3027–3036, 2013.

[200] Tevfik Metin Sezgin and Randall Davis. Hmm-based efficient sketch recognition. In *Proceedings of the 10th international conference on Intelligent user interfaces*, pages 281–283, 2005.

[201] Adwait Sharma, Joan Sol Roo, and Jürgen Steimle. Grasping microgestures: Eliciting single-hand microgestures for handheld objects. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–13, New York, NY, USA, 2019. Association for Computing Machinery.

[202] Alex Shaw and Lisa Anthony. Analyzing the articulation features of children's touchscreen gestures. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, pages 333–340, 2016.

[203] Lisa Skedung, Martin Arvidsson, Jun Young Chung, Christopher M Stafford, Birgitta Berglund, and Mark W Rutland. Feeling small: exploring the tactile perception limits. *Scientific reports*, 3(1):2617, 2013.

[204] Scott S. Snibbe, Karon E. MacLean, Rob Shaw, Jayne Roderick, William L. Verplank, and Mark Scheeff. Haptic techniques for media control. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology*, UIST '01, page 199–208. Association for Computing Machinery, 2001.

[205] Allan Stisen, Henrik Blunck, Sourav Bhattacharya, Thor Siiger Prentow, Mikkel Baun Kjærgaard, Anind Dey, Tobias Sonne, and Mads Møller Jensen. Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition. In *Proceedings of the 13th ACM conference on embedded networked sensor systems*, pages 127–140, 2015.

[206] Xing Su, Hanghang Tong, and Ping Ji. Activity recognition with smartphone sensors. *Tsinghua science and technology*, 19(3):235–249, 2014.

[207] Ke Sun, Wei Wang, Alex X Liu, and Haipeng Dai. Depth aware finger tapping on virtual displays. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, pages 283–295, 2018.

[208] Ke Sun, Ting Zhao, Wei Wang, and Lei Xie. Vskin: Sensing touch gestures on surfaces of mobile devices using acoustic signals. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, pages 591–605, 2018.

[209] Richard Swette, Keenan R May, Thomas M Gable, and Bruce N Walker. Comparing three novel multimodal touch interfaces for infotainment menus. In *Proceedings of the 5th International Conference on automotive user interfaces and interactive vehicular applications*, pages 100–107, 2013.

[210] Hong Z Tan, Charlotte M Reed, Yang Jiao, Zachary D Perez, E Courtenay Wilson, Jaehong Jung, Juan S Martinez, and Frederico M Severgnini. Acquisition of 500 english words through a tactile phonemic sleeve (taps). *IEEE Transactions on Haptics*, 13(4):745–760, 2020.

[211] Eugene M Taranta and Joseph J LaViola Jr. Penny pincher: a blazing fast, highly accurate $-family recognizer. In *Proceedings of the 41st Graphics Interface Conference*, pages 195–202, 2015.

[212] Jennifer L Tennison and Jenna L Gorlewicz. Toward non-visual graphics representations on vibratory touchscreens: Shape exploration and identification. In *Haptics: Perception, Devices, Control, and Applications: 10th International Conference, EuroHaptics 2016, London, UK, July 4-7, 2016, Proceedings, Part II 10*, pages 384–395. Springer, 2016.

[213] Jennifer L Tennison and Jenna L Gorlewicz. Non-visual perception of lines on a multimodal touchscreen tablet. *ACM Transactions on Applied Perception (TAP)*, 16(1):1–19, 2019.

[214] Jennifer L Tennison, P Merlin Uesbeck, Nicholas A Giudice, Andreas Stefik, Derrick W Smith, and Jenna L Gorlewicz. Establishing vibration-based tactile line profiles for use in multimodal graphics. *ACM Transactions on Applied Perception (TAP)*, 17(2):1–14, 2020.

[215] Bruce Thomas, Karen Grimmer, Joanne Zucco, and Steve Milanese. Where does the mouse go? an investigation into the placement of a body-attached touchpad mouse for wearable computers. *Personal and Ubiquitous computing*, 6:97–112, 2002.

[216] Garreth W Tigwell, David R Flatla, and Rachel Menzies. It's not just the light: understanding the factors causing situational visual impairments during mobile interaction. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*, pages 338–351, 2018.

[217] Garreth W Tigwell, Zhanna Sarsenbayeva, Benjamin M Gorman, David R Flatla, Jorge Goncalves, Yeliz Yesilada, and Jacob O Wobbrock. Addressing the challenges of situationally-induced impairments and disabilities in

mobile interaction. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–8, 2019.

[218] Hussain Tinwala and I. Scott MacKenzie. Eyes-free text entry on a touchscreen phone. In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*, pages 83–88, 2009.

[219] Giovanni Maria Troiano, Esben Warming Pedersen, and Kasper Hornbæk. User-defined gestures for elastic, deformable displays. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*, AVI '14, page 1–8, New York, NY, USA, 2014. Association for Computing Machinery.

[220] Hsin-Ruey Tsai, Te-Yen Wu, Da-Yuan Huang, Min-Chieh Hsiu, Jui-Chun Hsiao, Yi-Ping Hung, Mike Y Chen, and Bing-Yu 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*, pages 2164–2171, 2017.

[221] Theophanis Tsandilas. Fallacies of agreement: A critical review of consensus assessment methods for gesture elicitation. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 25(3):1–49, 2018.

[222] Huawei Tu, Xiangshi Ren, and Shumin Zhai. A comparative evaluation of finger and pen stroke gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1287–1296, 2012.

[223] Ying-Chao Tung, Chun-Yen Hsu, Han-Yu Wang, Silvia Chyou, Jhe-Wei Lin, Pei-Jung Wu, Andries Valstar, and Mike 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*, CHI '15, page 3327–3336, New York, NY, USA, 2015. Association for Computing Machinery.

[224] Matthew Turk. Multimodal interaction: A review. *Pattern recognition letters*, 36:189–195, 2014.

[225] Lawrence Van den Bogaert and David Geerts. User-defined mid-air haptic sensations for interacting with an ar menu environment. In Ilana Nisky, Jess Hartcher-O'Brien, Michaël Wiertlewski, and Jeroen Smeets, editors, *Haptics: Science, Technology, Applications*, pages 25–32, Cham, 2020. Springer International Publishing.

[226] Radu-Daniel Vatavu. Smart-pockets: Body-deictic gestures for fast access to personal data during ambient interactions. *International Journal of Human-Computer Studies*, 103:1–21, 2017.

[227] Radu-Daniel Vatavu. Visual impairments and mobile touchscreen interaction: state-of-the-art, causes of visual impairment, and design guidelines. *International Journal of Human–Computer Interaction*, 33(6):486–509, 2017.

[228] Radu-Daniel Vatavu. From natural to non-natural interaction: Embracing interaction design beyond the accepted convention of natural. In *Proceedings of the 25th International Conference on Multimodal Interaction*, pages 684–688, 2023.

[229] Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. Gestures as point clouds: A $p recognizer for user interface prototypes. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, page 273–280. Association for Computing Machinery, 2012.

[230] Radu-Daniel Vatavu, Lisa Anthony, and Jacob O Wobbrock. Gesture heatmaps: Understanding gesture performance with colorful visualizations. In *Proceedings of the 16th International Conference on Multimodal Interaction*, pages 172–179, 2014.

[231] Radu-Daniel Vatavu and Ovidiu-Ciprian Ungurean. Stroke-gesture input for people with motor impairments: Empirical results & research roadmap. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2019.

[232] Radu-Daniel Vatavu, Daniel Vogel, Géry Casiez, and Laurent Grisoni. Estimating the perceived difficulty of pen gestures. In Pedro Campos, Nicholas Graham, Joaquim Jorge, Nuno Nunes, Philippe Palanque, and Marco Winckler, editors, *Human-Computer Interaction – INTERACT 2011*, pages 89–106, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[233] Radu-Daniel Vatavu and Jacob O. Wobbrock. Formalizing agreement analysis for elicitation studies: New measures, significance test, and toolkit. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, page 1325–1334, New York, NY, USA, 2015. Association for Computing Machinery.

[234] Radu-Daniel Vatavu and Jacob O Wobbrock. Clarifying agreement calculations and analysis for end-user elicitation studies. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 29(1):1–70, 2022.

[235] Eric Vezzoli, Thomas Sednaoui, Michel Amberg, Frédéric Giraud, and Betty Lemaire-Semail. Texture rendering strategies with a high fidelity - capacitive visual-haptic friction control device. In *Proceedings, Part I, of the 10th International Conference on Haptics: Perception, Devices, Control, and Applications - Volume 9774*, EuroHaptics 2016, pages 251–260, Berlin, Heidelberg, 2016. Springer-Verlag.

[236] Yon Visell. Tactile sensory substitution: Models for enaction in hci. *Interacting with Computers*, 21(1-2):38–53, 2009.

[237] P. Viviani and C. Terzuolo. 32 space-time invariance in learned motor skills. In George E. Stelmach and Jean Requin, editors, *Tutorials in Motor Behavior*, volume 1 of *Advances in Psychology*, pages 525–533. North-Holland, 1980.

[238] Qianhui Wei, Min Li, Jun Hu, and Loe Feijs. Creating mediated touch gestures with vibrotactile stimuli for smart phones. In *Proceedings of the*

Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction, TEI '20, page 519–526, New York, NY, USA, 2020. Association for Computing Machinery.

[239] Daniel Wigdor and Ravin Balakrishnan. Tilttext: using tilt for text input to mobile phones. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 81–90, 2003.

[240] Jacob O Wobbrock. The future of mobile device research in hci. In *CHI 2006 workshop proceedings: what is the next generation of human-computer interaction*, pages 131–134, 2006.

[241] Jacob O Wobbrock, Krzysztof Z Gajos, Shaun K Kane, and Gregg C Vanderheiden. Ability-based design. *Communications of the ACM*, 61(6):62–71, 2018.

[242] Jacob O Wobbrock, Meredith Ringel Morris, and Andrew D Wilson. User-defined gestures for surface computing. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1083–1092, 2009.

[243] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, page 1083–1092, New York, NY, USA, 2009. Association for Computing Machinery.

[244] Jacob O. Wobbrock, Andrew D. Wilson, and Yang Li. Gestures without libraries, toolkits or training: A $1 recognizer for user interface prototypes. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*, UIST '07, page 159–168, New York, NY, USA, 2007. ACM.

[245] Jacob O Wobbrock, Andrew D Wilson, and Yang Li. Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, pages 159–168, 2007.

[246] Pui Chung Wong, Hongbo Fu, and Kening Zhu. Back-mirror: Back-of-device one-handed interaction on smartphones. In *SIGGRAPH ASIA 2016 Mobile Graphics and Interactive Applications*, pages 1–5. Association for Computing Machinery, 2016.

[247] Pui Chung Wong, Kening Zhu, Xing-Dong Yang, and Hongbo Fu. Exploring eyes-free bezel-initiated swipe on round smartwatches. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–11, 2020.

[248] Mike Wu, Chia Shen, Kathy Ryall, Clifton Forlines, and Ravin Balakrishnan. Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'06)*, pages 8–pp. IEEE, 2006.

[249] Te-Yen Wu, Zheer Xu, Xing-Dong Yang, Steve Hodges, and Teddy Seyed. Project tasca: Enabling touch and contextual interactions with a pocket-based textile sensor. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2021.

[250] Savita Yadav and Pinaki Chakraborty. Human-computer interaction as an important aspect of software: A tutorial. In *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, pages 40–44. IEEE, 2020.

[251] Yukang Yan, Chun Yu, Yingtian Shi, and Minxing Xie. Privatetalk: Activating voice input with hand-on-mouth gesture detected by bluetooth earphones. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pages 1013–1020, 2019.

[252] Hui-Shyong Yeo, Xiao-Shen Phang, Steven J Castellucci, Per Ola Kristensson, and Aaron Quigley. Investigating tilt-based gesture keyboard entry for single-handed text entry on large devices. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 4194–4202, 2017.

[253] Shumin Zhai, Per Ola Kristensson, Caroline Appert, Tue Haste Anderson, Xiang Cao, et al. Foundational issues in touch-surface stroke gesture design—an integrative review. *Foundations and Trends® in Human–Computer Interaction*, 5(2):97–205, 2012.

[254] Cheng Zhang, Anhong Guo, Dingtian Zhang, Yang Li, Caleb Southern, Rosa I Arriaga, and Gregory D Abowd. Beyond the touchscreen: an exploration of extending interactions on commodity smartphones. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 6(2):1–23, 2016.

[255] Maotian Zhang, Panlong Yang, Chang Tian, Lei Shi, Shaojie Tang, and Fu Xiao. Soundwrite: Text input on surfaces through mobile acoustic sensing. In *Proceedings of the 1st International Workshop on Experiences with the Design and Implementation of Smart Objects*, pages 13–17, 2015.

[256] Ruidong Zhang, Mingyang Chen, Benjamin Steeper, Yaxuan Li, Zihan Yan, Yizhuo Chen, Songyun Tao, Tuochao Chen, Hyunchul Lim, and Cheng Zhang. Speechin: A smart necklace for silent speech recognition. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 5(4), dec 2022.

[257] Yang Zhang and Chris Harrison. Quantifying the targeting performance benefit of electrostatic haptic feedback on touchscreens. In *Proceedings of the 2015 International Conference on Interactive Tabletops and Surfaces*, ITS '15, page 43–46. Association for Computing Machinery, 2015.

[258] Jian Zhao, Ricardo Jota, Daniel Wigdor, and Ravin Balakrishnan. Augmenting mobile phone interaction with face-engaged gestures. *arXiv preprint arXiv:1610.00214*, 2016.

[259] Siyan Zhao, Ali Israr, and Roberta Klatzky. Intermanual apparent tactile motion on handheld tablets. In *2015 IEEE World Haptics Conference (WHC)*, pages 241–247, 2015.

[260] Siyan Zhao, Ali Israr, Frances Lau, and Freddy Abnousi. Coding tactile symbols for phonemic communication. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2018.