Universität Stuttgart

Finger Orientation as an Additional Input Dimension for Touchscreens

Von der Fakultät für Informatik, Elektrotechnik und Informationstechnik
und dem Stuttgart Research Centre for Simulation Technology
der Universität Stuttgart zur Erlangung der Würde eines Doktors der
Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von
Sven André Mayer
aus Tübingen

Hauptberichter:	Prof. Dr. Niels Henze
Mitberichter:	Prof. Roderick Murray-Smith
Mitprüfer:	Prof. Dr. Jonas Kuhn

Tag der mündlichen Prüfung: 01.03.2019

Institut für Visualisierung und Interaktive Systeme
der Universität Stuttgart

2019

für Barbara und Harald

Zusammenfassung

Seit der Erfindung des ersten digitalen Computers im Jahr 1941 und später des ersten Personal Computer im Jahr 1975 hat sich die Art und Weise, wie Menschen mit Computern interagieren, radikal verändert. Die Tastatur ist zwar immer noch eines der beiden Hauptgeräte für Desktopcomputer, welche heute meistens mit einer Maus oder einem Trackpad kombiniert wird. Die Interaktion mit Desktop- und Laptopcomputer macht jedoch heute nur einen kleinen Prozentsatz der Interaktion zwischen Menschen und Computer aus. Heute interagieren Menschen meistens mit ubiquitären Computern. Während die ersten ubiquitären Computer ebenfalls über Tasten gesteuert wurden, änderte sich dies mit der Einführung des Touchscreens in Endverbrauchergeräte. Heute ist das Smartphone der verbreitetste ubiquitäre Computer, welcher stark auf die Interaktion durch direkte Berührung als die dominierende Eingabemethode setzt. Durch direkte Berührung können Benutzer mit der grafischen Benutzeroberfläche (GUI) interagieren. GUI-Steuerelemente können direkt durch einfaches Berühren gesteuert werde. Gegenwärtige ubiquitäre Computer mit Berührungsinteraktion reduzieren jedoch den Detailreichtum der Berührung auf eine zweidimensionale Position auf dem Bildschirm. In dieser Arbeit wird das Potenzial untersucht, wie eine einfache Berührung mit zusätzlichen Informationen über den Finger, der den Bildschirm berührt, bereichert werden kann. Im Detail, wird in dieser Arbeit die Verwendung der Fingerorientierung, als zwei zusätzliche Eingabe Dimensionen untersucht.

Es werden vier Schlüsselbereiche untersucht, welche die Grundlage bilden, um die Fingerorientierung als zusätzliche Eingabetechnik vollständig zu verstehen. Diese Erkenntnisse bieten Designern die Grundlage, neue Gesten und Anwendungsfälle zu entwerfen und dabei die Fingerorientierung zu berücksichtigen. Es werden zunächst Ansätze untersucht um die Eingabe der Fingerorientierung zu erkennen und Modelle bereitgestellt um die Orientierung des Fingers zu ermitteln. Zweitens werden Designrichtlinien vorgestellt, um eine bequeme Verwendung der Fingerorientierung sicherzustellen. Drittens wird eine Methode präsentiert, welche neue Interaktionstechniken in sozialen Umgebungen analysieren kann. So wird es ermöglicht, neue Anwendungsfälle auf eine mögliche Beeinträchtigung von Konversationen zu untersuchen. Zum Schluss werden drei Möglichkeiten vorgestellt, wie neue Interaktionstechniken wie zum Beispiel die Fingerorientierungseingaben dem Benutzer mitgeteilt werden können. Diese Arbeit trägt dazu bei, um die Fingerorientierung vollständig zu verstehen. Am Ende der Arbeit werden diese vier Schlüsselbereiche zusammengefasst um eine Basis zu schaffen in der Zukunft weitere Interaktionstechniken auf dieselbe Art und Weise zu untersuchen und bewerten.

Abstract

Since the first digital computer in 1941 and the first personal computer back in 1975, the way we interact with computers has radically changed. The keyboard is still one of the two main input devices for desktop computers which is accompanied most of the time by a mouse or trackpad. However, the interaction with desktop and laptop computers today only make up a small percentage of current interaction with computing devices. Today, we mostly interact with ubiquitous computing devices, and while the first ubiquitous devices were controlled via buttons, this changed with the invention of touchscreens. Moreover, the phone as the most prominent ubiquitous computing device is heavily relying on touch interaction as the dominant input mode. Through direct touch, users can directly interact with graphical user interfaces (GUIs). GUI controls can directly be manipulated by simply touching them. However, current touch devices reduce the richness of touch input to two-dimensional positions on the screen. In this thesis, we investigate the potential of enriching a simple touch with additional information about the finger touching the screen. We propose to use the user's finger orientation as two additional input dimensions. We investigate four key areas which make up the foundation to fully understand finger orientation as an additional input technique. With these insights, we provide designers with the foundation to design new gestures sets and use cases which take the finger orientation into account. We first investigate approaches to recognize finger orientation

input and provide ready-to-deploy models to recognize the orientation. Second, we present design guidelines for a comfortable use of finger orientation. Third, we present a method to analyze applications in social settings to design use cases with possible conversation disruption in mind. Lastly, we present three ways how new interaction techniques like finger orientation input can be communicated to the user. This thesis contributes these four key insights to fully understand finger orientation as an additional input technique. Moreover, we combine the key insights to lay the foundation to evaluate every new interaction technique based on the same in-depth evaluation.

Acknowledgements

Foremost, I would like to thank **Niels Henze**, for all his trust in me already back in 2014 when he offered me to become his first Ph.D. student. Since then we had endless discussion which shaped this thesis but also numerous other projects and even more importantly my character. I would like to also thank **Roderick Murry-Smith**, first of all for being on my Ph.D. committee, but also for hosting me as a Ph.D. intern in the summer of 2015. His background in computational interaction mainly shaped my path in the past years. I would like to thank my Ph.D. committee members **Jonas Kuhn** and **Stefan Funke** and my SimTech supervisor **Andrea Barth** for their valuable feedback and the valuable discussion.

Beyond my official supervisor and committee I got extensive support from **Albrecht Schmidt**. I thank him for his trust in my research projects and support to conduct all my studies. Additionally, in the last months, **Andreas Bulling** supported me to further pursue more research, I also want to thank him for this opportunity.

I thank **Lars Lischke** with whom I started my academic career back in 2008. In the 11 years we had endless discussions over research but also life in general, in the years he became a true friend who always supported me. In the first year of my Ph.D. I learned a great deal of how to conduct amazing research from **Markus Funk**. He deserves a special thanks for his support and his inspiration in all those

years. I want to thank **Huy Viet Le** for his inspiration and help in conducting the research present in this thesis, he had a major impact in the publication leading up to this thesis.

Beyond the core team who supported my I want to thank a number of people for their personal support **Stefan Schneegaß** for supporting me already back in the time before I started my Ph.D., he showed my how exciting academic research can and motivated me to start my Ph.D. A special thanks goes to **Alireza Sahami Shirazi** for teaching me a great deal about quantitative research and for being my first office mate; to **Paweł W. Woźniak** for teaching me so much about quantitative research; and to **Katrin Wolf** for her valuable feedback on so many of my projects. A huge thanks to my fellow Ph.D. brothers and sisters **Alexandra Voit**, **Dominik Weber**, **Rufat Rzayev**, and **Valentin Schwind** for always supporting me. A less work related thank you goes to **Francisco Kiss**, and **Tilman Dingler** for always finding the time to have a Whisky with me after a long day. Additionally, also a huge thank you to **Pascal Knierim** for his support but more importantly to for our hikes together, especially for hiking to Hallett Peak at 3.875m in Rocky Mountain National Park.

I believe that without the team of amazing people in Stuttgart I would not have had such a fantastic time in Stuttgart. Therefore, I want to say a huge thanks to all of you: **Anja Mebus**, **Bastian Pfleging**, **Céline Coutrix**, **Hyunyoung Kim**, **Jakob Karolus**, **Mariam Hassib**, **Matthias Hoppe**, **Mauro Avila**, **Miriam Greis**, **Murielle Naud-Barthelmeß**, **Norman Pohl**, **Passant El.Agroudy**, **Patrick Bader**, **Romina Poguntke**, **Thomas Kosch**, **Thomas Kubitza**, **Tonja Machulla**, and **Yomna Abdelrahman**. This is not only true for the Stuttgart team but also for all the collaboration. A huge thanks to all co-authors and collators, and especially to: **Ashley Colly**, **Giulio Jacucci**, **Lewis Chuang**, **Morten Fjeld**, **Nitesh Goyal**, and **Scott Greenwald**.

A special thanks goes to the student assistants who supported me mostly with conducting and implementing various experiments, namely: **Eric Hämmerle**, **Henrike Weingärtner**, **Jamie Ullerich**, **Johannes Wolf**, **Jonas Vogelsang**, **Max Weiß**, **Robin Schweigert**, and **Stephan Roth**. Beyond direct support by students, I hosted a number of interns, thanks to: **Alexandre Comeau-Vermeersch**, **Anna Wong**, **Brighten Jelke**, **El Hassan Makled**, **Jens Reinhard**,

Karen Lienemann, **Michael Mayer**, **Omar ElAzazy**, and **Perihan Gad**, not only for your support but also for your trust in me to be a great mentor. Finally, I supervised and co-supervised a great deal of Master and Bachelor students in their thesis, as well as practical projects, to these over 40 students, goes a big thank you, I have only excellent memories when looking back on the endless hours I spent with all of you.

Saving one last group of many individuals from the academic world which you would probably not know but all benefit from. I am talking about the countless Student Volunteers which whom I worked together to either run a conference or helped to run a conference. Thank you all so much it was an honor to work with all of you.

Beyond the amazing support form all my collaborators, colleagues, and friends in the academic world, I would like to also thank all friends outside of the academic world. Thanks to all of you for your patience when I had no time in the last 4.5 years, thanks for your understanding me, and thanks for supporting me on my way in pursuing this Ph.D. thesis.

Saving the group of people for last without whom nothing of this ever happened. Therefore, I owe my deepest gratitude to my family. Special thanks to both my sisters, **Anthea Mayer** and **Carlotta Mayer** for their love and support at all time, both in bad and good times. And finally, I would like to tank my parents, **Barbara Mayer** and **Harald Mayer**, for their trust, support, and love without I would have never become the person I am today. Their encouragement, understanding, security, and given freedom made it possible for me to present this Ph.D. thesis.

Thank you all.

Table of Contents

1 Introduction **17**
1.1 Definition of Finger Orientation 19
1.2 Research Questions 21
1.3 Research Context 22
1.4 Thesis Outline 23

2 Related Work **27**
2.1 Recognition 28
2.2 Ergonomic Constraints 30
2.3 Social Implications 31
2.3.1 Collocated Interactions 31
2.3.2 Evaluating In-Conversation Interactions 32
2.3.3 Disruptiveness in CSCW 34
2.4 Discoverability 35
2.5 Summary 36

3 Recognition **39**
3.1 Challenges in Recognizing Finger Orientation 40
3.2 Depth Camera Approach 40
3.2.1 Prototype 41
3.2.2 Study 41

3.2.3 Results 44
3.2.4 Discussion and Implications 48
3.3 Capacitive Sensor Approach 49
3.3.1 Study 50
3.3.2 Modeling 53
3.3.3 Results 60
3.3.4 Discussion 62
3.3.5 The Finger Orientation Data Set and Model 64
3.4 Summary 65

4 Ergonomic Constraints **67**
4.1 Ergonomic Challenges for Finger Orientation 68
4.2 Static Scenario 69
4.2.1 Hypotheses 69
4.2.2 Study 70
4.2.3 Results 74
4.2.4 Discussion 78
4.2.5 Design Considerations 79
4.2.6 Limitations 81
4.3 Two-Handed Interaction Scenario 82
4.3.1 Hypotheses 83
4.3.2 Study 83
4.3.3 Results 89
4.3.4 Discussion 95
4.4 Summary 97

5 Social Implications **99**
5.1 Evaluating the Disruptiveness of Mobile Interactions 100
5.2 Research Design 101
5.2.1 Requirements for a new approach 101
5.2.2 Blending existing approaches 102
5.2.3 Choice of Participants and Stimulus 103
5.2.4 Study Plan 104

5.3 Study 1: SurfaceSliding 105
5.3.1 Study 106
5.3.2 Results 111
5.3.3 Discussion 116
5.4 Study 2: Finger Orientation 120
5.4.1 Study 120
5.4.2 Results 122
5.4.3 Discussion 123
5.5 Advantages of the Mixed-Method Approach 125
5.5.1 Rapid Answers 125
5.5.2 Focused, on-the-spot Feedback 125
5.5.3 Designing with the Social Context in Mind 126
5.5.4 Potential for a Generative Role 126
5.6 Limitations of the Mixed-Method Approach 127
5.7 Summary 127

6 Discoverability **129**
6.1 State-of-the-Art Communication Approaches 130
6.2 Alternative Interaction Techniques 131
6.2.1 Finger Roll Interaction 131
6.2.2 Nail/Knuckle Interaction 132
6.2.3 Finger-Aware Interaction 132
6.3 Design Sessions 133
6.3.1 Procedure 133
6.3.2 Results 135
6.3.3 Discussion 140
6.4 Evaluation 141
6.4.1 Study Design 141
6.4.2 Apparatus 142
6.4.3 Procedure 144
6.4.4 Participants 144
6.5 Results 145
6.5.1 System Usability Scale (SUS) 145
6.5.2 AttrakDiff 146

6.5.3 Qualitative Results 148
6.6 Discussion 149
6.7 Design Implications 152
6.8 Summary 152

7 Conclusion **155**
7.1 Future Work 157

Bibliography **159**

List of Figures **187**

List of Tables **193**

List of Acronyms **196**

1

Introduction

Since the invention of the first digital computers, not only computing power, and memory had an exponential growth but also the number of systems with computing power. The first computer, the Zuse Z3 (1941) and its successor, had a different form than what we today envision as a computer. The Zuse Z3 was designed to serve multiple people. Later (1975) with the first true "personal computer," the Altair 8800 created by Micro Instrumentation and Telemetry Systems, the scenario shifted towards a one-to-one mapping. Here, the concept emerged that all humans have their own computing devices that serve the needs of one person. However, today we see a new mapping, computers are everywhere, often referred to as ubiquitous devices, and can be accessed by one or many humans at the same time. Thus, the new mapping is where one human can use multiple computing devices at once. This was envisioned by Mark Weiser [177] almost three decades ago. He envisioned tabs, pads, and boards which always surround us. Today, we see them as ubiquitous devices to fulfill one task or even to fulfill multiple tasks at the same time. With the increasing number of devices which can serve humans at any given time, the interaction between humans and computers becomes essential as the exchange of information between them is one major step in towards solving a problem.

In the past, we saw that humans make use of one or more computing devices to solve a problem. However, Licklider [101] and recent trends suggest that when humans and computing devices work together to solve a problem in symbiosis, they outperform both as individuals. Thus, the interaction between humans and computers is today a bidirectional interplay. Computing devices can alert the human, and thereby get their attention to inform new information, problems, and solutions. On the other hand, humans can provide the computer with information, problems, and solutions. Moreover, as we see an ever-increasing number of devices which surrounds us one of the limiting factors to utilize them is how we interact with these devices. While speech, and gesture interfaces have been developed over the last decade, touch interfaces are the dominant input for computing devices. New generations of laptops and stationary large high-resolution system incorporate direct touch interaction of laptops and stationary system incorporate direct touch interaction. While these devices start to incorporate touch their overall market share is small. On the other hand, in 2017, 5 billion mobile phones were in use and 66% of the world's population used one[1]. Thus, over the last decade, smartphones have not only become the primary device for mobile interaction but also serve as the primary computing device for many users. Therefore, the phone is the most ubiquitous computing devices and here we see that all phone companies are heavily relying on touch interaction as the dominant input mode.

Touch interaction is the most common mode of interaction with ubiquitous computers; however, today's implementation of touch interaction is mostly limited to a 2D point approximated through the fingers position on the touch surface. This 2D point neglects the diversity with which we use our hands in communication and to control physical tools. Through direct touch, users can intuitively interact with the graphical user interface (GUI). GUI controls can directly be selected by touching them. Here, the touchscreen relates merely the 2D touch point retrieved from the fingers position to the GUI control. However, the finger touching the surface delivers a wide range of information which could potentially be used to enrich the expressiveness of touch interaction, such as dwell time, pressure, or finger-aware interaction. Moreover, Holz and Baudisch further found that there is an offset between the point where the user assumes to have touched the

[1] `http://www.gsma.com/mobileeconomy/`

surface and their actual finger position [69, 70]. Holz and Baudisch concluded that touch is not a 2-dimensional interaction technique, but a 6-dimensional one [70], involving the finger position, orientation, and pressure. They showed that direct touch needs to be described by the 3D finger orientation relative to the touch surface. While touch can be improved using more dimensions, we argue that the limited expressiveness of a 2D touch point, as implemented today, can be enriched by used the finger orientation as additional input dimension.

To address the limited expressiveness of 2D touch interaction, commercial devices, as well as previous research, presented a wide range of on-screen interaction techniques to enhance the 2D touch point derived by the touching finger. Already in the first version of Android and iOS, they leveraged the time dimension (dwell time) to provide the long press. With the iPhone 6s in 2015, Apple introduced *3D Touch* which adds a pressure dimension to the interaction. In the current implementation, both methods are used to modify the touch input and alter the action. Researchers have investigated a wide range of touch input extensions: finger-aware interaction [21], finger shape [139], finger size [10], part of the hand [58], palm input [100], pressure [146], and finger roll [151]. Moreover, researchers have focused on how to utilize the finger orientation as an additional input dimension to enhance the expressiveness of a 2D touch point.

While we see a large body of work addressing the problem of limited expressiveness of 2D touch interaction, presented solutions have all been studied to tackle one of the many issues ranging from the recognition to the design of new applications. However, not the full impact of each of the new methods as a whole. Thus, in this thesis, we seek to understand one enhancement method in detail. In the following we focus on finger orientation as a new interaction method to enhance the expressiveness of 2D touch interaction.

1.1 Definition of Finger Orientation

The finger orientation can be defined by three rotations with respect to a defined coordinate system. We will use Tait–Bryan angles through the thesis. Moreover, we will use the names: pitch, roll, and yaw to describe the rotation around certain

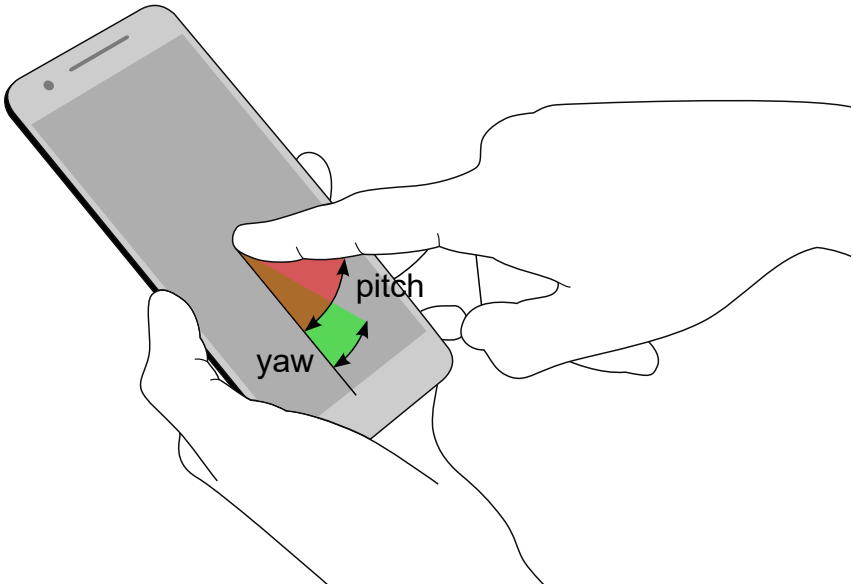

Figure 1.1: Examples of finger orientations in a mobile setup. Finger orientation input with pitch and yaw input can enlarge the input space for touch surfaces.

axes which roots in the naming for the aircraft rotations[1]. The yaw axis is defined to be perpendicular to the sides of the finger with its origin at the tip of the finger and directed towards the lower part of the finger, see Figure 1.1. The pitch axis is perpendicular to the yaw axis and is parallel to the sides of the finger with the origin at the tip of the finger and directed to the right side of the finger, see Figure 1.1. The roll axis is perpendicular to the other two axes with its origin at the tip of the finger, and is directed away from the root of the finger. We define yaw as 0° when the finger is parallel to the long edge of the touch surface (when in horizontal mode) and increases when the finger is rotated counterclockwise. We define pitch as 0° when the finger is parallel to the touch surface, i.e., the entire finger touches the surface. Finally, roll is 0° when the finger pitch axis is parallel to the touch surface and increases when the finger is rotated clockwise.

[1] `https://www.grc.nasa.gov/WWW/K-12/airplane/rotations.html`

Previous work in human-computer interaction (HCI) on the use of the finger orientation is as an additional input the pitch and yaw axis for interaction, cf. Xiao et al. [188] and Kratz et al. [83]. Moreover, research in HCI commonly uses the roll dimension as an separated input dimension known as roll interaction which is studied from pitch and yaw due to the raising complexity of the input, cf. Roudaut et al. [151].

In this thesis, we will use the mathematical definition for orientation as presented above. However, when referring to "finger orientation", we only use pitch and yaw as two additional input dimensions as additional input dimensions.

1.2 Research Questions

We identified four key areas which need to be addressed to fully understand the impact of a new interaction technique. We address these four key areas in this thesis concerning finger orientation as an additional input dimension to enhance the expressiveness of a 2D touch point. Therefore, in the following, we outline our four research questions upon which this thesis is built on and structured by, Table 1.1 presents an overview of all research questions. First, we identified that the recognition of an interaction method needs to be tackled. RQ1 therefore directly addressed the recognition problem. While RQ1 has been addressed in the past by, e.g., Xiao et al. [188] we see the need to improve the pitch and yaw recognition further. Second, to enable interface support for the new interaction technique ergonomic constraints needs to be identified. With RQ2 we raise the question to what extent can a human perform different finger orientation inputs. In detail, which are the comfortable orientations that humans can perform with their finger? Third, before releasing a new interaction technique, it is of the need to understand the social implications; thus, an evaluation of how the new interaction technique would affect society. RQ3 addresses the social component of our in-depth analysis of finger orientation, with this research question we aim to understand how finger orientation affects face-to-face conversations. Finally, with today's advanced interaction technique we see that they get implemented, but the way in which they get communicated to the user is unclear; we see this last area as an issue of discoverability. RQ4 is devoted to addressing the issue of

Table 1.1: Summary of the research questions in this thesis.

Topic	No.	Research Question	Chapter
Recognition	RQ1	How can finger orientation be detected?	3
Ergonomics	RQ2	What are the ergonomic constraints of using finger orientation as an additional input?	4
Social	RQ3	Which social implications has using finger orientation as an additional input?	5
Discoverability	RQ4	How should finger orientation input be communicated to the user?	6

discoverability for finger orientation in line with guidelines by both Shneiderman et al. [160] and Norman [138], as to essential use case it is important to understand how developers and researchers can communicate finger orientation to the human.

1.3 Research Context

The research that leads to this thesis was conducted at the University of Stuttgart in the years between 2014 and 2018 in the Socio-Cognitive Systems group.

Cluster of Excellence in Simulation Technology The main part of this thesis was conducted with the scope of the Cluster of Excellence in Simulation Technology at the University of Stuttgart as part of the project network "Reflexion and Contextualization." Within this context, this thesis was undergoing a mid-term evaluation examined by Prof. Dr. Niels Henze and Prof. Dr. Andrea Barth from the Institute for Computational Methods for Uncertainty Quantification.

Socio-Cognitive Systems Group, University of Stuttgart This thesis is built on the following six publications: [117, 118, 121, 124, 126, 132]. Particularly successful was the collaboration with Katrin Wolf and Paweł W. Woźniak as we

received an Honorable Mentions Award at the MobileHCI'17 conference for our first paper on how ergonomic constraints affect the feasibility of finger orientation as additional input [132].

The papers presented in this thesis, are inspired by collaborations with Huy Viet Le [88, 90, 91, 93–97, 100]. Moreover, we held a number of tutorials on machine learning for HCI research [92, 99, 119]. Furthermore, we published a number of papers in the domain of mid-air pointing [127, 129, 158]. Additional collaboration with research form the same team let to the following publications: [61, 62, 120, 123, 125, 170, 175].

Human-Computer Interaction Group, University of Stuttgart The six papers which lead to this thesis are mainly inspired by collaborations with Lars Lischke [102–107, 114], and Markus Funk [35–40]. Additional collaboration led to the following publications: [82, 131, 150, 156, 157, 181, 183, 187].

External Collaborations Work with collaborators around the world led to a set of publications: Jens Emil Grønbæk (Aarhus University, Denmark), Zhanna Sarsenbayeva (The University of Melbourne, Australia), Giulio Jacucci (University of Helsinki, Finland) [128], Markus Gärtner and Jonas Kuhn from the Institute for Natural Language Processing at the University of Stuttgart [51], Anna Kötteritzsch, Julian Fietkau and Michael Koch from the Universität der Bundeswehr München and Benjamin Weyers (RWTH Aachen University, Germany) [86], Morten Fjeld (Chalmers University of Technology, Sweden), Nitesh Goyal (Cornell University, US), Przemysław Kucharski (Lodz University of Technology, Poland) [185], Lewis L. Chuang, Alessandro Nesti and Heinrich H. Bülthoff, (Max Planck Institute for Biological Cybernetics, Tübingen, Germany) [130], and Ashley Colley (University of Lapland, Finland) [22].

1.4 Thesis Outline

This dissertation comprises seven chapters, the bibliography, and the enumerating lists. This work contains the results and evaluations of a total of eight empirical

studies conducted between summer 2014 and summer 2018 as well as an additional literature review in the related work chapter. This work is structured as follows:

Chapter 1: Introduction contains the description and motivation of this thesis, an overview of the research questions, and this outline.

Chapter 2: Related Work comprises the related work to put finger orientation in perspective with other on and off screen interaction techniques. Moreover, this Chapter also contains the specific related work to connect the four research questions with its specific related work.

Chapter 3: Recognition contains two approaches to extend existing recognition approaches for finger orientation detection. We first show how we can use static offset correction models to improve a depth camera recognition approach and moreover show that this approach not only works for the index finger but also for the middle, ring, and small finger. Second, we use Deep Neural Network (DNN) to improve the recognition when using capacitive images. This chapter addresses RQ1.

Chapter 4: Ergonomic Constraints investigates the feasibility of using finger orientation as an additional input. Here, we first investigate the input range for table tabletops in a static setup. Second, we extend this to a truly mobile scenario to understand how mobile scenarios influence the feasibility of finger orientation and, moreover, how this affects the orientation of the mobile device. In this chapter, we address RQ2.

Chapter 5: Social Implications deals with the question of how does finger orientation possible affect social settings. We first present and evaluate a new mixed-method approach which lets researchers evaluate any interaction technique in social settings. We, then, use this new approach to understand how finger orientation effects face-to-face conversations. Hereby, we answer RQ3

Chapter 6: Discoverability addresses the challenge on how finger orientation can be rendered to the user. Here we rendered an analysis which addresses not only finger orientation but a range of new interaction techniques to

understand how new interaction techniques, in general, should be rendered to the user. We identified a set of possible solutions which we validated to uncover their advantages as well as disadvantages. Thus, we address RQ4 concerning finger orientation but also present possible ways to communicate new interaction techniques in general.

Chapter 7: Conclusion summarizes the findings presented in this thesis. Moreover, we link the presented in-depth analysis of finger orientation to other interaction techniques and how they can benefit from the structured analysis presented in this thesis.

2

Related Work

New interaction techniques to enhance input for mobile computing devices can be divided into off-screen interaction techniques and on-screen interaction techniques. While one-screen interaction are performed on the touchscreen itself, off-screen interaction is performed on the device but not on the screen. For off-screen interaction techniques the most prominent are Back-of-Device (BoD) interactions. For example, Le et al. [95, 98] used a BoD touch panel to enhance one-handed smartphone interaction and further presented a novel smartphone prototype that registers touch input on the whole device surface to enable use cases such as grip recognition or touch input on the whole device surface [97]. On the other hand we see the enhancement using on-screen interaction techniques. Here, previous work presented a wide range of novel modalities to enhance touch input on touchscreen devices. Amongst others, these include using the finger shape [139] and size [10], parts of the hand [58], and finger pressure [146]. Moreover, Wang et al. [172] proposed the use of the finger orientation for interaction with tabletops. Wang and Ren [171] proposed use cases, such as selecting items in a pie menu by rotating the finger to make use of the new input dimension. Later, Xiao et al. [188] enlarged the set of use cases to the smartwatch domain. *Z-touch* by Takeoaka et al. [164] used finger pitch angle as an input source gained from a infrared camera beneath

the touch surface in a tabletop setting, for controlling Bézier curves in a drawing application. Rogers et al. [148] as well as Xiao et al. [188] proposed new GUI controls such as rolling context menus and circular sliders where the finger's yaw angle is mapped to a "twist" sensitive control.

In the following, we present the related work for our four sections which will lead to addressing the research questions (RQs). In line with RQ1 (see Chapter 3) we will present related work on detecting finger orientation. Second, we present work in the domain of RQ2 (see Chapter 4) by covering work in the domain of ergonomic constraints. Next, we focus on RQ3 (see Chapter 5) and present relevant work to support and outline the fundamentals of the research question. Lastly, we present related work to also link RQ4 (see Chapter 6) to previous work. Finally, we will summarize the related work and highlight potential shortcomings of the work that has been done in the domain of finger orientation as an additional input dimension to enhance the expressiveness of a 2D touch point.

2.1 Recognition

Evaluations of the input such as finger pressure [146] have shown that they are already suitable for frequent use. One input modality which was investigated in a wide range of prior work, and is still not usable in typical smartphones, is the use of the finger's pitch and yaw angle. A number of researchers [83, 148, 188, 189] presented different approaches to estimate the finger orientation on commodity smartphones. Being able to determine the finger orientation enables use cases such as increasing touch targeting accuracy using the pitch angle [70], manipulating 3D objects, zooming, or setting values on a small touchscreen (e.g., smartwatch) by changing the finger orientation. Over the last decade, multiples attempts have been made to add the finger orientation to the input space. Initial work on the use of finger orientation as an additional input channel [171, 172] was based on a tabletop setup with back projection and determined the finger orientation from the finger's contact area. Dang and André [26] followed up on the same approach and improved it further. Kratz et al. [83] attached a depth camera above the touchscreen to estimate the finger orientation and showed that users could consistently select and hold given target positions. While the estimation

works adequately, this approach requires an additional depth camera attached to the device which mainly is a prototyping tool. Rogers et al. [148] built a custom device based on conventional capacitive sensors to show the feasibility of estimating the finger orientation. Further approaches used touchscreens that sense the fingers proximity to the display (e.g., Hinckley et al. [65], and Samsung's AirView) to reconstruct the 3D finger position. Other approaches attached a mirror to a smartphones's front-camera to capture the finger orientation when touching the display similar to Wong et al.'s work [184]. However, they all have huge drawbacks, either they are bulky (depth camera, and mirror), are not mobile (tabletops), have no display (custom device), or lag in accuracy.

To enable finger orientation estimation on off-the-shelf smartphones without the need for additional sensors, researchers started to use the capacitive images provided by capacitive touchscreens. A capacitive image describes the differences in electrical capacitance between a baseline measurement when no finger is touching the screen, and a current measurement when a finger touches the screen. Amongst others, previous work used these images for biometric user identification [53, 71], hand grip recognition [16, 93], and envisioned a wide range of other use cases such as determining user's handedness, adaptive GUIs based on finger position, or predicting user actions [97]. Zaliva et al. [189] used a sliding window approach combined with an artificial neural network to estimate the finger's pitch and yaw orientation. However, this was done on a table top and used a sliding window approach to calculate the pitch and yaw angle of the finger. Due to the sliding window approach, an unavoidable latency is introduced while absolute input based on finger orientation is not possible. Similar to our work, Xiao et al. [188] used Gaussian process (GP) to estimation the pitch angle based features gained from the capacitive images of an off-the-shelf smartphone. However, for yaw, they used a simple heuristic. Their evaluation revealed an accuracy that is still not suitable for daily use.

While in 2011, Henze et al. [60] showed that touch screen offsets can be modeled using a polynomial function, Weir et al. [176] later showed that GPs are suitable to model the offsets for two-handed interaction and even improve the touch accuracy. Recently, Murray-Smith [135] proposed using Convolutional Neural Networks (CNNs) to improve touchscreens' capabilities further. With this

recent progress in machine learning, CNNs are now the state-of-the-art approach to train models based on images [84]. As CNNs require a large data set to be trained on [161], we conducted a study to collect a large number of capacitive images that are automatically labeled with pitch and yaw angles by a motion capture system as ground truth. Based on this data set, we train estimation models using different machine learning algorithms, starting from the recent work by Xiao et al. [188] over established machine learning algorithms such as *k*-nearest neighbor (*k*NN) and Random Forest (RF) and eventually showing the best results with CNNs.

2.2 Ergonomic Constraints

Ergonomic constraints have been observed in various prototypes using touch interfaces. Le et al. [95] argue that designers should consider ergonomic constraints when developing single-touch BoD interaction techniques and therefore studies are needed to understand how users interact with devices. Colley and Häkkilä [21] found that when using finger-specific interaction, it is necessary to pay attention to ergonomic limitation. They state, that the ring finger is not suitable for interaction. Le et al. [96] studied the range of the fingers when holding smartphones and areas that can comfortably be reached. They proposed design guidelines to ensure an ergonomic placement of interactive elements. Hoggan et al. [67] found that the feasibility of touch rotation depends on the rotation angle, and input becomes harder when the hand rotation increases. Xiao et al. [188] identified additional ergonomic problems when using enriched touch input. Long fingernails made a large pitch unfeasible to perform. Wolf et al. [182] further showed that the feasibility of pitch, yaw, drag, and finger lift gestures on hand-held devices depends on the grip and the touch location. They found that significant deviations from a natural grip cause ergonomic problems, especially for one-handed interaction. Beyond single-touch interactions, Lozano et al. [110] showed that designers need to consider ergonomic factors when designing new multitouch interaction techniques. Moreover, Goguey et al. [43] investigate which pitch and roll angles

occur in a tabletop scenario when performing atomic tasks such as tab and drag. However, in a mobile usage scenario, the user is not restricted to arm movements or touch surface orientation.

Overall, previous research highlighted the importance of extending the input space of touch interaction. In particular, determining the finger orientation was extensively studied by previous work. A growing body of work presented use cases for using the finger's orientation as an input technique. While ergonomic constraints have been studied for static tabletop-like scenarios, previous research did not consider two-handed interaction scenarios. This is especially surprising as two-handed interaction is much more common than tabletop interaction. Furthermore, being able to rotate the touchscreen with the hand that holds the device will likely make many finger angles much easier to perform. Therefore, results from previous work cannot be transferred to mobile interaction, the most common application of touchscreens.

2.3 Social Implications

As new interaction techniques can disrupt social settings, this section presents previous research which enable analyzing interaction techniques in social settings. First, we highlight the recent trend in designing for collocated interactions and showcase the social goals that recent systems have addressed. Secondly, we look at how past research has evaluated interaction techniques that were to be used during conversations. Lastly, we reflect on previous work addressing the disruptiveness of technologies in a social context.

2.3.1 Collocated Interactions

Designing interactions for settings where users are physically co-located and support social interactions through technology has received considerable attention in HCI. Lucéro et al. [111] explored how smartphones can support conversations and enable media sharing to enrich encounters for sitting around a table. They created the technical means to share photos easily during conversations. In a similar vein, Jarusriboonschai et al. [77] designed an application for icebreaking where users sitting in a group were prompted to interact with each other. Researchers have also

designed systems to specifically address the content of the conversation. Wang et al. [173] delivered pictures during conversations to stimulate brainstorming. While these past works explored how conversations can be stimulated, we look into how already established conversational setting can be less negatively affected by technology.

Previous work also investigated approaches for delivering messages without causing unnecessary disruptions. Woźniak et al. [186] investigated how amateur runners can communicate with their supporters while participating in a race without losing immersion, using ambient light feedback and vibration. Similarly, Chen and Abouzied [19] explored how strangers could be informed about their shared interests without disrupting the typical interactions during an academic conference. Again, ambient light feedback was used. Alternatively, tools by Goyal et al. [46, 50] devised the notion of "implicit sharing": an automated way to share insights socially between collaborators, requiring minimal effort to share. Our work is inspired by these approaches as we look for ways to minimize disruption in social settings.

Another strain of work has explored how users could become more aware of each other's activities in social settings. Social Displays [78] used an additional display on the back of the device to engage other users. Pearson et al. [143] investigated how the smartwatch screen can be used to display notifications to others close by. Fischer et al. [33] investigated how groups develop strategies to handle notifications while engaged in collaborative activities. Paay and Kjeldskov [142] looked into augmenting public places to better support social activities. Finally, Fjeld et al. [34] proposed a vision where civic environments could be optimized for meaningful discussion. All of these systems call for building extensive technical support for social interactions, but they do not investigate the risk of disruption that technology may generate.

2.3.2 Evaluating In-Conversation Interactions

There is no standard method for measuring engagement in conversations in HCI. However, research in Psychology has built on the understanding of gaze in social interaction for measuring engagement. Kendon [81] showed that gaze gives important visual feedback in conversation, both for the speaker and the

listener. Based on this, Vertegaal et al. [169] analyzed gaze behavior during discussions of three persons. Kendon [81] as well as Vertegaal et al. [169] analyzed conversations of about 8 minutes. Vertegaal et al. [169] showed that listeners look significantly more at the person talking than at others, while the speaker looks at the addressed persons equally. Bednarik et al. [2] proposed using gaze to indicate the engagement in video conferences. Shell et al. [159] proposed observing users' gaze to determine the attention on a system, and thereby cause different actions.

Jokinen et al. [79] show that human behavior in a three-party dialogue can be studied with the help of eye tracking; however when studying behavior of humans in a larger groups videotaping is favored over eye tracking. Chattopadhyay et al. [17] used this technique to study the impact of a collaborative presentation platform on the presenter and the listeners, while Rico et al. [147] used only interviews to evaluate the social acceptability of gestures after performing them in the wild.

Already in 1980, Goodwin [44] presented work using a video coding method to analyze face-to-face conversations. He transcribed the conversation down to phrasal break, false starts, long pauses, and isolated ungrammatical fragments and further they enriched the transcript with the gaze direction of both parties. In recent related work, Brown et al. [14] studied the impact of search results and phones on conversations. Therefore, the authors recorded 24 hours of video material, in which one researcher flagged 205 clips with a length of one to two minutes for further analyzes. These 205 videos have been coded and transcribed to understand the influence of phones in a face-to-face conversation. McMillan et al. [133] analyzed the effect of smartwatch use. They conducted a study in which they recorded 168 hours of material. They coded clips in which a smartwatch was present and extracted instances to identify effects of using a smartwatch. These studies show the large effort that has been put into studying this space qualitatively. However, researchers like Okamoto et al. [141] has raised questions of the validity of these coding techniques. They identified a influence of the coding person on the results.

2.3.3 Disruptiveness in CSCW

Borst et al. [11] reviewed how an interruption disputes a task and found that users could be interrupted in a low-problem state and maintain the problem state. On the other hand, Chen and Vertegaal [18] focused on reduction of interruptions whenever the user's mental load was excessive. Hudson and Smith [74] discussed the fundamental trade-off between awareness and disruption. In their work, they highlighted that the trade-off is unavoidable and needs to be studied independently for usage scenarios. Tolmie et al. [165] investigated interruptions in game play. Their solution tries to make interruptions visible and available to prevent them.

Trbovich and Harbluk [166] investigated the impact of cognitive distribution on driving behavior. They observed a change in driver visual behavior when using a speech-based interaction while driving. Bogunovich and Salvucci [8] investigated the direct effect of an interruption on the primary task. They found that the ringing duration in a phone ringing scenario had a significant impact on an email answering task. In contrast, Iqbal and Horvitz [76], in their analysis of a two-week study, found that desktop notifications created awareness but reduced task switching as explicit monitoring was not needed any more. Dabbish and Kraut [24] looked at awareness displays and social motivation for team members and showed that the timing of the interruptions by the awareness displays influences the performance. These works show that managing interruptions may play a crucial role in the success of a technical intervention in a social setting. Similarly in a social setting, Goyal [47] found that mobile interruptions by partners can prove dangerous towards collaborative data analytic challenges. The Author found that using acceleration of psycho-physiological sensors like GSR as a metric can help alleviate disruptions caused by such interruptions. This thesis is inspired by the above findings and continues to help identify ways we can design for better interruption management by understanding the disruptiveness of different interaction techniques.

Su and Wang [162] observed phone usage in pubs over three years. They found that phones could help enhance conversations but also cause disruption. This was confirmed in a similar study by Porcheron et al. [144]. Similarly, Ofek et al. [140] investigated on how to effectively deliver information to interlocutors during a conversation. Their study revealed that delivering batches of visual information

was the most effective. Newman and Smith [137] adopted a similar approach to study the influence of paper document and laptop usage in conversations. They concluded that providing assistance to keep the time short would help to cut time spent working on the laptop. Boyd et al. [12] developed SayWAT, a device that helped adults with autism to focus on face-to-face conversations. Moreover, Exposito el at. [32] proposed a system to reduce obtrusive note taking while collaborating remotely. They investigated selecting through eye tracking cues combined with foot-based gestures.

2.4 Discoverability

While Shneiderman et al. [160] and Norman [138] both argue for interaction *discoverability*, today's mobile devices look different. Apple, as one of the main players in the mobile market with over one billion active devices[1], use the "Tips" app[2] on all their iOS devices to introduce new features by triggering a notification and guiding the user through a tutorial. HTC's "Edge Sense" is communicated to users already during device setup and an additionally pop-up is shown whenever Edge Sense can be used.

The pinch-to-zoom gesture is available on all major smartphones, Microsoft Windows' touch interface, digital cameras (e.g. Sony Alpha a7 iii), and computer trackpads. However, an on-device communication concept was never developed. The two-finger gesture dates back to 1985 when Krueger et al. [85] used the index finger and the thumb to indicate the size of an ellipsis. One of the first occurrences where pinch-to-zoom is described is by Rubine in 1992 [152]. In 2005, it was used by Han [56] in a tabletop scenario. However, until 2007 it was not used for consumer devices nor was a strategy developed to communicate the pinch-to-zoom gesture to users. With the first iPhone, the gesture became available in a consumer product but a way to communicate the gesture on the device was not implemented. Instead, Apple used the presentation of the iPhone[3]

[1] `https://apple.com/newsroom/2018/02/apple-reports-first-quarter-results/`

[2] The "Tips" app by Apple `https://tips.apple.com/en-us/ios/iphone`

[3] Macworld San Francisco 2007 Keynote 2007-01-09: `https://youtube.com/watch?v=t40EsI0Sc_s`

to communicate the gesture live on stage by showcasing it twice, once for photos and later for maps. The presentation of the iPhone and subsequent ads by Apple explained the gesture to potential users which emerged as a cross-platform gesture in the following years.

Samsung's launcher shows a line on the side to indicate that a swipe to the center of the screen allows to open a shortcut menu and the iPhone X presents a swipeable line at the bottom of the screen as a replacement for the home button. A wide range of opportunities to use swipe interactions, for example in the Gmail app which allows swiping left or right to archive a mail, are not visually communicated. The iPhone's force touch allows to preview and open content and is also not visually communicated. Instead, it was presented in an Apple keynote and subsequent ads. The longpress in the Android eco-system is never communicated. Users must discover the input technique. Lastly, another gesture which became a cross-platform standard is the "pull-to-refresh" gesture, which is implemented by all major apps, e.g., Gmail, Facebook, and Instagram but never communicated to the user.

2.5 Summary

In this chapter, we presented the previous work related to using finger orientation. First, we present previous research in respect to the recognition of finger orientation as an additional input method on tabletops as well as mobile devices. We present approaches to recognize the orientation of the finger and show that while they improved over time, they are not ready to use for consumer electronics. Second, we present research done which link interaction to ergonomic influences as well as know ergonomic constraints. Here we stress that finger orientation was not evaluated from an ergonomic point of view yet. Third, we address the domain of social acceptability and descriptiveness though technologist. Here, we conclude that so far none of the presented approaches allows for a quick and easy analysis during one iteration of the human-centered design cycle. Lastly, we discuss that, while UX designers proposed a set of rules for good design practices,

today's interface designers do not necessarily follow them. Thus, we show that while rules exist, it is unclear how to introduce the user to novel interaction techniques.

3

Recognition

In this chapter, we investigate possible ways to recognize the finger orientation while touching a surface (RQ1). First, we investigate using a depth-camera over the touch surface, an approach proposed by Kratz et al. [83]. We show how to improve the accuracy of the proposed algorithm. Furthermore, we argue that this approach is suitable when investigating finger orientation in the early design stages, as the depth camera can turn any surface into a finger orientation aware surface. Second, we improve an approach first proposed by Xiao et al. [188] in which they use capacitive images to estimate the finger orientation directly from the touchscreen. This approach, in contrast, to the depth-camera approach keeps the form factor identical to a off-the-self device. Finally, we discuss the advantages of both approaches.

Parts of this chapter are based on the following publications:

S. Mayer, M. Mayer, and N. Henze. "Feasibility Analysis of Detecting the Finger Orientation with Depth Cameras." In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. MobileHCI'17. New York, NY, USA: ACM, 2017, 82:1–82:8. ISBN: 978-1-4503-5075-4. DOI: 10.1145/3098279.3122125

S. Mayer, H. V. Le, and N. Henze. "Estimating the Finger Orientation on Capacitive Touchscreens Using Convolutional Neural Networks." In: *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*. ISS '17. Brighton, United Kingdom: ACM, 2017, pp. 220–229. ISBN: 978-1-4503-4691-7. DOI: 10.1145/3132272.3134130

3.1 Challenges in Recognizing Finger Orientation

Over the last years, touchscreen input evolved to the main mechanism for mobile devices. Through direct touch, users can intuitively interact with the GUI. GUI elements can simply be selected by touching them. Recent capacitive touchscreens sense touch input by measuring a change in capacitance when a finger touches the display. These measurements are translated into a 2D point by the touchscreen controller. Since the measured capacitance is omitted afterward, touchscreen input is limited to 2D input.

Commercial devices, as well as previous research, presented a wide range of novel interaction techniques. Already in the first version of Android and iOS, they leveraged the time dimension to provide the long press. With the iOS 6s in 2015, Apple introduced *3D Touch* which adds a pressure dimension the interaction. Both methods are used to modify the touch input and alter the action. Roudaut et al. [151] presented a technique to use the roll of the finger to scroll through lists. Xiao et al. [188] and Zaliva [189] proposed using the finger orientation to increase the richness of the touch input. A larger input vocabulary enables a richer interaction and thereby enables new ways to manipulate potential applications.

3.2 Depth Camera Approach

In the following, we improve approach by Krazer et al. [83] who used a depth-camera over the touch surface to determine the finger orientation. We first, built a prototype to recreate the original setup using new components. We than measured the accuracy of the prototype with the algorithm by Krazer et al. [83]. We found a systematic error in the original approach which we reduce by using offset-shift

models for both pitch and yaw orientation. Finally, we discuss potential use cases and argue that the depth-camera approach while being bulky has its advantages in the early prototyping stages.

3.2.1 Prototype

For our prototype, shown in Figure 3.3, we use a Samsung Galaxy Tab Pro 8.4 which offers $2560 \times 1600px$ on an 8.4-inch screen resulting in $359.39PPI$. As a depth camera, we use an Intel RealSense F200. The camera has a minimum sensing distance of $20cm$ and a resolution of $640 \times 480px$ at $120FPS$. We use the RealSense F200 due to its small minimum distance in comparison to other available depth sensors. However, we needed to overcome the $20cm$ between the tablets screen and the depth sensor. Therefore, we 3D printed a mount for the tablet and laser cutted a connection plate to attach the camera to the tablet. We firmly connected the parts using metal screws.

3.2.2 Study

To evaluate the accuracy of our setup with the algorithm proposed by Kratz et al. [83] we collected ground-truth data by conducting an experiment. The ground-truth was determined by three RGB cameras to always get a clear view.

3.2.2.1 Apparatus

We recorded the ground-truth data using three RBG cameras which we fixed on a wooden frame. We mounted one camera on top of the tablet, one on the left and one on the right (see Figure 3.2). The top camera was used to determine the fingers yaw while the left and right were used to determine the fingers pitch. We needed two cameras to determine pitch because when we insert extreme yaw angle one camera was always covered by the rest of the hand. We used three Microsoft Lifecam HD 3000 which recorded with $1280 \times 720px$ at $3FPS$. We replaced the flexible parts of the original camera mount with a non-flexible plastic connector (see Figure 3.2). The three camera streams were used to later determine the finger orientation through a manual labeling process. We developed an Android

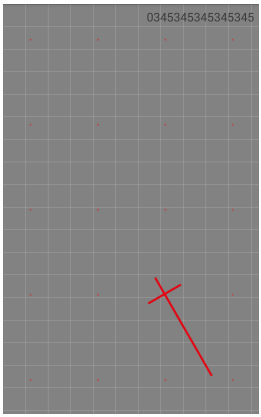

Figure 3.1: The study app is showing instructions to perform a 30° yaw input at one specific position.

application, which displays red crosshairs indicating the touch position. The crosshair further indicated which finger yaw angle the participant should perform, see Figure 3.1.

3.2.2.2 Design and Task

We designed the study using a repeated-measures design with four independent variables: TARGETS, YAW, PITCH, and FINGER. We randomized the order of FINGER, and within FINGER we randomized TARGETS, and YAW. To cover a broad range of possible positions, we used 20 TARGETS arranged in a 4×5 grid on the tablet. The targets further represented five PITCH input angles: 15°, 30°, 45°, 60°, and 75°. Xiao et al. [188] found that a pitch of 90° cannot be detected with long nails. Thus we did not investigate angles steeper than 75°. Further, we used five YAW input angles: −60°, −30°, 0°, 30°, and 60°. As we present in Chapter 4 the comfort zone of yaw input for the right hand is ranging from −33.75° to 101.25°. To not stress the participates too much we limited the range to −60° and to built a symmetric model to 60° on the other extreme. Further, all tasks were performed with four FINGERS: index, middle, ring and little finger. Thus, we had a design with $20 \times 5 \times 5 \times 4 = 2,000$ conditions.

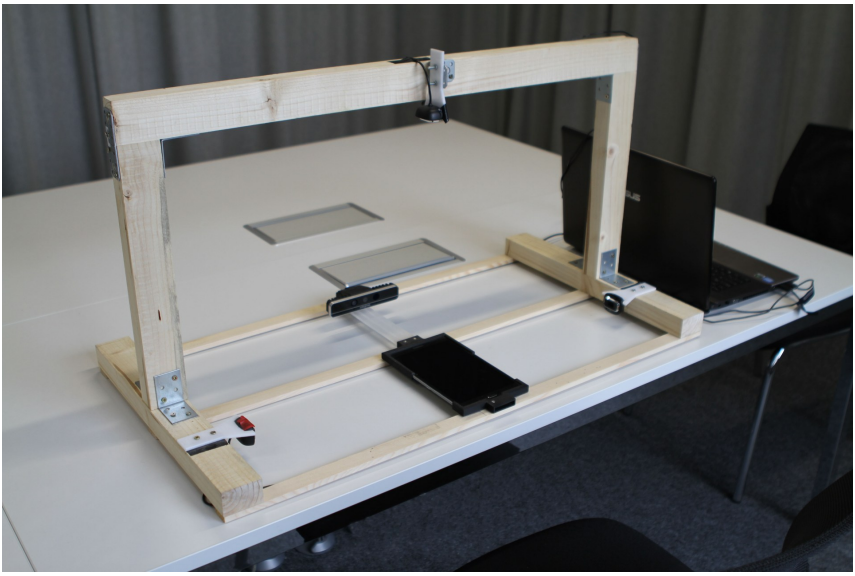

Figure 3.2: The wooden frame with the attached web cameras which we used for ground truth recording in our study.

Performing a specific pitch angle is not easy. To overcome this challenge Xiao et al. [188] used laser cutted stabilizers which they placed below the participants' finger. However, this is not possible using the camera based approach, as the stabilizers would influence the depth image. Thus we decided to let participants perform a movement and determine the PITCH angle in the post processing. Therefore we asked half of the participants to start with a pitch close to 0° and then change the pitch of the finger up to a steep angle close to 90°. The other half was asked to move from 90° down to 0° pitch. We specified these two movements to reduce an effect of the finger moving only in one direction.

3.2.2.3 Procedure

First, we welcomed our participants and informed them about the procedure of the study. Second, we asked them to fill in a consent form and a questionnaire with demographic data. Afterward, the experimenter marked each finger with two

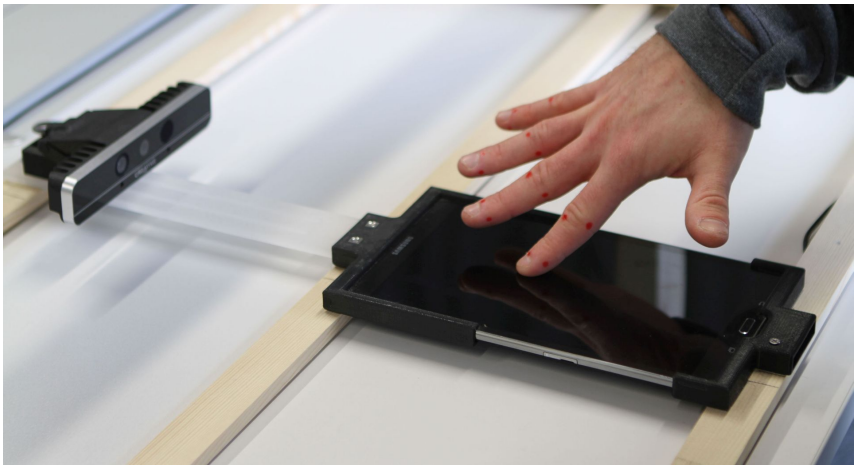

Figure 3.3: A participant while performing the task.

red dots on the left and right side as well as on top of the finger to later determine the finger orientation. Then, we explained to the participants that they have to touch the center of the red crosshair while aligning the finger with the longer red line indicating the yaw angle. Afterward, we asked participants to move the finger slowly up or down to input several pitch angles (see Figure 3.3).

3.2.2.4 Participants

We recruited participants from an internal university self-volunteer pool. We recruited 12 participants (9 male, and 3 female) which were aged between 22 and 35 ($M = 25.83$, $SD = 3.31$). All participants used their right hand. We reimbursed them with 5 € for their participation.

3.2.3 Results

First, we corrected the camera lens distortion for the three recorded RGB-camera streams. Second, we manually labeled the finger posture with the help of the red markers on the finger for each of the five PITCH angles. Due to the continuous change of the pitch angle, we were able to label accurate PITCH angles. However,

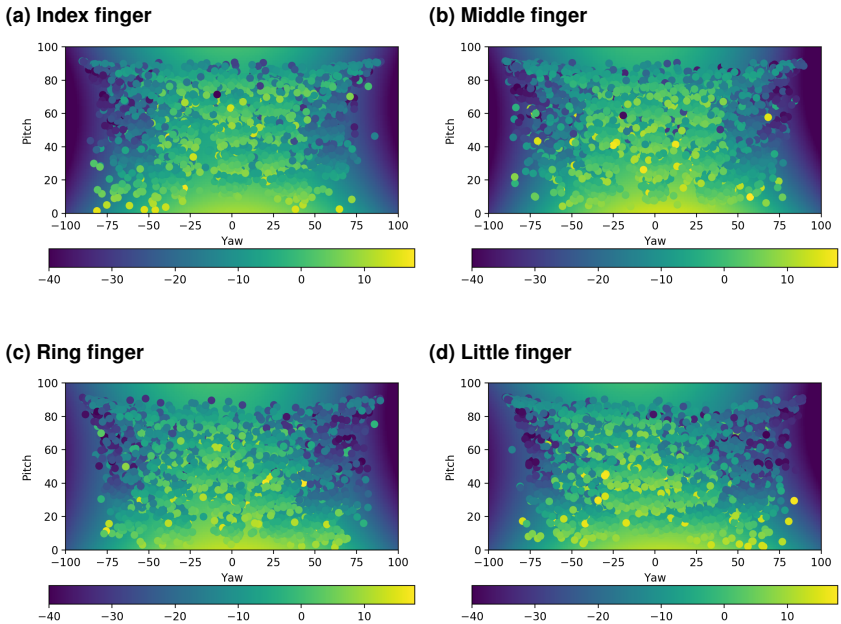

Figure 3.4: The scatter plots are showing the points where we gained data samples from the study. The underlining plane represents the correction model for the pitch correction based on pitch and yaw of the depth camera.

for the YAW angles, we were bound to the participants' accuracy (M = 3.1°, SD = 9.9°). For the modeling, we used the yaw angles actually performed by the participants, not the initial categories.

Using the depth images, we determined the pitch and yaw with the PointPose algorithm [83]. Due to the manual labeling and noise in the depth data, we removed outliers where the distance between ground-truth and predicted angles is more than two standard deviations away from the average. This was done for pitch and yaw individually. In total, we removed 8.2% of the data. Then, we calculated the root mean squared error (RMSE) for each finger, see Table 3.1. The average RMSE is 15.4° for pitch and 12.2° for yaw.

The PointPose algorithm [83] was evaluated regarding precession over time. The evaluation of Kratz et al. [83] used an alignment task where the target was presented and the participant had to move a cursor to overlap with the target. Whereby the cursor could be manipulated through either pitch or yaw input. Accuracy was determined by measuring the variation of a 7.5 sec recording where the participants had to hold the alignment.

Comparing our results with the results reported by Kratz et al. [83] is not trivial because Kratz et al. averaged over 7.5 sec whereby we used a concrete error not the variance while holding the finger. Further, Kratz et al. [83] used 7 steps for yaw ranging from $-30°$ to $30°$ and 5 steps for pitch ranging from $50°$ to $75°$. On average they reported a change in variation of $M = -.92°$ ($SD = 6.36°$) for pitch and $M = -2.52°$ ($SD = 14.67°$) for yaw.

3.2.3.1 Modelling

In the following, we present our model to reduce the error through offset correction. We modeled the offset with a full second order two-dimensional polynomial, as in Equation 3.1. We choose Equation 3.1 after a one-dimensional polynomial fitted less accurate and the visual inspection suggested a more complex underlying behavior. Furthermore, a more complex function led to overfitting. The pitch and yaw offset corrections are modeled independently from each other. Thus, we fitted 8 functions, 4 fingers $\times$ 2 degrees of orientation (pitch/yaw). However, the correction model for pitch and yaw is based on both pitch and yaw angles gained

Table 3.1: The RMSE and standard division for pitch and yaw per finger.

	Pitch			**Yaw**		
	RMSE	**M**	**SD**	**RMSE**	**M**	**SD**
Index	15.7	−10.8	11.4	11.9	2.8	11.7
Middle	17.	−10.5	13.4	14.7	3.4	14.3
Ring	13.8	−7.4	11.7	11.2	3.5	10.7
Little	14.8	−9.8	11.1	10.8	3.	10.4
Mean	15.4	9.6	11.9	12.2	3.2	11.8

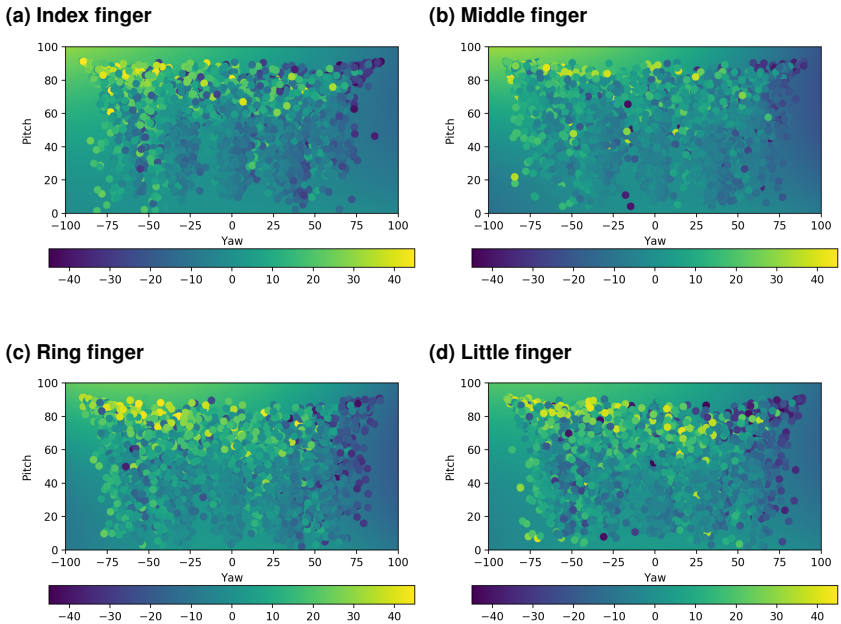

Figure 3.5: The scatter plots are showing the points where we gained data samples from the study. The underlining plane represents the correction model for the yaw correction based on pitch and yaw of the depth camera.

from the depth camera as the α and β input for the Equation (3.1). Whereby we used the predicted angles by the depth camera for α and β, results are shown in Figure 3.4 for the pitch correction and in Figure 3.5 for the yaw correction. We validated the improvements for all functions by the use of leave-3-participants-out cross-validation, which is a split of 75% : 25% for train and test.

$$f(\alpha,\beta) = a\alpha^2 + b\beta^2 + c\alpha\beta + d\alpha + e\beta + f \tag{3.1}$$

For the pitch correction, we achieved an average reduction of the RMSE of 41.7%, all results are listed in Table 3.2. The overall remaining pitch error

improved from $M = -9.6°$ without correction model to $M = -.9°$ with correction model. For the yaw correction, we achieved an average reduction of the RMSE of 14.7%, all results are listed in Table 3.2. The overall remaining yaw error improved from $M = 3.2°$ without correction model to $M = .2°$ with correction model.

For the final model we used the training and test data to fit the model, we achieved an RMSE reduction by 45.4% for pitch and 21.83% for yaw. Further the fitness of the pitch correction functions for the four fingers functions is $R^2 = [.50 \;\; .45 \;\; .49 \;\; .54]$ (see Figure 3.4), and for the four yaw functions the fitness is $R^2 = [.63 \;\; .58 \;\; .67 \;\; .76]$ (see Figure 3.5).

3.2.4 Discussion and Implications

In a first step, we recorded ground truth data pitch and yaw to determine the accuracy of PointPose. In a second step, we applied offset models to correct the mean error of the PointPose method.

We showed that the root mean squared error without correction is 11.75° for pitch. This results in an offset of 13.1% of the possible pitch input range which is from 0° to 90°. Further, in our study, we explored the yaw range between −60° and 60° resulting in an RMSE of 8.74° and an average offset of 7.3%. Thus high precision input is not possible with the proposed method. Even for imprecise input, we see a lack of feasibility to use this method.

We also show that the predicted results are more accurate close to the center of the observed input space (pitch = 70° and yaw = 0°), see Figures 3.4 and 3.5.

Table 3.2: The reduction of RMSE when applying the correction models to pitch and yaw.

	Pitch in %	**Yaw in %**
Index finger	41.8	15.2
Middle finger	43.	16.5
Ring finger	40.4	13.9
Little finger	41.5	13.2
Mean	41.7	14.7

For the pitch correction, we can see an overall trend of a larger pitch error with yaw values away from the center. Further, we can observe that this is radially symmetric (see Figure 3.4). Also for the yaw correction, we see an overall drift in the mean data (see Figure 3.5). To correct the drift and improve the predicted accuracy, we applied one offset model per pitch/yaw and finger and thereby reduction of RMSE for pitch by 45.4%, and for yaw by 21.83%.

When comparing our results with correction and the results reported by Xiao et al. [188], we achieve a similar pitch error and smaller yaw error. Their method leads to a pitch error of 9.7° while our method achieved 11.75°. For the yaw error, Xiao et al. [188] reported 26.8° while our method achieved a three times smaller yaw error of 8.74°.

We used the tablet to display targets, not for the actual recognition nor the model. Thus, the touch position was not taken into account in the analysis nor the offset correction. Doing so allows using the depth camera also without a tablet. Thus, mounting the depth camera onto a not touch sensitive surface is possible; this can be useful for system prototyping when building a first fully functional apparatus. We envision using our approach even in earlier stages e.g. when designing new GUI interfaces using paper prototypes. Here, the behavior of the finger orientation can be observed, and the GUI can be designed adaptive to the input.

3.3 Capacitive Sensor Approach

Previous work proposed algorithms to determine pitch and yaw to use them as an additional input modality. However, determining a finger's orientation using off-the-shelf devices and existing algorithms is still not precise [188]. Due to the high potential of using finger orientation as an additional input, we aim to improve the algorithms proposed by previous work. In this Section, we present a range of machine learning models to estimate the fingers' pitch and yaw angle.

In the following sections, we then present the details of our study design and the data set we collected in our experiment. Next, we interpret and discuss the implications of the data obtained. Finally, we present how we built our new finger orientation estimation using machine learning.

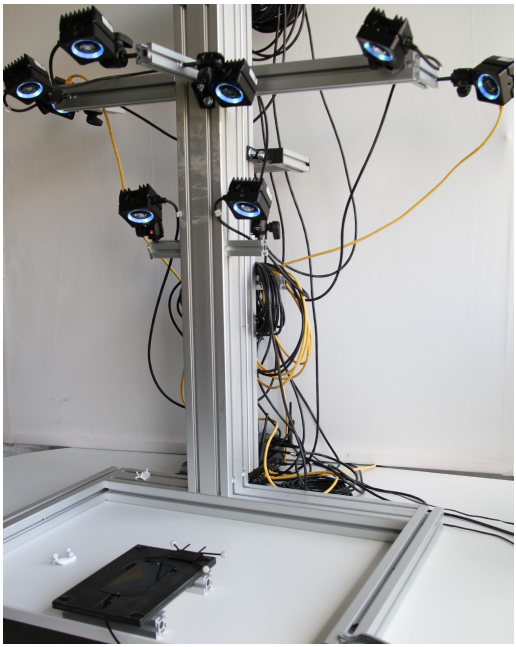

Figure 3.6: The study setup showing the Nexus 5 and the aluminum profiles where the cameras are firmly mounted to.

3.3.1 Study

In this study, we collect capacitive images and respective finger orientation angles as ground truth using a motion capture system. We followed the approach by Holz and Baudisch [69] to collect ground truth data of the finger orientation. Specifically, a finger orientation consists of a pitch angle and a yaw angle. We define pitch as the angle between the finger and the horizontal touch surface. The pitch is 0° when the finger is parallel to the touch surface, i.e., the entire finger touches the surface. The yaw angle represents the angle between the finger and the vertical axis. Yaw is 0° when the finger is parallel to the long edge of the phone and increases when the finger is rotated counterclockwise.

3.3.1.1 Apparatus

The apparatus, shown in Figure 3.6, includes an LG Nexus 5 smartphone, eight *OptiTrack* Prime 13W motion capture cameras, and one PC to operate the *OptiTrack*. We modified the Android kernel of the LG Nexus 5 as described by Le et al. [93] to gain access to the capacitive images. Using our Android application, we recorded the 15×27 pixel images at $20FPS$ as well as the respective 2D touch point provided by the Android SDK. While recording, the Android application instructs participants to touch on a red $2 \times 2cm$ crosshair as shown in Figure 3.7. To further record the respective 3D finger orientation in relation to the orientation of the smartphone, we attached a rigid body with 3 markers each onto the participant's index finger and the smartphone. This enables the motion capture system to reconstruct the pitch and yaw angle of the finger, and record them at $240FPS$. Based on this information, the experimenter could monitor all orientations that were covered by the participants live on the PC.

3.3.1.2 Design and Task

The experiment consists out of two phases, the *tapping phase* in which participants tapped the screen repeatedly, and the *moving phase* in which participants altered the finger tip without removing the finger from the screen. Since the resolution of the capacitive image is low, we hypothesized that the blob representing the finger could look different depending on whether the touch is performed in the center of a pixel, or on the pixel borders. Thus, touches in each phase were performed on a *pixel center* and on *pixel borders*, which was represented by the red cross-hair. In total, this results in 2×2 tasks which are counterbalanced using a Latin square.

3.3.1.3 Procedure

After signing the consent form and filling out a demographic questionnaire, we attached the reflective rigid body markers to the participant's index finger. Participants were then instructed to touch the display with their index finger, while slowly altering the pitch and yaw angle of the finger. In the *moving phase*, they were instructed not to remove the finger while altering the orientation. In the *tapping phase*, they were instructed to alter the orientation when touching

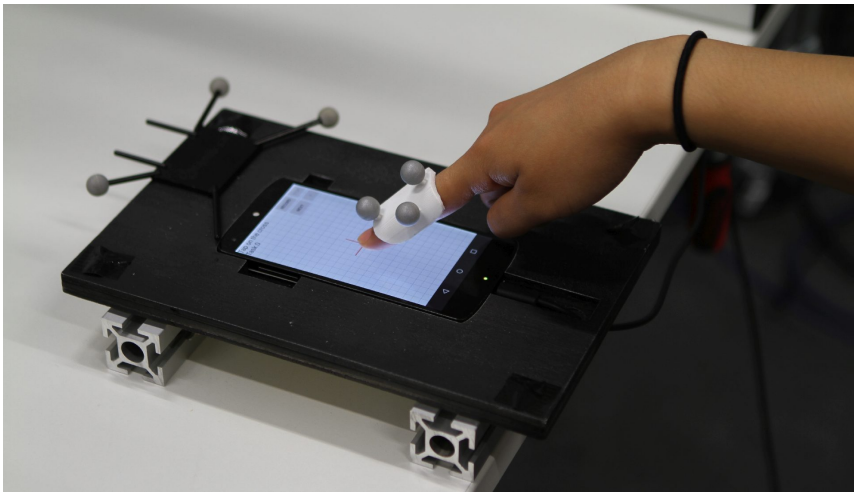

Figure 3.7: A close up of a participants hand while performing the study. On the participants index finger we attacked the ridget body to track the finger orientation.

the screen, then lift off the finger and repeat the procedure. This were done until the experimenter confirms that all angles between 0° and 90° for pitch and -90° to 90° degrees for yaw was covered. A degree counts as covered when at least 20 capacitive images for that degree were recorded. Using the live monitor application, the experimenter instructed the participants to cover all angles and ensured that the recordings were complete.

3.3.1.4 Participants

We recruited 35 participants (28 male, and 7 female) through our university's mailing list. Due to technical issues we excluded two participants. For the remaining 33 participant (26 male, and 7 female) the age ranged from 20 to 33 years ($M = 22.9$, $SD = 3.4$). All of them had either no visual impairment or corrected to normal vision. None of the participants had any locomotor disabilities. Further, all participants were right handed. Only participants with short fingernails were invited to participate, as this was stated to be an issue by Xiao et al. [188]. Therefore participants were able to cover the full pitch range from 0° and 90°.

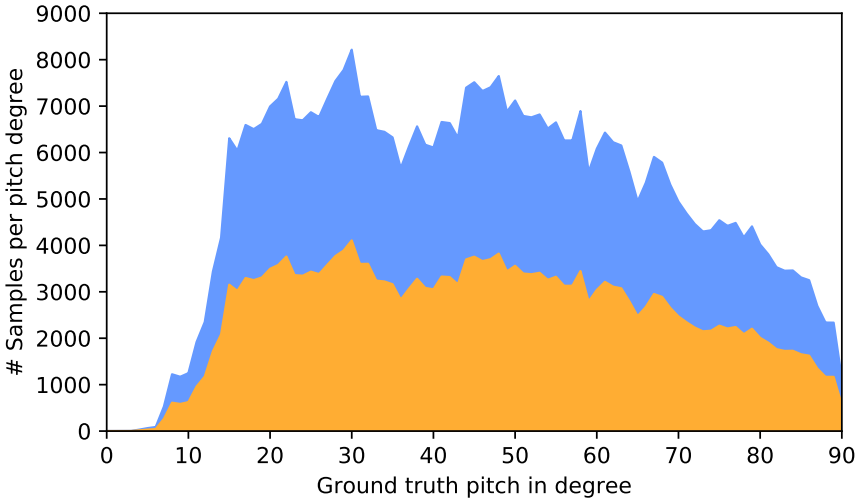

Figure 3.8: The blue counts are representing the distribution of pitch samples which we used for the modeling. The yellow area represents the distribution of pitch samples we recorded in our study. The are in between in obtained by flipping the yaw data.

3.3.2 Modeling

In a pre-processing step, we mapped the capacitive data record on the phone with the OptiTack recorded on the PC. As the capacitive images are recorded at $20FPS$, and the OptiTack samples are recorded at $240FPS$, we used the closest OptiTack sample for each capacitive image. This resulted in an average offset of $25\mu sec$ ($SD = 162\mu sec$).

As the first step, we removed all samples from the tapping condition. The sampling rate $20FPS$ did often only capture the finger while moving but not the finger fully touching the touchscreen. As the second step, we followed Xiao et al. [188] and filtered noise below $3pF$ (picofarad). The sample distribution is shown in Figure 3.9 for yaw and the pitch distribution in Figure 3.8. Results in Chapter 4 show that yaw input beyond $-33.75°$ falls into a non-comfort zone, we recorded fewer samples towards $yaw = -90°$. To compensate for this effect, we added vertically flipped versions of all initial capacitive images to our data set to balance the yaw samples. We then performed a blob detection on the capacitive

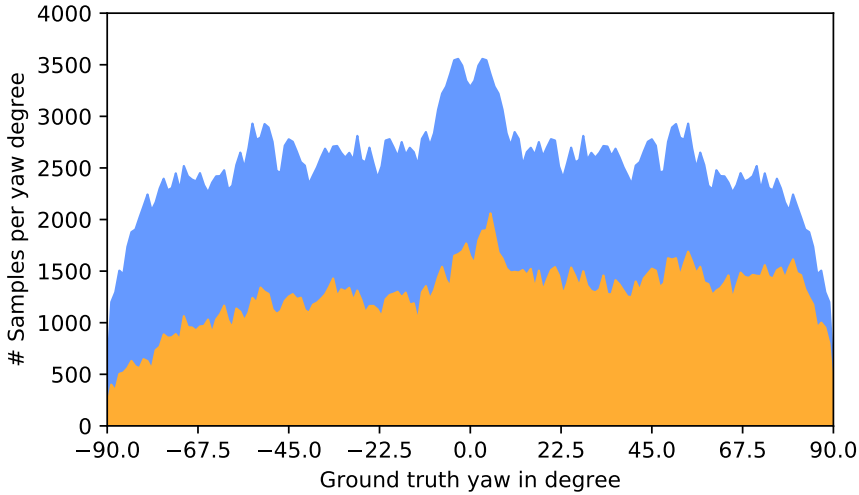

Figure 3.9: The blue counts are representing the distribution of yaw samples which we used for the modeling. The yellow area represents the distribution of yaw samples we recorded in our study. The area in between in obtained by flipping the yaw data.

images; the biggest blob was 15×22. We cropped all blobs and pasted the blobs into the upper left corner of an empty 15×22 image (referred to as *blob image*). The blob detection omitted all blobs that were not greater than one pixel of the image ($4.1 \times 4.1mm$) as these can be considered as the noise of the capacitive touchscreen. We used the pixel values of the blob images as input features for the model. In total, our data set consists of $457,268$ blob images.

For all models, we derived training and test sets using an $80\% : 20\%$ split by participants respectively. As we had 33 participants in total, we split them into the train and test set. The first 26 participants were used for training, and remaining 7 participants were used for testing. For all models, we used an optimizer function to reduce the root mean squared error (RMSE).

3.3.2.1 Feature-Based Approach

The recent feature-based approach by Xiao et al. [188] uses 42 features extracted from the capacitive image to estimate the pitch and yaw angle. For the pitch angle,

they used a GP regression. We reimplemented these 42 features. However, when feeding them to the GP, we hit the limits of GP due to a training time of $\mathcal{O}(n^3)$ and memory scaling of $\mathcal{O}(n^2)$ [89]. Since Xiao et al. recorded 1,224 samples to train their model, a GP regression worked for their in comparatively small data set to estimate pitch. However, with our data set, which is 373 times larger, training GPs on the full data set is not feasible anymore.

To train a GP we used a subset of our data set. To not vary the number of reference points, we used only samples which have the properties of the original implementation, pitch: 0° to 90° in 15° steps and yaw: -60° to 60° in 15° steps. Since no $pitch = 0^\circ$ samples have been recorded in our study the following comparison is not validated for 0° pitch. Further, we used $\pm 1^\circ$ for pitch and yaw to create a data set to implement the approach presented by Xiao et al. [188]. This resulted in a $4,977$ samples large data set. We divided the data set in a train- and a test- data set. For each reference point, 75% are used for training and 25% for testing. This ensured that each original reference point was trained and tested, resulting in $3,711$ training samples and $1,266$ test samples. As Xiao et al. [188] did not report how many trainings samples they recorded, we can not ensure the same size. However, they reported that the test set contained $1,224$ samples, which is roughly the amount we use for testing. Due to the proprietary implementation[1], Xiao et al. [188] did not report which kernel or parameters they used. Thus we used the trial-and-error method [23] combined with a grid search to find parameters for the GP reimplementation of Xiao et al. [188]. To train our reimplemented GP, we used scikit-learn[2].

To make use of the rich data set we collected, we additionally decided to also test apseudo-implementation of the approach by Xiao et al. [188]. We use a simple k-nearest neighbor (kNN) approach as well as a DNN approach. For yaw Xiao et al. used the ellipsoid of the blob with a 90° correction when the pitch is larger than 50°.

[1] `http://qeexo.com/`

[2] We used the Gaussian process (GP) version available in the python package scikit-learn: `http://scikit-learn.org/stable/modules/gaussian_process.html`

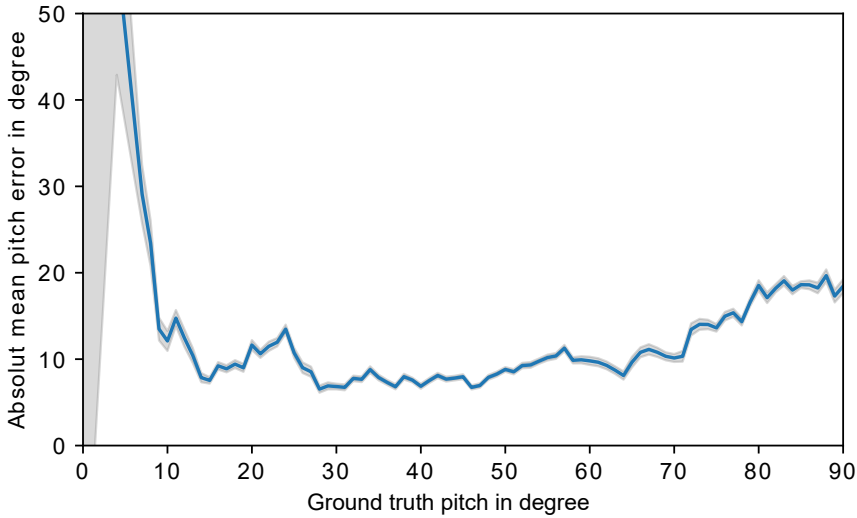

Figure 3.10: The remaining pitch MAE when using out CNN + L2 model. The gray area shows the 95% CI.

3.3.2.2 Pitch Estimation Using Features

For the *GP reimplementation* of Xiao et al. [188] using their 42 features we found that a RationalQuadratic[1] kernel performed best with $Alpha = .01$ and $LengthScale = 100$.

For the *pseudo implementation* we replaced the GP with a kNN, and a DNN. (1) For the kNN estimation approach using the features by Xiao et al. we achieved the best results using $k = 129$. We used a change RMSE smaller than $\varepsilon = .001$ as a stopping criteria. (2) For the DNN we used a two ReLu [136] layer structure with 100 and 50 neurons respectively. To train our model we used an Adagrad Optimizer [28] with an exponential decaying learning rate ($LearningRate = .01$ and $DecayRate = .2$). We initialized the weights using the Xavier initialization scheme [42] while the biases were initialized with .01.

[1] `http://scikit-learn.org/stable/modules/generated/sklearn.gaussian_process.kernels.RationalQuadratic.html`

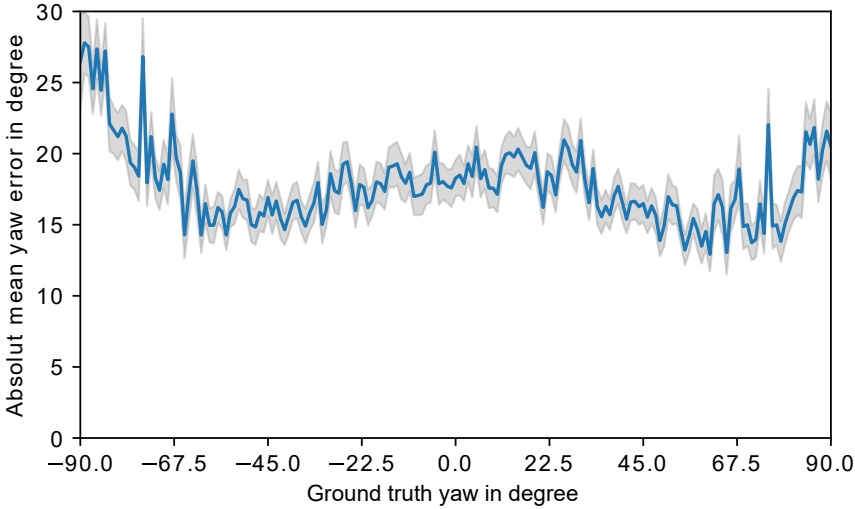

Figure 3.11: The remaining yaw error when using out CNN + L2 model. The gray area shows the 95% CI.

3.3.2.3 Yaw Estimation Using Features

For the *GP reimplementation* we used the S_1 ellipsoid feature with the heuristic described by Xiao et al. [188] to estimate yaw.

The simple heuristic is based on the estimated pitch however we also want to take full advantage of the large data set. Thus, for the *pseudo implementation* we used the simple heuristic using S_1 ellipsoid feature as well a kNN and a RF model for the yaw estimation. We performed a grid search for kNN and RF. We again used an $\varepsilon = .001$ as an early stopping in RMSE change. For kNN, the best results were achieved with a $k = 109$ while $Estimators = 79$ performed the best for RF .

3.3.2.4 Representation Learning Approaches

We propose a new way to determine the pitch and yaw of the touching finger. This method uses DNNs with the raw capacitive blob values as input also known

as *representation learning* [3]. Thus we first applied a blob detection to identify the touching finger and then directly feed the 15×22 sized blob into the DNN to estimate the pitch and the yaw of the finger.

Beyond the feature-based baseline, we implemented two baselines that use the raw pixels of the blob to estimate pitch and yaw. Therefore, we trained kNN and RF models to estimate pitch and yaw independent from each other. We used the implementations of scikit-learn[1].

We implemented a neural network for regressing using *TensorFlow*[2] and tested different network configurations by varying the amount of neurons and layers, activation functions, and optimizers provided by *TensorFlow*. We trained 6 different neural network structures: two DNNs one for pitch and one for yaw and one DNN and three CNNs which estimate pitch and yaw at the same time. Further, we used early stopping to prevent overfitting for all neural networks [15]. *TensorFlow* has a large amount of parameter for their functions if the parameter is not reported in the following section we used the standard parameter of *TensorFlow* version 1.2.1. We applied the trial-and-error method [23] to find the best parameters for our models.

***k*-nearest neighbor (*k*NN):** We started using kNN regression as a baseline estimation for pitch and yaw interdependently. We performed a grid search to identify the best k for the two models. For pitch, we found that $k = 180$ for pitch and $k = 278$ for yaw performed best.

Random Forest (RF): As a more advanced model than kNN, we used two RFs as a second baseline to estimate pitch and yaw interdependently. We performed a grid search to identify the best i number of trees in the forest for the two models. For pitch, we found that $i = 85$ for pitch and $i = 17$ for yaw performed best.

Deep Neural Network (DNN): We used two DNNs, one to estimate pitch and one for yaw. We achieved the best results using a 3-layer DNN. We used a ReLu-ReLu-Sigmoid structure, with 500, 300, 200 neurons respectively. To train our model, we used an AdaGrad Optimizer [28] with a exponential

[1]The scikit-learn package version v0.19.0: `http://scikit-learn.org`

[2]TensorFlow is an open source software library for machine learning which allows easy deployment on various platforms e.g., CPS and GPU. `http://tensorflow.org`

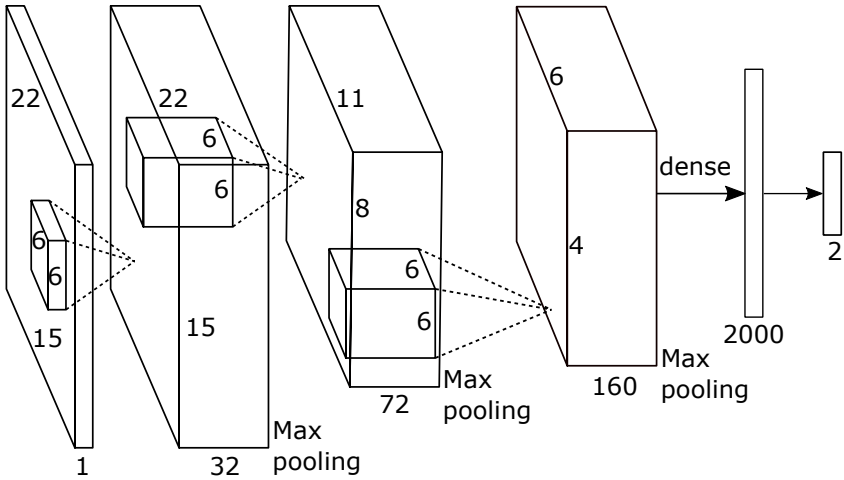

Figure 3.12: The layer structure for our best performing model: CNN + L2 model. Max-pooling is used with a Stride of 2 and pedding is set to same. The dense layers use a softplus activation function ($softplus(x) = log(1 + exp(x))$).

decay learning rate ($LearningRate = .01$ and $DecayRate = .2$). We initialized our weights using Xavier initialization scheme [42] while the biases were initialized with .01.

Combined DNN: For a DNN with 2 output neurons to estimate pitch and yaw at the same time, we achieved the best results using a 3-layer DNN. As layers, we used a ReLu-Relu-Sigmoid structure, with 1200, 800, and 400 neurons respectively. We initialized the weights using the Xavier initialization scheme [42] while the biases were initialized with .01. As optimizer we used a Adagrad Optimizer [28] combined with a exponential decay learning rate ($LearningRate = .01$ and $DecayRate = .2$).

Convolutional Neural Network (CNN): Next, we used a CNN with 2 output neurons to also estimate pitch and yaw at the same time. As CNNs are designed for image-like data, this was the next obvious step for us. We used 3 convolution layer each with 2×2 max-pooling and a ReLu activation function. All convolutional layers have a filter size of 7×7 and

32, 72 and 160 filter banks respectively for the three layer, followed by 2 fully connected layers (FCLs). The first FCL uses a softplus[1] activation function. We used 2000 output neurons for the first FCL and initialized the weights using Xavier initialization scheme [42] while the biases were initialized with .01. As optimizer we used a Momentum Optimizer [145] using a momentum of .9 combined with a exponential decay learning rate ($LearningRate = .02$ and $DecayRate = .1$).

CNN + L2: To improve our first CNN, we applied L2 Regularization [6]. An L2 Regularization of .015 for the two FCLs performed best with an extra change in the network structure. We changed the filter size from 7×7 of all convolutional layers in comparison to the previous model to 6×6 as is yield better accuracy, see Figure 3.12 for a model structure.

CNN + L2 + BatchNorm: In the last model we added batch normalization [75] to CNN model with L2 Regularization. We used the same structure as the CNN with L2 Regularization model with enabled scale as well as an optimizer, exponential decay learning rate, and L2 Regularization.

3.3.3 Results

All results of our 2 baseline approaches (*k*NN and RF), 5 Neural Network (NN) approaches as well as the results of the best estimation for *reimplementation* and *pseudo implementation* using the features proposed by Xiao et al. [188] are presented in Table 3.3.

3.3.3.1 Feature-Based Approaches

We used a subset of our data set to train and test the *reimplementation* due to the limitations of GPs and the full data sat for the *pseudo implementation*.

[1] The softplus function is defined as $softplus(x) = log(1 + exp(x))$

	Pitch			Yaw			Overall
Method	**RMSE**	**MAE**	**SD**	**RMSE**	**MAE**	**SD**	**RMSE**
GP reimplementation of Xiao et al. [188]*	14.74	11.78	14.38	56.58	40.51	39.51	–
pseudo implementation of Xiao et al. [188]**	14.19	11.58	8.21	44.53	33.39	29.46	–
*k*NN	13.96	11.25	8.27	33.07	23.06	23.7	–
RF	12.99	10.24	7.99	28.55	20.89	19.46	–
DNN	13.05	10.25	8.07	27.10	19.53	18.79	–
Combined DNN	13.44	10.71	8.13	26.98	19.51	18.4	29.74
CNN	12.80	10.03	7.96	24.5	17.6	17.04	27.43
CNN + L2	12.8	10.09	7.88	24.19	17.62	16.58	27.16
CNN + L2 + BatchNorm	12.75	9.99	7.92	24.48	18.33	16.24	27.59

Table 3.3: The best results for all tested estimation models. Errors are reported in angular degree error. *) These results have been achieved with a small subset of the original data set (1.4%). **) For the reported values we used a DNN instated of a GP regression for the pitch estimation as the data set was to big for a GP.

3.3.3.2 Reimplementation

For the *reimplementation* we achieved a RMSE of 14.74° ($MAE = 11.78°$, $SD = 14.38°$) for pitch usning a GP and a RMSE of 56.58° ($MAE = 40.51°$, $SD = 39.51°$) for yaw using the simple heuristic.

3.3.3.3 Pseudo implementation

For the *pseudo implementation* we achieved a RMSE of 16.14° ($MAE = 13.2°$, $SD = 9.28°$) using *k*NN. The DNN estimator using features, which replaced the original GP, achieved a RMSE of 14.19°. Thus, the DNN using the features by Xiao et al. [188] outperforms the *k*NN with features by 13.7%.

The simple heuristic to estimate the yaw of the finger proposed by Xiao et al. [188] achieved a RMSE of 44.53°. Further, we achieved a RMSE of 31.77° ($MAE = 22.96°$, $SD = 21.96°$) when we use a kNN as an estimator. And a RF model achieved a RMSE of 31.75° ($MAE = 22.95°$, $SD = 21.94°$). Thus our baseline comparisons using kNN and RF both performed better than the simple heuristic by 28.7%.

3.3.3.4 Representation Learning Approaches

Next, we again used kNN and RF as a baseline estimator. However, now we are using the raw blob values. For the kNN baseline, we achieved a remaining RMSE of 13.96° for pitch and 33.07° for yaw. For the Random Forest (RF) baseline, we achieved a remaining RMSE of 12.99° for pitch and 28.5° for yaw. The RF using the raw blob outperforms the simple heuristic using features by 8.5% for pitch and by 36.% for yaw.

With our first with two separate DNNs, one for pitch and one for yaw, we achieved a RMSE of 13.05° for pitch and 27.10° for yaw. We further achieved similar results when using a DNN to predict pitch and yaw at the same time this resulted in an overall RMSE of 29.74°.

Then we used three different types of CNN. A simple CNN outperformed the DNN by 7.8%. We further were able to reduce the RMSE to 27.16° when using a CNN with L2 Regularization. However, when applying batch normalization to the previous model, the overall result dropped to a RMSE of 27.59°.

Thus the estimator with CNN and L2 Regularization performed best with an overall RMSE error of 27.16°. The error distribution for pitch is shown in Figure 3.10 and for yaw in Figure 3.11.

3.3.4 Discussion

In this work, we collected a data set automatically labeled by a motion capture system. In total, we used $457,268$ labeled samples to train our models. However, Figures 3.8 and 3.9 indicate an unequal distribution for pitch and yaw samples. The results we will present in Chapter 4 indicate that performing low pitch angles can be hard even in the yaw range observed in this section. Further, particapnts

stated that performing yaw angles outside of the range $-33.75°$ to $101.25°$ is significantly harder for right-handed people than performing yaw angles within the range. Since we had 31 right-handed participants, this explains the unequal sample distribution for the yaw samples.

Using our labeled data set, we evaluated the feature-based approach for the *GP reimplementation* as well as the *pseudo implementation*, and further presented multiple models including two baseline approaches (k-nearest neighbor (kNN) and Random Forest (RF)), and five different Neural Networks (NNs). In contrast to Xiao et al. [188], we used the raw capacitive image instead of feature engineering. Even the two baseline approaches using *representation learning* yield a lower estimation error than the two feature-based approachs.

In contrast to all other models, we trained the *GP reimplementation* with a subset of the data set which makes a real comparison hard. However, the RMSE for the GP pitch estimator is in the same range of the other models. On the other hand, the *SD* is 175% larger than the second worst pitch *SD*. Further, the simple heuristic for yaw performed worth for the GP reimplementation throughout all other yaw estimations. Additionally, the *SD* is the highest which is 134% larger than the second worst yaw *SD*.

A comparison of our results with the *pseudo implementation* of Xiao et al. [188] revealed that the *pseudo implementation* of the feature-based approach performed worse by 16.2% for pitch and 19.7% for yaw. Since our data set consists out of $457,268$ labeled samples, we have a large variance compared to their data set which consists of only 1,224 test samples. Further, they trained and evaluated their model in $15°$ steps while our model was trained and evaluated on a floating point level of precision. As shown in Figures 3.8 and 3.9, we cover the full pitch (from $0°$ to $90°$ degrees) and yaw range (from $-90°$ to $90°$) in $1°$ steps.

Both our baselines (kNN and RF) using the raw capacitive image outperformed our implementations of the feature-based approach proposed by Xiao et al. [188]. Using five different NNs, we showed that we could further improve the estimation accuracy. We started with two separate DNNs to predict pitch and yaw. We achieved similar estimation results using one combined DNN. Eventually, we used Convolutional Neural Networks (CNNs) to further improve the finger orientation

estimation accuracy of the combined DNN by 7.8% in RMSE. Overall, we reduced the pitch RMSE by 9.8% and the yaw RMSE by 45.7% in comparison to the best feature-based approach.

While this is a step towards a precise estimation of the finger orientation, there is still a remaining error in both pitch and yaw which could result in jitter. This could limit the usability and restrict the usage of finger orientation to non-precise input. One reason includes the limitation of the touch sensor. With a pixel size of $4.1 \times 4.1mm$, the capacitive image still has a low-resolution which restricts the performance of the estimation. While we removed the majority of the touchscreen noise, the remaining noise still affects the estimation precision negatively. This could be improved by using a more precise high-resolution touch sensor. Further, Williamson [179] showed that increasing the sensing range above the display surface enables, for instance, to detect if two fingers belong to the same hand. Hinckley et al. [65] used an increasing the sensing range to sense a finger before the accentual touch and thereby enables new interaction techniques. Both enables detection of the whole finger without actually touching the display which can be used to model the finger shape and thus also orientation. This technology is already available in commercial smartphones, such as the Samsung Galaxy S4 which has the Air View feature. A better sensing range was well as a higher resolution could improve the accuracy.

One limitation of our current model is that it is only trained with samples where the whole finger was captured by the touch sensor. Thus, we assume a drop in accuracy when touching close to the screen edges where only a part of the finger is visible. This should be investigated in further developments. Our data set could be used to train models which take edge inputs into account by cropping the images and thereby simulating edge inputs.

3.3.5 The Finger Orientation Data Set and Model

The data set collected in this section is freely available under a GPL license[1] and available on GitHub[2]. The data set contains the full capacitive image as well as

[1] `https://gnu.org/licenses/gpl-3.0.en.html`

[2] `https://github.com/interactionlab/Capacitive-Finger-Orientation-Estimation`

the labels for pitch and for yaw which we automatically labeled using a motion capture system. Further, we provide the scripts for prepossessing, training, and testing on GitHub. The prepossessing scripts include the blob detections using scikit-image's implementation[1] of *find contours* by Lorensen and Cline [109]. Training and testing scripts of the method using a CNN with L2 Regularization are also published. Finally, the model which performs best is released together with data set and code. The model can directly be deployed using TensorFlow Mobile[2] directly to mobile devices (Android and iOS) and prototyping platforms such as the Raspberry Pi. The data is published in CSV files, and the code is written in Python 3.6 using *TensorFlow* version 1.2.1. We later added a *Keras* version of the final model for easier deployment. However, this was not used to calculate the accuracies. The implementation is using the *TensorFlow* back end.

3.4 Summary

In this chapter, we presented approaches to answer RQ1: "How can finger orientation be detected?" We presented two approaches which both can determine the finger orientation when a finger touches a surface. Both approaches are complementary. While the depth camera approach is bulky, it can turn any flat surface into a surface with finger orientation input. The capacitive sensor approach is small; however, needs a specific sensor which allows to acquire the capacitive images. Moreover, when comparing the performance of both approaches, they perform similar in terms of pitch estimation. In contrast, the performance for yaw is more accurate when using the depth camera approach. However, the Machine Learing (ML) model of the capacitive sensor approach can be further improved retraining the model with more data. Therefore, we argue that both approaches are useful; however, in different scenarios. We see the depth camera approach outperform the other approach in prototyping scenarios while the capacitive sensor approach has advantages when it comes to portability and in-the-wild evaluation as well as consumer market mass-production.

[1] `http://scikit-image.org/docs/dev/api/skimage.measure.html#skimage.measure.find_contours`

[2] `https://www.tensorflow.org/mobile/`

4

Ergonomic Constraints

In this chapter, we investigate the ergonomic constraints when using finger orientation as additional input dimensions (RQ2). First, we study ergonomic constraints in a static tabletop scenario to understand the fundamental constrains of one-handed interaction. Second, we study ergonomic constraints in a two-handed interaction scenario. Here, we aim to understand how humans counteract their own ergonomic constraints by using both hands and arms to enable a wider range of finger orientations. Our findings indicate that pitch and yaw do significantly affect perceived feasibility to perform a touch action. Further, we found that there is a subset of yaw angles at which users are comfortable performing touch actions, the *comfort* zone. Outside of this zone is the *non-comfort* zone, where the touch interaction is perceived to require significantly more effort.

Parts of this chapter are based on the following publications:

S. Mayer, P. Gad, K. Wolf, P. W. Woźniak, and N. Henze. "Understanding the Ergonomic Constraints in Designing for Touch Surfaces." In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '17. Vienna, Austria: ACM, 2017, 33:1–33:9. ISBN: 978-1-4503-5075-4. DOI: 10.1145/3098279.3098537

S. Mayer, H. V. Le, and N. Henze. "Designing Finger Orientation Input for Mobile Touchscreens." In: *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services.* MobileHCI '18. Barcelona, Spain: ACM, 2018, 29:1–29:9. ISBN: 978-1-4503-5898-9. DOI: 10.1145/3229434.3229444

4.1 Ergonomic Challenges for Finger Orientation

While previous work has suggested compelling ways to use the finger orientation as input and technology that can sense the orientation is now available. However, commercial systems do not yet use finger orientation as part of the interaction. One reason is that the orientation of the finger is restricted by the physiological constraints of the user's hand. When a user touches a flat surface, there are finger orientations in which touch is uncomfortable or even impossible to perform. As not all finger orientations can be used for the input, it is essential to learn about users' restrictions when designing orientation-aware input techniques. We propose readdressing the ergonomics of single-finger touch on interactive surfaces by investigating how systems can effectively use different finger orientations. In this chapter, we look closely at different finger orientations to understand the limitations they impose on the touch action. We systematically vary pitch and yaw configurations (see Figure 1.1) to determine which finger orientations are optimal for touch interactions and when users find it feasible to perform the touch action. To that end, we conducted a study where participants rated the perceived feasibility of the touch action when the finger pitch and yaw were varied independently. In a controlled experiment, we used pitch stabilizers to test 4 pitch values and 16 equally spaced yaw angles. Our findings indicate that pitch and yaw do significantly affect perceived feasibility to perform a touch action. Further, we found that there is a subset of yaw angles at which users are comfortable performing touch actions, the *comfort* zone. Outside of this zone is the *non-comfort* zone, where the touch action is perceived to require significantly more effort, and some touch actions were found to be impossible to perform. Based on these results, we discuss design considerations for using finger orientation input in future applications.

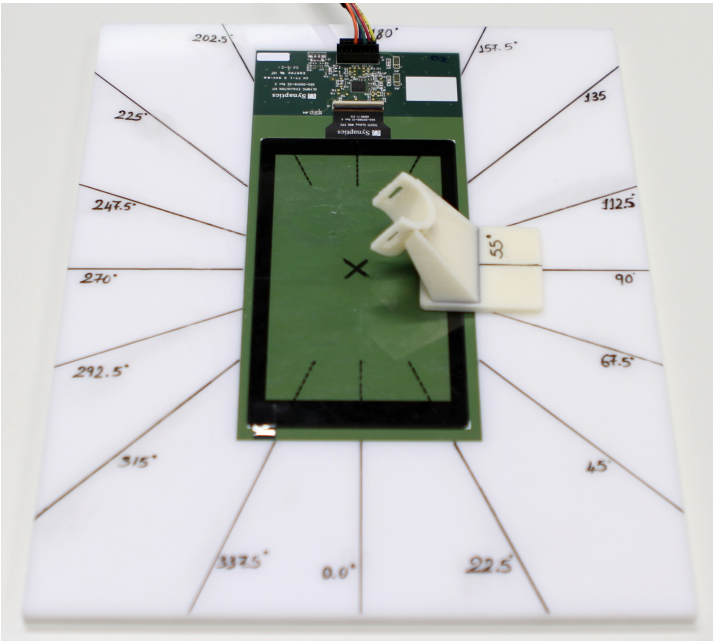

Figure 4.1: The study apparatus with the 3D printed 55° pitch stabilizer and the 16 yaw positions drawn on the touch surface.

4.2 Static Scenario

In the following sections, we investigate the static scenario where the touch surface is mounted flat in front of the user. Here, we present the details of our study design and the results of the experiment. Next, we interpret and discuss the implications of the data obtained. Finally, we present how our findings help to design future applications.

4.2.1 Hypotheses

Our study investigates the ergonomics of approaching a touch point with different finger orientations and is guided by the following three hypotheses:

Hypothesis 1 (**H1**): Even though users often use their non-dominant hand, e.g. in encumbered situations, users perceive touch actions performed with their dominant HAND as more feasible. Users prefer operating devices with their dominant hand, and we expected this will influence our results.

Hypothesis 2 (**H2**): Changes in finger PITCH would affect feasibility RATINGS. We decided to explore this hypothesis as past work provided little evidence on how pitch values affected input feasibility.

Hypothesis 3 (**H3**): The more the finger YAW would diverge from the direction parallel to the user's arm, the lower the feasibility RATING would be. We noted that increased twist in the wrist was expected to decrease feasibility. While verifying this hypothesis, we endeavored to identify how much twist was allowed while still producing a feasibility RATING suitable for designing interaction techniques.

4.2.2 Study

In our study, we systematically manipulated the finger pitch and yaw while performing a touch action. To study our three hypotheses, we conducted the study in a controlled environment with a number of constraints to ensure the validity of the study. Our goal was to observe the touch action as an atomic task. Therefore, we artificially restricted the participants' finger posture to prevent them from subconsciously adjusting their hand, which would result in a larger input range. Moreover, movements of the participant's body would have caused a larger input range. Therefore, participants were not allowed to move either the apparatus or their chair.

To investigate the effect of finger orientation on the feasibility of a touch action, as an atomic task we explore the full potential input range. We used 4 pitch angles and 16 yaw angles, each with a step size of 22.5°.

4.2.2.1 Study Design

In a repeated measures experiment, we asked participants to perform touch actions with their index finger. We asked them to rate the feasibility of the touch action resulting in the dependent variable RATING. Feasibility, in this context, was

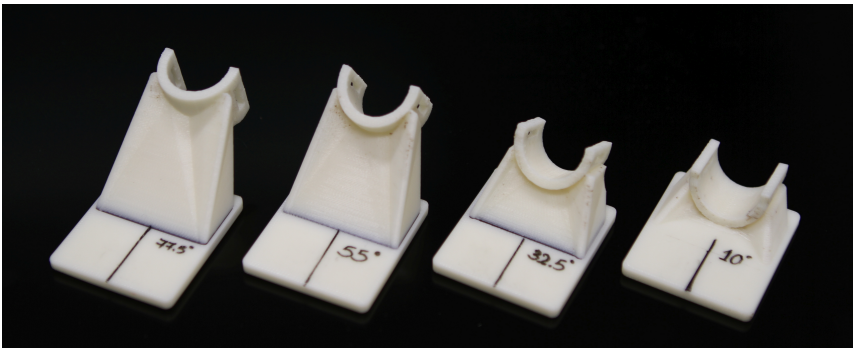

Figure 4.2: The four pitch stabilizers we used in the study to limit PITCH to 77.5°, 55°, 32.5°, and 10° presented from left to right.

defined as the effort required to perform the touch action. The experiment was conducted with three independent variables: PITCH and YAW of the index finger, as well as HAND. We used 10°, 32.5°, 55°, and 77.5° for PITCH. We did not investigate angles steeper than 77.5°, due to findings by Xiao et al. [188] who stated that a pitch of 90° cannot be detected and performed with long nails. For YAW, we covered the full 360° range resulting in 0.0° to 337.5° with 22.5° steps. All combinations were tested with the index finger of the right and the left HAND. Thus, we used a PITCH $\times$ YAW $\times$ HAND $= 4 \times 16 \times 2$ study design resulting in 128 conditions.

4.2.2.2 Apparatus

Our apparatus design aimed to maximize the control over the independent variables. Xiao et al. [188] stated that it is difficult to reliably ensure that participants can touch a screen with a particular pitch. In their study, Xiao et al. used laser-cut plastic wedges to align the finger at a particular pitch. The wedges, however, were removed during the recording process, which influenced the accuracy. To ensure that participants perform the touch actions with a particular pitch, we used 3D-printed pitch stabilizers. We manufactured pitch stabilizers with a PITCH of 10°, 32.5°, 55°, and 77.5° as presented in Figure 4.2. Participants had to place the finger on the pitch stabilizer while performing a touch action. The four pitch

stabilizers ensured that participants performed a touch action with a given pitch. Further, the pitch stabilizers ensured that participants did not vary the roll of the finger during touch acquisition.

We used a touch-sensitive sensor by Synaptics, to ensure that the participants touched the surface. The touch layer was surrounded by a white plastic frame to level the area around the touch layer (see Figure 4.3). This resulted in a flat surface that enabled secure positioning the pitch stabilizer on the sensor. We marked the center of the touch sensor with a permanent marker. We further marked the 16 input yaw angles with a line on the surface and wrote the angle next to the line (see Figure 4.1). The touch sensor was fixed on a desk to ensure that participants could not move it.

We employed a tablet to guide participants through the study. During the study, an application running on the tablet showed the hand, the pitch, and the yaw that should be used for the next trial. The application randomized the order of yaw and pitch. Participants were asked to rate the feasibility of the performed touch action with a slider control on a scale with 100 steps from "*easy*" to "*hard*". Using continuous rating scales that have a long history in psychophysical measurement and enables a robust evaluation [167]. Further, we choose a slider with no ticks as Matejka et al. [116] showed that ticks influence the distribution of the results. Additionally, the application gave the opportunity to tick a checkbox indicating that the input was not feasible. The checkbox enabled distinguishing between very hard but possible and physically impossible touch actions.

4.2.2.3 Procedure

After welcoming participants, we explained the study and asked them to give consent. We then asked them to take a seat in a chair which was aligned with the center of the apparatus. We fixed the position of the chair and asked participants to neither move the chair nor the touch layer of the apparatus during the study. We further explained to them how to place and use the pitch stabilizer. After the participants felt comfortable using the apparatus, we explained how to use the rating scale and the tick box, that is if they could not perform the touch action they should tick the box. Further, we explained that they should rate the effort required to perform the touch action and then explained the labels on the scale

Figure 4.3: The apparatus we used in our study, showing the tablet, the touch layer and one of the pitch stabilizer while one participant touches the touch surface.

in detail. We explicitly mentioned that *easy* meant that little to no effort was required to perform that touch action whereas *hard* described an action that was near impossible to complete.

Next, we started the main part of the study. The tablet showed PITCH, YAW, and HAND that should be performed next. Participants were asked to perform the touch action in the center of the sensor (the center was marked as shown in Figure 4.3) three times using the given PITCH, YAW, and HAND. We asked the participants to slide the finger in the guiding rail provided by the stabilizer. At the end of each condition, participants had to provide a feasibility RATING on the slider control on the tablet. Participants first performed the touch actions using all combinations of PITCH and YAW with one hand followed by the other hand, with the order of HAND being counter-balanced. Within HAND, the PITCH condition was randomized, and within the PITCH condition, the YAW condition was randomized to avoid participants from changing the stabilizer often. After the participants had performed all conditions, we thanked them for their volunteer participation.

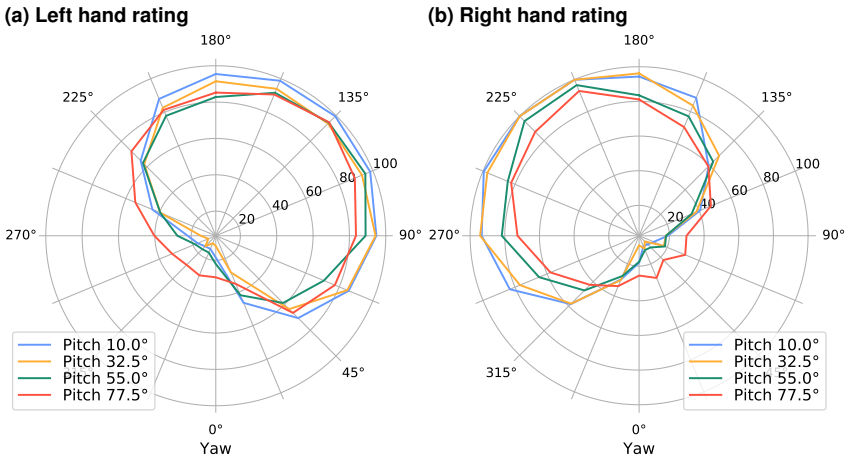

Figure 4.4: The average feasibility RATING (from 0 = "*easy*" to 100 = "*hard*") for the different PITCH inputs.

4.2.2.4 Participants

We recruited participants from an internal university self-volunteer pool. 10 male and 9 female participants agreed to take part in the study. These participants were between 22 and 44 years old ($M = 25.9$, $SD = 2.7$). Of all participants 16 participants were right-handed, 3 left-handed and none of the participants were ambidextrous. One of the right-handed participants did not follow the procedure of the study. Therefore, we discarded the data collected from this participant.

4.2.3 Results

We collected 2304 ratings from 18 participants. Out of the 2304 rated conditions, 485 (21.1%) were marked by the participants as not feasible to perform. All inputs that the participants marked to be not feasible where considered to be "*hard*" (100 points) for the analysis.

We applied the Aligned Rank Transform (ART) [180] procedure to the feasible RATINGS, using the ARTool toolkit[1] to align and rank our data.

We conducted a three-way analysis of variance (ANOVA) to determine whether the independent variables significantly influenced the perceived feasibility of performing the touch action. Our analysis revealed significant main effects for PITCH, YAW, and HAND on feasibility ($F_{3,2176} = 5.413$, $p < .005$; $F_{15,2176} = 196.194$, $p < .001$.; $F_{1,2176} = 22.701$, $p < .001$, respectively). Further, we found significant two-way interactions between PITCH $\times$ HAND and YAW $\times$ HAND ($F_{3,2176} = 3.027$, $p = .028$; $F_{15,2176} = 147.566$, $p < .001$, respectively). However, there was no significant two-way interaction between PITCH $\times$ YAW ($F_{45,2176} = 1.179$, $p = .194$). Lastly, we found a significant three-way interaction between PITCH, YAW, and HAND ($F_{45,2176} = 2.361$, $p < .001$). Figure 4.4 presents the distribution of feasibility RATINGS for all YAWS and both HANDS. Consequently, we employed further comparisons to investigate how the different variables influenced the results.

We calculated a Sine regression to predict the RATING based on YAW. We found a regression equation with $R^2 = .991$ for the right index finger and $R^2 = .978$ for the left index finger. The predicted RATING is equal to

$$\text{RATING} = 54.8 - 44.5\sin(\text{YAW} + 0.9) \tag{4.1}$$

for the right hand and

$$\text{RATING} = 58.4 + 43.1\sin(\text{YAW} - 0.8) \tag{4.2}$$

for the left hand with YAW in radians, see Figure 4.5.

Next, we investigated which YAW angles produced touch actions that are perceived as impossible to perform. For the right index finger, the participants stated 244 out of 1152 (21.2%) times that touch was not feasible using the given orientation and 241 out of 1152 (20.9%) times for the left index finger. For the right hand, 99.18% of trials that were perceived to be impossible fell into the range from 112.5° to 315.0°. In the case of the left hand, 100% of the impossible

[1]The Aligned Rank Transform (ART) tool by Wobbrock et al.
`http://depts.washington.edu/madlab/proj/art/index.html`

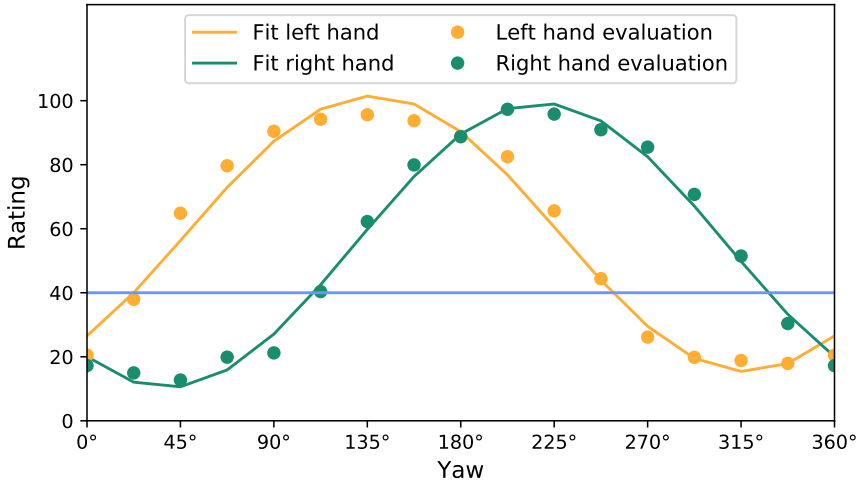

Figure 4.5: The average feasibility RATING (from 0 = "*easy*" to 100 = "*hard*") for the different YAW inputs averaged over all PITCHes. The figure also shows the fitted sin curve representing the RATINGS. The blue line indicates the threshold between *comfort* and *non-comfort* zones.

trials were reported in the range from 45.0° to 247.5°. Considering that the RATING is harder to perform in some input zones, we defined a threshold of 40 to mark the range where the trail was rated as impossible from the rest, as explained next. Consequently, we observed that the YAW space could be divided into two zones, which we named the *comfort* and *non-comfort* zones, as shown in Figure 4.6.

Table 4.1: One-way RM-ANOVAs to determine if the RATING is depended on PITCH within zones and HAND.

Zone	HAND	df	F	p
comfort	right	3, 428	9.385	<.001
comfort	left	3, 428	9.436	<.001
non-comfort	right	3, 716	9.539	<.001
non-comfort	left	3, 716	6.049	<.001

Further, we noted that for the right HAND, the *comfort* zone ($M = 51.89$, $SD = 36.65$) was rated significantly different from the *non-comfort* zone ($M = 62.95$, $SD = 36.29$) by conducting a Welch Two Sample t-test ($t_{900.61} = -4.98$, $p < .001$). This was confirmed for the left hand (*comfort* zone: $M = 45.04$, $SD = 36.69$; *non-comfort* zone: $M = 60.92$, $SD = 38.91$) as well ($t_{949.77} = -6.954$, $p < .001$). Therefore, we choose the threshold of 40 for the RATING to divide the two zones.

The *comfort* zone for the right HAND ranges from 326.25° to 101.25° and the *comfort* zone for the left HAND ranges from 258.75° to 33.75°. Therefore the span of both *comfort* zones is equal to 135.0° for both hands and two *comfort* zones overlap by 67.5° Thus the *non-comfort* zones are 225.0° wide.

We used four one-way RM-ANOVAs to investigate whether PITCH significantly affected the feasibility RATING in the two zones and HAND. As Table 4.1 shows, we found significant effects; the ratings are presented in Figure 4.4. Further, we did the same for YAW; and results are presented in Table 4.2.

4.2.3.1 Left-handed Participants

We also analyzed the data produced by the 3 left-handed participants. We collected 384 ratings from 3 left-handed participants. Out of the 384 ratings, 50 (13.0%) were rated not feasible. Figure 4.7 compares the average RATING for all YAW conditions between left- and right-handed participants. The data suggests that left-handed participants reported RATING similar to right-handed participants. Thus this indicates that the findings are valid irrespective of the dominant hand.

Table 4.2: One-way RM-ANOVAs to determine if the RATING is depended on YAW within zones and HAND.

Zone	HAND	df	F	p
comfort	right	5, 426	6.439	<.001
comfort	left	5, 426	8.505	<.001
non-comfort	right	9, 710	55.513	<.001
non-comfort	left	9, 710	49.397	<.001

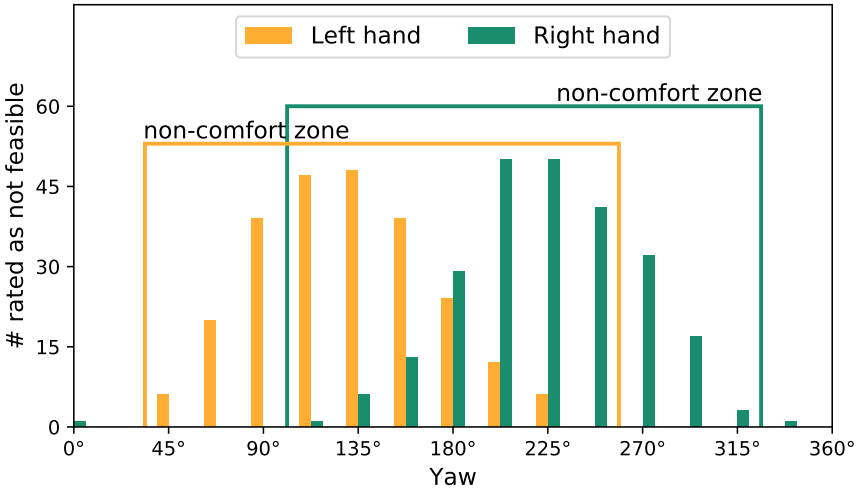

Figure 4.6: The bars represent how often a yaw angle was rated as not feasible to perform.

4.2.4 Discussion

Our results show that finger orientation has a significant effect on the perceived feasibility of touch actions. As expected, participants perceived actions performed with the dominant HAND as more feasible than those performed with the non-dominant hand. Thus, the result of the initial three-way ANOVA confirms **H1**.

Our analysis revealed a significant effect of PITCH on the feasibility RATING. This indicates that the feasibility of performing touch actions is influenced by finger PITCH confirming **H2**. This is in contrast to Wang and Ren [171] who found no difference in accuracy between vertical and oblique touch. Furthermore, the results indicate that flat angles are preferred when touching in the *comfort* zone while steep angles are overall rated to be easier when operating in the *non-comfort* zones. Owing to a significant requirement to twist the finger, higher ratings in *comfort* zone to otherwise are understandable.

Further, our analysis revealed a significant effect of YAW on the feasibility RATING. In particular, we found that the distribution of RATINGS can be approximated by a sine curve. This shows that the perceived feasibility of the touch

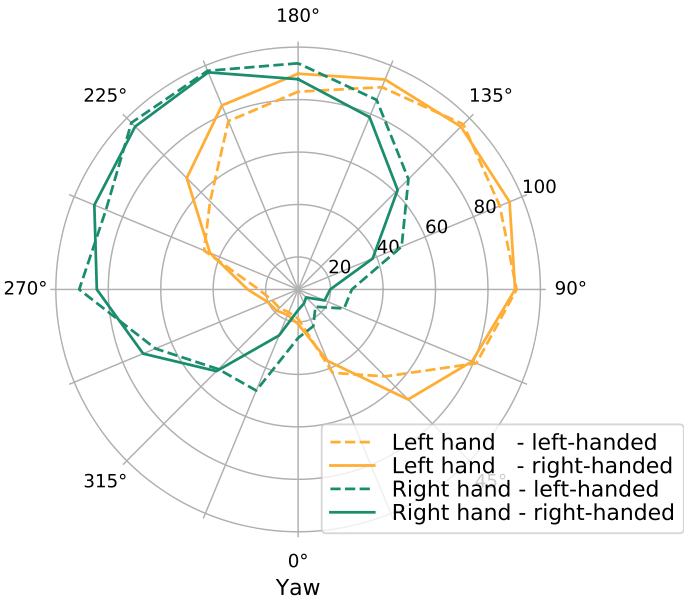

Figure 4.7: Comparison of HAND in respect to the handedness of the participants, showing the average value per YAW.

action increases steeply while the finger diverts from the parallel-to-arm direction. We also observed that most YAW values could render the touch action impossible (as evidenced by the existence of *comfort* and *non-comfort* zones). Consequently, the range within which yaw input is feasible is highly restricted and a larger YAW results in decreased feasibility RATINGS which confirms **H3**.

4.2.5 Design Considerations

Here, we chart how our findings influence the design of future single-finger input techniques for interactive surfaces in the form of five design considerations.

Avoid entering the *non-comfort* yaw zone: The *non-comfort* zones cover 225° out of 360° of the possible input space for both hands and are therefore much larger than the *comfort* zones. The *comfort* and the *non-comfort* zones significantly differ in perceived feasibility when touching a surface

with different finger orientations. Consequently, future designs of input techniques should not require the user to use orientations that fall into the *non-comfort* zones. Requiring input in the *non-comfort* zone creates a possibility for the task to be perceived as impossible. Thus, tasks like widgets that require rotating with a single finger should be avoided at all costs.

Range for effective yaw input depends on the hand: While many interactive surfaces can detect from which angle the user's hand is approaching, future designs must take that into account while designing yaw gestures. The yaw rotation possibilities depend on the hand used. If the interactive surface cannot detect which hand is being used, yaw gestures should be limited to the 45°-wide overlap in the *comfort* zones of the left and right hands to ensure that the gesture is feasible to perform with both hands.

Make use of pitch-based techniques for contextual features: We have shown that touch at different pitch angles is perceived as varying in feasibility. Previous work reported influence on accuracy. In contrast to yaw, the range of feasible pitch input is the same for the left and the right hand. This suggests that there is a design space for designing interactions based on finger pitch for interactive surfaces. Similarly to touch pressure techniques (e.g. Apple's 3D Touch), finger pitch could be used to activate additional functions such as contextual menus.

Make use of pitch-based techniques for modal interaction: As different pitch angles can be perceived and differentiated well, using these for interaction can affords different modes in touch-based reading devices like ebook readers and tablets. For example, pitch based techniques could offer an alternative mode (to time or pressure) when one needs to parse complex textual data. Most common techniques for parsing text include note-taking [48], annotating [49], and insight generation from these notes as a solo or collaborative activity [50]. Varying the pitch angles can activate the mode to highlight text, or annotate it with notes.

Use edges of the comfort zone for safety-critical input: Our results show that the perceived feasibility rating rises as the finger divert from the parallel-

to-arm direction. Future designs could exploit this observation by using higher yaw angles for sensitive input. For example, when confirmation to restore factory settings is required, the user could be asked to perform a 67.5° yaw rotation. While the task would still fall in the *comfort* zone (and thus be feasible), it would require more effort than a single tap thus limiting possible slips.

A combination of pitch and yaw can also be used to offer a second dimension to afford sharing or disclosure. For example, sharing digital notes has been shown to improve performance in critical tasks [46, 185]. Setting the mode for digital notes to be private or transparent for public consumption could be done by varying pitch (simultaneously or in succession) with an angular yaw movement. Further work is required to address the opportunities and limitations resulting from this approach.

Explore the benefits of pitch when unsure about yaw: Our results show a potential for future designs to use pitch input when yaw values may fall outside of the *comfort* zone. This may be the case when multiple users use a single touch device e.g. when collaboratively browsing photos on a tablet lying on a table. Further, yaw is often limited when users are likely to use one-handed input e.g. while shopping. Given that appropriate sensing is available, pitch input may enable effective two-dimensional navigation even when the finger approaches the touch surface at a yaw angle outside of the *comfort* zone. Consequently, we suggest enabling pitch-based navigation in scenarios when yaw-based techniques are possibly restricted.

4.2.6 Limitations

The study used a highly controlled setting, which ensured that neither the participant nor the device was moved. Thus, our results can only be directly applied for iteraction with to stationary devices. Allowing the user to move the device or allowing the user to move around the device would increase the range of feasible inputs, but would increase the complexity of the interaction and the time required to interact. We, therefore, believe that it is advisable that users should not be required to interact in the *non-comfort* zone. However, as this study is a first

attempt to investigate the feasibility of a single touch action with varies pitch and yaw, the aim was to investigate the core limitations of using pitch and yaw as an input dimension.

Our study mainly focused on right-handed participants. A larger number of left-handed participants would be required for conducting statistical analysis of data from left-handed users. However, we assume that the *comfort* zone is similar for left- and for right-handed users. This is supported by the results of the three left-handed participants. While we cannot be certain that there are no differences, the similarity between left- and right-handed participants suggests that potential differences are small.

Our investigation is limited to pitch and yaw angles. We explicitly limited roll variation by using pitch stabilizers. Existing interaction techniques already use roll as an input source. For instance, in Apple's iOS, it is possible to roll the finger for precise text cursor manipulation. *Fat thumb* interaction by Boring et al. [10] used the pitch and the roll of the finger for pan and zoom operations. While the roll range is highly limited by the arm's kinematic chain, it still requires further investigation.

4.3 Two-Handed Interaction Scenario

To enable finger orientation input on the go, we investigate the use of finger orientation in a two-handed smartphones scenario, see Figure 4.8. In detail, we study the ergonomic constraints of finger orientation input for mobile devices. While we in Section 4.2 investigated finger orientation in a static, restricted tabletop scenario, we extend the work to study how users move the device and how this affects what can comfortably be used for a two-handed interaction scenario. We conducted a second study and asked 20 participants to rate the comfort and feasibility of touch actions. Participants aligned their index finger with given pitch and yaw angles while holding the device with their second hand. They were allowed to freely move their finger and the device while we ensured that they could still perceive content on the screen.

4.3.1 Hypotheses

Our study investigates the ergonomics of approaching a touch point with different finger orientations and is guided by the hypotheses described below.

In the first study presented in Section 4.2 participants rated the feasibility of finger orientation as input from "*easy*" to "*hard*". They found that the *comfort* zone is smaller than the *non-comfort* zone using the finger orientations which were feasible. In the study, the touch surface was flat on a table, and by allowing the user to move and rotate the touch surface, we expect the participants to compensate exhausting body movements by moving and rotating the device. Thus, we formed the following hypotheses:

Hypothesis 1a (**H1a**): Finger orientation input for two-handed smartphone interaction is easier than for tabletop interaction.

Hypothesis 1b (**H1b**): Finger orientation input for two-handed smartphone interaction has a larger *comfort* zone than for tabletop interactions.

Hypothesis 1c (**H1c**): No finger orientation is infeasible when using both hands to interact with a smartphone.

The finger orientation movement will affect the orientation of the smartphone. Consequently, we infer the following:

Hypothesis 2a (**H2a**): The smartphone orientation varies more in the *comfort* zone than in the *non-comfort* zone.

Hypothesis 2b (**H2b**): The smartphone orientation is reflectively symmetric based on the hand of interaction.

4.3.2 Study

To investigate the ergonomic constraints of using the finger orientation as an input for mobile devices we conducted a study with 20 participants. To test the hypotheses, participants were asked to perform touch actions while we systematically manipulated PITCH_{Finger} and YAW_{Finger} of the orientation of the finger in relation to the touch surface. Participants were asked to perform the touch action with

Figure 4.8: A participants performing a 32.5° pitch and 45° yaw input with the left hand while being equipped with our 3D printed tracking parts.

one hand while holding the touch surface, a smartphone, with the other hand. We recorded the orientation of the participant's finger, and the phone with a high precision motion tracking system.

4.3.2.1 Study Design

We used a within-subject design. We asked participants to perform touch actions with their index finger and rate the feasibility of the touch action with the dependent variable RATING. Our overall study design follows the design described in Section 4.2. We use the same independent variables with exactly the same levels. We used the same three independent variables: PITCH_{Finger}, YAW_{Finger}, and HANDS. We used 10°, 32.5°, 55°, and 77.5° for PITCH. For YAW, we covered the full 360° range resulting in 0.0° to 337.5° with 22.5° steps. All combinations of PITCH_{Finger} and YAW_{Finger} were tested with the index finger of both HANDS. Thus, we used a $\text{PITCH}_{Finger} \times \text{YAW}_{Finger} \times \text{HAND} = 4 \times 16 \times 2$ study design

Figure 4.9: The setup with the 6DOF tracking system, the four pitch stabilizer on the left and the Motorola Nexus 6 with markers.

resulting in 128 conditions. In addition to the dependent variable RATING, we recorded the smartphone orientation (PITCH_{Phone}, ROLL_{Phone}, and YAW_{Phone}) as additional dependent variables.

In contrast to the first study, participants were allowed to freely move the touch surface, their finger and their body to perform the touch action.

4.3.2.2 Apparatus

The apparatus consists of a Motorola Nexus 6 running the study application, a 6DOF tracking system (see Figure 4.9), and four pitch stabilizers (see Figure 4.11). The application shows the next to perform touch action as well as a rating scale where participants had to rate the feasibility of the touch action. To ensure that participants were able to read content on the screen we presented them a word

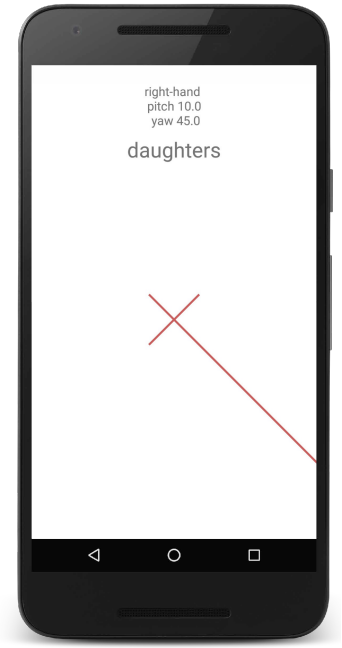

Figure 4.10: The study app is showing instructions to perform a 10° pitch and 45° yaw input at the red crosshair while remembering the word daughters which is displayed in the upper half.

which they had to remember and make a one out of three choices when the rating scale was presented. We used all nouns out of the phrase set by MacKenzie and Soukoreff [112]. However, we removed the plural form of the noun if the singular version was also in the phrase set.

The study application first presented a red crosshair in the center of the screen with one longer line to indicate the yaw orientation participants had to perform, see Figure 4.10. Then, the application presented a rating scale where participants had to rate the feasibility of the performed touch action with a slider control on a scale with 100 steps from "*easy*" to "*hard*". As in Section 4.2 we also added

(a) Stabilizers **(b) Rendering**

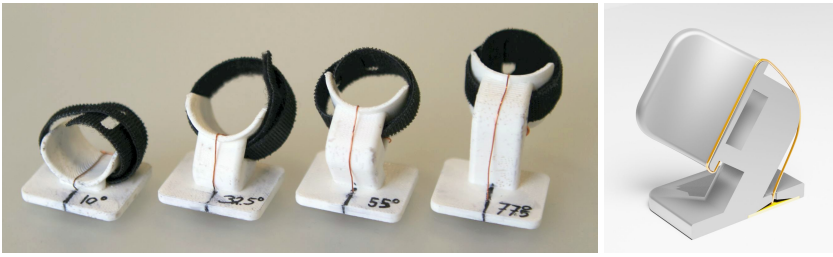

Figure 4.11: (a) The four pitch stabilizers with the copper plate and the wire, we used in the study to limit PITCH to 77.5°, 55°, 32.5° and 10° presented from left to right. (b) A CAD model of a pitch stabilizer, revealing the wiring and the copper plate in the base.

the opportunity to tick a checkbox indicating that the input was not feasible. The checkbox enabled the participants to distinguish between very hard but possible and physically impossible touch actions.

To track the phone and finger, we used a high precision marker-based 6DOF tracking system. The system consisted of 8 *OptiTrack Prime 13W* cameras. After calibration, the system reported a residual mean error of *.2mm*.

To guarantee an accurate pitch angle we manufactured pitch stabilizers similar to the ones used in Section 4.2, with a PITCH_{Finger} of 10°, 32.5°, 55°, and 77.5° as presented in Figure 4.11. However, we further improved the first design to enable tracking of the stabilizer through the touchscreen. Therefore we integrated a copper plate into the base of the stabilizer and an electric wire from the copper plate to the slide of the stabilizer where the participant's finger touches the wire. This generated a touch event underneath the stabilizer similar to the WebClip by Kubitza et al. [87] but without any electronic circuit similar to Wolf et al. [183]. Additionally, we added a velcro fastener to allow free movements of the participants (see Figure 4.11).

4.3.2.3 Procedure

We followed the instructions and procedure which we used in Section 4.2. After welcoming a participant, we explained the procedure of the study and asked them

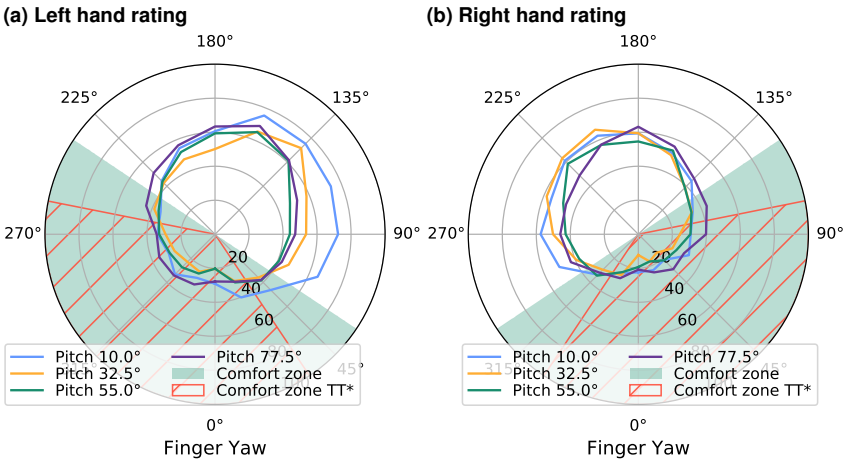

Figure 4.12: The average feasibility RATING (from 0 = "*easy*" to 100 = "*hard*") for the different PITCH$_{Finger}$ inputs. The green areas represent the *comfort* zone in a two-handed smartphone scenario. * the red striped areas represent the *comfort* zone for tabletops as presented in Section 4.2.

to fill an informed consent. Then we introduced them to the system. We further explained that they had to hold the phone with one hand while touching the screen with the other. We explained that they had to touch the red crosshair and align the finger yaw orientation with the long red line. Participants had to touch on the red crosshair three times. To ensure they had visual contact we extended the procedure to also read one word on the screen. They had to remember the word and then rate the input feasibility. Here we explain in detail how to understand the scale to match it. The application presented the question *How feasible was it to perform the touch action?* Additionally, we explained the meaning of "*easy*" and "*hard*" as defined in Section 4.2 as *the effort required to perform the touch action*. After participants were familiar with the procedure, we started the app to collect demographic data and initialize the randomization. Then we equipped the participants with the finger marker and the pitch stabilizer needed for the

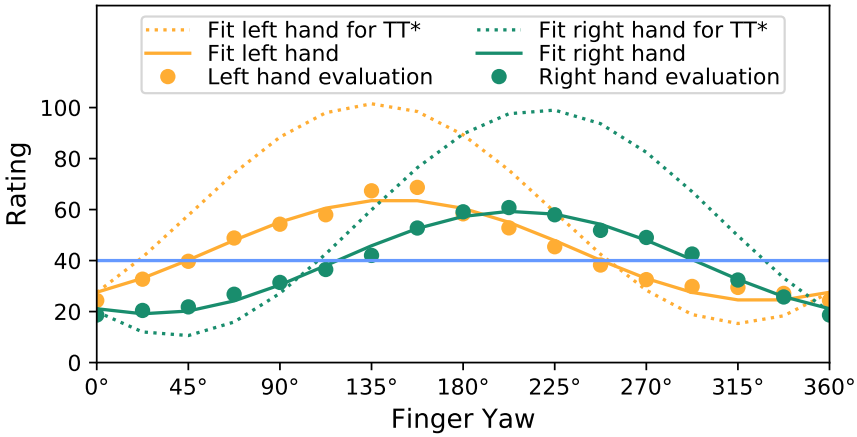

Figure 4.13: The average feasibility RATING (from 0 = "*easy*" to 100 = "*hard*") for the different YAW_{Finger} inputs averaged over all PITCH_{Finger}. The figure also shows the fitted sin curve representing the RATINGS. The blue line indicates the threshold between *comfort* and *non-comfort* zones as defined in Section 4.2. * approximated rating for tabletops as presented in Section 4.2.

condition, see Figure 4.8. After each condition, a pop-up told the participants to change the condition settings, and here the experimenter helped to switch the stabilizer.

4.3.2.4 Participants

We recruited participants from our university's volunteer pool. In total, 20 participants took part in the study (14 male, and 6 female). The age range was between 20 and 27 years ($M = 23.7$, $SD = 1.9$). All participants were right-handed, and none had locomotor coordination problems. We reimbursed the participants with 10 €.

4.3.3 Results

We collected 2,560 ratings from 20 participants. The average RATING was 41.8 ($SD = 24.7$). From our 2,560 ratings, none was marked by the participants as

not feasible to perform. Further, 41 (1.6%) of the words were wrongly selected by participants. The RATING for these wrongly selected words was $M = 56.4$ ($SD = 24.2$). To ensure that only samples were going into the analysis where the participants had been able to read the text, we removed all samples with wrongly selected words.

4.3.3.1 Rating

To conduct a RM-ANOVA, we applied the ART [180] procedure to the feasible RATINGS, using the ARTool toolkit[1] to align and rank our data.

We conducted a three-way RM-ANOVA to determine whether the independent variables significantly influenced the perceived feasibility of performing the touch action. Our analysis revealed significant main effects for PITCH_{Finger} , YAW_{Finger}, and HAND on feasibility ($F_{3,2371} = 18.15$, $p < .001$; $F_{15,2371} = 81.17$, $p < .001$; $F_{1,2371} = 38.45$, $p < .001$, respectively). Further, we found significant two-way interaction between $\text{YAW}_{Finger} \times$ HAND ($F_{15,2371} = 29.37$, $p < .001$). However, there were no significant two-way interactions between $\text{PITCH}_{Finger} \times$ HAND and $\text{PITCH}_{Finger} \times \text{YAW}_{Finger}$ ($F_{3,2371} = 1.598$, $p = .19$; $F_{45,2371} = .942$, $p = .58$, respectively). Lastly, we found a significant three-way interaction between PITCH_{Finger} , YAW_{Finger}, and HAND ($F_{45,2371} = 2.08$, $p < .001$).

Figure 4.12 presents the distribution of feasibility RATINGS for all finger YAWS and both HANDS. We employed further comparisons to investigate how the different variables influenced the results.

We calculated a sine regression to model the RATING based on YAW_{Finger}. Therefore we can model the rating for the right hand using

$$\text{RATING} = 39.16 - 20.17 \sin(\text{YAW}_{Finger} + 1.11) \tag{4.3}$$

and the left hand using:

$$\text{RATING} = 43.91 + 19.7 \sin(\text{YAW}_{Finger} - .98) \tag{4.4}$$

[1]The Aligned Rank Transform (ART) tool by Wobbrock et al.
`http://depts.washington.edu/madlab/proj/art/index.html`

The fitness for the right hand is $R^2 = .98$ and for the left hand $R^2 = .96$. We compared our functions with the function in Section 4.2 using t-tests. For the left hand model functions there was a significant difference in the modeled RATING for our new function ($M = 42.1$, $SD = 13.6$) and the tabletop function ($M = 56.2$, $SD = 29.5$); ($t_{15} = -3.2$, $p = .006$). For the right hand model functions there was also a significant difference in the modeled RATING for our new function ($M = 38.9$, $SD = 15.5$) and the tabletop function ($M = 54.8$, $SD = 34.4$); ($t_{15} = -3.2$, $p = .005$).

In Section 4.2 we divided the YAW_{Finger} input space into a *comfort* zone and a *non-comfort* zone. We argued for their split based on input rated as not feasible to perform. None of our participants rated a single input as not feasible; however, our results as presented in Figures 4.12 and 4.13 follow the same trend. Consequently, we used the same threshold of 40 to divide the *comfort* zone and a *non-comfort* zone.

The *comfort* zone for the right HAND ranges from 303.75° to 123.75° and the *comfort* zone for the left HAND ranges from 236.25° to 56.25°. Therefore the span of both *comfort* zones is equal to 180.0° for both hands and two *comfort* zones overlap by 112.5°. Thus the *non-comfort* zones are also 180.0° wide.

4.3.3.2 Phone Orientation

We matched the motion tracking data and the touch data using the timestamps of the touch events and the motion tracking data. We filtered all samples where the

Table 4.3: One-way RM-ANOVAs to determine if the phones' orientation is depended on ZONE within HAND.

Axes	HAND	df	F	p
PITCH_{Phone}	right	1, 19	38.47	<.001
PITCH_{Phone}	left	1, 19	30.21	<.001
ROLL_{Phone}	right	1, 19	18.31	<.001
ROLL_{Phone}	left	1, 19	25.38	<.001
YAW_{Phone}	right	1, 19	7.61	.012
YAW_{Phone}	left	1, 19	14.54	<.002

time distance was larger than 30ms. This resulted in 177 (2.36%) filtered samples. The remaining time difference was $M = .36ms$ with an SD of 1.98. Thus, the following analysis is based on the remaining 7,324 touch samples.

In the following the 0° orientation of the phone for all 3 axes (PITCH_{Phone}, ROLL_{Phone}, and YAW_{Phone}) is defined as the phone laying flat in portrait mode with the screen up in front of the participant. Further, the rotation direction of the axes are defined to be positive when rotating clockwise and negative when rotating counterclockwise.

We conducted three one-way RM-ANOVAs for each HAND to determine whether ZONE within HAND significantly influenced the orientation of the phone (PITCH_{Phone}, ROLL_{Phone}, and YAW_{Phone}). As Table 4.3 shows, we found significant effects for all six one-way RM-ANOVAs. The orientations are presented in Figure 4.14.

Lastly, we mirrored the data of the phone orientation for the left-handed interaction, resulting in a dataset that mimics right-hand usage data, as shown in Figure 4.15. We first modeled the orientation of the phone using a sine function, resulting in an average R^2 of .83, for PITCH_{Phone}, ROLL_{Phone}, and YAW_{Phone} R^2 is .83, .79, and .83 respectively. We then modeled the orientation with a skewed Sinus function, a Clausen function [20].

$$S_n(x) = \sum_{k=0}^{\infty} \frac{sin(kx)}{k^n} \tag{4.5}$$

To fit the skewed Sinus function to the data, we added fitting parameters to stretch or compress the function if needed. We again used ordinary least squares to estimate the fitting parameters a to e for our fitting function:

$$fit(x) = aS_b(c(x-d)) + e \tag{4.6}$$

Using a skewed Sinus function we achieved an average fit of $R^2 = .89$, for PITCH_{Phone}, ROLL_{Phone}, and YAW_{Phone} R^2 is .85, .86, and .96 respectively. The fitted functions are presented in Figure 4.15.

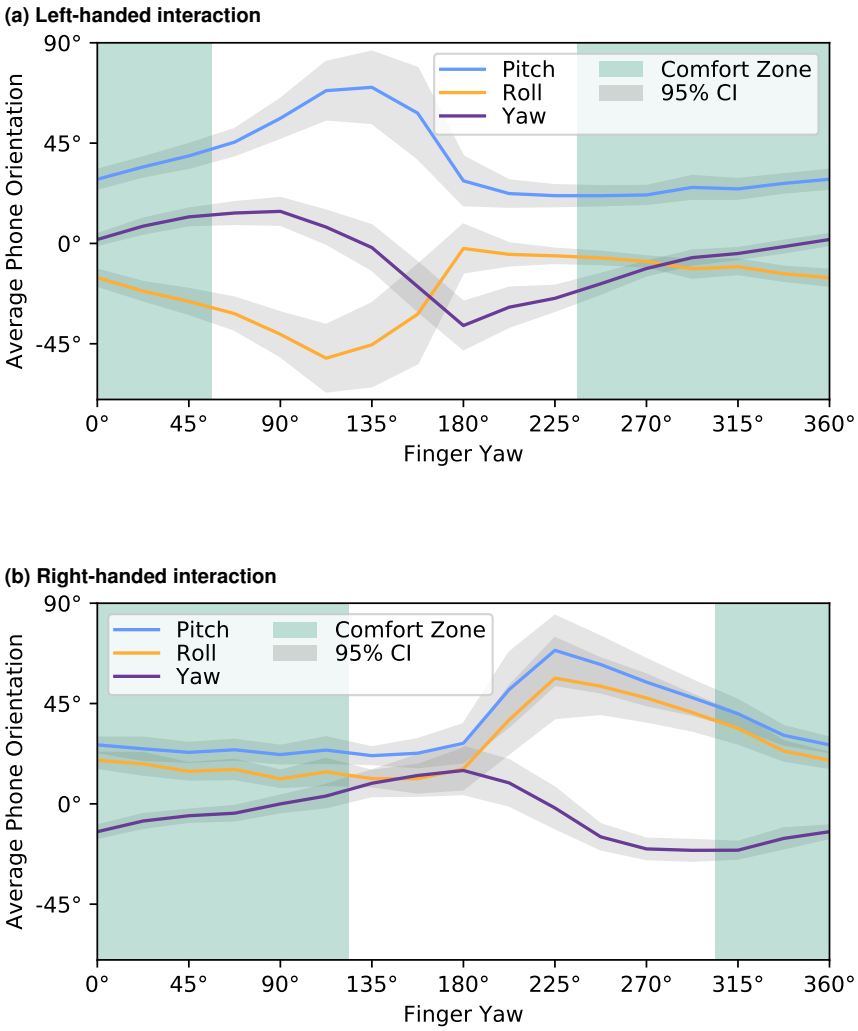

Figure 4.14: The average phone orientation in a two-handed smartphone interaction scenario.

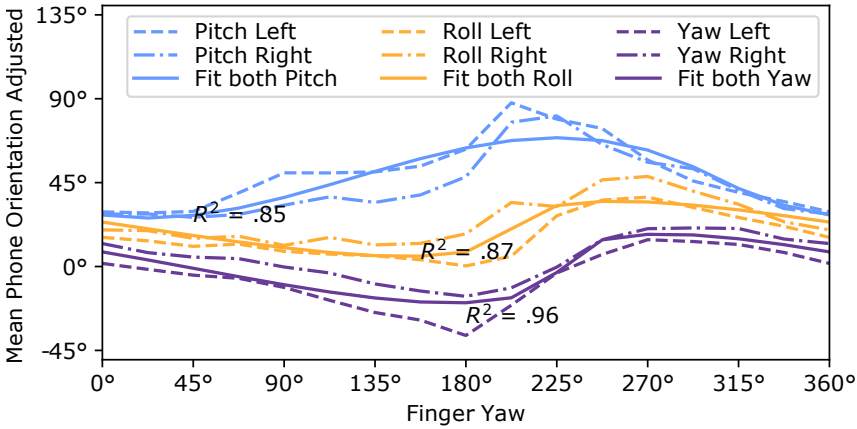

Figure 4.15: The average phone orientation adjusted to be hand invariant.

4.3.3.3 Pointing Accuracy

We first filtered 54 (respective .7%) of the 7,680 touch events where the distance to the center is larger than the mean plus three times the SD. The remaining average distance to the target was $M = 2.9mm$ ($SD = 1.75$). We conducted a three-way RM-ANOVA to determine whether the independent variables significantly influenced the touch accuracy. Our analysis revealed significant main effects for PITCH_{Finger} and YAW_{Finger} on distance ($F_{3,2363} = 483.493$, $p < .001$; $F_{15,2363} = 10.03$, respectively), however, not for HAND ($p < .001$; $F_{1,2363} = 1.243$, $p = .264$). Further, we found significant two-way interactions between $\text{PITCH}_{Finger} \times$ HAND and $\text{PITCH}_{Finger} \times \text{YAW}_{Finger}$ ($F_{3,2363} = 15.659$, $p =< .001$; $F_{45,2363} = 1.724$, $p < .003$, respectively). However, there was no significant two-way interaction between HAND $\times$ YAW_{Finger} ($F_{15,2363} = .564$, $p = .904$). Lastly, we found a significant three-way interaction between PITCH_{Finger}, YAW_{Finger}, and HAND ($F_{45,2363} = 1.512$, $p = .016$).

A Tukey's HSD post hoc test on PITCH_{Finger} with Bonferroni correction applied only showed a significant difference (all p < .05) between 10° and all

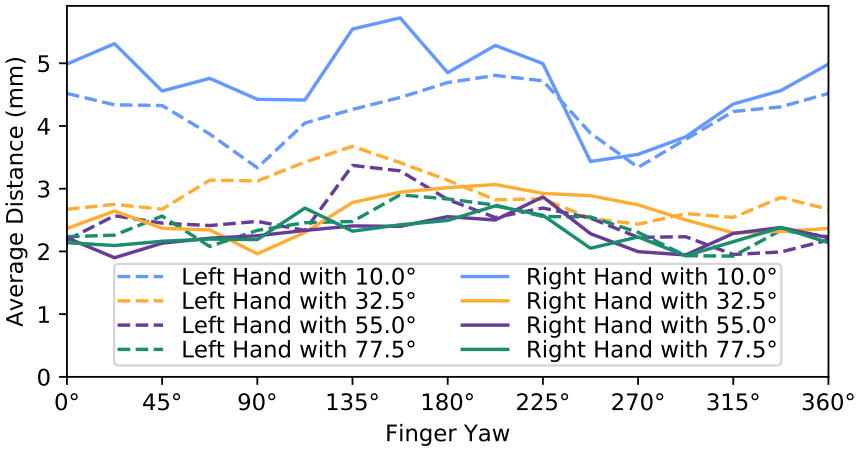

Figure 4.16: The average diatnce bewtween the touch point of the finger tip and the cross hair on the screen.

other pitch values. We did not conduct the post hoc test for YAW_{Finger} due to the number of comparisons which likely lead to no insights which we support by a visual analysis of Figure 4.16.

4.3.4 Discussion

We first modeled the RATING using a Sine wave. We conclude that the overall trend of how the rating correlates with the finger yaw orientation is in line with the findings made in Section 4.2. However, the modeled ratings for both hands are significantly easier than in the tabletop scenario. Therefore we confirm **H1a**. Thus, allowing the users to move the device and the fingers makes the input easier to perform, despite the fact that the participants had to control more degrees-of-freedom.

We used the same rating threshold as in Section 4.2 of 40 to distinguish between the *comfort* zone and the *non-comfort* zone . The *comfort* zone for two-handed interaction was 180.0° and therefore 33.3% larger than in the tabletop scenario. Thus, we confirm **H1b**. This allows designers to use a larger input range

for yaw inputs. Furthermore, the overlap of the left and the right hand's *comfort* zones is 66.7% larger, enabling designers to implement yaw without adjusting for handedness.

We carried out our study in the same ways as described in Section 4.2. However, in contrast to the tabletop results, for our two-handed scenario, we did not observe infeasible input. Therefore, we can confirm **H1c**. For tabletop scenarios the *non-comfort* zone describes inputs which might not be feasible to perform by users, in the two-handed scenario the *non-comfort* zone can be used to gain attention for safety-critical input without making the input too hard.

Our analysis revealed a significant effect of YAW on the smartphone's orientation. In detail, we found that the smartphone orientation changes more in the *non-comfort* zones than in the *comfort* zones confirming **H2a**. While we expect that the screen was readable at all time due to the low error rate when selecting words, the change in orientation shows that the screen was not always perfectly facing the participant. Therefore, while reading a single word remains possible, designers need to be aware that using the finger orientation as the input changes the orientation of the display. Especially when exceeding the *comfort* zone, the readability will decrease.

We modeled the smartphone orientation using a Clausen function and achieved an average R^2 of .89 when the left-hand data is mirrored. This shows that the smartphone orientation can be modeled for both hands with one function for each degree-of-freedom. Thus, we confirm **H2b**. The function enables us to model the smartphone orientation for each finger orientation. Further, this allows designers to understand how possible inputs would affect the smartphone orientation and thus influence the readability of the content displayed on the smartphone.

Lastly, our analysis revealed that the offset between the input and the target is significantly different for a pitch of 10° compared to all other conditions. Holz and Baudisch [69] found significant differences between all of their 5 conditions ranging from 15° to 90°. However, they studied only 2 levels of YAW_{Finger}, while our analysis also revealed significant main effects for 16 levels of YAW_{Finger} and both HANDS. Thus, we conclude that today's touchscreens are not well suited

for 50% of the finger orientations. While Henze et al. [60] presented a model to improve touch input, taking the finger orientation into account would further improve single touch input.

In line with Section 4.2, we showed that the feasibility RATING for finger orientation with two-hands can also be modeled using a Sine function. However, we found that finger orientation is easier when the user is allowed to move and rotate the phone with the second hand than in the tabletop scenario. As a consequence, this leads to a phone orientation which is not perfectly in sight of the user.

In our study, we did not control if participants bent their finger on the Proximal Interphalangeal (PIP) joint or the Metacarpophalangeal (MCP) joint of the index finger. However, the two 3D printed parts controlled for the distal interphalangeal (DIP) joint of the index finger. As the DIP has a limited movement range, we assume that our results can be transferred to situations where users can bend all joints.

4.4 Summary

In this chapter, we investigated RQ2: "Which ergonomic contains occur when using finger orientation as an additional input?" In two controlled lab studies, we first showed how ergonomic constraints affect the feasibility of finger orientation as an additional input in a tabletop setup. We then showed how finger orientation get's affected in a mobile scenario. In detail, we conducted two study to investigate the feasibility of using finger orientation where the user is preforming different finger orientations. We systematically varied the finger orientations using 4 different pitch and 16 different yaw angles while ensuring that the screen was visible to the participant during the interaction. Further, participants were asked to perform all combinations with the index finger of both the right and the left hand. The results show that not all orientations are equally feasible, with some orientations being infeasible to perform. Furthermore, we show that the input space for each hand can be divided into two zones; the *comfort* and *non-comfort* zones. While there are differences between the two scenarios the overall trend remains the same. Moreover, the feasible input range is large in the mobile

scenario due to the fact that users start to turn the phone. We showed that the feasibility rating correlated to the phone orientation. A harder feasibility rating, therefore, results in a phone orientation tilted away from the user. Thus, the larger input range in the mobile phone scenario comes at the cost of a in the worst case scenario up side down flipped smartphone.

5

Social Implications

The proliferation of mobile devices has rendered mobile notifications ubiquitous. However, researchers are only slowly beginning to understand how these technologies affect everyday social interactions. In particular, the negative social influence of mobile interruptions remains unexplored from a methodological perspective. In the following chapter, we, first, present a new mixed-method approach to evaluate interaction techniques regarding their disruptiveness in social settings. To showcase our new approach, we compare the standard Android caller against a new method called *SurfaceSliding* where the phone is slid over the table instead of using the touchscreen. In the second part of the chapter, we evaluate finger orientation interaction in a face-to-face conversation using our new mixed-method evaluation approach (RQ3).

Parts of this chapter are based on the following publications:

S. Mayer, L. Lischke, P. W. Woźniak, and N. Henze. "Evaluating the Disruptiveness of Mobile Interactions: A Mixed-Method Approach." In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: ACM, 2018, 406:1–406:14. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3173980[a]

Planned publication: S. Mayer, H. V. Le, M. Weiß, and N. Henze. *Effect of Finger Orientation on Social Settings.* 2019

[a] A video of this study is available under:
`https://www.youtube.com/watch?v=6-HGXW6bLPw`

5.1 Evaluating the Disruptiveness of Mobile Interactions

In the following section, we contribute a mixed-method evaluation approach for assessing the disruptiveness of mobile interaction techniques. We postulate an approach that uses a conversation task for two participants. We employ eye tracking as a quantitative measure and combine it with qualitative evaluation based on semi-structured final interviews. In traditional coding techniques for face-to-face conversations used in past work [14, 44, 133], hour-long encoding of video and audio material is needed. In contrast, our work offers an alternative approach that highlights how eye-tracking and interviews can offer complementary yet different insights. Our method is different from past approaches as it is designed to offer quick answers that result in actionable insight during a design process.

Further, we present an example study where we analyze the disruptiveness of two techniques to dismiss calls, namely the standard Android incoming call dialog and *SurfaceSliding*; a new method which allows the user to slide the phone on the table to decline a call. Using the rapid eye-tacking analysis, we found that participants were indeed significantly objectively less distracted by the one dismiss technique. Further, the interviews revealed a set of factors which further help improve the design: (a) the need for empower participants to fine-tune their interruption acceptance levels, (b) the impact of the social setting on interruption acceptance (c) the importance of exceptions and emergencies, and (d) the influence of group dynamics. The objective and subjective results together present a holistic perspective on user behavior and reasoning.

5.2 Research Design

We present the considerations and choices involved in creating our evaluation approach to assess the disruptive impact of mobile interruptions in conversation. We discuss the driving research questions and the alternatives that we considered while designing our approach.

5.2.1 Requirements for a new approach

As part of a larger project on designing novel ways to interact with mobile devices in social settings, we encountered a key issue early in our design process. Once we started developing early prototypes, we required ways to rapidly gather feedback from participants and decide which ideas needed to be rejected early. However, we found it difficult to determine which of our prototypes were potentially disruptive to conversations. The method we required, needed to: (a) allow for work with low-fidelity prototypes and Wizard-of-Oz studies; (b) provide clear answers on which interaction techniques are best among a set; (c) generate enough qualitative user feedback to play a generative role for later stages of the design process; (d) be applicable for casual social environments.

As shown in our review of related work in Section 2.3, CSCW and HCI literature did not present any viable solutions. In our search for methods, we later investigated other literature sources for inspiration. Disruption and interruption in conversations is studied in communication science and it is sometimes used to understand interaction between users and interactive artefacts. For example, Karsvall [80] used a dialogical approach to understand the team dynamics of operating theatres. While such an approach is well grounded and offers an in-depth look into the details of the social dynamic involved, the analysis requires an amount of time that is not manageable for an interaction design process. Further, conversation analysis may not yield results that could motivate decisions on which prototypes to choose.

Another field that has a history of studying interruptions is Human Factors and Ergonomics (HFE). Models of disruptions and interruptions are often used in HFE (e.g. Endsley and Jones [30]), so one could expect that suitable methods could be translated from HFE. This is not the case, however, for two reasons. First,

most studies of disruptions focus on well-defined controlled tasks. Hodgetts and Jones [66] studied the impact of mobile notifications, which appears to be relevant for our inquiry. However, as appropriate in the HFE tradition, the notifications were studies while performing a highly artificial task (the Tower of London task). This limits the applicability of the methods used in HFE for application in an interaction design process. Secondly, and more fundamentally, the entire field of HFE focuses on non-discretionary use [52], which limits the validity of whatever possibly adaptable method when used in the context of casual social interactions.

Consequently, we opted for re-appropriating some of the methods used in the past in the HCI field and combining them into a new approach. This way, we could ensure a new mixed-method approach would be ecologically valid for social settings and potentially offer input to the design process.

5.2.2 Blending existing approaches

Faced with the challenge of designing a new approach to meet our needs, we decided to adopt a mixed-method approach. This was motivated by the need for a decision of which early prototypes were worth pursuing, and assuring that enough generative user feedback was produced. Our approach consists of a quantitative eye tracking metric, an in-depth semi-structured interview and a generic discussion task designed for strangers.

To qualitatively measure the disruptiveness a given interaction technique produces, we decided to augment the method used by Vertegaal et al. [169]. In this work, the authors show that gaze features are directly coupled to conversational attention. Specifically, they show that looking at one's interlocutor is a significant indicator that one is engaged in a conversation. In our approach, we use this property by investigating the differences in gaze directed towards the interlocutor. Based on the findings of Vertegaal et al. [169] we assume that participants direct a participant- and interlocutor-specific fraction of their gaze time to their interlocutor. Consequently, we investigate how different interaction techniques affect the time spent looking at the potentially disruptive technology and the interlocutor. From a technical perspective, the method requires employing two eye trackers as shown in Figure 5.2. As continuous timings are measured, data analysis is performed using fast and simple statistical methods such as analysis of variance.

While using quantitative eye tracking metrics can reveal how disruptive a method is, the quantitative metrics are unlikely to stimulate further development in a design process. Thus, we decided to employ semi-structured interviews [17, 45], a staple method of interaction design, to assure that enough user feedback can be gathered to not only eliminate possible prototypes, but also stimulate the design of new solutions. In our approach, we use an entry interview to establish the participants' initial attitudes towards disruptions caused by technology in conversations. This is not only done to generate possible design inspiration, but also to ascertain whether participants find any disruptions acceptable. As shown later in this section, some participants consider state-of-the-art technology-based disruptions offensive and they differentiate between the disruptiveness levels of different interaction techniques. A final interview is performed to qualitatively assess the disruptiveness caused by a given technique. This also enables gathering suggestions for prototype improvement. Finally, we employ a simplified version of qualitative coding with affinity diagramming [57] for interview analysis as this offers a rapid way to analyze and understand the feedback provided by interviews. As our approach is not intended to build a structured understanding of disruption, we believe that a more advanced qualitative analysis method is not required.

Our proposed mixed-method approach combines quantitative eye tracking metrics and qualitative interview feedback. The eye tracking data allows determining which technique is less disruptive statistically. However, a better understanding of how the technologies influenced the conversation can only be gained from qualitative interviews. The interviews allow determining further design opportunities and identify drawbacks which lead to future design improvements and better interaction concepts.

5.2.3 Choice of Participants and Stimulus

As we endeavored to design an approach that would allow for rapid evaluation, we considered flexibility in terms of participant choice as a key feature. Participants are often hard to recruit, especially for studies early in the design process, when rapid feedback is needed. Our approach can be used with a wide variety of participants as it uses a generic conversation stimulus. Participants are grouped in pairs randomly. Further, the experimental design allows for large within-pair

variability in eye tracking metrics and thus it does not put any restrictions on the participants' gender, age, race, native language etc. Previous work showed that gaze fixations are an effective way to evaluate two systems [5]. Moreover Okamoto et al. [141] showed eye movements are affected by conversations. Thus, when using eye tracking in conversations, this should be taken into account in the design. As a result measuring the time of eye contact or time spent looking at the conversation partner is not efficient as a disruption measurement. Therefore we propose time spent on the disruption as the measurement. Therefore, a between-groups design for the experiments is necessary as sequence effects are likely to appear when a conversation is prolonged.

We use a generic discussion task in our approach. We decided to use a task designed by language experts specifically for conversations between strangers with a particular focus on paying attention to the other party in the discussions. Consequently, we use the discussion task from the University of Cambridge's English for Speakers of Other Languages Certificate in Advanced English Speaking Test [29]. The task provides a stimulus that is both manageable for advanced non-native speakers and engaging enough for native speakers. The task also includes a shorter introductory segment that can be used as an icebreaker for the discussion. An additional advantage of using a speaking exam task is the fact that analogous tasks exist and can be easily found for other languages and, thus, our approach is not specific to English. Again, using such a task necessitates a between-groups design, as there is no possibility to assure that different discussion topics are equally stimulating.

5.2.4 Study Plan

The final study procedure in our approach is shown in Figure 5.1. At the beginning of each study session, the facilitators conduct individual semi-structured interviews (Entry Interview, see Figure 5.1). The interview serves as a means of collecting demographic data on the participants. Further, it introduces the conversation task. The purpose of the study is not revealed to the participants until the end of the study. Informing the participants about the focus on mobile disruptiveness may cause potential bias due to increased awareness to interruptions. Thus only one participant will be introduced to the additional disruptive task.

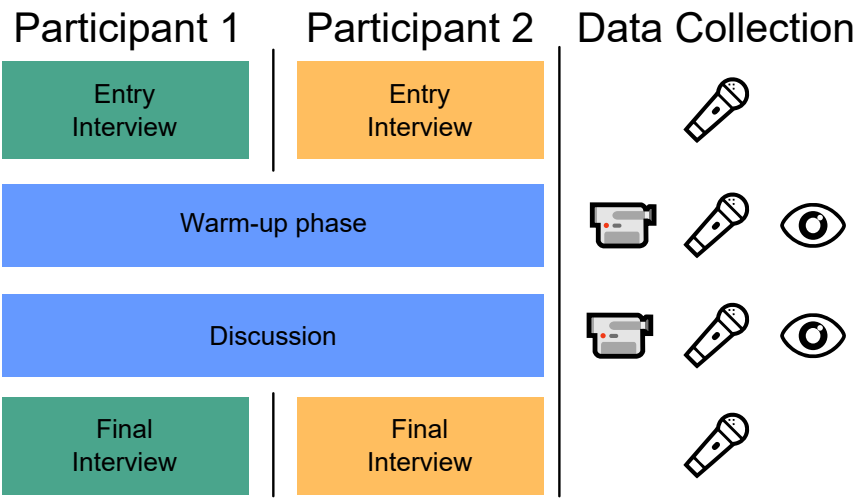

Figure 5.1: The study procedure in our new mixed-method approach. Data collection methods are: video, audio and eye tracking.

After the instruction and training phace, the participants are introduced to each other and start wearing the eye trackers. The facilitator then presents the experimental task. First, the participants run through a warm-up phase to get confident with their conversation afterward also the disruption is taking place to observe the participants reactions and potential behavior change. After the discussion is concluded (we recommend a time of 10*min*, based on language examination experience [29]), individual debriefings (Final Interview, see Figure 5.1) are performed. All interviews are audio recorded. As a safety precaution, we also recommend video recording the conversation. After the study, eye tracking data is analysed using inferential significance testing and simplified qualitative analysis with affinity diagramming is performed on the interview data.

5.3 Study 1: SurfaceSliding

Next, we present a user study that illustrates a practical application of our approach implementing the proposed study plan. We contribute a standard study

description here to showcase how our approach can be used effectively to provide a concise report if required. Our presented mixed-method approach allows analyzing any device or concept, which can potentially disrupt ongoing conversations. Ubiquitous mobile devices such as smartphones or smartwatches have high potential to cause disruption [154], and hence we present a scenario where a mobile phone causes disruption in a face-to-face conversation. In the example study, we compared two techniques to decline incoming calls.

In particular, we investigate *SurfaceSliding*, a new method to decline incoming calls and compare it with the standard technique provided by the Android operating system. The interaction technique is mostly inspired by the work of Wiese et al. [178] who showed that participants tend to have their phones lying on a table when at home or in the office. Work by Porcheron et al. [144] and Su and Wang [162] show that phones also play an important role in social settings, and highlights that phones in a pub setting are often positioned visibly on the table. We chose an incoming call scenario as a disruption example inspired by Bogunovich and Salvucci [8]. Declining a call is an action that requires attention, but the attention should be minimal. Thus, this task is a good candidate for designing for limiting disruptions. This has the advantage since all participants were equally familiar with the scenario and most likely have been in a similar scenario before. *SurfaceSliding*, an intermediate prototype in our process for building less disruptive mobile phone interactions, enables the participant to decline a call by dragging the phone along the table (instead of dragging their finger over the touch screen).

5.3.1 Study

The study had a between-groups design with a single independent variable, TECHNIQUE. In the *Touch* condition, the participant with the phone got a standard Android incoming call interface, the interaction is presented in Figure 5.3. In the *SurfaceSliding* condition, the phone was modified so that the participant could use *SurfaceSliding* to interact with it. The interaction technique is illustrated in Figure 5.4. Further, as only one participant needed to interact with the phone, we will refer to them as *HasPhone* and *NoPhone*.

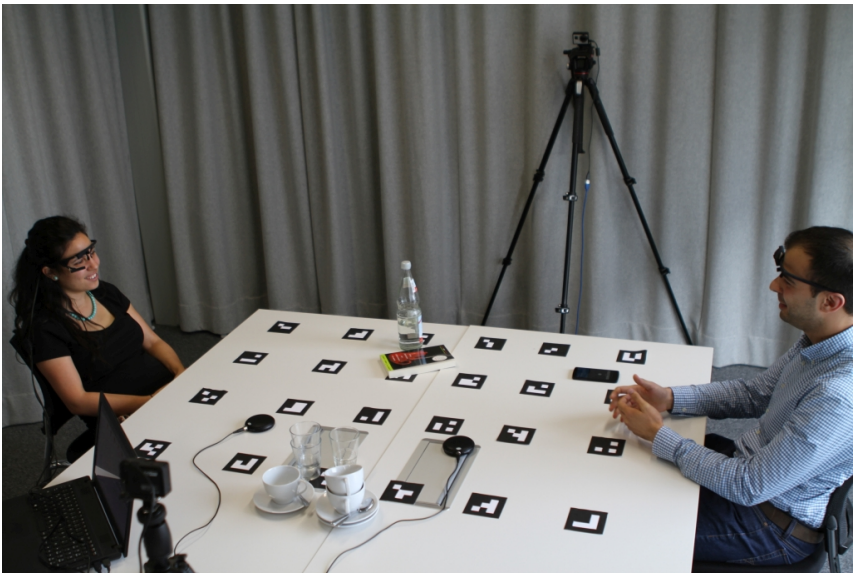

Figure 5.2: The study setup showing two participants in a conversation wearing mobile eye tracker.

5.3.1.1 Apparatus

To investigate the effect of using *SurfaceSliding* for declining calls we integrated *SurfaceSliding* in the standard GUI of Android v5.0. We implemented *SurfaceSliding* on a Nexus 5, in combination with the incoming call prototype, which had been modified to show an incoming call eight times within 10*min* of discussion. However, the first two minutes of the conversation were left free of disruption to ensure a fluent discussion. After minute two, eight incoming calls were scheduled each within one of the remaining 8 minutes. Further, we made sure that two calls were always at least 20*sec* apart. The study was controlled by a separate laptop where it was possible to start the study or turn the phone into testing mode to show participants the declining methods. The phone was in silent mode the whole time, so ringtone and vibration were turned off. Further, the screen was black by default. The only indicator for an incoming call was the light when the phone displayed the incoming call. After declining a call, the screen turned black again.

We implemented a recognition algorithm which was able to detect moving the device in any direction on the table. For detecting the movement, we used the front facing camera combined with OpenCV's optical flow algorithm. Further, we used the phones' built-in sensors to make sure that the phone was not picked up from the table. The detection algorithm was fully implemented on the phone itself. To trigger the decline action the phone needed to be moved by $5cm$ to the left.

We used two GoPro Hero 3+ and a Zoom H6 audio recorder with two table microphones to capture the content of the discussion. Further, we placed three coffee mugs, four glasses, one bottle and one book on the table to simulate a realistic discussion environment as would occur in an office or café scenario. A Zoom H1 audio recorder was used to record the second interview. Both conversation partners were equipped with PupilLabs mobile eye trackers to determine how long each participant looked at the phone. To be able to automatically detect where the participants were looking, we used table-mounted QR codes (see Figure 5.2) for establishing a reference coordinate. For both eye-tracking cameras, we used an IR camera with IR illumination (dark pupil tracking) recording with $640 \times 480px$ at $30FPS$. The participant without a phone (*NoPhone*) had a world camera with a field of view (FoV) of $90°$ recording with $1920 \times 1080px$ at $30FPS$. The other participant (*hasPhone*) had a world camera with a FoV of $100°$ diagonal recording with $1910 \times 1080px$ at $30FPS$.

5.3.1.2 Task

To ensure that participants did not know about any involvement of technology we invited the participants to a study with the title *Stress in Conversations*. We used 5 questions from the CAE speaking test by [29] for the conversation. We played the 5 questions from the exam DVD included with the book to provide the same stimulus for all participant pairs. The questions used were as follows: (1) "Here are some pictures showing different ways in which the computer affects our lives. First, talk to each other about how these pictures show the role of computers nowadays. Then decide which picture best reflects the difference computers have made to our lives." (2) "Some people say that computers are helping to create a generation without social skills. What's your opinion?" (3) "What are advantages

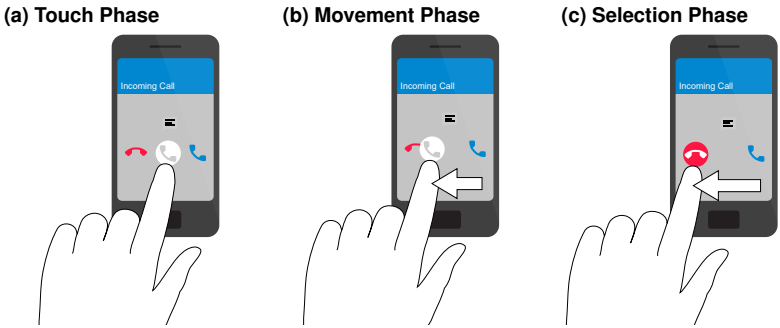

Figure 5.3: Declining an incoming call selection phase using the standard touch interface. In the first step (a) the participant taps the center icon and then (b) moves it over to the decline symbol, finally (c) the release of the finger will trigger the highlighted action.

and disadvantages of shopping by computer?" (4) "How far do you agree that computer is the greatest invention of modern times?" (5) "A lot of personal information about all of us is now kept on computers. Do you find this worrying?"

5.3.1.3 Procedure

The participants were guided through the whole study by two researchers. When both participants arrived at our study room, one researcher asked the first participant to enter the room while the second stayed outside to interview the two participants independently. Entry interviews were then performed, during which the participants filled in a consent form and a demographics questionnaire.

The participant who arrived first was always the participant interacting with the phone (*HasPhone*). Therefore *HasPhone* was asked to enter the study room first while the other participant (*NoPhone*) was interviewed in front of the study room. We first interviewed *HasPhone*. We then handed the phone to *HasPhone* and told them that they had to interact with the phone during the study. However, we also told them that the other participant was unaware that a phone would be involved in the study. We then showed how declining a call worked using the test

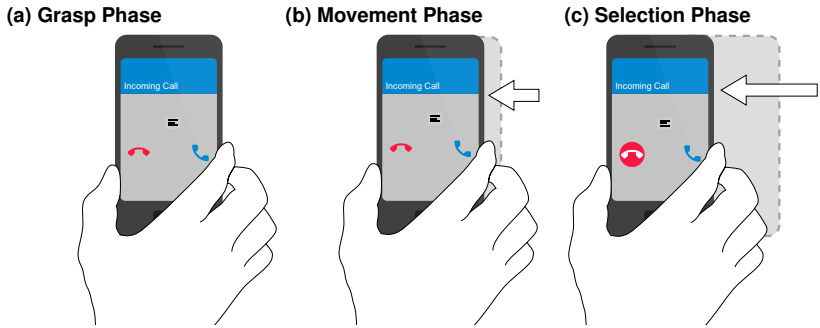

Figure 5.4: Declining an incoming call using *SurfaceSliding*. In a first step (a) the participant grasps the phone. Then the moves the phone in the direction of the decline symbol in respect to the center of the phone (b). After the movement (c) the decline call action is triggered.

mode of the app. We let them try and familiarize with the interaction techniques until they were comfortable declining calls. After the trial phase, we then set the phone to study mode where the screen turned black and was waiting for simulated incoming calls. After the introduction phase, we invited the *NoPhone* participant into the study room.

Then, we asked them to put on the eye-trackers and started the calibration process; starting all recoding, video, audio, and eye-tracking. Then we played back the discussion instructions from CAE [29]. We used the first instruction as an icebreaker to start the conversation. The first instruction came with images to help the discussion develop. The question was discussed for about two minutes, then we presented the second question followed by more questions if needed. When presenting the second question, we removed the images from the table. This was when the data recording that was later analyzed started. Whenever the discussion was less active, the facilitator asked an additional question from the task sheet.

After the discussion phase, participants were interviewed individually. First, we asked them about the overall outcome of the discussion to reflect on the discussion. Followed by questions about their personal discussion behavior. We

then inquired whether they noticed the presence of the phone and whether it affected discussion. We then asked whether the declining of the phone call was appropriate. The *HasPhone* participants were asked how they felt on two levels: how it felt to observe the phone and what was the experience of declining the call.

Finally, both participants were asked how they deal with and where they store their phone in the following four situations: (1) in a private face-to-face conversation, (2) in a private group conversation (3) face-to-face in a business situation and (4) in a business group setting. We wrapped up the interview with an open question for additional comments

5.3.1.4 Participants

We recruited 24 participants (17 female, and 7 male) through our university's mailing list. The participants were aged from 19 to 33 years ($M = 24.9$, $SD = 3.9$). All of them had either no visual impairment or corrected to normal vision by wearing contact lenses. Three of the pairs had known each other beforehand. We reimbursed the participants with 10 €.

5.3.2 Results

In this section, we present the results of our example study in which we compared a standard decline a call technique against *SurfaceSliding*. We show how engaging participants in a conversation task allowed us to statistically determine which interaction technique was less disruptive and how the interview results allow us identified possible future design improvements.

5.3.2.1 Eye Tracking Data

Due to technical problems of the eye tracking software, we excluded two groups from the eye tracking analysis. We first used the built-in QR code plug-in of the eye tracker to recognize where the participants looked. However, in contrast to pre-study results, the eye tracking accuracy was insufficient, most likely due to participant not sitting fully upright. Therefore, we manually labeled the remaining 20 eye tracking videos. For further analysis we treated participants independently from each other by using the variable *HasPhone* with two levels

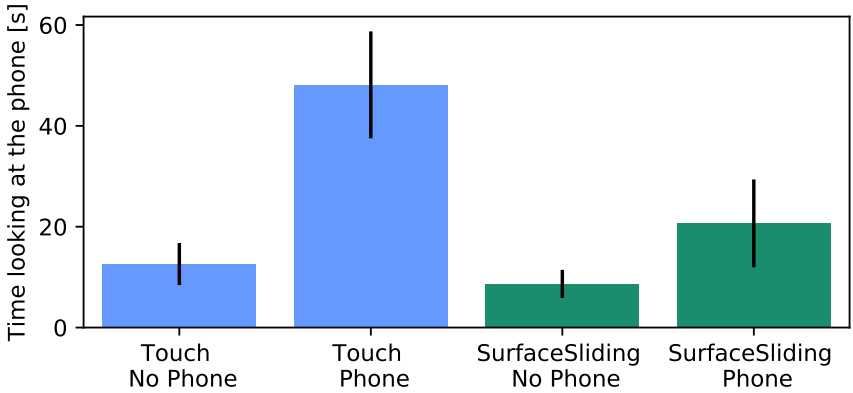

Figure 5.5: The average values and standard error for TECHNIQUE × PHONE.

— either the participant had the phone or not. The results of how long the participants looked on average at the phone is presented in Figure 5.5. On average, participants spent $M = 22.5s$, $SD = 21.8$ with looking at the phone. Participants in the *Touch–Phone* condition looked at the phone for the longest period of time, $M = 48.1s$, $SD = 23.7$. Those in the *SurfaceSliding–NoPhone* condition exhibited the shortest time looking at the phone, $M = 8.6s$, $SD = 8.6$. A two-way ANOVA revealed a statistically significant difference between the *Phone* and the *NoPhone* conditions $F_{1,16} = 10.587$, $p = .005$. We also found a significant difference between the *Touch* and the *SurfaceSliding* conditions $F_{1,16} = 4.623$, $p = .047$. We found no significant interaction effect ($p = .139$).

5.3.2.2 Qualitative Data

Interview recordings were transcribed for analysis, the total recording time was $4.31h$. We conducted initial filtering with two researchers to only leave data about in-conversation interactions. Next, three researchers coded 15% of the material to establish an initial coding tree. A single researcher coded the rest of the data. Once the initial coding was finished, we used affinity diagramming with printed quotes to establish higher-level themes through constant comparison. We labeled quotes of the participants with group name, the condition, (S) for *SurfaceSlding*

or (T) for the *Touch* condition, and a *P* if they had a phone or *NP* if they had no phone. We first present general comments about the technique studied and then discuss the themes of disruption by technology in conversation.

5.3.2.3 General Feedback

As our participants were unaware that we were interested in studying phone interactions, we asked how their conversation was influenced by the incoming calls. Independent of TECHNIQUE, none of the participants without a phone realized that the phone was part of the study. Two participants did not acknowledge the presence of a device on the table until we asked them explicitly (*5–S–NP* and *12–T–NP*). One participant believed that the other person was checking the time on the phone *(1–T–NP)*. All others expressed their awareness of the interruption quite vividly:

> *"(My discussion partner) has repeatedly interrupted (the conversation). (2–S–NP)"*

Most of the negative statements concerned the lack of familiarity with the method:

> *"I'd rather not use (SurfaceSliding), but this may be entirely due to my current habits. (12–T–P)"*

On the other hand, participants reflected that the method was easy to use, even without attributing visual attention to the phone:

> *"(With SurfaceSliding) rejecting calls was easy, also without looking [at the phone]. (5–S–P)"*

Another participant noted that *SurfaceSliding* did not have a negative influence on the conversation:

> *"(SurfaceSliding) is clearly unobtrusive. (11–S–P)"*

Four participants provided additional suggestions on how to improve declining incoming calls. Participants suggested that one should always slide away from the body to decline calls, irrespective of the location of the phone or the hand used. A left-handed participant suggested:

> *"(The system) should account for right- and left-handed people. (3–S–P)"*

Further, by analyzing the recorded videos, we can conclude that none of the participants tried to pick up the phone from the table; even those who had the *Touch* condition and had not been instructed to keep the phone on the table. It is notable that all calls were successfully declined.

5.3.2.4 Levels of Phone Acceptance in Conversation

When asked about the role that smartphones may play in conversations, the participants reported many different stances. Our analysis showed two equally large groups of participants. The first group was strict about eliminating any phone interaction from the conversation. They believed a smartphone was only an unnecessary distractor. One participant remarked:

> *"Irrespective of the situation, if one wants to participate in a discussion, the phone belongs in the pocket. (8–T–NP)"*

The second group expressed that the acceptance of a smartphone in a conversation was highly dependent on the context of the interaction. These participants mentioned that the topic and the people present highly affected what was acceptable. They also believed that the acceptability depended on how the interactions were handled:

> *"It's only okay when the calls are rejected and the influence on the conversation is as small as possible. (9–S–P)"*

We observed that some of the participants were happy to accept a phone placed on the table during conversation, especially in more casual settings. The position of the phone on the table also appeared to be important

> *"In a meeting, (it's okay to put it) on the table, but a bit on the side. (7–S–P)"*

5.3.2.5 Private and Professional Settings

Most participants expressed that different acceptance levels were valid for professional and casual settings. Many recommended ways where suggested to handle incoming calls specifically in a business environment. Participants expressed that while participants are often expected to be aware of the state of their smartphone, they should not interact with it during meetings:

> *"In business meetings, the phone can lay on the table, but it should not be used. (8–T–NP)"*

In contrast, participants showed more flexible attitudes when interacting with groups of friends or family members. Attitudes ranged from simply being more lenient towards interruptions to not perceiving the presence of a phone as a factor influencing the conversation:

> *"(When talking to) friends, the phone is always on the table, but not in the case of family. (7–S–P)"*

> *"[...] with friends, the phone can always be in your hand. (11–S–P)"*

Notably, participants were stricter in home environments where they saw no need to pay attention to their phone when having family conversations.

5.3.2.6 Handling Interruptions, Exceptions and Emergencies

Participants commented extensively on situations where exceptions are possible and smartphone use during conversation may be acceptable. They suggested a number of cultural codes that could be used in special circumstances, including placing the phone on the table as an indicator that one is expecting a call:

> *"When you have company, [put the phone] on mute in your pocket. On the table is acceptable when expecting extremely important calls. (3–S–P)"*

Interrupting conversation was perceived as highly problematic, even in exceptional situations:

"When one picks up, one should apologize immediately. (12–T–NP)"

Moreover, participants indicated that one should verify if the call is important for a possible interruption to be acceptable.

"There are things that are important and need to be dealt with immediately... work, family stuff. (9–S–P)"

5.3.2.7 Influence on Group Dynamics

Our participants reported that the composition of the conversation group and its size highly influenced the handling of possible interruptions. Conversations involving only two participants, as the one explored in our study, were considered most sensitive to interruptions and prompted immediate reactions:

"It was very impolite when the conversation partner looked at the phone. I wanted to say something. (5–S–P)"

Larger groups seemed to offer more leeway in interacting with a phone. One participant commented on how a bigger group enabled a limited amount of interruption:

"It's also not okay in a group, but a bit more okay (than in a one-to-one conversation). (9–S–NP)"

Finally, participants saw putting one's phone away as a sign of appreciating the other parties in the conversation and deeming the social activity interesting:

"If one's having a good conversation, the phone should be put away entirely. (9–S–P)"

5.3.3 Discussion

We investigated whether *SurfaceSliding* was less intrusive and how the design can be improved in the future. The presented findings will help to improve *SurfaceSliding* as well as designers to build less disruptive ubiquitous systems. As we discuss the quantitative eye-tracking data showed a reduced disruption

when using *SurfaceSliding*. The qualitative results discussed highlights further improvements. Furthermore, the lessons learned to foster discussions about the social acceptability of disruptions through technologies, in general.

5.3.3.1 Around Device interaction

The eye tracking data indicates that participants paid significantly more attention to their interlocutors when using *SurfaceSliding*. Therefore, we can conclude that our technique provides an effective and non-interrupting way to decline calls in-conversation. In the light of the analysis by Vertegaal et al. [169], it can be inferred that participants in the *SurfaceSliding* condition were more engaged in the conversations than the participants using the current default method. This fact suggests that design space for creating new non-disruptive techniques for managing attention should be explored further. *SurfaceSliding* provides a working example of how in-conversation interruptions can be effectively reduced. The fact that none of the participants grabbed the phone in their hand during the study suggests that they were comfortable with interacting with the device while it was lying on table. This confirms previous work on spatial interaction pattern for on-surface mobile devices [186]. While past work has shown that sliding mobile devices on surfaces can be effective for complex task and long interaction periods, *SurfaceSliding* illustrates that moving along a horizontal surface can also be effective for short, atomic tasks.

5.3.3.2 Addressing Complete Designs

As it surfaced in the interviews, the participants expected they would be able to choose the direction of the sliding. We recognize that this improvement may render our technique more effective. However, this require extensive sensing of the user's body position or an explicit set-up phase before using *SurfaceSliding*. Further, participants may confuse directions, which increase error rates. These facts showcase that a more explorative approach not limited to two techniques may have generated more feedback. However, that providing the participants with an experience of declining the call, irrespective of the details of the technique, triggered valuable on-the-spot feedback.

5.3.3.3 Strict Users

We observed that there is a user group strictly opposed to phones entering conversations. These participants find placing the phone on the table unacceptable and answering calls during conversation is offensive to them. We believe that future systems should offer more support for this user group. For instance, a user should conveniently be able to inform other participants in the conversation that they do not wish that phone to be part of the experience. This could be achieved by a setting in the user's device that is communicated to other devices. Furthermore, that specific user group requires an easy means to deactivate phone output during a conversation. If future mobile devices are able to sense conversations, this user group would require all disturbing features to be deactivated. Overall, future mobile devices should provide participants with effective means to both deactivate disturbing features in their phones and inform other participants that they do not wish to be disturbed. However, as the strict users only represent part of the user base future systems will also need to provide opportunities to negotiate behaviors, especially in intercultural settings. Our approach, however, is not suited to designing with and for those participants.

5.3.3.4 Emergency Cases

Nearly all of the participants reflected that, despite their wish to limit interruptions, designs should account for emergency cases. Participants noted that there are exceptions where interruptions are acceptable. If future systems are to offer better management of in-conversation interruptions, they must incorporate the means to prioritize emergencies. While current systems do include some means to deal with emergencies (e.g. in iOS, the do not disturb mode can be deactivated when the participants receive repeated calls), more extensive features are required. Again, our study participants reflected that determining what constituted an emergency highly dependents on the context. For instance, while on holiday, calls from work are likely to be emergencies, and a mid-day call from the school may require the immediate attention of the parent. Designers of future mobile devices should help participants define possible emergencies and use context sensing efficiently to ensure interruptions are efficient exceptional circumstances.

5.3.3.5 Accounting for Changing Context

We noticed that our participants often remarked that the decision on whether attending to an event or information was worth interrupting a conversation depends on many factors. In a way, deciding whether to answer a call or message during a conversation is an economic decision. There is a certain social cost associated with answering the call, which is dependent on the context of the conversation. There is also the social cost of possibly ignoring an important call or missing a social interaction. Each time the user decides to accept or decline, they make a conscious decision, yet current mobile devices only offer the identity of the caller as a potential aid in making an informed choice. Further, participants reported that they not only consider who is calling them when deciding whether to answer the call; they also considered the context of the conversation and the possible purpose of the incoming call. There is an emergent need for mobile devices to adapt to the richer context. Our approach uses static measures that require stationary equipment and thus it is not particularly suited for studying how disruptiveness is affected by change of usage context. We recognise that the majority of the participants in our study expressed that their acceptance of interruptions varied depending on the context of usage. The composition of the conversational group or the purpose of the conversation affected the level of interruption that was deemed acceptable. While our approach allows the experimenter to freely introduce a context, the static setting of the discussion task renders switching social contexts difficult.

5.3.3.6 Summary

We presented two studies where we investigated techniques for declining calls in a face-to-face conversation. We were able to revisit the conversation and draw conclusions from the participants' behavior using video and audio recordings. Using our approach enabled us to understand the impact of a new interaction technique on disruptiveness. Eye tracking revealed a significant drop in time spent looking on the phone when using the new technique. Interviews provided evidence for underlining social mechanics that affect disruptions.

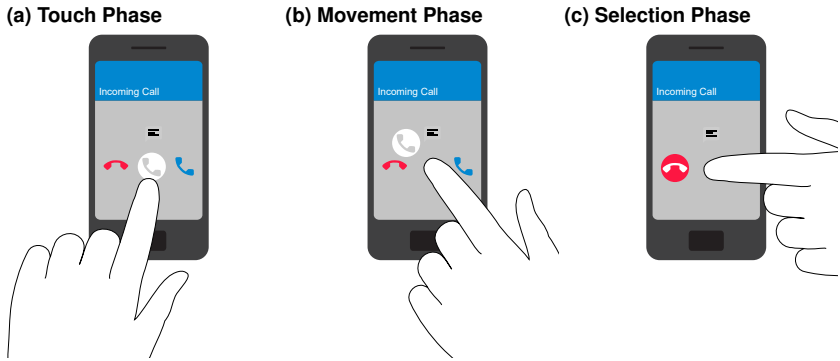

Figure 5.6: Declining an incoming call selection phase using finger orientation interface. In the first step (a) the user taps the center icon and then (b) changing the yaw of the finger in order to move the selection icon twards the decline button, finally (c) the release of the finger will trigger the highlighted action.

5.4 Study 2: Finger Orientation

In the following, we investigate how finger orientation affects social settings. Therefore, we use again our mixed-method approach in a face-to-face conversation, see Figure 5.7. Again we used the incoming call scenario to distract the conversion and make user of finger orientation to decline to calls. Thus, we present a second showcase on how the new mixed method works and therefore evaluate finger orientation.

5.4.1 Study

We compared a standard decline a call technique against a finger orientation declining method using the same setup as described in Section 5.3. The study had a between-groups design with a single independent variable, TECHNIQUE. Again, in the *Touch* condition, the participant with the phone got a standard Android incoming call interface, the interaction is presented in Figure 5.3. In the *FingerOrientation* condition, the phone was modified so that the participant could use the finger orientation to reject the calls. The interaction technique is

Figure 5.7: The study setup showing one participant wearing mobile eye tracker.

illustrated in Figure 5.6. In line with our example study see Section 5.3, only one participant needed to interact with the phone, we will refer to them as *HasPhone* and *NoPhone*.

5.4.1.1 Apparatus

The hardware setup we used, was the same as described in Section 5.3. We used the same implementation of the *Touch* condition. However, for the *FingerOrientation* condition, we used a Wizard-of-Oz approach to not influence the results by an imperfect implementation. Therefore, one of the two experimenter used an extra tablet to input the yaw values. In contrast to the example study (Section 5.3), we reduced the distance between the participants to improve the recognition rate of the QR code markers.

5.4.1.2 Procedure

The procedure was in line with the one reported in Section 5.3, participants were guided through the whole study by two researchers. First, they where interview

separately. Second, participants where played back the discussion instructions from CAE [29]. Lastly, they where again interview separately about how the discussion went and how the incoming called and the interaction influenced the discussion.

5.4.1.3 Participants

We recruited participants from our university's volunteer pool. In total, 24 participants took part in the study (16 male, and 8 female). The age range was between 20 and 35 years ($M = 23.3$, $SD = 3.7$). All of them had either no visual impairment or corrected to normal vision by wearing contact lenses. We reimbursed the participants with 10 €.

5.4.2 Results

In this section, we present the results of our example study in which we compared a standard decline a call technique with *FingerOrientation*. We show how engaging users in a conversation task allowed us to statistically determine which interaction technique was less disruptive and how the interview results allow us identified possible future design improvements.

5.4.2.1 Eye Tracking Data

We first used the built-in QR code plug-in of the eye tracker to recognize where the participants looked. The results of how long the participants looked on average at the phone in relation to the conversation length is presented in Figure 5.5. A two-way ANOVA revealed a statistically significant difference between the *Phone* and the *NoPhone* conditions $F_{1,20} = 7.155$, $p = .015$. However, we found no significant difference between the *Touch* and the *FingerOrientation* conditions $F_{1,20} = .007$, $p = .934$. Further, we found no significant interaction effect ($F_{1,20} = .066$, $p = .799$).

5.4.2.2 Phone Log Data

To understand if the reaction time as well as the interaction time changes when using *FingerOrientation* as declining method we conducted two one-way ANOVAs.

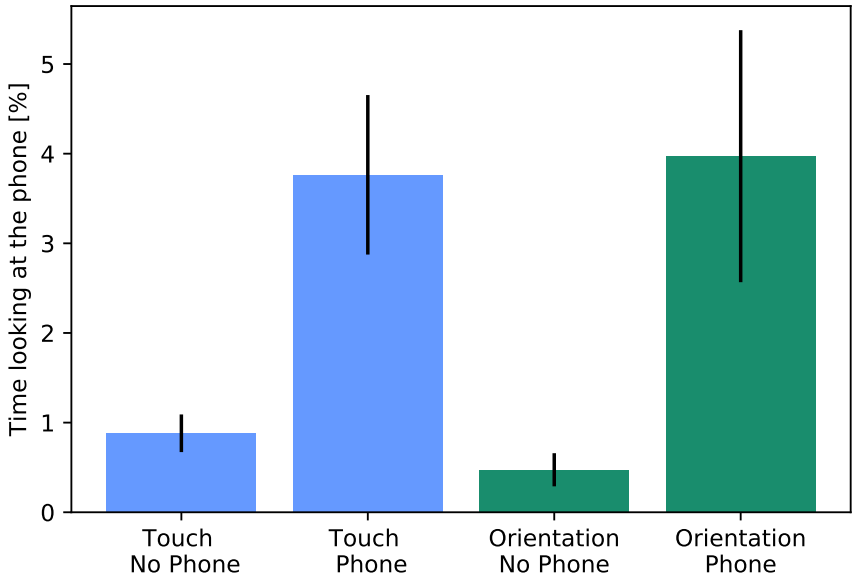

Figure 5.8: The average percentages of looking at the phone during the discussion and standard error for TECHNIQUE × PHONE.

We found no significant difference between the *Touch* and the *FingerOrientation* conditions for the reaction time $F_{1,10} = .001$, $p = .999$, see Figure 5.9a. We found no significant difference between the *Touch* and the *FingerOrientation* conditions for the interaction time $F_{1,10} = .118$, $p = .739$. Participants in the *Touch* condition interacted with the phone for the shorter period of time ($M = 1.1s$, $SD = .2$) than the *FingerOrientation* condition for $M = 1.2s$ ($SD = .7$), see Figure 5.9b.

5.4.3 Discussion

We conducted a study with 12 pairs of participants in a simulated coffee house scenario to investigate if there is a change in conversation disruption when finger orientation instead of normal touch interaction is used. We conducted the study as proposed in Section 5.2. Moreover, the overall study setup was similar to the study presented in Section 5.3.

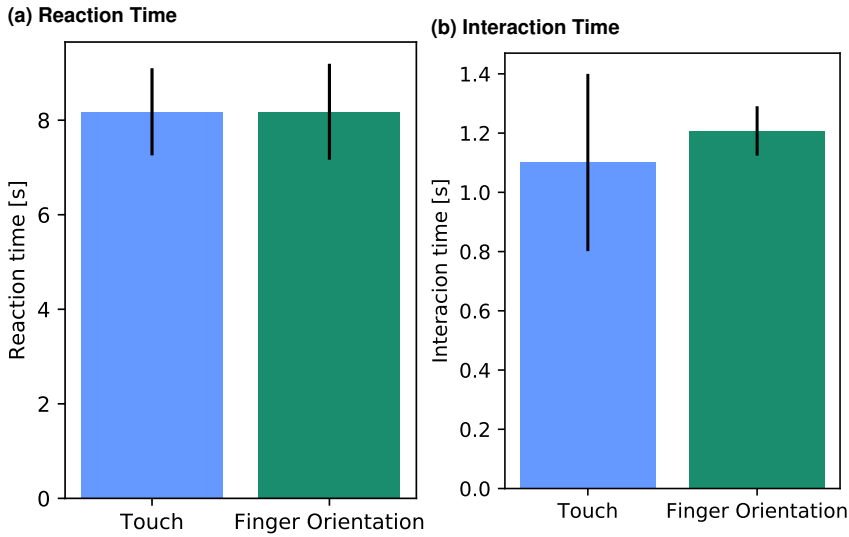

Figure 5.9: (a) Showing the average reaction time (TCT-R) between incoming call highlight and a participant touched the phone. (b) Showing the average interaction time to decline a call.

In line with the results of the study presented in Section 5.3 the participant without a phone looked significantly less on the phone then the participant with the phone. However, finger orientation interaction did not significantly decrees or increases the time looked on the phone. Thus, this indicated that there is no difference in disruptiveness. The further stress this with analyzing the phone logs. Here, we revealed no difference in reaction time and no difference in interaction time with the phone. Thus, we argue that in cases where designers need more dimensions to interact with the phone, they can used finger orientation without risking to disrupt conversions with there implementation.

5.5 Advantages of the Mixed-Method Approach

Participants interpreted the phone as an unnecessary distraction to the conversation, showing the dilemma of notifications caused by technologies in general. However, today's smart device participants rely on notifications. Conducting the two studies enables us to reflect on the properties of our mixed-method approach. In the following, we discuss the applicability of our mixed-method approach for new technologies and new interaction techniques. We further discuss how eye tracking enables a rapid decision on which technology is less disruptive while the interviews complement the choice of technology and point out further improvements.

5.5.1 Rapid Answers

Our first study showed a clear result in terms of which studied interaction technique is superior. The eye tracking data showed a significant effect which strongly indicates that *SurfaceSliding* causes less disruption than the baseline technique. This shows that our approach is well suited to fit in a design process. The fast quantitative analysis can be used for quick A/B testing during design sprints for rapid evaluation of interaction techniques. This is in contrast to past approaches such as conversation analysis. Once a superior technique is determined, designers can use the qualitative feedback to stimulate further refinements of the interaction technique. A study that offers meaningful feedback can be conducted and analyzed within a day for a number of participants.

5.5.2 Focused, on-the-spot Feedback

The conducted studies showed that our approach managed to trigger extensive reactions in the participants. We believe that contextualizing the interview questions enabled the participants to reflect extensively on the disruptiveness caused by interruption techniques. The discussion task served as a kind of priming before the final interviews. We theorize that the extensive material we obtained was partly provoked by the direct experience of being engaged in a discussion. We believe that the fact that participants are interviewed directly after a discussion facilitates focusing specifically on the disruption caused by the technology and

imagine how the discussion could be shaped had the technology been different. As only one of each pair of participants is tasked with using the phone, designers are able to obtain a first-hand account of how one attempts to limit the disruption caused to the interlocutor while using a particular interaction technique. This contributes to building the ecological validity of our method.

5.5.3 Designing with the Social Context in Mind

It is worth noting that some participants were willing to sacrifice some degree of their social engagement for better awareness of the events communicated through their phone. For instance, more interruptions can be acceptable in a personal conversation when one intends to be informed about the current score in an important football match. Such situations produce opportunities for designers to intervene and create tools for personalized interruption rules, which could act like the mobile version of an email filter. While we used a neutral discussion task without any distractions, our approach easily enables modifying the context of the disruption. One can alter the pre-study briefings provided to the user or introduce additional distractors in the study room. Further, we recognize that while we suggest a quick qualitative analysis and subsequent iteration, our approach can potentially offer deeper insights if more detailed analysis is performed.

5.5.4 Potential for a Generative Role

As the participants were engaged directly in a discussion, the reactions triggered by the prototypes studied can be reflected upon immediately. In the final interview, participants can directly reflect on particular disruption caused by particular actions. In later stages of the design process, such accounts enable designers to understand the details of the disruptiveness caused by the prototype and stimulate participants to provide generative feedback. As the participants can easily recall the discussion task, our approach facilitates discussing alternative techniques or usage contexts. Further, for techniques that require bodily movements the direct engagement with the mobile device makes participants more likely to imagine real-life scenarios. Past work has shown that direct physical engagement with

low-fidelity prototypes in social contexts can be particularly beneficial in a design process [163]. Thus, we believe our method can be applied to a variety of design contexts and constitute a rapid and versatile formative evaluation tool.

5.6 Limitations of the Mixed-Method Approach

We also recognize that the new technologies which are studied by the researchers may effect the participant's behavior. In our studies, modifying the user's smartphones would have rendered the study technically unfeasible. Furthermore, as our study investigates a new interaction technique, using multiple smartphone models may play a role in the social acceptance of the interactions. However, using one model guaranteed consistency between the studied groups. Furthermore, the required measurement infrastructure in the study setup itself might influence the participant's behavior. In future work, the influence of wearing an eye tracker and the surrounding infrastructure should be analyzed.

The proposed mixed-method approach is designed to evaluate the distinctiveness in social settings. In detail, we focus on two-people, face-to-face discussions. We believe that the proposed method is also applicable for social settings involving more persons. However, applying our mixed-method approach to a setting with more people needs to be investigated first.

5.7 Summary

In this chapter, we first did an analysis to understand how to study new interaction techniques in social settings. We then developed a study plan based on the requirements. In an example study we showed how the new mixed-method approach can deliver insights on two level. First the mixed-method approach provides the experimenter with a quick A-B test to understand if one interaction technique is less disruptive using eye-tracking data. Second the approach delivers new insights about the interaction and general behavior of the participants within the interaction. This can foster a new iteration in the design circle.

In our example study, we studied a new call declining method in a face-to-face conversation. The new call declining method *SurfaceSliding* works by

sliding the whole phone in the direction of the smartphone in contracts to the standard Android interface where a icon needs to be dragged on the screen itself. Participants where asked to have a conversation for 10*min* while one participant had the modified phone and received 8 calls during the 10*min*. The analysis first relieved that *SurfaceSliding* is less disruptive to both the participant with the phone was well without the phone. Moreover, the interviews reviled how *SurfaceSliding* can be improved further in a next design iteration, but also uncovered how humans today are mixed between how they handle there phones in conversations in general.

In the second half of the chapter, we finally where able to address RQ3: "Which social implications has using finger orientation as an additional input?" Here, we used the new mixed-method approach to study how finger orientation interaction would effect face-to-face conversations. Therefore we implemented a caller where the yaw of the finger is mapped to the selection icon, similar to a pie menu. We used the same study setup and set of interview questions as in the example study so compare them. We found no difference in disruptiveness, which is positive as one can hypothesize that finger orientation input in general leads to more arm movements. However, due to the interviews the we got feedback and insights on how to improve finger orientation in a pie menu selection task in the future.

As we recognize that our studies are constrained by the fact that it was conducted in a lab setting, we hope that using our approach will be complemented by other studies that use alternative methods such as in-the-wild deployments of new interaction techniques. We also believe that an ethnographic study of the social acceptability of smartphone interruptions in public settings such as cafés or libraries will produce interesting insights for design. We hope that our work will inspire further developments and the creation of enhanced evaluation methods for future interaction techniques.

6

Discoverability

To understand how finger orientation input should be presented to the user (RQ4), we investigate how to communicate novel input techniques for smartphones in the following chapter. Through interviews with 12 UX experts, we identified three potential approaches: *Depiction* uses an icon to visualize the input technique, *Pop-up* shows a modal dialog when the input technique is available, and *Tutorial* explains all available input techniques in a centralized way. To understand which approach is most preferred by users we conducted a study with 36 participants that introduced novel techniques using one of the communication methods. While *Depiction* was preferred, we found that the approach should be selected based on the complexity of the interaction, novelty to the user, and the device size.

Parts of this chapter are based on the following publications:

S. Mayer, L. Lischke, A. Lanksweirt, H. V. Le, and N. Henze. "How to Communicate New Input Techniques." In: *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*. NordiCHI '18. Oslo, Norway: ACM, 2018, pp. 460–472. ISBN: 978-1-4503-6437-9. DOI: 10.1145/3240167.3240176

6.1 State-of-the-Art Communication Approaches

A number of input techniques beyond a single touch including force touch and gesture shortcuts are possible on touchscreen devices. However, they are not widely used, often unknown to users, and not well communicated. Recent research proposed further input techniques to enlarge the input space of today's touch devices, including finger-aware input [21], phone squeeze input [68], BoD interaction [27], finger orientation input. While some of these techniques are already available for commercial devices, none have become widely used.

As system's functions have to be learned they are not always obvious. As shown by Müller et al. [134], a visual cue that highlights input possibilities significantly increases how often people interact with a system. Moreover, both Shneiderman et al. [160] and Norman [138] argue for the *discoverability* of interaction and indeed we see many ways to help users to understand new input techniques. Hover effects are, for example, a common way to communicate the possibility to click a button. More complex interactions are harder to communicate. With Word 1997, Microsoft introduced Clippy, a virtual assistant that provided in-situ help for text processing by highlighting possible actions. Clippy was removed six years later and is considered a classic example of how not to foster discoverability [149].

As the affordance of input techniques for touchscreens that go beyond simple touch interaction is limited, novel input techniques for touchscreens must be communicated. The most common approach to introduce novel input techniques is through the GUI. Today, Apple use the "Tips" app to explain how all features of the iOS eco-system work. In cases of an update, Apple triggers notifications to advise users that they can learn about new features in the "Tips" app. On their U11 smartphones, HTC informs users about "Edge Sense" during the device setup and additionally shows a pop-up whenever edge sense can be used within an app. While Apple's "Tips" app and HTC's device setup enable users to understand how to use new input techniques, true discoverability in the sense of Shneiderman et al. [160] and Norman [138] is not achieved. They both argue that a function should be self-explanatory and new input techniques should seamlessly be learned while using the device.

This chapter seeks to understand how user experience (UX) experts envision to communicate input techniques beyond a single touch. Moreover, we aim to understand which communication method is preferred by users. Therefore, we conducted design sessions with UX experts. We asked them how they envision enabling discoverability. We found that designers were split between three different approaches to communicate new input techniques: a) *Depiction*, an approach similar to Shneiderman et al. [160] that highlights available input technique through icons; b) a *Pop-up* which informs users about available input techniques whenever a new one is available; and c) the *Tutorial* which explains all input techniques in a centralized way. We evaluated the three approaches using five different tasks. In each task, the user needed to use a different novel input technique, namely: *Finger Orientation Interaction*, *Finger Roll Interaction*, *Nail/Knuckle Interaction*, and *Finger-Aware Interaction*. We found that participants preferred *Depiction* over both *Pop-ups* and the *Tutorial*.

6.2 Alternative Interaction Techniques

In the follwoing, we highlight four novel input techniques which potentially will make it into consumer devices in the near future. In our studies, the four techniques form the foundation to study ways for communicating novel input techniques. All four have been studied in detail in previous work but are not available for consumer devices: *Finger Orientation Interaction*, *Finger Roll Interaction*, *Nail/Knuckle Interaction*, and *Finger-Aware Interaction*.

6.2.1 Finger Roll Interaction

Roudaut et al. [151] proposed using the roll of the finger for input. They envision a circular clockwise / counterclockwise input by rolling the finger to the side. They argue that the circular gesture can be used to access hidden menus. Huang et al. [73] used the finger roll to implement a keyboard on a smartwatch.

Roudaut et al. [151] distinguish between taps, strokes and roll inputs by analyzing the trajectory of the touch input. Hernandez-Rebollar et al. [63, 64] used six dual-axis accelerometers attached to the fingers to track the position and

the roll of the fingers. Huang et al. [73] also used inertial measurement sensors to implement a keyboard that assigns different characters to different areas of users' finger pads contacting the touchscreen.

6.2.2 Nail/Knuckle Interaction

The most prominent work regarding nail/knuckle interaction is by Harrison et al. [58]. They envision using a normal tap as one input and further distinguish between knuckle, nail and fingertip. Lopes et al. [108] use different hand gestures for actions such as copying, pasting and deleting objects on a tabletop. Lastly, Hsiu et al. [72] used nail deformation as an indirect measurement to estimate the "force" on the touchscreen.

Harrison et al. [58] identify the different inputs based on changes in the acoustical spectrogram retrieved from conventional medical stethoscope with an electret microphone. In contrast, Lopes et al. [108] use the sound of the gesture for input identification. They used the the characteristics of the amplitude envelope and the fundamental frequency to detect different interactions.

6.2.3 Finger-Aware Interaction

Finger-aware interaction is mostly used with a specific finger as a modifier of a touch event, allowing different fingers to be responsible for different actions. Colley and Häkkilä [21] used finger-aware interaction to map different functions onto the fingers themselves. For instance, they envisioned navigating the contact app with different fingers, e.g., opening a contact using the index finger and making a call by tapping the contact with the thumb. Gupta and Balakrishnan [55] implemented a smartwatch keyboard which makes uses of finger-aware interaction by mapping two characters to one key, and depending on the finger used one of the two characters is sent to the application layer. Gupta et al. [54] proposed "Porous Interfaces". Two applications are stacked on top of each other with a semi-transparent front layer. They envisioned an interaction where one finger can interact with the front application and another with the application in the background.

Finger identification approaches that attach sensors to the user generally yield the best recognition rate. A large body of work applied infrared sensing from

beneath a tabletop for finger-aware interaction [1, 31, 41]. Gupta et al. [54, 55] used infrared sensors mounted on different fingers to identify touches made by the index and middle finger. Similarly, Masson et al. [115] based their recognition on touchpads using vibration sensors attached to the user's finger. Further approaches include using electromyography [4], gloves [113] and RFID tags [168]. Another approach uses cameras to identify touches from different fingers. Researchers predominantly used a combination of RGB cameras and computer vision for detection [174, 190].

6.3 Design Sessions

To explore ways to communicate new input techniques to the user, we conducted an interview series with 12 UX experts. We recruited the experts (9 male and 3 female) from two leading design universities and one institute focusing on HCI. All interviews were audio recorded for later analysis. For the assessment of the four input techniques, we used a Latin square design to balance the order.

6.3.1 Procedure

After the experts were welcomed, they were asked to sign a consent form and fill in a questionnaire about demographics. Then we introduced them to the interview and explained its overall intent: "How should a touchscreen system introduce new input techniques?" Participants had the chance to ask questions throughout the study. After the general introduction, we informed the participants about the four input techniques using a slideshow, namely: *Nail/Knuckle Interaction*, *Finger Orientation Interaction*, *Finger Roll Interaction*, and *Finger-Aware Interaction*, see Figure 6.1. For each input technique we had an idea creation phase where we asked the experts to imagine how the input techniques could be used in the mobile devices' most popular types of applications [9], such as instant messaging, browsing, and email apps.

After the idea creation session we interviewed the experts in depth on each of the four input techniques. Following this they each chose one of their use-cases for a more in-depth interview comprising 13 questions to ensure good designs as laid out in the "Eight Golden Rules" by Shneiderman et al. [160] and the "Seven

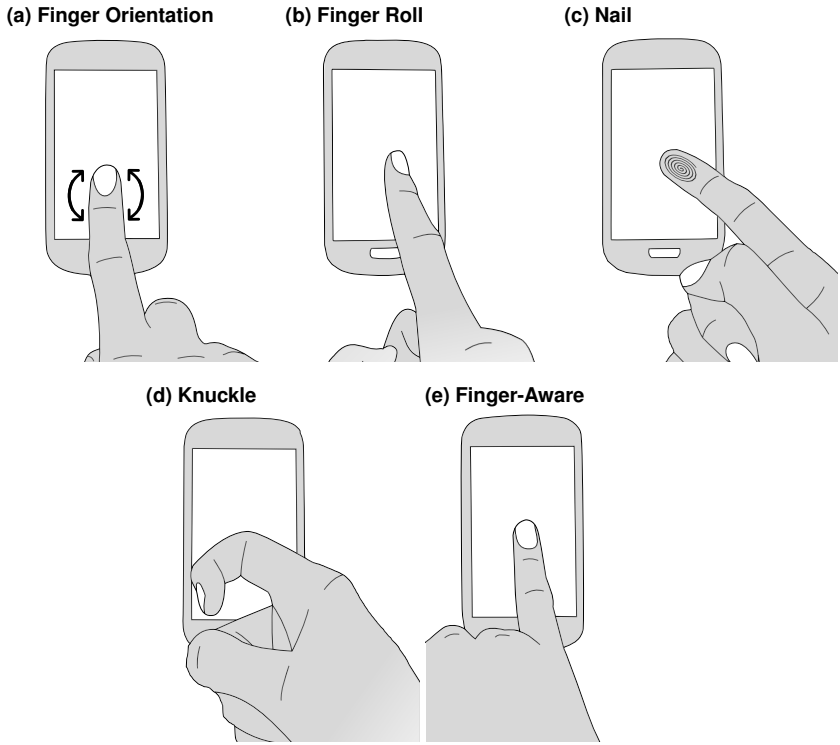

Figure 6.1: The input techniques which were used to study possible communication patters for novel input techniques: *Finger Orientation Interaction*, *Finger Roll Interaction*, *Nail/Knuckle Interaction*, and *Finger-Aware Interaction*.

Fundamental Design Principles" by Don Norman [138]. For each input technique, we gave the experts a sheet of paper with five designated sections for drawings, labeled (1) pre interaction, (2) interaction possibilities, (3) during the interaction, (4) after the interaction, and (5) possible error stats. We asked the experts to use the sections they needed to sketch their ideas.

We wrapped up the interview with final remarks and answered remaining questions. Lastly, we thanked the experts for their participation in our expert interview and reimbursed them with 10 €.

6.3.2 Results

We conducted 12 expert interviews with a total length of $1,005min$ ($M = 83.3min$, $SD = 7.2$, $Min = 60$, $Max = 120$). We transcribed all interviews and coded them using Atlas.ti[1]. We transcribed the interview literally while not summarizing or transcribing phonetically. However, we transcribed pauses of greater than one second to understand the conversation further. This technique is known to offer a subjective experience [7]. Next, three researchers coded one interview of the material to establish an initial coding tree. A single researcher coded the rest of the data. Finally, we employed a simplified version of qualitative coding with affinity diagramming [57] for interview analysis as this offers a rapid way to analyze and understand the feedback provided by interviews. In the following, we first present insightful comments from the idea creation session and then about the four discussed input techniques. To relate opinions, we name the experts E1 to E12. A set of sketches drawn by the experts is shown in Figure 6.2.

Summarizing how the experts rated the intuitiveness of the input techniques, only 3 experts considered *Nail/Knuckle Interaction* as the most intuitive input technique, followed by the *Finger Roll Interaction* where 6 experts found them to be generally intuitive. Lastly, both *Finger Orientation* and *Finger-Aware Interaction* was found to be generally intuitive by 7 experts.

6.3.2.1 Finger Orientation Interaction

As discussed earlier, finger orientation input has, unlike the other input techniques, two dimensions, which can be changed at the same time. Further, in previous research, finger orientation has often been studied as a single input technique. This is reflected in the interviews. Experts either used it as a combined input techniques where two parameters can be changed at the same time or as two independent operations.

The experts envisioned using *Finger Orientation Interaction* for several use cases. They generally considered the input technique to be mainly useful for manipulating views. Manipulations such as zooming, which today is typically realized using two fingers, can be substituted using the orientation of the finger.

[1] `http://atlasti.com/de/produkt/v7-windows/`

Here, zooming (E1, E3, E6, E11) and scrolling (E4, E5, E6, E9) were named as examples for fundamental input techniques. Further, manipulating a 3D view as a more complex use case was envisioned (E2, E5, E7, E8). It could be, for example, used to manipulate an object or to zoom and rotate a map at the same time. Furthermore, E10 imagined changing values by changing the orientation of a finger. Thereby, the user could select dates in a calendar using the pitch of the finger. Similarly, E1-E4 envisioned setting the time or a timer using the yaw of the finger. The experts also imagined accessing different shortcuts with each angle of the finger (E3, E6, E9, E12) or mapping it to a brush type or a brush size (E5, E8) in a drawing application. E1 and E10 proposed using uncomfortable finger orientations for safety-critical actions, e.g., factory reset.

Eight experts considered pop-ups to be an appropriate way for communicating the input technique to users (E1, E2, E4-E6, E8, E11, E12). Moreover, E10 suggested a more intuitive way to communicate the input technique, where the user is guided by an interactive animation to learn how the new input technique works. Furthermore, E5 and E11 suggested using a tutorial to explain the input techniques. Using icons to visualize the new input technique, thus following the depiction method, was mentioned by E3.

The experts generally agreed that smartphones are well-suited for implementing finger orientation input. Five highlighted that finger orientation is also well suited for input on smartwatches; on the other hand, finger orientation on tablets was only highlighted three times. Additionally, E9 stated that finger orientation input should always be implemented as a relative input, as performing absolute angles is difficult for users.

6.3.2.2 Finger Roll Interaction

Experts considered *Finger Roll Interaction* to be useful for switching between views (E1, E4-E9); either to switch between apps or in an app switch between views. As in-app use cases, the experts proposed moving between one messenger conversation and another or to flip pages in an ebook. Switching between views using roll input could also be used to manipulate GUI elements such as a “Switch” or toggle button (E1, E4-E9). This switching function could also be implemented as a scrolling function according to 3 experts (E2, E9, E12). On the other hand,

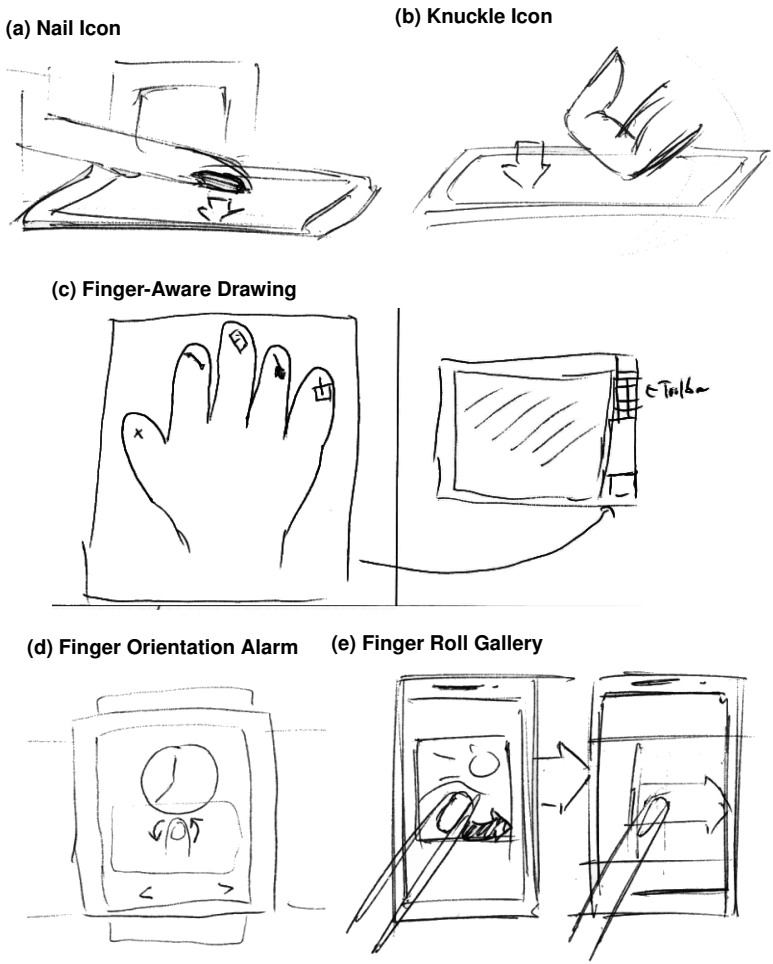

Figure 6.2: Sketches drawn by the experts during the interview to underline their strategies for their use cases. (a) and (b) present possible *depiction* icons to guide the user to use their nail or knuckle as input. (c) - (d) present three different use cases each for one input technique.

again experts made use of rolling as a continuous input for GUI elements such as adjusting a thermostat (E10) or to set a position on a slider as used for music and video player manipulation (E2). Further, two experts (E3 and E8) envisioned the *Finger Roll Interaction* to control games. Lastly, a shortcut menu similar to Roudaut et al. [151] was mentioned by E6.

The experts proposed two basic approaches for communicating *Finger Roll Interaction* to users: (1) using a pop-up and (2) using an icon that depicts the interaction. Here, E1, E4-E9, E11, and E12 suggested using pop-ups. E2, E3, E8, E10 suggested depiction to communication the interaction. The experts envisioned using an icon combined with a specific way of guiding the user to the interaction. For the guidance, the experts envisioned a transformation of the touched object whenever a *Finger Roll Interaction* is possible. For instance, E2 suggested transforming the "play" button in a music app into a slider when skimming through the song is possible using *Finger Roll Interaction*.

The experts generally envisioned *Finger Roll Interaction* to be used on all screen sizes. However, E5 and E8 had concerns in regards to using *Finger Roll Interaction* on smartwatches.

6.3.2.3 Nail/Knuckle Interaction

In contrast to *Finger Orientation* and *Finger Roll*, this interaction uses categorical input rather than continuous input dimensions. This led to two different types of actions in the interviews. However, the actual use of nail or knuckle can easily be applied to the other input techniques. Most of the experts stated that input technique could be implemented for system-wide actions.

All experts saw nail and knuckle input as a perfect solution for shortcuts, such as taking a screenshot (E1, E4-E8, E11, E12), undo (E2, E3, E8), marking mail as spam (E9), snoozing of the alarm (E7), and within music applications (E10). Furthermore, the input technique could be used to select multiple objects and for scrolling, similar to finger-aware interaction (E2). Nail and knuckle input was further envisioned for unlocking or turning on the screen using a knock (E8, E9) and opening the context menu (E1, E4). E3 would use the input for safety-critical input like a factory reset. E4 had the idea to replace already existing functions like long-press replacement.

Experts generally agreed on two ways to communicate the new interaction; first, by showing a pop-up, when the interaction is available for the first time. However, as most of the proposed use cases are system-wide operations, the experts also proposed explaining the input technique during the setup of the device in a tutorial.

Four experts stated that whenever a special action is triggered visual feedback to the user would be beneficial. Four experts proposed a growing wave similar to the pattern a drop produces on a water surface. Furthermore, experts see the usefulness of nail and knuckle interaction as rather limited. Four experts considered the input technique to be useful for all touchscreen devices, two only for smartphones, and one for tablets. Lastly, two experts (E2, E9) argued that there might be problems in using the nail input with long nails and that this should be studied independently.

6.3.2.4 Finger-Aware Interaction

Experts proposed finger-specific shortcuts (E2-E4, E8, E10, E12) for certain apps such as calendars (E2, E4, E10) or to stop an alarm (E2). They also proposed different tones for each finger in a piano application (E10). Three experts (E3, E10, E12) saw a benefit for drawing apps. They envisioned two different approaches, either to map a different color to each finger or to map different brushes to each finger. Another area was the text editing domain. E1 and E6 envisioned copy and paste using two dedicated fingers, and E7 proposed enhancing caret positioning using finger-aware input. A specific finger could be used to select whole words, unlike today's implementation of caret manipulation. E8 and E9 envisioned using a specific finger open a system-wide context menu. On the other hand, multi-finger shortcuts have been proposed for app switching similar to the iOS implementation (E4). E4 and E5 see a benefit for finger-aware interaction on keyboards, where for example italic text could be realized using one finger, or one finger used to enter the second layer of characters on each key to substitute the long-press. Both E2 and E11 proposed a GUI element with a maximum of five options, one per finger. They envision this to be similar to a slider, without taking up the space on the screen to fit a long slider widget.

The majority of the experts drew a hand like a symbol to communicate the different option per finger to the user. However, they again used the representations in different ways to explain the finger-aware interaction to the user. Six experts (E1, E4, E6, E9, E11, E12) stated they would use icons with text to communicate the interaction, with the two options of when the device is getting set up or when the interaction is available for the first time pop-ups. E2, E3, E5, E7, E8, and E10 preferred a depiction as the form of communication. Moreover, three of the experts stated that they would see the benefit of finger-aware interaction for larger screens (E5, E7, E9).

6.3.3 Discussion

To understand how UX experts would design ways to communicate new input techniques, we asked them which use cases they envisioned and how they would communicate the input techniques to users. We asked them to envision use cases for the following four input techniques: *Nail/Knuckle Interaction*, *Finger Orientation Interaction*, *Finger Roll Interaction*, and *Finger-Aware Interaction*. They did so, then each elaborated on their favorite use case in-depth. They envisioned how this use case would work with the new input technique and how they would communicate this to users.

We found that experts are split between three methods to introduce a new input technique. The most common method was to use *Pop-ups* whenever a new input technique is available. Second, we found that for interaction techniques which they found to be intuitive they suggested using less obtrusive *Depiction* (e.g. icons) to communicate a new technique. Last, the experts suggested using an introduction during device setup using a *Tutorial* where the user is guided through a process and the option to revisit the tutorial as in the iPhone's "Tips" app. The results of the design session showed that the experts envisioned a wide variety of use cases but focused on three different methods to communicate new input techniques to users. They would choose a given method on the basis of how intuitive they considered the input technique to be. In the following, we compare the three communication patterns: *Depiction*, *Pop-up*, and *Tutorial* using a study where users are asked to learn and to perform the new input techniques.

Depiction: a small icon next to the element of interest in the GUI depicting the available input techniques. The depiction is intended to work without additional textual explanations.

Pop-up: a modal dialog which appears the first time an input technique is available in the next view. The pop-up contains a textual description and visual depiction.

Tutorial: an introduction into all new input techniques at once, either when the input technique becomes available through an update or when setting up the device, again, using a combination of textual description and depiction.

6.4 Evaluation

Based on the findings from the interview series, we designed a lab study in which we compared the three communication methods *Depiction*, *Pop-up*, and *Tutorial* with regards to their UX.

6.4.1 Study Design

We conducted a lab study to compare the three methods for communicating new input techniques proposed by the UX experts. Namely we compare the COMMUNICATIONPATTERNS: *Depiction*, *Pop-up*, and *Tutorial*, see Figure 6.3. We prototyped five different TASKS: *Alarm*, *Chat*, *Drawing*, *Gallery*, and *Map*, see Figure 6.4. To minimize the influence of unreliable novel implementations of the discussed touchscreen-based input techniques we used a Wizard-of-Oz study design [25]. We conducted the study with COMMUNICATIONPATTERNS as a between-subjects variable while TASKS was a within-subjects variable. This ensures that no participants had experience of new input techniques with multiple COMMUNICATIONPATTERNS. We used the system usability scale (SUS) [13], the AttrakDiff [59] questionnaire, and three open questions as depended variables.

In the *Chat* task, the participant had the option to use *Nail/Knuckle Interaction* to enrich the interaction. To cover the *Finger Orientation Interaction* we added two separate tasks to enable the wizard to recognize the movement accurately. In the *Alarm* task, participants had to rotate the finger around the yaw axis to change

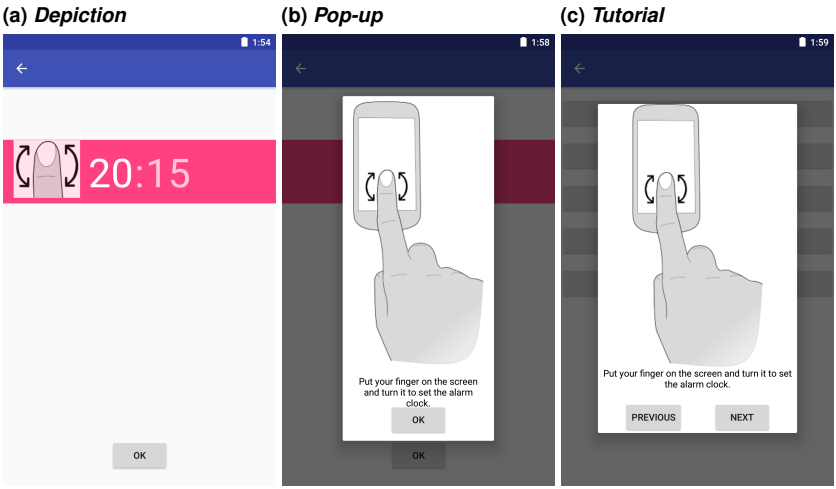

Figure 6.3: The three different COMMUNICATIONPATTERNS which were proposed by the experts in the design session.

the time. In the *Map* task, the pitch of the finger manipulates a map view. In the *Gallery* task, *Finger Roll Interaction* is used to scroll through images. Finally, *Finger-Aware Interaction* is used for a *Drawing* application, where each finger is mapped to a different color.

6.4.2 Apparatus

We used a Nexus 5X Android smartphone for the participants to learn and perform the new input techniques and a Nexus 7 for the wizard. Bluetooth was used to send the commands from the wizard to the smartphone used by the participants. We audio recorded the participants' responses to the open questions. Further, we recorded the whole study using a GoPro Hero3+.

Alarm task: participants were asked to set five different times by changing the yaw orientation of their finger while touching the screen: Clockwise rotation increased the time. The input technique was realized as a relative input always starting from the last value. For the *Depiction* condition, we displayed an icon with two curved arrows around the finger as proposed by the experts, see Figure 6.4a.

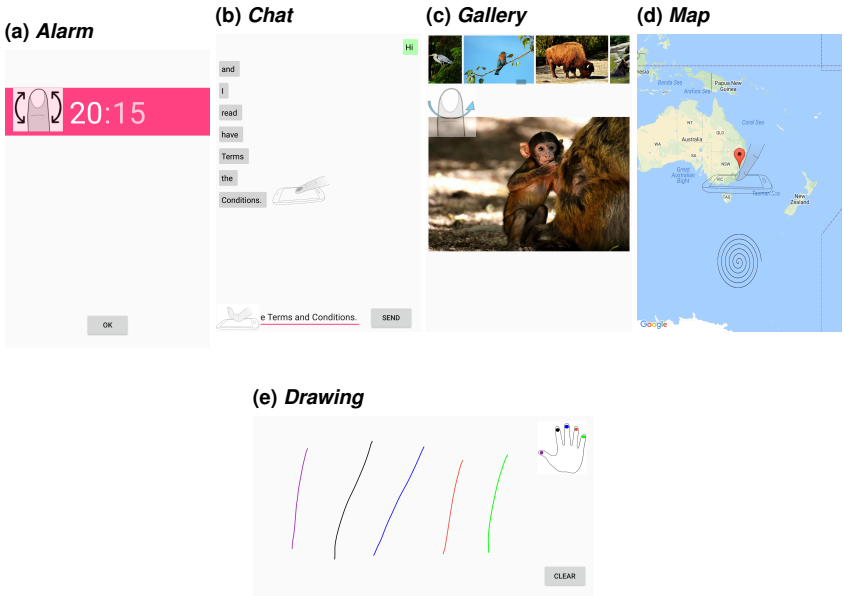

Figure 6.4: The five different TASKS used in the evaluation study.

Chat task: we implemented shortcuts as proposed by the experts. Touching a text using the nail copied the text and touching with the knuckle pasted the text from the clipboard. The task was to agree to terms and conditions by pasting "I have read the Terms and Conditions" into a text field word by word. Experts proposed depicting the nail and knuckle, see Figure 6.4b for the icons used in this task.

Drawing task: participants were asked to draw a scene from their last vacation, a meal, a car, a pet and an island. Participants were further asked to use at least three different colors. Each color was assigned to one finger; the color assignment being shown by a small hand icon, see Figure 6.4e. By touching the hand participants were able to remap and change colors.

Gallery task: participants were asked to find five specific images in a gallery containing the 100 image[1] using *Finger Roll Interaction*. Scrolling through the images was possible by rolling the finger and visualized with an arrow over an fingertip, see Figure 6.4c. The position of the *Finger Roll Interaction* was not taken into account. The target images were printed on paper.

Map task: participants were asked to use a map for finding six cities, each on a different continent. Moving the map was possible through panning with the finger, while zooming in and out of the map was realized by changing the pitch of the finger while still touching the screen. This again was visualized by an icon representing the finger and its pitch in relation to the device, see Figure 6.4d.

6.4.3 Procedure

After welcoming the participants, we explained the purpose and the procedure of the study. Afterward, we asked them to fill out a consent form and a demographics questionnaire. During the whole study, the participants were seated on a chair, the wizard (experimenter) was sitting directly opposite to the participant, with a table in between. The study started by handing the smartphone to the participant. In the *Tutorial* condition, the participant first learned about all input techniques using the tutorial and then started with the TASKS. In the other conditions, the participants directly started with the tasks A pop-up informed them about the input technique in the *Pop-up* condition and an icon representing the input technique was displayed in the *Depiction* condition. The order of the tasks was randomized. No further information was given by the experimenter; however, after each task, participants were asked three questions: (1) Did you feel comfortable performing the input? (2) Did you like the method introducing the input technique? and (3) Do you have suggestions for improving the introduction method?

6.4.4 Participants

We recruited 36 participants (23 male and 13 female). The participants were aged from 20 to 29 years ($M = 24.2$, $SD = .38$). The majority (21) of them were

[1] All images used in the study are under Creative Commons CC0 available at: `pixabay.com`

Android users, 13 were iOS users, and only 2 were Windows Phone users. In total, the study took between 30 and 40*min* per participant. We reimbursed the participants with 5 €.

6.5 Results

In total 36 participants rated 180 interactions, each using an SUS and an AttrakDiff. The UX of each of the three COMMUNICATIONPATTERNS was assessed by 12 participants in a between-subjects design. Thus each participant filled in five SUS and five AttrakDiff, one for each TASK. Additionally, they answered a set of three questions regarding the COMMUNICATIONPATTERN for each TASK. The audio recordings were transcribed by two researchers and we performed a simplified qualitative analysis with affinity diagramming on the interview data [57].

6.5.1 System Usability Scale (SUS)

To conduct a two-way mixed model ANOVA, we applied the Aligned Rank Transform (ART) [180] to the SUS scores, using the ARTool toolkit[1] to align and rank our data.

We conducted a two-way mixed model ANOVA to determine whether TASK and COMMUNICATIONPATTERN significantly influenced the usability of the interaction, see Figure 6.5. For all means and standard deviations see Table 6.1. Our analysis revealed significant main effects for TASK and COMMUNICATIONPATTERN on the SUS score ($F_{4,132} = 5.975$, $p < .001$; $F_{2,33} = 7.783$, $p < .002$, respectively). However, there were no significant two-way interactions between TASK $\times$ COMMUNICATIONPATTERN ($F_{8,132} = 1.276$, $p = .261$). Pairwise post-hoc comparisons using Tukey's method for p-value adjustment within the levels of the main factor COMMUNICATIONPATTERN revealed significant differences of the SUS score between *Depiction* vs. *Pop-up* ($t_{147.78} = 3.142$, $p < .006$) and *Depiction* vs. *Tutorial* ($t_{147.78} = 3.637$, $p < .002$). However, the pairwise comparisons did not reveal a significant difference for *Pop-up* vs. *Tutorial* ($t_{147.78} = .495$, $p = .874$).

[1] `http://depts.washington.edu/madlab/proj/art/index.html`

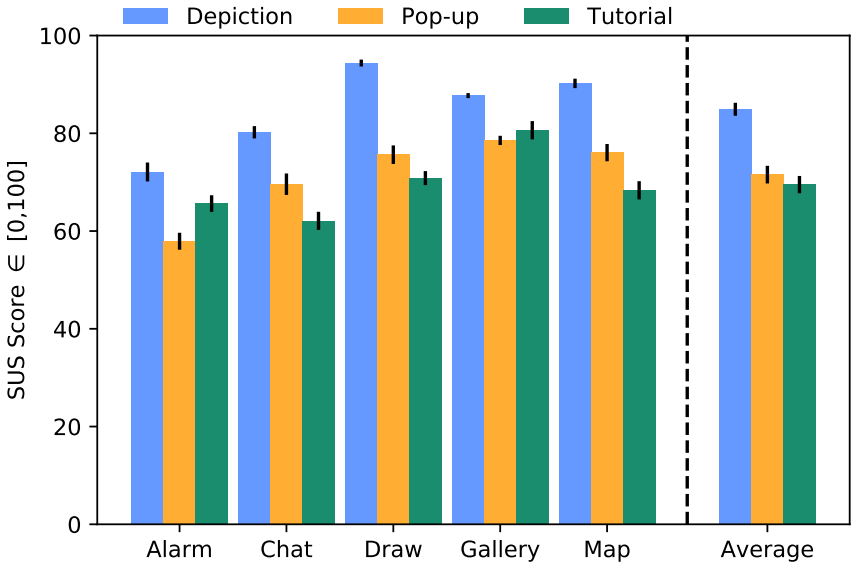

Figure 6.5: The system usability scale (SUS) results of COMMUNICATIONPATTERN × TASK. Error bars are showing the standard error.

6.5.2 AttrakDiff

To conduct a two-way mixed model ANOVAs, we again applied the Aligned Rank Transform (ART) [180] to the three scores of the AttrakDiff, using the ARTool toolkit to align and rank our data. We performed four two-way mixed model ANOVAs one for each scale: Pragmatic Quality (PQ), Hedonic Quality-Identity (HQ-I), Hedonic Quality-Simulation (HQ-S), and Attractiveness (ATT). For all means and standard deviations see Table 6.2.

We conducted a two-way mixed model ANOVA to determine whether TASK and COMMUNICATIONPATTERN significantly influenced the Pragmatic Quality (PQ), see Table 6.2 and Figure 6.7. Our analysis revealed significant main effects of TASK and COMMUNICATIONPATTERN on the PQ score ($F_{4,132} = 10.045$, $p < .001$; $F_{2,33} = 5.553$, $p < .01$, respectively). However, there were no significant two-way interactions between TASK × COMMUNICATIONPATTERN ($F_{8,132} = 1.3$, $p = .249$). Pairwise post-hoc comparisons using Tukey's method for p-value

adjustment within the levels of the main factor COMMUNICATIONPATTERN revealed significant differences of the PQ score between *Depiction* vs. *Tutorial* ($t_{125.79} = 3.256$, $p < .005$). However, the pairwise comparisons did not reveal significant differences for *Depiction* vs. *Pop-up* ($t_{125.79} = 2.244$, $p = .068$) and *Pop-up* vs. *Tutorial* ($t_{125.79} = 1.012$, $p = .571$).

We conducted a second ANOVA to determine whether TASK and COMMUNICATIONPATTERN significantly influenced the Hedonic Quality-Stimulation (HQ-S), see Table 6.2. Our analysis revealed no significant main effects nor a significant two-way interaction ($p > .05$), see Table 6.2.

We conducted a third ANOVA to determine whether TASK and COMMUNICATIONPATTERN significantly influenced the Hedonic Quality-Identity (HQ-I), see Table 6.2 and Figures 6.6 and 6.7. Our analysis revealed significant main effects for TASK on the HQ-I score ($F_{4,132} = 4.071$, $p < .004$). However, there were no significant main effect for COMMUNICATIONPATTERN and no significant two-way interaction between TASK $\times$ COMMUNICATIONPATTERN ($F_{2,132} = 1.129$, $p = .336$, $F_{8,132} = .851$, $p = .56$, respectively).

Lastly, we conducted a fourth ANOVA to determine whether TASK and COMMUNICATIONPATTERN significantly influenced the Attractiveness (ATT), see Table 6.2. Our analysis revealed significant main effects for TASK on the ATT score ($F_{4,132} = 9.275$, $p < .001$). However, there were no significant main effect

	Depiction		**Pop-up**		**Tutorial**	
	M	**SD**	**M**	**SD**	**M**	**SD**
Alarm	72.1	23.3	57.9	20.8	65.6	20.4
Chat	80.2	15.1	69.6	26.3	62.1	22.3
Drawing	94.4	8.5	75.6	22.7	70.8	17.2
Gallery	87.7	6.2	78.5	11.4	80.6	22.5
Map	90.2	11.6	76.	21.2	68.3	22.4
Mean	84.9	15.9	71.5	21.7	69.5	21.3

Table 6.1: The system usability scale (SUS) results of COMMUNICATIONPATTERN $\times$ TASK, SUS score translate in letter grades as follows: 65.0 - 71.0 = “C”, 71.1 - 72.5 = “C+”, and 84.1 - 100.0 = “A+” [155].

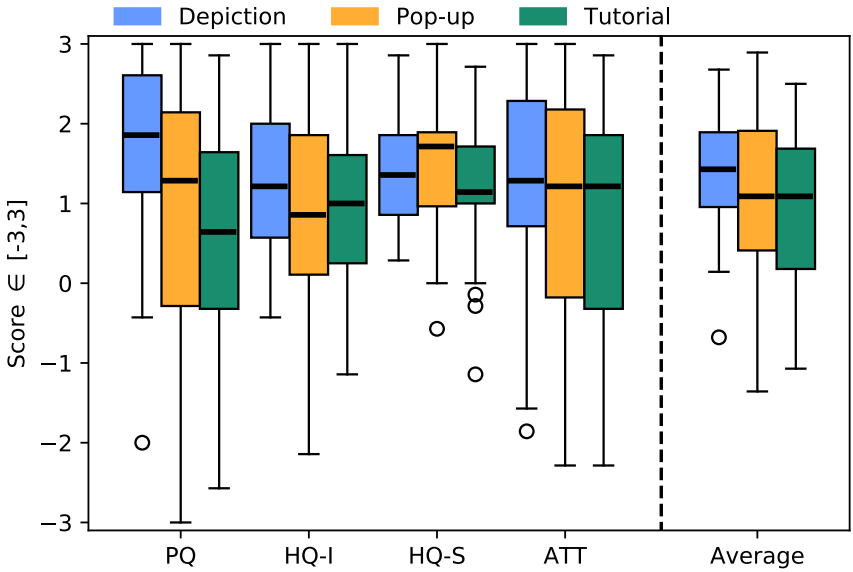

Figure 6.6: The AttrakDiff results of the four categories Pragmatic Quality (PQ), Hedonic Quality-Identity (HQ-I), Hedonic Quality-Simulation (HQ-S), and Attractiveness (ATT) for the three COMMUNICATIONPATTERNS.

for COMMUNICATIONPATTERN and no significant two-way interaction between TASK × COMMUNICATIONPATTERN ($F_{2,132} = 1.129$, $p = .434$, $F_{8,132} = .885$, $p = .531$, respectively).

6.5.3 Qualitative Results

Asked if they felt comfortable performing the input techniques, participants provided generally positive feedback. However, the *Alarm* task stood out with 17 out of 36 (47.2%) participants considering this interaction uncomfortable. All other tasks were considered uncomfortable by fewer than 10 participants. The *Drawing* tasks seemed to be the most comfortable tasks as they only received negative comments by four participants.

Next, participants were asked to comment on the communication method. Here, we found that the GALLERY task was the most criticized across all COM-

	Depiction		Pop-up		Tutorial	
	M	**SD**	**M**	**SD**	**M**	**SD**
PQ	1.71	.14	.9	.21	.69	.18
HQ-I	1.3	.11	.86	.17	.95	.12
HQ-S	1.38	.08	1.53	.1	1.27	.1
ATT	1.35	.15	1.01	.19	.86	.17
Mean	1.44	.18	1.08	.17	.94	.14

Table 6.2: The AttrakDiff results of the four categories Pragmatic Quality (PQ), Hedonic Quality-Identity (HQ-I), Hedonic Quality-Simulation (HQ-S), and Attractiveness (ATT) of COMMUNICATIONPATTERN × TASK. All scales range between -3 and 3.

MUNICATIONPATTERNS (6 × *Depiction*, 3 × *Pop-up*, and 5 × *Tutorial*). On the other hand, in the *Drawing* task, only the *Pop-up*, and *Tutorial* were criticized. All other 164 comments were positive.

Participants provided several comments improving the input techniques. However, in regards to the COMMUNICATIONPATTERNS participants had two major suggestions. First, participants asked for an animation instead of static icons 50 of the 180 (27.8%) times (16 × *Depiction*, 17 × *Pop-up*, and 17 × *Tutorial*). Second, 16 times participants recommended a video to explain the input techniques (1 × *Depiction*, 7 × *Pop-up*, and 8 × *Tutorial*).

6.6 Discussion

We conducted a mixed design study with 36 participants. Each participant performed five different TASKS, each with a different input technique. The novel input techniques were communicated in three different ways either through *Depiction*, *Pop-up*, or *Tutorial*. Each participant was only subject to one of the three COMMUNICATIONPATTERNS. We are mainly interested in how the different COMMUNICATIONPATTERNS influenced the participants' ratings rather than how the TASKS performed against each other. Thus, the discussion focuses on comparing the COMMUNICATIONPATTERNS.

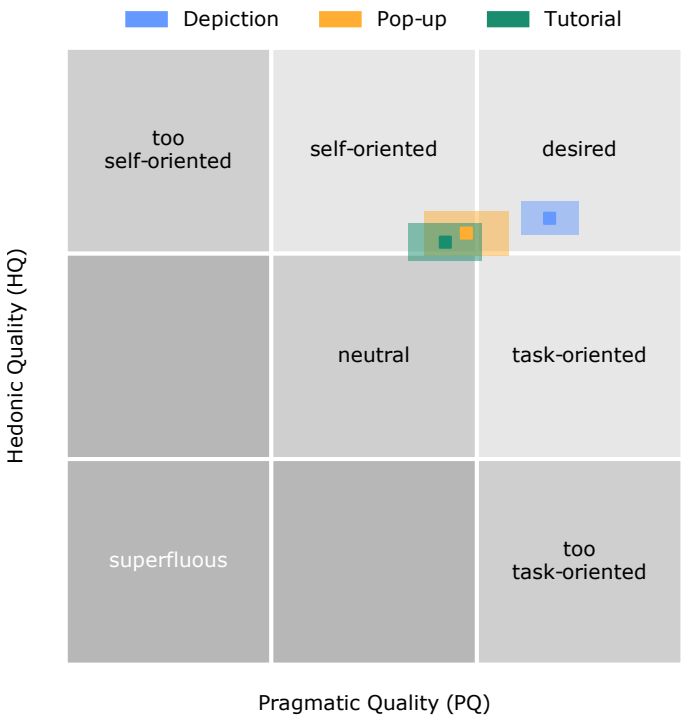

Figure 6.7: Portfolio presentation graph comparison of the AttrakDiff, with Hedonic Quality (HQ) = Hedonic Quality-Identity (HQ-I) + Hedonic Quality-Simulation (HQ-S).

Looking at the SUS results, our analysis revealed that the *Depiction* method to communicate new input techniques outperformed both the *Pop-up* and the *Tutorial* in terms of overall usability of the techniques. Moreover, the portfolio presentation of the AttrakDiff charted the *Depiction* in the "desired" area while the other COMMUNICATIONPATTERNS were positioned in the less attractive "self-oriented" area. However, only the Pragmatic Quality (PQ) is significantly different for *Depiction* vs. *Tutorial*.

A number of participants commented on the icon for visualizing the available input technique. Across all COMMUNICATIONPATTERNS, they asked for an animation. Moreover, for the *Pop-up* and *Tutorial* they would have liked a video to guide them through the procedure of the new input technique.

Summarizing our results we found that users prefer the *Depiction* approach using icons over both *Pop-up* and *Tutorial* with regards to the SUS, the Pragmatic Quality (PQ) of the AttrakDiff and the qualitative feedback. Therefore, our results are in line with the design recommendation by Shneiderman et al. [160] and Norman [138]. On the other hand, today's consumer devices provide features that lack easy and intuitive discoverability. Thus, they need to use tutorials while setting up a new device or using pop-ups. This is not only true for new devices but also for new in-app features. As a result of our studies, we conclude that *Depiction* is generally preferred by users. However, we also see advantages of the other methods which would suggest that using *Pop-up* or *Tutorial* can in some cases also be beneficial.

Depiction offers an in-situ visualization of the "simple" interactions [153] directly within the GUI. While this has the advantage that the user is informed about the input technique right on the spot where the technique is used, the representation is limited to a small visual footprint, similar to the fingerprint icon for unlocking the phone. Therefore, long explanations cannot be embedded within a *Depiction* and the representation always uses display space not only when the interaction is new to the user. Moreover, while animating the *Depiction* is possible, this will guide the users' attention away from the content towards the interaction where the GUI should enable to perform a task and not distract the user.

Pop-ups enable developers and designers to a communicate "compound" interactions [153] (multiple gestures as one single input) in different levels of detail. A simple icon combined with text is one option; however, animations or even videos can also be used to communicate input techniques to users. The drawback of *Pop-ups* is that they disrupt the interaction flow and force users to switch the context whenever the *Pop-ups* show up to teach a new input technique.

Tutorials are similar to *Pop-ups* as they can communicate "compound" interactions, but also enable developers and designers to communicate more conditional "compound" interactions and even multiple input techniques at the same time. While the workflow of the user is not interrupted by *Tutorials*, the user is asked to learn multiple input techniques at once which increases the workload and can be confusing.

6.7 Design Implications

We derived the following design implications for the three approaches *Depiction*, *Pop-up*, and *Tutorial* to communicate novel input techniques to users.

Interaction complexity dependent communication. "Simple" input techniques should be explained through *Depiction*. "Compound" input techniques should be explained through *Pop-ups* and conditional "compound" input techniques through a *Tutorial*.

Animate if possible. *Pop-ups* and *Tutorials* should be animated and presented in a visually compelling way. However, *Depiction* should be only animated when an input technique is available for the first time; later no animation should be used to avoid distracting the user.

Make use of the screen space. *Pop-ups* are preferable to *Depiction* for small screen sizes to save the space for displaying content. For large screens *Tutorials* are preferable to *Pop-ups* as an extra side view can present all information without cutting down on the user's content.

6.8 Summary

In this chapter, we investigated RQ4: "How should finger orientation input be communicated to the user?" We first conducted design sessions with 12 UX experts and found that in general there are three approaches for communicating new input techniques, namely: *Depiction*, *Pop-up*, and *Tutorial*. To understand each approach, we conducted a study in which 36 participants were taught new input techniques to perform five different tasks using one of the three approaches. Based on the findings of both studies we derived three design implications for how to communicate new input techniques. In particular, we found that the approach should be selected based on the complexity of the interaction, novelty to the user, and the device size.

While we derived a set of three concrete design implications to introduce users to new input techniques, future research should investigate the long-term effects of each approach as our study was conducted in a lab environment. Here,

future research should focus on long term memory effects. Especially when using *Pop-ups* and *Tutorials*, new input techniques might be forgotten over time. As our study was conducted in a lab setting, this possibly influenced the participants' ability to identify the new interaction. Thus, the input techniques should be deployed in real-life tasks which would enable in-the-wild evaluation.

7

Conclusion

In this thesis, we investigated how interaction with touchscreens can be enhanced through finger orientation. We first investigated different recognition approaches to extract both the pitch and yaw angle of the finger to enable devices to implement new gestures and use cases (RQ1, see Chapter 3). We first presented a camera-based approach which can turn any flat surface into a surface with finger orientation recognition capabilities. We highlight that this helps especially in the early stages of the human-centered design circle where low fidelity prototypes come to play. Additionally, we presented a method to recognize the finger orientation for high fidelity prototypes and consumer devices where we used machine learning and Convolutional Neural Networks (CNNs) to estimate the finger orientation.

While the recognition of new interaction techniques is an important topic, in this thesis we highlight a set of other domains which are equally important to understand before designers can implement new use cases. In detail, we highlighted three key areas: ergonomic constraints (RQ2, see Chapter 4), social implications (RQ3, see Chapter 5), and discoverability (RQ4, see Chapter 6). While these areas are important for all new interaction techniques, in the thesis, we address them with respect to touchscreens and finger orientation.

Overall, we showed that ergonomic constraints need to be considered when designing new use cases. In respect to finger orientation, our main finding here is that the range of comfortably usable finger orientations is limited. Therefore we defined the comfort and non-comfort zones. These zones are influenced by the interacting hand which therefore also needs to be considered. Finally, we showed that the user's situation is in needs to be considered. Tabletop-like scenarios are leading to a smaller range of comfortably usable finger orientation inputs in contracts to mobile scenarios where the user can also rotate the screen to a certain extent. Thus, before implementing new use cases for finger orientation designers need to be aware of these ergonomic constraints.

To understand and evaluate the social implications of new interaction techniques, we first presented a new mixed method approach to study the influence on social settings and then evaluate how finger orientation performs. Our new approach combines traditional conversation analysis with new technologies. We substitute manual gaze decoding with eye trackers worn by all participants. Moreover, we employ entry and final interviews to better understand how the participants felt and interacted during the conversation as well as how they perceive the interaction technique. Lastly, these interviews are not only employed to understand the interaction technique but also to give new insights to improve the interaction for the next iteration of the design circle. The last part of this thesis focuses on the discoverability of new interaction techniques. Numerous interaction techniques have been proposed in the past and they certainly all work different. Thus, informing the user about them in a key challenge in order to promote new interaction techniques and eventually making them state-of-the-art. We showed that three communication methods are envisioned by designers: depiction, pop-up, and tutorial. Moreover, we show that from a design point of view they all have their advantages and disadvantages; however, users prefer depiction over the others. Thus, we present design implications which help designers to inform the user.

Research presented in the past which investigates new interaction techniques mainly focus on the recognition or implementation and the use cases. Thus, past research often highlights how a new interaction technique functions and how it can be utilized to improve interaction. Beyond the recognition, we present arguments

that the characteristics of new interaction techniques are also important to consider. We argue that, ergonomic constraints, social implications, and discoverability are equally important areas when investigating new interaction techniques. Moreover, we argue that these areas need to be investigated even before designers can design new use cases to embrace the new interaction techniques. Therefore, this thesis is the foundation for designers who want to enable touch surfaces to be reactive to the finger orientation.

With this thesis, we contribute insights to enable designers developing new use cases but also want to highlight that there is more than the recognition part to new interaction techniques. Therefore, this thesis presents a new approach on how to study new interaction technique which is necessary to fully understand an interaction technique. In the thesis we present an in-depth assessment of finger orientation as additional input dimensions for touchscreens. We call this type of in-depth evaluation a vertical evaluation where each facet of interaction is analyzed to understand the design space of a new interaction technique better. The results gained from a vertical evaluation will help to design better use cases as they are known to very stakeholder even before considering and designing new use cases. Vertical evaluation is in contrast to horizontal evaluation where different interaction techniques are compared or implemented using the same strategies. Here we want to highlight that the same methods which were used to recognize the finger orientation can easily be applied to recognize other properties of the finger. Using motion tracking data and user rating can help to understand other interaction techniques better. The method presented to study social implications can adapt to analyze any interaction techniques with a conversation partner in mind.

7.1 Future Work

With this thesis, we lay the foundation to understand finger orientation as additional input dimensions for touchscreens. With this newly gained knowledge of how to incorporate finger orientation as additional interaction technique for a new generation of devices, we see the next step clearly in designing and implementing new use cases. While related work already suggested some use cases, e.g. Xiao et

al. [188], we think a fundamental question is not answered yet. Finger orientation can be used as a relative and absolute input. Here, relative input would change the view in a GUI in relation to the change of the finger orientation, while in an absolute input, the view would adopt the same angle as the finger on the surface. However, we see evidence that they both could be viable for different use cases. For instance, turning a knob or rotating an image could be done relative in contrast to 3D navigation where it might be better to interact absolute. This is a major challenge when designing new use cases which enable finger orientation manipulation.

Beyond further investigation on how to design new use cases for finger orientation as additional input dimensions, we see the need to understand interaction techniques which have been proposed in the past but also new propped ones better. Here, we clearly emphasis to use the vertical approach even before designing new use cases as more work needs to be invested in the early stages of investigations of new interaction techniques. We argue for long-term benefits in well designed use cases when fostering awareness of the design challenges in the early stages.

Finally, in this thesis, we show that recognition, ergonomic constraints, social implications, and discoverability need to be considered as a whole when trying to fully understand a new interaction technique. Thus, this thesis only presents a vertical evaluation of finger orientation. However, we presented a second dimension, the horizontal evaluation, which links the key areas of different interaction techniques. In this thesis, we link to other interaction techniques and use method form other fields to assess finger orientation. However, a in-depth horizontal linking will foster a better understanding of all key areas. Therefore, future work should not only investigate the vertical evaluation of a single interaction technique but also horizontal evaluation.

Bibliography

[1] O. K.-C. Au, C.-L. Tai. "Multitouch Finger Registration and Its Applications." In: *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*. OZCHI '10. Brisbane, Australia: ACM, 2010, pp. 41–48. ISBN: 978-1-4503-0502-0. DOI: 10.1145/1952222.1952233 (cit. on p. 133).

[2] R. Bednarik, S. Eivazi, M. Hradis. "Gaze and Conversational Engagement in Multiparty Video Conversation: An Annotation Scheme and Classification of High and Low Levels of Engagement." In: *Proceedings of the 4th Workshop on Eye Gaze in Intelligent Human Machine Interaction*. Gaze-In '12. Santa Monica, California: ACM, 2012, 10:1–10:6. ISBN: 978-1-4503-1516-6. DOI: 10.1145/2401836.2401846 (cit. on p. 33).

[3] Y. Bengio, A. Courville, P. Vincent. "Representation Learning: A Review and New Perspectives." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.8 (2013), pp. 1798–1828. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2013.50 (cit. on p. 58).

[4] H. Benko, T. S. Saponas, D. Morris, D. Tan. "Enhancing Input on and Above the Interactive Surface with Muscle Sensing." In: *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ITS '09. Banff, Alberta, Canada: ACM, 2009, pp. 93–100. ISBN: 978-1-60558-733-2. DOI: 10.1145/1731903.1731924 (cit. on p. 133).

[5] S. Bertel, T. Dressel, T. Kohlberg, V. von Jan. "Spatial Knowledge Acquired from Pedestrian Urban Navigation Systems." In: *Proceedings of the 19th International*

Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '17. Vienna, Austria: ACM, 2017, 32:1–32:6. ISBN: 978-1-4503-5075-4. DOI: 10.1145/3098279.3098543 (cit. on p. 104).

[6] B. Bilgic, I. Chatnuntawech, A. P. Fan, K. Setsompop, S. F. Cauley, L. L. Wald, E. Adalsteinsson. "Fast image reconstruction with L2-regularization." In: *Journal of Magnetic Resonance Imaging* 40.1 (2014), pp. 181–191 (cit. on p. 60).

[7] A. Blandford, D. Furniss, S. Makri. *Qualitative Hci Research: Going Behind the Scenes*. Morgan & Claypool Publishers, 2016, pp. 1–115. ISBN: 1627057595, 9781627057592. DOI: 10.2200/S00706ED1V01Y201602HCI034 (cit. on p. 135).

[8] P. Bogunovich, D. Salvucci. "The Effects of Time Constraints on User Behavior for Deferrable Interruptions." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. Vancouver, BC, Canada: ACM, 2011, pp. 3123–3126. ISBN: 978-1-4503-0228-9. DOI: 10.1145/1978942.1979404 (cit. on pp. 34, 106).

[9] M. Böhmer, B. Hecht, J. Schöning, A. Krüger, G. Bauer. "Falling Asleep with Angry Birds, Facebook and Kindle: A Large Scale Study on Mobile Application Usage." In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI '11. Stockholm, Sweden: ACM, 2011, pp. 47–56. ISBN: 978-1-4503-0541-9. DOI: 10.1145/2037373.2037383 (cit. on p. 133).

[10] S. Boring, D. Ledo, X. A. Chen, N. Marquardt, A. Tang, S. Greenberg. "The Fat Thumb: Using the Thumb's Contact Size for Single-handed Mobile Interaction." In: *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services*. MobileHCI '12. San Francisco, California, USA: ACM, 2012, pp. 39–48. ISBN: 978-1-4503-1105-2. DOI: 10.1145/2371574.2371582 (cit. on pp. 19, 27, 82).

[11] J. P. Borst, N. A. Taatgen, H. van Rijn. "What Makes Interruptions Disruptive?: A Process-Model Account of the Effects of the Problem State Bottleneck on Task Interruption and Resumption." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: ACM, 2015, pp. 2971–2980. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702156 (cit. on p. 34).

[12] L. E. Boyd, A. Rangel, H. Tomimbang, A. Conejo-Toledo, K. Patel, M. Tentori, G. R. Hayes. "SayWAT: Augmenting Face-to-Face Conversations for Adults with Autism." In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, 2016, pp. 4872–4883. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858215 (cit. on p. 35).

[13] J. Brooke et al. "SUS-A quick and dirty usability scale." In: *Usability evaluation in industry* 189.194 (1996), pp. 4–7 (cit. on p. 141).

[14] B. Brown, M. McGregor, D. McMillan. "Searchable Objects: Search in Everyday Conversation." In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. CSCW '15. Vancouver, BC, Canada: ACM, 2015, pp. 508–517. ISBN: 978-1-4503-2922-4. DOI: 10.1145/2675133.2675206 (cit. on pp. 33, 100).

[15] R. Caruana, S. Lawrence, C. L. Giles. "Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping." In: *Advances in neural information processing systems*. 2001, pp. 402–408 (cit. on p. 58).

[16] W. Chang, K. E. Kim, H. Lee, J. K. Cho, B. S. Soh, J. H. Shim, G. Yang, S.-J. Cho, J. Park. "Recognition of grip-patterns by using capacitive touch sensors." In: *IEEE International Symposium on Industrial Electronics*. Vol. 4. IEEE. 2006, pp. 2936–2941 (cit. on p. 29).

[17] D. Chattopadhyay, K. O'Hara, S. Rintel, R. Rädle. "Office Social: Presentation Interactivity for Nearby Devices." In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, 2016, pp. 2487–2491. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858337 (cit. on pp. 33, 103).

[18] D. Chen, R. Vertegaal. "Using Mental Load for Managing Interruptions in Physiologically Attentive User Interfaces." In: *CHI '04 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '04. Vienna, Austria: ACM, 2004, pp. 1513–1516. ISBN: 1-58113-703-6. DOI: 10.1145/985921.986103 (cit. on p. 34).

[19] J. Chen, A. Abouzied. "One LED is Enough: Catalyzing Face-to-face Interactions at Conferences with a Gentle Nudge." In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '16. San Francisco, California, USA: ACM, 2016, pp. 172–183. ISBN: 978-1-4503-3592-8. DOI: 10.1145/2818048.2819969 (cit. on p. 32).

[20] T. Clausen. “Über die Function $sin(\varphi) + \frac{1}{2^2} sin(2\varphi) + \frac{1}{3^2} sin(3\varphi) +$ etc.” In: *Journal für die reine und angewandte Mathematik*. Berlin, Germany: Duncker und Humblot, 1832, pp. 298–300. ISBN: 978-1-4503-2968-2 (cit. on p. 92).

[21] A. Colley, J. Häkkilä. “Exploring Finger Specific Touch Screen Interaction for Mobile Phone User Interfaces.” In: *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design*. OzCHI ’14. Sydney, New South Wales, Australia: ACM, 2014, pp. 539–548. ISBN: 978-1-4503-0653-9. DOI: 10.1145/2686612.2686699 (cit. on pp. 19, 30, 130, 132).

[22] A. Colley, S. Mayer, N. Henze. “Investigating the Effect of Orientation and Visual Style on Touchscreen Slider Performance.” In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI ’19. Glasgow, Scotland UK: ACM, 2019, 406:1–406:14. ISBN: 978-1-4503-5970-2. DOI: 10.1145/3290605.3300419 (cit. on p. 23).

[23] P. Coulibaly, F. Anctil, B. Bobée. “Daily reservoir inflow forecasting using artificial neural networks with stopped training approach.” In: *Journal of Hydrology* 230.3 (2000), pp. 244 –257. ISSN: 0022-1694. DOI: http://dx.doi.org/10.1016/S0022-1694(00)00214-6 (cit. on pp. 55, 58).

[24] L. Dabbish, R. E. Kraut. “Controlling Interruptions: Awareness Displays and Social Motivation for Coordination.” In: *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*. CSCW ’04. Chicago, Illinois, USA: ACM, 2004, pp. 182–191. ISBN: 1-58113-810-5. DOI: 10.1145/1031607.1031638 (cit. on p. 34).

[25] N. Dahlbäck, A. Jönsson, L. Ahrenberg. “Wizard of Oz studies — why and how.” In: *Knowledge-Based Systems* 6.4 (1993). Special Issue: Intelligent User Interfaces, pp. 258 –266. ISSN: 0950-7051. DOI: https://doi.org/10.1016/0950-7051(93)90017-N (cit. on p. 141).

[26] C. T. Dang, E. André. “Usage and Recognition of Finger Orientation for Multi-Touch Tabletop Interaction.” In: *Human-Computer Interaction – INTERACT 2011*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 409–426. ISBN: 978-3-642-23765-2 (cit. on p. 28).

[27] A. De Luca, E. von Zezschwitz, N. D. H. Nguyen, M.-E. Maurer, E. Rubegni, M. P. Scipioni, M. Langheinrich. “Back-of-device Authentication on Smartphones.”

In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. Paris, France: ACM, 2013, pp. 2389–2398. ISBN: 978-1-4503-1899-0. DOI: 10.1145/2470654.2481330 (cit. on p. 130).

[28] J. Duchi, E. Hazan, Y. Singer. "Adaptive Subgradient Methods for Online Learning and Stochastic Optimization." In: *Journal of Machine Learning Research* 12 (July 2011), pp. 2121–2159. ISSN: 1532-4435 (cit. on pp. 56, 58, 59).

[29] C. ESOL. *Speaking Test Preparation Pack for CAE Paperback with DVD*. Speaking Test Preparation Pack. Cambridge University Press, 2008. ISBN: 9781906438395 (cit. on pp. 104, 105, 108, 110, 122).

[30] M. R. Endsley, D. G. Jones. "Disruptions, Interruptions and Information Attack: Impact on Situation Awareness and Decision Making." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 45.2 (2001), pp. 63–67. DOI: 10.1177/154193120104500214 (cit. on p. 101).

[31] P. Ewerling, A. Kulik, B. Froehlich. "Finger and Hand Detection for Multi-touch Interfaces Based on Maximally Stable Extremal Regions." In: *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*. ITS '12. Cambridge, Massachusetts, USA: ACM, 2012, pp. 173–182. ISBN: 978-1-4503-1209-7. DOI: 10.1145/2396636.2396663 (cit. on p. 133).

[32] M. Exposito, V. Zeamer, P. Maes. "Unobtrusive Note Taking: Enriching Digital Interpersonal Interactions Using Gestures." In: *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. CSCW '17 Companion. Portland, Oregon, USA: ACM, 2017, pp. 167–170. ISBN: 978-1-4503-4688-7. DOI: 10.1145/3022198.3026319 (cit. on p. 35).

[33] J. E. Fischer, S. Reeves, S. Moran, C. Greenhalgh, S. Benford, S. Rennick-Egglestone. "Understanding Mobile Notification Management in Collocated Groups." In: *Proceedings of the 13th European Conference on Computer Supported Cooperative Work*. ECSCW 2013. London: Springer London, 2013, pp. 21–44. ISBN: 978-1-4471-5346-7. DOI: 10.1007/978-1-4471-5346-7_2 (cit. on p. 32).

[34] M. Fjeld, P. Woźniak, J. Cowls, B. Nardi. "Ad hoc encounters with big data: Engaging citizens in conversations around tabletops." In: *First Monday* 20.2 (2015). DOI: http://dx.doi.org/10.5210/fm.v20i2.5611 (cit. on p. 32).

[35] M. Funk, L. Lischke, S. Mayer. “Neue Impulse für visuelle Kommissionierassistenzsysteme aus der Mensch-Computer-Interaktion.” In: *Warehousing 4.0*. Lauda-Königshofen, Germany: B+G Wissenschaftsverlag, Jan. 1, 2017, pp. 223–236 (cit. on p. 23).

[36] M. Funk, S. Mayer, A. Schmidt. “Using In-Situ Projection to Support Cognitively Impaired Workers at the Workplace.” In: *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*. ASSETS ’15. Lisbon, Portugal: ACM, Jan. 1, 2015, pp. 185–192. ISBN: 978-1-4503-3400-6. DOI: 10.1145/2700648.2809853 (cit. on p. 23).

[37] M. Funk, T. Kosch, K. Wolf, P. Knierim, S. Mayer, A. Schmidt. “Automatic Projection Positioning Based on Surface Suitability.” In: *Proceedings of the 5th ACM International Symposium on Pervasive Displays*. PerDis ’16. Oulu, Finland: ACM, Jan. 1, 2016, pp. 75–79. ISBN: 978-1-4503-4366-4. DOI: 10.1145/2914920.2915014 (cit. on p. 23).

[38] M. Funk, S. Mayer, M. Nistor, A. Schmidt. “Mobile In-Situ Pick-by-Vision: Order Picking Support Using a Projector Helmet.” In: *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. PETRA ’16. Corfu, Island, Greece: ACM, Jan. 1, 2016, 45:1–45:4. ISBN: 978-1-4503-4337-4. DOI: 10.1145/2910674.2910730 (cit. on p. 23).

[39] M. Funk, A. Sahami Shirazi, S. Mayer, L. Lischke, A. Schmidt. “Pick from Here!: An Interactive Mobile Cart Using In-situ Projection for Order Picking.” In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp ’15. Osaka, Japan: ACM, Jan. 1, 2015, pp. 601–609. ISBN: 978-1-4503-3574-4. DOI: 10.1145/2750858.2804268 (cit. on p. 23).

[40] M. Funk, L. Lischke, S. Mayer, A. Sahami Shirazi, A. Schmidt. “Teach Me How! Interactive Assembly Instructions Using Demonstration and In-Situ Projection.” In: Singapore: Springer Singapore, Jan. 1, 2018, pp. 49–73. ISBN: 978-981-10-6404-3. DOI: 10.1007/978-981-10-6404-3_4 (cit. on p. 23).

[41] E. Ghomi, S. Huot, O. Bau, M. Beaudouin-Lafon, W. E. Mackay. “ArpèGe: Learning Multitouch Chord Gestures Vocabularies.” In: *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*. ITS ’13. St. Andrews, Scotland, United Kingdom: ACM, 2013, pp. 209–218. ISBN: 978-1-4503-2271-3. DOI: 10.1145/2512349.2512795 (cit. on p. 133).

[42] X. Glorot, Y. Bengio. "Understanding the difficulty of training deep feedforward neural networks." In: *In Proceedings of the International Conference on Artificial Intelligence and Statistics. Society for Artificial Intelligence and Statistics*. Vol. 9. AISTATS'10. Chia Laguna Resort, Sardinia, Italy: JMLR.org, 2010, pp. 249–256 (cit. on pp. 56, 59, 60).

[43] A. Goguey, G. Casiez, D. Vogel, C. Gutwin. "Characterizing Finger Pitch and Roll Orientation During Atomic Touch Actions." In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: ACM, 2018, 589:1–589:12. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3174163 (cit. on p. 30).

[44] C. Goodwin. "Restarts, Pauses, and the Achievement of a State of Mutual Gaze at Turn-Beginning." In: *Sociological Inquiry* 50.3-4 (1980), pp. 272–302. ISSN: 1475-682X. DOI: 10.1111/j.1475-682X.1980.tb00023.x (cit. on pp. 33, 100).

[45] N. Gorm, I. Shklovski. "Participant Driven Photo Elicitation for Understanding Activity Tracking: Benefits and Limitations." In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. CSCW '17. Portland, Oregon, USA: ACM, 2017, pp. 1350–1361. ISBN: 978-1-4503-4335-0. DOI: 10.1145/2998181.2998214 (cit. on p. 103).

[46] N. Goyal, S. R. Fussell. "Effects of Sensemaking Translucence on Distributed Collaborative Analysis." In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '16. San Francisco, California, USA: ACM, 2016, pp. 288–302. ISBN: 978-1-4503-3592-8. DOI: 10.1145/2818048.2820071 (cit. on pp. 32, 81).

[47] N. Goyal, S. R. Fussell. "Intelligent Interruption Management Using Electro Dermal Activity Based Physiological Sensor for Collaborative Sensemaking." In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1.3 (Sept. 2017), 52:1–52:21. ISSN: 2474-9567. DOI: 10.1145/3130917 (cit. on p. 34).

[48] N. Goyal, G. Leshed, S. R. Fussell. "Effects of Visualization and Note-taking on Sensemaking and Analysis." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. Paris, France: ACM, 2013, pp. 2721–2724. ISBN: 978-1-4503-1899-0. DOI: 10.1145/2470654.2481376 (cit. on p. 80).

[49] N. Goyal, G. Leshed, S. R. Fussell. “Leveraging Partner’s Insights for Distributed Collaborative Sensemaking.” In: *Proceedings of the 2013 Conference on Computer Supported Cooperative Work Companion*. CSCW ’13. San Antonio, Texas, USA: ACM, 2013, pp. 15–18. ISBN: 978-1-4503-1332-2. DOI: 10.1145/2441955.2441960 (cit. on p. 80).

[50] N. Goyal, G. Leshed, D. Cosley, S. R. Fussell. “Effects of Implicit Sharing in Collaborative Analysis.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’14. Toronto, Ontario, Canada: ACM, 2014, pp. 129–138. ISBN: 978-1-4503-2473-1. DOI: 10.1145/2556288.2557229 (cit. on pp. 32, 80).

[51] M. Gärtner, S. Mayer, V. Schwind, E. Hämmerle, E. Turcan, F. Rheinwald, G. Murawski, L. Lischke, J. Kuhn. “NLATool: An Application for Enhanced Deep Text Understanding.” In: *Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations*. COLING’18. Aug. 20, 2018, p. 4 (cit. on p. 23).

[52] J. Grudin. “From Tool to Partner: The Evolution of Human-Computer Interaction.” In: *Synthesis Lectures on Human-Centered Interaction* 10.1 (2017), pp. i–183. DOI: http://dx.doi.org/10.2200/S00745ED1V01Y201612HCI035 (cit. on p. 102).

[53] A. Guo, R. Xiao, C. Harrison. “CapAuth: Identifying and Differentiating User Handprints on Commodity Capacitive Touchscreens.” In: *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*. ITS ’15. Madeira, Portugal: ACM, 2015, pp. 59–62. ISBN: 978-1-4503-3899-8. DOI: 10.1145/2817721.2817722 (cit. on p. 29).

[54] A. Gupta, M. Anwar, R. Balakrishnan. “Porous Interfaces for Small Screen Multitasking Using Finger Identification.” In: *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. UIST ’16. Tokyo, Japan: ACM, 2016, pp. 145–156. ISBN: 978-1-4503-4189-9. DOI: 10.1145/2984511.2984557 (cit. on pp. 132, 133).

[55] A. Gupta, R. Balakrishnan. “DualKey: Miniature Screen Text Entry via Finger Identification.” In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI ’16. San Jose, California, USA: ACM, 2016, pp. 59–70. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858052 (cit. on pp. 132, 133).

[56] J. Y. Han. “Low-cost Multi-touch Sensing Through Frustrated Total Internal Reflection.” In: *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*. UIST ’05. Seattle, WA, USA: ACM, 2005, pp. 115–118. ISBN: 1-59593-271-2. DOI: 10.1145/1095034.1095054 (cit. on p. 35).

[57] G. Harboe, E. M. Huang. “Real-World Affinity Diagramming Practices: Bridging the Paper-Digital Gap.” In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI ’15. Seoul, Republic of Korea: ACM, 2015, pp. 95–104. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702561 (cit. on pp. 103, 135, 145).

[58] C. Harrison, J. Schwarz, S. E. Hudson. “TapSense: Enhancing Finger Interaction on Touch Surfaces.” In: *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*. UIST ’11. Santa Barbara, California, USA: ACM, 2011, pp. 627–636. ISBN: 978-1-4503-0716-1. DOI: 10.1145/2047196.2047279 (cit. on pp. 19, 27, 132).

[59] M. Hassenzahl, M. Burmester, F. Koller. “AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität.” In: *Mensch & Computer 2003*. Springer, 2003, pp. 187–196 (cit. on p. 141).

[60] N. Henze, E. Rukzio, S. Boll. “100,000,000 Taps: Analysis and Improvement of Touch Performance in the Large.” In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI ’11. Stockholm, Sweden: ACM, 2011, pp. 133–142. ISBN: 978-1-4503-0541-9. DOI: 10.1145/2037373.2037395 (cit. on pp. 29, 97).

[61] N. Henze, H. V. Le, S. Mayer, V. Schwind. “Improving Software-Reduced Touchscreen Latency.” In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. mobileHCI’17. Sept. 4, 2017. DOI: 10.1145/3098279.3122150 (cit. on p. 23).

[62] N. Henze, U. Martin, M. A. Rieger, L. Lischke, S. Mayer, C. von Molo, B. Steinhilber, F. Wagenblast. “Konzept zur Entwicklung moderner Bedienformen für Betriebszentralen.” In: *ETR - Eisenbahntechnische Rundschau* 1 (Jan. 1, 2017), pp. 26–30 (cit. on p. 23).

[63] J. L. Hernandez-Rebollar, N. Kyriakopoulos, R. W. Lindeman. “The AcceleGlove: A Whole-hand Input Device for Virtual Reality.” In: *ACM SIGGRAPH 2002 Conference Abstracts and Applications*. SIGGRAPH ’02. San Antonio, Texas: ACM, 2002, pp. 259–259. ISBN: 1-58113-525-4. DOI: 10.1145/1242073.1242272 (cit. on p. 131).

[64] J. L. Hernandez-Rebollar, R. W. Lindeman, N. Kyriakopoulos. "A Multi-Class Pattern Recognition System for Practical Finger Spelling Translation." In: *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces*. ICMI '02. Washington, DC, USA: IEEE Computer Society, 2002, pp. 185–. ISBN: 0-7695-1834-6. DOI: 10.1109/ICMI.2002.1166990 (cit. on p. 131).

[65] K. Hinckley, S. Heo, M. Pahud, C. Holz, H. Benko, A. Sellen, R. Banks, K. O'Hara, G. Smyth, W. Buxton. "Pre-Touch Sensing for Mobile Interaction." In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, 2016, pp. 2869–2881. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858095 (cit. on pp. 29, 64).

[66] H. M. Hodgetts, D. M. Jones. "Interruptions in the Tower of London Task: Can Preparation Minimise Disruption?" In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 47.8 (2003), pp. 1000–1004. DOI: 10.1177/154193120304700810 (cit. on p. 102).

[67] E. Hoggan, J. Williamson, A. Oulasvirta, M. Nacenta, P. O. Kristensson, A. Lehtiö. "Multi-touch Rotation Gestures: Performance and Ergonomics." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. Paris, France: ACM, 2013, pp. 3047–3050. ISBN: 978-1-4503-1899-0. DOI: 10.1145/2470654.2481423 (cit. on p. 30).

[68] D. Holman, A. Hollatz, A. Banerjee, R. Vertegaal. "Unifone: Designing for Auxiliary Finger Input in One-handed Mobile Interactions." In: *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*. TEI '13. Barcelona, Spain: ACM, 2013, pp. 177–184. ISBN: 978-1-4503-1898-3. DOI: 10.1145/2460625.2460653 (cit. on p. 130).

[69] C. Holz, P. Baudisch. "The Generalized Perceived Input Point Model and How to Double Touch Accuracy by Extracting Fingerprints." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. Atlanta, Georgia, USA: ACM, 2010, pp. 581–590. ISBN: 978-1-60558-929-9. DOI: 10.1145/1753326.1753413 (cit. on pp. 19, 50, 96).

[70] C. Holz, P. Baudisch. "Understanding Touch." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. Vancouver, BC, Canada: ACM, 2011, pp. 2501–2510. ISBN: 978-1-4503-0228-9. DOI: 10.1145/1978942.1979308 (cit. on pp. 19, 28).

[71] C. Holz, S. Buthpitiya, M. Knaust. “Bodyprint: Biometric User Identification on Mobile Devices Using the Capacitive Touchscreen to Scan Body Parts.” In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI ’15. Seoul, Republic of Korea: ACM, 2015, pp. 3011–3014. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702518 (cit. on p. 29).

[72] M.-C. Hsiu, C. Wang, D.-Y. Huang, J.-W. Lin, Y.-C. Lin, D.-N. Yang, Y.-p. Hung, M. Chen. “Nail+: Sensing Fingernail Deformation to Detect Finger Force Touch Interactions on Rigid Surfaces.” In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI ’16. Florence, Italy: ACM, 2016, pp. 1–6. ISBN: 978-1-4503-4408-1. DOI: 10.1145/2935334.2935362 (cit. on p. 132).

[73] D.-Y. Huang, M.-C. Tsai, Y.-C. Tung, M.-L. Tsai, Y.-T. Yeh, L. Chan, Y.-P. Hung, M. Y. Chen. “TouchSense: Expanding Touchscreen Input Vocabulary Using Different Areas of Users’ Finger Pads.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’14. Toronto, Ontario, Canada: ACM, 2014, pp. 189–192. ISBN: 978-1-4503-2473-1. DOI: 10.1145/2556288.2557258 (cit. on pp. 131, 132).

[74] S. E. Hudson, I. Smith. “Techniques for Addressing Fundamental Privacy and Disruption Tradeoffs in Awareness Support Systems.” In: *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work*. CSCW ’96. Boston, Massachusetts, USA: ACM, 1996, pp. 248–257. ISBN: 0-89791-765-0. DOI: 10.1145/240080.240295 (cit. on p. 34).

[75] S. Ioffe, C. Szegedy. “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.” In: *CoRR* abs/1502.03167 (2015) (cit. on p. 60).

[76] S. T. Iqbal, E. Horvitz. “Notifications and Awareness: A Field Study of Alert Usage and Preferences.” In: *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*. CSCW ’10. Savannah, Georgia, USA: ACM, 2010, pp. 27–30. ISBN: 978-1-60558-795-0. DOI: 10.1145/1718918.1718926 (cit. on p. 34).

[77] P. Jarusriboonchai, A. Malapaschas, T. Olsson. “Design and Evaluation of a Multi-Player Mobile Game for Icebreaking Activity.” In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI ’16. Santa Clara, California, USA: ACM, 2016, pp. 4366–4377. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858298 (cit. on p. 31).

[78] P. Jarusriboonchai, A. Malapaschas, T. Olsson, K. Väänänen. “Increasing Collocated People’s Awareness of the Mobile User’s Activities: A Field Trial of Social Displays.” In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW ’16. San Francisco, California, USA: ACM, 2016, pp. 1691–1702. ISBN: 978-1-4503-3592-8. DOI: 10.1145/2818048.2819990 (cit. on p. 32).

[79] K. Jokinen, H. Furukawa, M. Nishida, S. Yamamoto. “Gaze and Turn-taking Behavior in Casual Conversational Interactions.” In: *ACM Trans. Interact. Intell. Syst.* 3.2 (Aug. 2013), 12:1–12:30. ISSN: 2160-6455. DOI: 10.1145/2499474.2499481 (cit. on p. 33).

[80] A. Karsvall. “Intercreativity in Surgical Practice : A Dialogical Approach to Interaction & Technology.” PhD thesis. Linköping University, The Tema Institute, 2011, p. 167 (cit. on p. 101).

[81] A. Kendon. “Some functions of gaze-direction in social interaction.” In: *Acta Psychologica* 26 (1967), pp. 22–63. ISSN: 0001-6918. DOI: 10.1016/0001-6918(67)90005-4 (cit. on pp. 32, 33).

[82] F. Kiss, K. Kucharski, S. Mayer, L. Lischke, P. Knierim, A. Romanowski, P. W. Woźniak. “RunMerge: Towards Enhanced Proprioception for Advanced Amateur Runners.” In: *Proceedings of the 2017 ACM Conference Companion Publication on Designing Interactive Systems*. DIS ’17 Companion. New York, NY, USA: ACM, 2017, pp. 192–196. DOI: 10.1145/3064857.3079144 (cit. on p. 23).

[83] S. Kratz, P. Chiu, M. Back. “PointPose: Finger Pose Estimation for Touch Input on Mobile Devices Using a Depth Sensor.” In: *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*. ITS ’13. St. Andrews, Scotland, United Kingdom: ACM, 2013, pp. 223–230. ISBN: 978-1-4503-2271-3. DOI: 10.1145/2512349.2512824 (cit. on pp. 21, 28, 39–41, 45, 46).

[84] A. Krizhevsky, I. Sutskever, G. E. Hinton. “Imagenet classification with deep convolutional neural networks.” In: *Proceedings of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe, NV, USA: Curran Associates, Inc., 2012, pp. 1097–1105 (cit. on p. 30).

[85] M. W. Krueger, T. Gionfriddo, K. Hinrichsen. “VIDEOPLACE–an Artificial Reality.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’85. San Francisco, California, USA: ACM, 1985, pp. 35–40. ISBN: 0-89791-149-0. DOI: 10.1145/317456.317463 (cit. on p. 35).

[86] A. Kötteritzsch, B. Weyers, L. Lischke, S. Mayer, P. W. Woźniak, J. Fietkau, M. Koch. "Workshopband Urban und Digital – Gemeinsam auf interaktiven Wegen." In: *Mensch und Computer 2016 – Workshopband*. Ed. by B. Weyers, A. Dittmar. Aachen: Gesellschaft für Informatik e.V., Jan. 1, 2016. DOI: 10.18420/muc2016-ws14-0000 (cit. on p. 23).

[87] T. Kubitza, N. Pohl, T. Dingler, A. Schmidt. "WebClip: A Connector for Ubiquitous Physical Input and Output for Touch Screen Devices." In: *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp '13. Zurich, Switzerland: ACM, 2013, pp. 387–390. ISBN: 978-1-4503-1770-2. DOI: 10.1145/2493432.2493520 (cit. on p. 87).

[88] A. Kumar, A. Radjesh, S. Mayer, H. V. Le. "Improving the Input Accuracy of Touchscreens using Deep Learning." In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI'19 EA. Glasgow, Scotland UK: ACM, May 4, 2019. DOI: 10.1145/3290607.3312928 (cit. on p. 23).

[89] N. Lawrence, M. Seeger, R. Herbrich. "Fast Sparse Gaussian Process Methods: The Informative Vector Machine." In: *Proceedings of the 15th International Conference on Neural Information Processing Systems*. NIPS'02. Cambridge, MA, USA: MIT Press, 2002, pp. 625–632 (cit. on p. 55).

[90] H. V. Le, S. Mayer, N. Henz. "InfiniTouch: Finger-Aware Interaction on Fully Touch Sensitive Smartphones." In: *Proceedings of the 31th Annual ACM Symposium on User Interface Software and Technology*. UIST'18. New York, NY, USA: ACM, Oct. 14, 2018. DOI: 10.1145/3242587.3242605 (cit. on p. 23).

[91] H. V. Le, S. Mayer, N. Henze. "Investigating the Feasibility of Finger Identification on Capacitive Touchscreens using Deep Learning." In: *Proceedings of the 24th International Conference on Intelligent User Interfaces*. IUI '19. Marina del Ray, CA, USA: ACM, Mar. 17, 2019. DOI: 10.1145/3301275.3302295 (cit. on p. 23).

[92] H. V. Le, S. Mayer, N. Henze. "Machine Learning with Tensorflow for Mobile and Ubiquitous Interaction." In: *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*. MUM '17. Stuttgart, Germany: ACM, Jan. 1, 2017, pp. 567–572. ISBN: 978-1-4503-5378-6. DOI: 10.1145/3152832.3156559 (cit. on p. 23).

[93] H. V. Le, S. Mayer, P. Bader, N. Henze. "A Smartphone Prototype for Touch Interaction on the Whole Device Surface." In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '17. Vienna, Austria: ACM, 2017, 100:1–100:8. ISBN: 978-1-4503-5075-4. DOI: 10.1145/3098279.3122143 (cit. on pp. 23, 29, 51).

[94] H. V. Le, S. Mayer, T. Kosch, N. Henze. "Demonstrating PalmTouch: The Palm as An Additional Input Modality on Commodity Smartphones." In: *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. MobileHCI'18. New York, NY, USA: ACM, Sept. 3, 2018 (cit. on p. 23).

[95] H. V. Le, S. Mayer, K. Wolf, N. Henze. "Finger Placement and Hand Grasp During Smartphone Interaction." In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '16. Santa Clara, California, USA: ACM, 2016, pp. 2576–2584. ISBN: 978-1-4503-4082-3. DOI: 10.1145/2851581.2892462 (cit. on pp. 23, 27, 30).

[96] H. V. Le, S. Mayer, P. Bader, N. Henze. "Fingers' Range and Comfortable Area for One-Handed Smartphone Interaction Beyond the Touchscreen." In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: ACM, 2018, 31:1–31:12. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3173605 (cit. on pp. 23, 30).

[97] H. V. Le, S. Mayer, P. Bader, F. Bastian, N. Henze. "Interaction Methods and Use Cases for a Full-Touch Sensing Smartphone." In: *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '17. Denver, Colorado, USA: ACM, 2017, pp. 2730–2737. ISBN: 978-1-4503-4656-6. DOI: 10.1145/3027063.3053196 (cit. on pp. 23, 27, 29).

[98] H. V. Le, P. Bader, T. Kosch, N. Henze. "Investigating Screen Shifting Techniques to Improve One-Handed Smartphone Usage." In: *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. NordiCHI '16. Gothenburg, Sweden: ACM, 2016, pp. 27–37. ISBN: 978-1-4503-4763-1. DOI: 10.1145/2971485.2971562 (cit. on p. 27).

[99] H. V. Le, S. Mayer, A. E. Ali, N. Henze. "Machine Learning For Intelligent Mobile User Interfaces using Keras." In: *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. MobileHCI'18. New York, NY, USA: ACM, Sept. 3, 2018 (cit. on p. 23).

[100] H. V. Le, T. Kosch, P. Bader, S. Mayer, N. Henze. "PalmTouch: Using the Palm As an Additional Input Modality on Commodity Smartphones." In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: ACM, 2018, 360:1–360:13. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3173934 (cit. on pp. 19, 23).

[101] J. C. R. Licklider. "Man-Computer Symbiosis." In: *IRE Transactions on Human Factors in Electronics* HFE-1.1 (1960), pp. 4–11. ISSN: 0099-4561. DOI: 10.1109/THFE2.1960.4503259 (cit. on p. 18).

[102] L. Lischke, S. Mayer, J. Hoffmann, P. Kratzer, S. Roth, K. Wolf, P. W. Woźniak. "Interaction Techniques for Window Management on Large High-resolution Displays." In: *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*. MUM '17. Stuttgart, Germany: ACM, Jan. 1, 2017, pp. 241–247. ISBN: 978-1-4503-5378-6. DOI: 10.1145/3152832.3152852 (cit. on p. 23).

[103] L. Lischke, S. Mayer, K. Wolf, N. Henze, H. Reiterer, A. Schmidt. "Screen Arrangements and Interaction Areas for Large Display Work Places." In: *Proceedings of the 5th ACM International Symposium on Pervasive Displays*. PerDis '16. Oulu, Finland: ACM, Jan. 1, 2016, pp. 228–234. ISBN: 978-1-4503-4366-4. DOI: 10.1145/2914920.2915027 (cit. on p. 23).

[104] L. Lischke, S. Mayer, K. Wolf, A. Sahami Shirazi, N. Henze. "Subjective and Objective Effects of Tablet's Pixel Density." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: ACM, Jan. 1, 2015, pp. 2769–2772. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702390 (cit. on p. 23).

[105] L. Lischke, S. Mayer, A. Preikschat, M. Schweizer, B. Vu, P. W. Woźniak, N. Henze. "Understanding Large Display Environments: Contextual Inquiry in a Control Room." In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI EA '18. New York, NY, USA: ACM, Apr. 21, 2018, LBW134:1–LBW134:6. DOI: 10.1145/3170427.3188621 (cit. on p. 23).

[106] L. Lischke, S. Mayer, K. Wolf, N. Henze, A. Schmidt, S. Leifert, H. Reiterer. "Using Space: Effect of Display Size on Users' Search Performance." In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '15. Seoul, Republic of Korea: ACM, Jan. 1, 2015, pp. 1845–1850. ISBN: 978-1-4503-3146-3. DOI: 10.1145/2702613.2732845 (cit. on p. 23).

[107] L. Lischke, P. W. Woźniak, S. Mayer, A. Preikschat, M. Fjeld. “Using Variable Movement Resistance Sliders for Remote Discrete Input.” In: *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*. ISS ’17. Brighton, United Kingdom: ACM, 2017, pp. 116–125. ISBN: 978-1-4503-4691-7. DOI: 10.1145/3132272.3134135 (cit. on p. 23).

[108] P. Lopes, R. Jota, J. A. Jorge. “Augmenting Touch Interaction Through Acoustic Sensing.” In: *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ITS ’11. Kobe, Japan: ACM, 2011, pp. 53–56. ISBN: 978-1-4503-0871-7. DOI: 10.1145/2076354.2076364 (cit. on p. 132).

[109] W. E. Lorensen, H. E. Cline. “Marching Cubes: A High Resolution 3D Surface Construction Algorithm.” In: *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH ’87. New York, NY, USA: ACM, 1987, pp. 163–169. ISBN: 0-89791-227-6. DOI: 10.1145/37401.37422 (cit. on p. 65).

[110] C. Lozano, D. Jindrich, K. Kahol. “The Impact on Musculoskeletal System During Multitouch Tablet Interactions.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’11. Vancouver, BC, Canada: ACM, 2011, pp. 825–828. ISBN: 978-1-4503-0228-9. DOI: 10.1145/1978942.1979062 (cit. on p. 30).

[111] A. Lucero, J. Holopainen, T. Jokela. “Pass-them-around: Collaborative Use of Mobile Phones for Photo Sharing.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’11. Vancouver, BC, Canada: ACM, 2011, pp. 1787–1796. ISBN: 978-1-4503-0228-9. DOI: 10.1145/1978942.1979201 (cit. on p. 31).

[112] I. S. MacKenzie, R. W. Soukoreff. “Phrase Sets for Evaluating Text Entry Techniques.” In: *CHI ’03 Extended Abstracts on Human Factors in Computing Systems*. CHI EA ’03. Ft. Lauderdale, Florida, USA: ACM, 2003, pp. 754–755. ISBN: 1-58113-637-4. DOI: 10.1145/765891.765971 (cit. on p. 86).

[113] N. Marquardt, J. Kiemer, D. Ledo, S. Boring, S. Greenberg. “Designing User-, Hand-, and Handpart-aware Tabletop Interactions with the TouchID Toolkit.” In: *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ITS ’11. Kobe, Japan: ACM, 2011, pp. 21–30. ISBN: 978-1-4503-0871-7. DOI: 10.1145/2076354.2076358 (cit. on p. 133).

[114] U. Martin, C. von Molo, N. Henze, L. Lischke, S. Mayer, M. A. Rieger, B. Steinhilber, F. Wagenblast. "Ethnographic Analysis of the Dispatchers' Workplace in a Rail-Based Transport Control Center." In: *Third German Workshop on Rail Human Factors 2018*. Vol. 3. RHF'18. Apr. 18, 2018, p. 7. published (cit. on p. 23).

[115] D. Masson, A. Goguey, S. Malacria, G. Casiez. "WhichFingers: Identifying Fingers on Touch Surfaces and Keyboards Using Vibration Sensors." In: *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. UIST '17. Québec City, QC, Canada: ACM, 2017, pp. 41–48. ISBN: 978-1-4503-4981-9. DOI: 10.1145/3126594.3126619 (cit. on p. 133).

[116] J. Matejka, M. Glueck, T. Grossman, G. Fitzmaurice. "The Effect of Visual Appearance on the Performance of Continuous Sliders and Visual Analogue Scales." In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, 2016, pp. 5421–5432. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858063 (cit. on p. 72).

[117] S. Mayer, H. V. Le, N. Henze. "Designing Finger Orientation Input for Mobile Touchscreens." In: *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '18. Barcelona, Spain: ACM, 2018, 29:1–29:9. ISBN: 978-1-4503-5898-9. DOI: 10.1145/3229434.3229444 (cit. on pp. 22, 68).

[118] S. Mayer, H. V. Le, N. Henze. "Estimating the Finger Orientation on Capacitive Touchscreens Using Convolutional Neural Networks." In: *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*. ISS '17. Brighton, United Kingdom: ACM, 2017, pp. 220–229. ISBN: 978-1-4503-4691-7. DOI: 10.1145/3132272.3134130 (cit. on pp. 22, 40).

[119] S. Mayer, H. V. Le, N. Henze. "Machine Learning for Intelligent Mobile User Interfaces Using TensorFlow." In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '17. Vienna, Austria: ACM, 2017, 62:1–62:5. ISBN: 978-1-4503-5075-4. DOI: 10.1145/3098279.3119915 (cit. on p. 23).

[120] S. Mayer, L. Lischke, P. W. Woźniak. "Drones for Search and Rescue." In: *Proceedings of the 2019 International workshop on Human-Drone Interaction*. iHDI'19. May 4, 2019 (cit. on p. 23).

[121] S. Mayer, M. Mayer, N. Henze. "Feasibility Analysis of Detecting the Finger Orientation with Depth Cameras." In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. MobileHCI'17. New York, NY, USA: ACM, 2017, 82:1–82:8. ISBN: 978-1-4503-5075-4. DOI: 10.1145/3098279.3122125 (cit. on pp. 22, 39).

[122] S. Mayer, H. V. Le, M. Weiß, N. Henze. *Effect of Finger Orientation on Social Settings*. 2019 (cit. on p. 100).

[123] S. Mayer, V. Schwind, H. V. Le, D. Weber, J. Vogelsang, J. Wolf, N. Henze. "Effect of Orientation on Unistroke Touch Gestures." In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland UK: ACM, 2019, 406:1–406:14. ISBN: 978-1-4503-5970-2. DOI: 10.1145/3290605.3300928 (cit. on p. 23).

[124] S. Mayer, L. Lischke, P. W. Woźniak, N. Henze. "Evaluating the Disruptiveness of Mobile Interactions: A Mixed-Method Approach." In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: ACM, 2018, 406:1–406:14. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3173980 (cit. on pp. 22, 99).

[125] S. Mayer, P. Knierim, P. W. Woźniak, M. Funk. "How Drones Can Support Backcountry Activities." In: *Proceedings of the 2017 natureCHI workshop, in conjunction with ACM mobileHCI'17*. Vol. 2. NatureCHI'17. Sept. 3, 2017, p. 6 (cit. on p. 23).

[126] S. Mayer, L. Lischke, A. Lanksweirt, H. V. Le, N. Henze. "How to Communicate New Input Techniques." In: *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*. NordiCHI '18. Oslo, Norway: ACM, 2018, pp. 460–472. ISBN: 978-1-4503-6437-9. DOI: 10.1145/3240167.3240176 (cit. on pp. 22, 129).

[127] S. Mayer, K. Wolf, S. Schneegass, N. Henze. "Modeling Distant Pointing for Compensating Systematic Displacements." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: ACM, 2015, pp. 4165–4168. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702332 (cit. on p. 23).

[128] S. Mayer, L. Lischke, J. E. Grønbæk, Z. Sarsenbayeva, J. Vogelsang, P. W. Woźniak, N. Henze, G. Jacucci. "Pac-Many: Movement Behavior when Playing Collaborative and Competitive Games on Large Displays." In: *Proceedings of the*

2018 CHI Conference on Human Factors in Computing Systems. CHI '18. Montreal QC, Canada: ACM, 2018, 539:1–539:10. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3174113 (cit. on p. 23).

[129] S. Mayer, V. Schwind, R. Schweigert, N. Henze. "The Effect of Offset Correction and Cursor on Mid-Air Pointing in Real and Virtual Environments." In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: ACM, 2018, 653:1–653:13. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3174227 (cit. on p. 23).

[130] S. Mayer, H. V. Le, A. Nesti, N. Henze, H. H. Bülthoff, L. L. Chuang. "The Effect of Road Bumps on Touch Interaction in Cars." In: *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '18. Toronto, ON, Canada: ACM, 2018, pp. 85–93. ISBN: 978-1-4503-5946-7. DOI: 10.1145/3239060.3239071 (cit. on p. 23).

[131] S. Mayer, L. Lischke, P. Kucharski, P. W. Woźniak. "Towards Enhancing Data Exploration with Multiple Mobile Devices." In: *Proceedings of 2015 Cross-Surface workshop, in conjunction with ACM ITS'15*. Vol. 1. Nov. 15, 2015, p. 4 (cit. on p. 23).

[132] S. Mayer, P. Gad, K. Wolf, P. W. Woźniak, N. Henze. "Understanding the Ergonomic Constraints in Designing for Touch Surfaces." In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '17. Vienna, Austria: ACM, 2017, 33:1–33:9. ISBN: 978-1-4503-5075-4. DOI: 10.1145/3098279.3098537 (cit. on pp. 22, 23, 67).

[133] D. McMillan, B. Brown, A. Lampinen, M. McGregor, E. Hoggan, S. Pizza. "Situating Wearables: Smartwatch Use in Context." In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: ACM, 2017, pp. 3582–3594. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025993 (cit. on pp. 33, 100).

[134] J. Müller, R. Walter, G. Bailly, M. Nischt, F. Alt. "Looking Glass: A Field Study on Noticing Interactivity of a Shop Window." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '12. Austin, Texas, USA: ACM, 2012, pp. 297–306. ISBN: 978-1-4503-1015-4. DOI: 10.1145/2207676.2207718 (cit. on p. 130).

[135] R. Murray-Smith. "Stratified, Computational Interaction Via Machine Learning." In: *Eighteenth Yale Workshop on Adaptive and Learning Systems*. New Haven, CT, USA, 2017, pp. 95–101 (cit. on p. 29).

[136] V. Nair, G. E. Hinton. “Rectified linear units improve restricted boltzmann machines.” In: *Proceedings of the 27th international conference on machine learning*. ICML’10. Omnipress, 2010, pp. 807–814 (cit. on p. 56).

[137] W. Newman, E. L. Smith. “Disruption of Meetings by Laptop Use: Is There a 10-second Solution?” In: *CHI ’06 Extended Abstracts on Human Factors in Computing Systems*. CHI EA ’06. Montréal, Québec, Canada: ACM, 2006, pp. 1145–1150. ISBN: 1-59593-298-4. DOI: 10.1145/1125451.1125667 (cit. on p. 35).

[138] D. Norman. *The design of everyday things: Revised and expanded edition*. Basic Books (AZ), 2013 (cit. on pp. 22, 35, 130, 134, 151).

[139] I. Oakley, C. Lindahl, K. Le, D. Lee, M. R. Islam. “The Flat Finger: Exploring Area Touches on Smartwatches.” In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI ’16. Santa Clara, California, USA: ACM, 2016, pp. 4238–4249. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858179 (cit. on pp. 19, 27).

[140] E. Ofek, S. T. Iqbal, K. Strauss. “Reducing Disruption from Subtle Information Delivery During a Conversation: Mode and Bandwidth Investigation.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’13. Paris, France: ACM, 2013, pp. 3111–3120. ISBN: 978-1-4503-1899-0. DOI: 10.1145/2470654.2466425 (cit. on p. 34).

[141] D. G. Okamoto, L. S. Rashotte, L. Smith-Lovin. “Measuring Interruption: Syntactic and Contextual Methods of Coding Conversation.” In: *Social Psychology Quarterly* 65.1 (2002), pp. 38–55. ISSN: 01902725 (cit. on pp. 33, 104).

[142] J. Paay, J. Kjeldskov. “Understanding Situated Social Interactions: A Case Study of Public Places in the City.” In: vol. 17. CSCW ’08 2. 2008, pp. 275–290. DOI: 10.1007/s10606-007-9072-1 (cit. on p. 32).

[143] J. Pearson, S. Robinson, M. Jones. “It’s About Time: Smartwatches As Public Displays.” In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI ’15. Seoul, Republic of Korea: ACM, 2015, pp. 1257–1266. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702247 (cit. on p. 32).

[144] M. Porcheron, J. E. Fischer, S. Sharples. “Using Mobile Phones in Pub Talk.” In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooper-*

ative Work & Social Computing. CSCW '16. San Francisco, California, USA: ACM, 2016, pp. 1649–1661. ISBN: 978-1-4503-3592-8. DOI: 10.1145/2818048.2820014 (cit. on pp. 34, 106).

[145] N. Qian. "On the momentum term in gradient descent learning algorithms." In: *Neural Networks* 12.1 (1999), pp. 145 –151. ISSN: 0893-6080. DOI: http://dx.doi.org/10.1016/S0893-6080(98)00116-6 (cit. on p. 60).

[146] G. Ramos, M. Boulos, R. Balakrishnan. "Pressure Widgets." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '04. Vienna, Austria: ACM, 2004, pp. 487–494. ISBN: 1-58113-702-8. DOI: 10.1145/985692.985754 (cit. on pp. 19, 27, 28).

[147] J. Rico, S. Brewster. "Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. Atlanta, Georgia, USA: ACM, 2010, pp. 887–896. ISBN: 978-1-60558-929-9. DOI: 10.1145/1753326.1753458 (cit. on p. 33).

[148] S. Rogers, J. Williamson, C. Stewart, R. Murray-Smith. "AnglePose: Robust, Precise Capacitive Touch Tracking via 3D Orientation Estimation." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. Vancouver, BC, Canada: ACM, 2011, pp. 2575–2584. ISBN: 978-1-4503-0228-9. DOI: 10.1145/1978942.1979318 (cit. on pp. 28, 29).

[149] Y. Rogers, H. Sharp, J. Preece. *Interaction design: beyond human-computer interaction*. John Wiley & Sons, 2011 (cit. on p. 130).

[150] A. Romanowski, S. Mayer, L. Lischke, K. Grudzień, T. Jaworski, I. Perenc, P. Kucharski, M. Obaid, T. Kosizski, P. W. Woźniak. "Towards Supporting Remote Cheering During Running Races with Drone Technology." In: *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '17. Denver, Colorado, USA: ACM, 2017, pp. 2867–2874. ISBN: 978-1-4503-4656-6. DOI: 10.1145/3027063.3053218 (cit. on p. 23).

[151] A. Roudaut, E. Lecolinet, Y. Guiard. "MicroRolls: Expanding Touch-screen Input Vocabulary by Distinguishing Rolls vs. Slides of the Thumb." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09. Boston, MA, USA: ACM, 2009, pp. 927–936. ISBN: 978-1-60558-246-7. DOI: 10.1145/1518701.1518843 (cit. on pp. 19, 21, 40, 131, 138).

[152] D. Rubine. “Combining Gestures and Direct Manipulation.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’92. Monterey, California, USA: ACM, 1992, pp. 659–660. ISBN: 0-89791-513-5. DOI: 10.1145/142750.143072 (cit. on p. 35).

[153] J. Ruiz, Y. Li, E. Lank. “User-defined Motion Gestures for Mobile Interaction.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’11. Vancouver, BC, Canada: ACM, 2011, pp. 197–206. ISBN: 978-1-4503-0228-9. DOI: 10.1145/1978942.1978971 (cit. on p. 151).

[154] A. Sahami Shirazi, N. Henze, T. Dingler, M. Pielot, D. Weber, A. Schmidt. “Large-scale Assessment of Mobile Notifications.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’14. Toronto, Ontario, Canada: ACM, 2014, pp. 3055–3064. ISBN: 978-1-4503-2473-1. DOI: 10.1145/2556288.2557189 (cit. on p. 106).

[155] J. Sauro, J. R. Lewis. *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann, 2016 (cit. on p. 147).

[156] S. Schneegass, S. Mayer, T. Olsson, K. V. Laerhoven. “From Mobile to Wearable: Using Wearable Devices to Enrich Mobile Interaction.” In: *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. MobileHCI ’15. Copenhagen, Denmark: ACM, Jan. 1, 2015, pp. 936–939. ISBN: 978-1-4503-3653-6. DOI: 10.1145/2786567.2795396 (cit. on p. 23).

[157] S. Schneegass, T. Olsson, S. Mayer, K. van Laerhoven. “Mobile Interactions Augmented by Wearable Computing: A Design Space and Vision.” In: *Int. J. Mob. Hum. Comput. Interact.* 8.4 (Jan. 1, 2016), pp. 104–114. ISSN: 1942-390X. DOI: 10.4018/IJMHCI.2016100106 (cit. on p. 23).

[158] V. Schwind, S. Mayer, A. Comeau-Vermeersch, R. Schweigert, N. Henze. “Up to the Finger Tip: The Effect of Avatars on Mid-Air Pointing Accuracy in Virtual Reality.” In: *The Annual Symposium on Computer-Human Interaction in Play Extended Abstracts*. CHI PLAY ’18. Melbourne, VIC, Australia: ACM, 2018, pp. 477–488. ISBN: 978-1-4503-5624-4. DOI: 10.1145/3242671.3242675 (cit. on p. 23).

[159] J. S. Shell, R. Vertegaal, A. W. Skaburskis. “EyePliances: Attention-seeking Devices That Respond to Visual Attention.” In: *CHI ’03 Extended Abstracts on*

Human Factors in Computing Systems. CHI EA '03. Ft. Lauderdale, Florida, USA: ACM, 2003, pp. 770–771. ISBN: 1-58113-637-4. DOI: 10.1145/765891.765981 (cit. on p. 33).

[160] B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmqvist, N. Diakopoulos. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. 6th. Pearson, 2016. ISBN: 013438038X, 9780134380384 (cit. on pp. 22, 35, 130, 131, 133, 151).

[161] P. Y. Simard, D. Steinkraus, J. C. Platt. "Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis." In: *Proceedings of the Seventh International Conference on Document Analysis and Recognition*. Vol. 3. ICDAR '03. Washington, DC, USA: IEEE Computer Society, 2003, pp. 958–962. ISBN: 0-7695-1960-1 (cit. on p. 30).

[162] N. M. Su, L. Wang. "From Third to Surveilled Place: The Mobile in Irish Pubs." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: ACM, 2015, pp. 1659–1668. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702574 (cit. on pp. 34, 106).

[163] D. Svanaes, G. Seland. "Putting the Users Center Stage: Role Playing and Low-fi Prototyping Enable End Users to Design Mobile Systems." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '04. Vienna, Austria: ACM, 2004, pp. 479–486. ISBN: 1-58113-702-8. DOI: 10.1145/985692.985753 (cit. on p. 127).

[164] Y. Takeoka, T. Miyaki, J. Rekimoto. "Z-touch: An Infrastructure for 3D Gesture Interaction in the Proximity of Tabletop Surfaces." In: *ACM International Conference on Interactive Tabletops and Surfaces*. ITS '10. Saarbrücken, Germany: ACM, 2010, pp. 91–94. ISBN: 978-1-4503-0399-6. DOI: 10.1145/1936652.1936668 (cit. on p. 27).

[165] P. Tolmie, A. Crabtree, T. Rodden, S. Benford. "Are You Watching This Film or What?: Interruption and the Juggling of Cohorts." In: *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*. CSCW '08. San Diego, CA, USA: ACM, 2008, pp. 257–266. ISBN: 978-1-60558-007-4. DOI: 10.1145/1460563.1460605 (cit. on p. 34).

[166] P. Trbovich, J. L. Harbluk. "Cell Phone Communication and Driver Visual Behavior: The Impact of Cognitive Distraction." In: *CHI '03 Extended Abstracts on*

Human Factors in Computing Systems. CHI EA '03. Ft. Lauderdale, Florida, USA: ACM, 2003, pp. 728–729. ISBN: 1-58113-637-4. DOI: 10.1145/765891.765954 (cit. on p. 34).

[167] H. Treiblmaier, P. Filzmoser. "Benefits from Using Continuous Rating Scales in Online Survey Research." In: *Proceedings of the International Conference on Information Systems*. ICIS '11. Shanghai, China: Association for Information Systems, 2011, pp. 1–15. ISBN: 978-0-615-55907-0 (cit. on p. 72).

[168] K. Vega, H. Fuks. "Beauty Tech Nails: Interactive Technology at Your Fingertips." In: *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*. TEI '14. Munich, Germany: ACM, 2013, pp. 61–64. ISBN: 978-1-4503-2635-3. DOI: 10.1145/2540930.2540961 (cit. on p. 133).

[169] R. Vertegaal, R. Slagter, G. van der Veer, A. Nijholt. "Eye Gaze Patterns in Conversations: There is More to Conversational Agents Than Meets the Eyes." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '01. Seattle, Washington, USA: ACM, 2001, pp. 301–308. ISBN: 1-58113-327-8. DOI: 10.1145/365024.365119 (cit. on pp. 33, 102, 117).

[170] A. Voit, S. Mayer, V. Schwind, N. Henze. "Online, VR, AR, Lab, and In-Situ: Comparison of Research Methods to Evaluate Smart Artifacts." In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland UK: ACM, 2019, 406:1–406:14. ISBN: 978-1-4503-5970-2. DOI: 10.1145/3290605.3300737 (cit. on p. 23).

[171] F. Wang, X. Ren. "Empirical Evaluation for Finger Input Properties in Multi-touch Interaction." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09. Boston, MA, USA: ACM, 2009, pp. 1063–1072. ISBN: 978-1-60558-246-7. DOI: 10.1145/1518701.1518864 (cit. on pp. 27, 28, 78).

[172] F. Wang, X. Cao, X. Ren, P. Irani. "Detecting and Leveraging Finger Orientation for Interaction with Direct-touch Surfaces." In: *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology*. UIST '09. Victoria, BC, Canada: ACM, 2009, pp. 23–32. ISBN: 978-1-60558-745-5. DOI: 10.1145/1622176.1622182 (cit. on pp. 27, 28).

[173] H.-C. Wang, D. Cosley, S. R. Fussell. "Idea Expander: Supporting Group Brainstorming with Conversationally Triggered Visual Thinking Stimuli." In: *Proceed-*

ings of the 2010 ACM Conference on Computer Supported Cooperative Work. CSCW '10. Savannah, Georgia, USA: ACM, 2010, pp. 103–106. ISBN: 978-1-60558-795-0. DOI: 10.1145/1718918.1718938 (cit. on p. 32).

[174] J. Wang, J. Canny. "FingerSense: Augmenting Expressiveness to Physical Pushing Button by Fingertip Identification." In: *CHI '04 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '04. Vienna, Austria: ACM, 2004, pp. 1267–1270. ISBN: 1-58113-703-6. DOI: 10.1145/985921.986040 (cit. on p. 133).

[175] D. Weber, S. Mayer, A. Voit, R. V. Fierro, N. Henze. "Design Guidelines for Notifications on Smart TVs." In: *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*. TVX '16. Chicago, Illinois, USA: ACM, Jan. 1, 2016, pp. 13–24. ISBN: 978-1-4503-4067-0. DOI: 10.1145/2932206.2932212 (cit. on p. 23).

[176] D. Weir, S. Rogers, R. Murray-Smith, M. Löchtefeld. "A User-specific Machine Learning Approach for Improving Touch Accuracy on Mobile Devices." In: *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*. UIST '12. Cambridge, Massachusetts, USA: ACM, 2012, pp. 465–476. ISBN: 978-1-4503-1580-7. DOI: 10.1145/2380116.2380175 (cit. on p. 29).

[177] M. Weiser. "The Computer for the 21st Century." In: *Scientific american* 265.3 (1991), pp. 94–105. ISSN: 0036-8733. DOI: 10.1038/scientificamerican0991-94 (cit. on p. 17).

[178] J. Wiese, T. S. Saponas, A. B. Brush. "Phoneprioception: Enabling Mobile Phones to Infer Where They Are Kept." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. Paris, France: ACM, 2013, pp. 2157–2166. ISBN: 978-1-4503-1899-0. DOI: 10.1145/2470654.2481296 (cit. on p. 106).

[179] J. Williamson. "Fingers of a Hand Oscillate Together: Phase Syncronisation of Tremor in Hover Touch Sensing." In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, 2016, pp. 3433–3437. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858235 (cit. on p. 64).

[180] J. O. Wobbrock, L. Findlater, D. Gergle, J. J. Higgins. "The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. Vancouver, BC, Canada: ACM, 2011, pp. 143–146. ISBN: 978-1-4503-0228-9. DOI: 10.1145/1978942.1978963 (cit. on pp. 75, 90, 145, 146).

[181] K. Wolf, S. Mayer, S. Meyer. "Microgesture Detection for Remote Interaction with Mobile Devices." In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. MobileHCI '16. Florence, Italy: ACM, Jan. 1, 2016, pp. 783–790. ISBN: 978-1-4503-4413-5. DOI: 10.1145/2957265.2961865 (cit. on p. 23).

[182] K. Wolf, M. McGee-Lennon, S. Brewster. "A Study of On-device Gestures." In: *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services Companion*. MobileHCI '12. San Francisco, California, USA: ACM, 2012, pp. 11–16. ISBN: 978-1-4503-1443-5. DOI: 10.1145/2371664.2371669 (cit. on p. 30).

[183] K. Wolf, S. Schneegass, N. Henze, D. Weber, V. Schwind, P. Knierim, S. Mayer, T. Dingler, Y. Abdelrahman, T. Kubitza, M. Funk, A. Mebus, A. Schmidt. "TUIs in the Large: Using Paper Tangibles with Mobile Devices." In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '15. Seoul, Republic of Korea: ACM, 2015, pp. 1579–1584. ISBN: 978-1-4503-3146-3. DOI: 10.1145/2702613.2732863 (cit. on pp. 23, 87).

[184] P. C. Wong, H. Fu, K. Zhu. "Back-Mirror: Back-of-device One-handed Interaction on Smartphones." In: *SIGGRAPH ASIA 2016 Mobile Graphics and Interactive Applications*. SA '16. Macau: ACM, 2016, 10:1–10:5. ISBN: 978-1-4503-4551-4. DOI: 10.1145/2999508.2999522 (cit. on p. 29).

[185] P. W. Woźniak, N. Goyal, P. Kucharski, L. Lischke, S. Mayer, M. Fjeld. "RAMPARTS: Supporting Sensemaking with Spatially-Aware Mobile Interactions." In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, 2016, pp. 2447–2460. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858491 (cit. on pp. 23, 81).

[186] P. W. Woźniak, K. Knaving, S. Björk, M. Fjeld. "RUFUS: Remote Supporter Feedback for Long-Distance Runners." In: *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '15. Copenhagen, Denmark: ACM, 2015, pp. 115–124. ISBN: 978-1-4503-3652-9. DOI: 10.1145/2785830.2785893 (cit. on pp. 32, 117).

[187] P. W. Woźniak, L. Lischke, S. Mayer, A. Preikschat, M. Schweizer, B. Vu, C. von Molo, N. Henze. "Understanding Work in Public Transport Management Control Rooms." In: *Companion of the 2017 ACM Conference on Computer Supported Cooperative*

Work and Social Computing. CSCW '17 Companion. Portland, Oregon, USA: ACM, Jan. 1, 2017, pp. 339–342. ISBN: 978-1-4503-4688-7. DOI: 10.1145/3022198.3026341 (cit. on p. 23).

[188] R. Xiao, J. Schwarz, C. Harrison. "Estimating 3D Finger Angle on Commodity Touchscreens." In: *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*. ITS '15. Madeira, Portugal: ACM, 2015, pp. 47–50. ISBN: 978-1-4503-3899-8. DOI: 10.1145/2817721.2817737 (cit. on pp. 21, 27–30, 39, 40, 42, 43, 49, 52–57, 60–63, 71, 158).

[189] V. Zaliva. "3D finger posture detection and gesture recognition on touch surfaces." In: *12th International Conference on Control Automation Robotics Vision*. ICARCV 2012 December. IEEE. 2012, pp. 359–364. ISBN: 9781467318716. DOI: 10.1109/ICARCV.2012.6485185 (cit. on pp. 28, 29, 40).

[190] J. Zheng, D. Vogel. "Finger-Aware Shortcuts." In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. San Jose, California, USA: ACM, 2016, pp. 4274–4285. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858355 (cit. on p. 133).

All URLs cited were checked in March 2019.

List of Figures

1.1 Examples of finger orientations in a mobile setup. Finger orientation input with pitch and yaw input can enlarge the input space for touch surfaces. 20

3.1 The study app is showing instructions to perform a 30° yaw input at one specific position. 42
3.2 The wooden frame with the attached web cameras which we used for ground truth recording in our study. 43
3.3 A participant while performing the task. 44
3.4 The scatter plots are showing the points where we gained data samples from the study. The underlining plane represents the correction model for the pitch correction based on pitch and yaw of the depth camera. 45
3.5 The scatter plots are showing the points where we gained data samples from the study. The underlining plane represents the correction model for the yaw correction based on pitch and yaw of the depth camera. 47
3.6 The study setup showing the Nexus 5 and the aluminum profiles where the cameras are firmly mounted to. 50

3.7 A close up of a participants hand while performing the study. On the participants index finger we attacked the ridget body to track the finger orientation. 52
3.8 The blue counts are representing the distribution of pitch samples which we used for the modeling. The yellow area represents the distribution of pitch samples we recorded in our study. The are in between in obtained by flipping the yaw data. 53
3.9 The blue counts are representing the distribution of yaw samples which we used for the modeling. The yellow area represents the distribution of yaw samples we recorded in our study. The area in between in obtained by flipping the yaw data. 54
3.10 The remaining pitch MAE when using out CNN + L2 model. The gray area shows the 95% CI. 56
3.11 The remaining yaw error when using out CNN + L2 model. The gray area shows the 95% CI. 57
3.12 The layer structure for our best performing model: CNN + L2 model. Max-pooling is used with a Stride of 2 and pedding is set to same. The dense layers use a softplus activation function ($softplus(x) = log(1 + exp(x))$). 59

4.1 The study apparatus with the 3D printed 55° pitch stabilizer and the 16 yaw positions drawn on the touch surface. 69
4.2 The four pitch stabilizers we used in the study to limit PITCH to 77.5°, 55°, 32.5°, and 10° presented from left to right. 71
4.3 The apparatus we used in our study, showing the tablet, the touch layer and one of the pitch stabilizer while one participant touches the touch surface. 73
4.4 The average feasibility RATING (from 0 = "*easy*" to 100 = "*hard*") for the different PITCH inputs. 74

4.5 The average feasibility RATING (from 0 = "*easy*" to 100 = "*hard*") for the different YAW inputs averaged over all PITCHes. The figure also shows the fitted sin curve representing the RATINGS. The blue line indicates the threshold between *comfort* and *non-comfort* zones. 76
4.6 The bars represent how often a yaw angle was rated as not feasible to perform. 78
4.7 Comparison of HAND in respect to the handedness of the participants, showing the average value per YAW. 79
4.8 A participants performing a 32.5° pitch and 45° yaw input with the left hand while being equipped with our 3D printed tracking parts. 84
4.9 The setup with the 6DOF tracking system, the four pitch stabilizer on the left and the Motorola Nexus 6 with markers. 85
4.10 The study app is showing instructions to perform a 10° pitch and 45° yaw input at the red crosshair while remembering the word daughters which is displayed in the upper half. 86
4.11 (a) The four pitch stabilizers with the copper plate and the wire, we used in the study to limit PITCH to 77.5°, 55°, 32.5° and 10° presented from left to right. (b) A CAD model of a pitch stabilizer, revealing the wiring and the copper plate in the base. 87
4.12 The average feasibility RATING (from 0 = "*easy*" to 100 = "*hard*") for the different PITCH_{Finger} inputs. The green areas represent the *comfort* zone in a two-handed smartphone scenario. * the red striped areas represent the *comfort* zone for tabletops as presented in Section 4.2. 88
4.13 The average feasibility RATING (from 0 = "*easy*" to 100 = "*hard*") for the different YAW_{Finger} inputs averaged over all PITCH_{Finger}. The figure also shows the fitted sin curve representing the RATINGS. The blue line indicates the threshold between *comfort* and *non-comfort* zones as defined in Section 4.2. * approximated rating for tabletops as presented in Section 4.2. 89

4.14 The average phone orientation in a two-handed smartphone interaction scenario. 93
4.15 The average phone orientation adjusted to be hand invariant. . . 94
4.16 The average diatnce bewtween the touch point of the finger tip and the cross hair on the screen. 95

5.1 The study procedure in our new mixed-method approach. Data collection methods are: video, audio and eye tracking. 105
5.2 The study setup showing two participants in a conversation wearing mobile eye tracker. 107
5.3 Declining an incoming call selection phase using the standard touch interface. In the first step (a) the participant taps the center icon and then (b) moves it over to the decline symbol, finally (c) the release of the finger will trigger the highlighted action. . . . 109
5.4 Declining an incoming call using *SurfaceSliding*. In a first step (a) the participant grasps the phone. Then the moves the phone in the direction of the decline symbol in respect to the center of the phone (b). After the movement (c) the decline call action is triggered. 110
5.5 The average values and standard error for TECHNIQUE × PHONE. 112
5.6 Declining an incoming call selection phase using finger orientation interface. In the first step (a) the user taps the center icon and then (b) changing the yaw of the finger in order to move the selection icon twards the decline button, finally (c) the release of the finger will trigger the highlighted action. 120
5.7 The study setup showing one participant wearing mobile eye tracker. 121
5.8 The average percentages of looking at the phone during the discussion and standard error for TECHNIQUE × PHONE. 123
5.9 (a) Showing the average reaction time (TCT-R) between incoming call highlight and a participant touched the phone. (b) Showing the average interaction time to decline a call. 124

6.1 The input techniques which were used to study possible communication patters for novel input techniques: *Finger Orientation Interaction*, *Finger Roll Interaction*, *Nail/Knuckle Interaction*, and *Finger-Aware Interaction*. 134
6.2 Sketches drawn by the experts during the interview to underline their strategies for their use cases. (a) and (b) present possible *depiction* icons to guide the user to use their nail or knuckle as input. (c) - (d) present three different use cases each for one input technique. 137
6.3 The three different COMMUNICATIONPATTERNS which were proposed by the experts in the design session. 142
6.4 The five different TASKS used in the evaluation study. 143
6.5 The system usability scale (SUS) results of COMMUNICATIONPATTERN × TASK. Error bars are showing the standard error. . . 146
6.6 The AttrakDiff results of the four categories Pragmatic Quality (PQ), Hedonic Quality-Identity (HQ-I), Hedonic Quality-Simulation (HQ-S), and Attractiveness (ATT) for the three COMMUNICATIONPATTERNS. 148
6.7 Portfolio presentation graph comparison of the AttrakDiff, with Hedonic Quality (HQ) = Hedonic Quality-Identity (HQ-I) + Hedonic Quality-Simulation (HQ-S). 150

List of Tables

1.1 Summary of the research questions in this thesis. 22

3.1 The RMSE and standard division for pitch and yaw per finger. . . 46
3.2 The reduction of RMSE when applying the correction models to pitch and yaw. 48
3.3 The best results for all tested estimation models. Errors are reported in angular degree error. *) These results have been achieved with a small subset of the original data set (1.4%). **) For the reported values we used a DNN instated of a GP regression for the pitch estimation as the data set was to big for a GP. 61

4.1 One-way RM-ANOVAs to determine if the RATING is depended on PITCH within zones and HAND. 76
4.2 One-way RM-ANOVAs to determine if the RATING is depended on YAW within zones and HAND. 77
4.3 One-way RM-ANOVAs to determine if the phones' orientation is depended on ZONE within HAND. 91

6.1 The system usability scale (SUS) results of COMMUNICATION-PATTERN × TASK. 147

6.2 The AttrakDiff results of the four categories Pragmatic Quality (PQ), Hedonic Quality-Identity (HQ-I), Hedonic Quality-Simulation (HQ-S), and Attractiveness (ATT) of COMMUNICATIONPATTERN × TASK. All scales range between -3 and 3. 149

List of Acronyms

ANOVA	analysis of variance
ART	Aligned Rank Transform
ATT	Attractiveness
BoD	Back-of-Device
CNN	Convolutional Neural Network
CSCW	Computer-Supported Cooperative Work
DNN	Deep Neural Network
FCL	fully connected layer
FoV	field of view
GP	Gaussian process
GUI	graphical user interface
HCI	human-computer interaction
HFE	Human Factors and Ergonomics
HQ-I	Hedonic Quality-Identity
HQ-S	Hedonic Quality-Simulation
*k*NN	*k*-nearest neighbor
ML	Machine Learing
NN	Neural Network
PQ	Pragmatic Quality
RF	Random Forest

RMSE	root mean squared error
SUS	system usability scale
UX	user experience