SINGLE-HANDED GESTURE INPUT USING FINGER-TO-FINGER TOUCH AND HAND POSE

Bachelor's thesis

to obtain the academic degree Bachelor of Science submitted
to the Faculty Mathematics and Computer Science
of Saarland University

on June 28th, 2017 by

Lena Hegemann

Submission date:	June 28th, 2017
First reviewer:	Prof. Dr. Jürgen Steimle *Human-Computer Interaction Group*
Second reviewer:	Prof. Dr. Antonio Krüger *Innovative Retail Laboratory*
Supervisor:	Mohamed Soliman *Human-Computer Interaction Group*

ABSTRACT

We present an extended design space for novel hand gestures using thumb-to-finger touch, finger-to-thumb touch as well as hand pose. The touch can be performed on multiple sides of the fingers and can be discrete or continuous. We investigate the properties of the gesture space and structure it into 14 easy to understand classes, based on three gesture primitives. From the gesture space, we derive a concrete set of gestures for the use in interactive user interfaces and propose possible mappings to actions in applications. We evaluate the comfort of the gesture set and the demand of the gesture space in two user studies. We show among other results that it is easier for fingers close to the thumb to initiate finger-to-thumb touch, that there are significant differences between finger sides with the trend of the side facing to the touching finger to be less demanding and that the segments closer to the tip of the fingers and the thumb are preferred by users. As a technical enabler for the gesture recognition, we use a state of the art hand tracker, which can be mounted on the head or shoulder. We describe how the technical novelties of the hand tracker were evaluated. We demonstrate the usability of the gestures with two interactive user interfaces, a photo application, and a video player.

ACKNOWLEDGEMENTS

First and foremost, I want to thank Prof. Dr. Jürgen Steimle, who supervised my thesis and guided and supported my work. My thank also goes to Prof. Dr. Antonio Krüger who was so kind to accept the role as the second reviewer.

I also want to acknowledge my original supervisor Dr. Martin Weigel and Mohamed Soliman who took over for the second half. I am very grateful for their advice, remarks, and feedback. Special thanks to Franziska Müller who set up and maintained the hand tracking system at my working place(s) and was always available for questions about it. I want to thank the people from Human-Computer-Interaction Group and Graphics, Vision and Video Group that gave me valuable input.

The user studies, as well as the evaluation of the gesture recognition pipeline, would not have been possible without the anonymous subjects of the studies. Thank you for participating.

I want to thank Guillermo Reyes Fuentes for helping with the finishing touches and handing in my thesis at the examination office. Last but not least I want to thank Michael Hedderich who gave me moral support and proofread the thesis.

CONTENTS

LIST OF FIGURES

If not noted otherwise figures are designed and edited by the author.

1 Introduction

In this work, we present a novel type of hand gestures for input to interactive systems. Gestures are movements of the body or of body parts with the intent to communicate. They can be used as a natural and intuitive mean to interact with computers [15]. Hand gestures are one very promising type of gestures because the high dexterity of the human hand allows for a great variety of poses and movements [25].

Besides the poses, hand gestures can also leverage touches between fingers. We call this finger-to-finger touch input. It is an interaction technique where touch events between any two fingers of the same hand, including the thumb, occur. Primarily, this can be the thumb touching one of the fingers or a finger touching the thumb.

Finger-to-finger touch is promising because it uses the tactile cues that the skin provides. This helps to perform very fine-grained and precise touch gestures even without looking at the hand. The eyes stay free for a second task or visual output. Furthermore, this one-handed input strategy consists of small and subtle finger movements. For this reason, is can be used discreetly. Since it avoids spacious arm movements, it is also less tiring. Finally, touch has the property to always be clearly present or not. Using this property, users can accurately time their interactions.

There has been a growing interest on finger-to-finger touch in recent years [2, 3, 5, 13, 14, 18, 21, 22, 29, 30, 33]. These prior works showed promising interaction techniques and contributed to the technical feasibility of hand gesture input. However, their gesturing mainly stayed limited to thumb-to-finger touch on the inside of the hand or used only a small area of the finger. Consequently, the applications stayed limited as well.

We aim to make fuller use of the dexterity of the human hand. By extending the gesture space we allow for more gestures and hence more possible mappings between gestures and user controlled functions. This results in more expressive input even for highly complex applications. To achieve this we extend limited thumb-to-finger touch towards the full finger-to-finger touch space by adding finger-to-thumb touch. We increase the overall touch area by using multiple sides and segments of the fingers for touch. Additionally, we combine finger-to-finger touch with the pose of the hand to allow for more expressivity.

The goal of this thesis is to

1. Explore the gesture design space of finger-to-finger touch under consideration of the hand pose. Gaining knowledge about the possible gestures and mappings is important to leverage their full potential.

2. Develop a high-level definition for gestures in this space, in order to facilitate the understanding of finger-to-finger touch gestures and the implementation of a gesture classifier that recognizes them from sensor input. The challenge is to describe gestures in the most simple way that is still exhaustive.

3. Propose a concrete set of gestures for application scenarios. For this, it is necessary that the gestures cover the actions required by applications in a way that makes sense to users. At the same time, gestures need to be distinctive enough to avoid confusion or false classification by the system.

4. Evaluate the properties of finger-to-finger touch gestures with users and test the gesture set in real-life application scenarios to proof its usability.

The remaining parts of this thesis start with a review of related work (2). After that, we continue with a detailed description of the gesture design space (3.1), a proposition on how to describe gestures in it (3.2) and the introduction of a concrete gesture set (3.3). We proceed with two user studies: a pre-study asking participants about the comfort of our gesture set (4.1) and an in-depth follow-up study on the workload of the gesture space (4.2). After that, two application scenarios follow (5), which we implemented with a novel hand tracker using a depth camera. We include a description of how the technical evaluation of the tracker worked (5.1). Finally, we conclude with a discussion of limitations and future work (6).

2 Related Work

This work has been informed by prior research in the field of single hand gestures, covering implementations, taxonomies as well as user studies.

2.1 Mid-air Single Hand Gesture Interfaces

Diverse mid-air gestures performed in a one-handed fashion have been investigated by prior work. Gestures relying on the pose of the hand are a common approach. Cording gestures [4, 12, 20, 25], distinguish gestures by the combination of fingers that are open or bent. A further type of gestures forms a shape or sign with the hand [16, 24, 28]. Digits [16] introduces a system to reconstruct the shape of the hand and showcases application examples with a variety of gestures. Additional to the shape, they also leverage the movement of fingers as well as the in-air movement of the whole hand. Soli [18] proposes a class of gestures they call action gestures, which focus on the motion of the fingers.

There has also been work on gestures using touch between fingers. [23] and [5] developed interaction techniques based on pinch gestures. For such a pinch gesture, the thumb touches one of the four fingertips. A very subtle and still expressive form of interaction between a finger and thumb was presented in FingerPad [3] and NailO [14]. FingerPad uses the outer segment of the index finger as a touchpad, on which the thumb draws stroke gestures. In NailO, a similar interaction takes place the other way around. The index finger interacts on a device covering the thumbnail.

But not only the outer finger segment has so far been considered for touch input. Other parts of the fingers are also suitable, e.g. for taps by the thumb [1,13,22,29,33]. DigiTap [22] uses the positions of the joints and the fingertip for tapping gestures with the thumb and the authors of DigitSpace [13] distributed buttons along the front of the finger segments that they found to be the most comfortable. For another type of thumb-to-finger touch gestures, users convert their fingers into functional sliders by sliding their thumb along them [2, 18, 21, 30]. TIMMi [33] and Ubi-finger [29] presented interactive finger gloves that enabled touch at two to three positions and also considered whether the finger was bent or straight.

These previous works show the variety of possibilities for single hand gestures, be they pose-based or involving touch between fingers. However, the touch events between fingers mainly focus on thumb-to-finger touch or are limited to small areas of the fingers and the thumb. Furthermore, the combination of the hand pose with touch gestures stays limited. We only found examples that considered the pose of the finger which was touched [29,33] or where pinch gestures delimited movement-based gestures [16]. Otherwise, it remains widely unexplored. Therefore the design space is limited and with it the possible application.

2.2 Gesture Taxonomies

There exist different approaches to classify gestures to groups and describe them. Sets of hand gestures that were derived from user elicited data have been divided by the ways the hand holds something while gesturing [32] and the type of action like tapping or swipe [1]. Karam et. al. [15] have presented a classification of gestures by their styles, the application domains as well as input and output technologies. Krupka et. al. [17] presented a language to describe hand gestures with the aim to facilitate the development of hand gesture interfaces. It characterizes hand gestures by the position and poses of the palm and each finger.

While previous work has structured gestures and therefore contributed to a better understanding of single hand gestures and facilitated implementations of gesture recognition systems, we are not aware of any work that includes a high-level description of gestures involving touch between thumb and fingers. Therefore they are not sufficient to describe our gestures.

2.3 User Studies on Hand Gestures and Touch

Our work has also been informed by prior user studies. Huang et. al. [13] showed that tap and stroke gestures are most comfortable on the finger segments situated closest to the thumb and fingertips. They also found that users can distinguish at least 16 buttons on the comfort zones. Another study measured the subjective user ratings on their ability to tap on finger joints and came to similar results [22].

Furthermore, there are findings that point to the advantages of the hand as a touch interface. Visual and tactile cues provided by the skin improve the accuracy of pointing [8] and help users to orient themselves [6,9].

Concerning hand pose, it is important to understand the dexterity of the fingers. Sridhar et. al. [25] investigated how independent each finger can move (finger individuation). In a user study, they found that the middle and ring finger have the least ability to move without also moving the neighboring fingers. Wolf et. al. [32] report a qualitative explanation for the differing ability of fingers to move independently derived from expert elicitation.

User studies revealed the dexterity of the hand. In terms of touch, they located comfort regions and investigate how tactile and visual cues help to improve the performance of touch interfaces on the palm and fingers. For other parts of the hand than the inner side, such findings are still missing.

3 Gestures

In this chapter, we will introduce the new extended design space and propose a gesture language to describe this kind of gestures. We will also propose a concrete gesture set.

3.1 Gestures Design Space

Our extended design space for one-handed gestures includes a bigger scope for discrete as well as continuous finger-to-finger touch and combines it with mid-air gestures based on pose and finger motion. This brings several advantages.

While previous research has mainly focused on thumb-to-finger interaction, where the thumb touches the fingers, we also include finger-to-thumb touch, in which the fingers initiate a touch on the thumb. This form of interaction has new characteristics due to the special properties of the thumb.The thumb is particularly mobile, which is why it can meet the fingers in differing positions. By folding it inwards its backside becomes reachable for the fingers. This allows three-dimensional touch interaction.For finger-to-thumb touch, the same touch gesture can be performed by several fingers or even multiple fingers at the same time. This interesting feature makes it possible to design mappings for several similar actions.At the same time, the thumb is the finger with the largest diameter and the biggest segments. This makes it a large touch area for the fingers. The thumb thus offers more space for touch gestures, which could be beneficial for the accuracy of the user and the detection system.

Not only the consideration of the thumb as a touch surface for the fingers enlarges the overall touch area. Previous work mainly took the front side of the fingers into account. However, to a great extent, the other sides of the fingers are also reachable for the thumb. The greater touch area allows for a broader variety of touch gestures, including strokes in several directions. They can be around or along the finger, staying within a segment or crossing joints and wrinkles. The broader variation of touch actions adds possible physical or spatial mappings and various affordances.

Touch has the property to have two states: either there is a touch present or not. This very clear feature allows users to time their interaction. However, sometimes more than these two states are necessary. A common approach to solving this problem is temporal overloading of touch gestures, e.g. distinguishing between a short and long version of a tap or introducing double taps [?]. However, slow actions make the interface less efficient. Furthermore, they are less robust to track and problematic because of the different perception of speed between users. Combining instead finger-to-finger touch with hand pose adds further expressivity to the gesturing. It removes the need of temporal overloading and allows to work only with quick-to-perform gestures.

Another advantage of having a bigger gesture space (and therefore having more possible gestures) is that it allows for designing interfaces with many well-defined gestures instead of interacting with sequences of gestures.

3.1.1 Possibilities within the Design Space

Considering the finger-to-finger touch gesture space in combination with hand pose provides designers with a huge amount of possibilities. We found the following gesture properties to lay within the design space:

- Simple touches from thumb-to-finger and finger-to-thumb with the exact location of the touch. This can be on any side or segment of the finger reachable for the thumb or fingers. We refer to the finger parts as shown in Figure 3.1.
- Movement of the position of the touch, resulting in a sliding motion or draw gesture
- The pose of each finger, given by the angles of the joints.
- Movements of the fingers not involved in a touch
- Movement of fingers while they touch
- Combinations of the above

However, there are other considerations and constraints that need to be taken into account when designing gestures.

First of all mental demand plays a role. Some of the combinations above are too complicated to be used effectively e.g. two different movements at the same time are hard to coordinate.

Furthermore, some gestures are hard to perform because of anatomical constraints. e.g. the tip of the index finger cannot reach every part of the pinkie and individuation of fingers has limits. [25] reported that the middle and ring finger cannot move without coactivating their neighbor fingers. [1] found in an elicitation study that despite the pinkie has a higher individuation than its neighbors, participants avoided using it. The authors of the study assume that this is because the pinkie is relatively weak. This weakness is another anatomical constraint, we need to take into account.

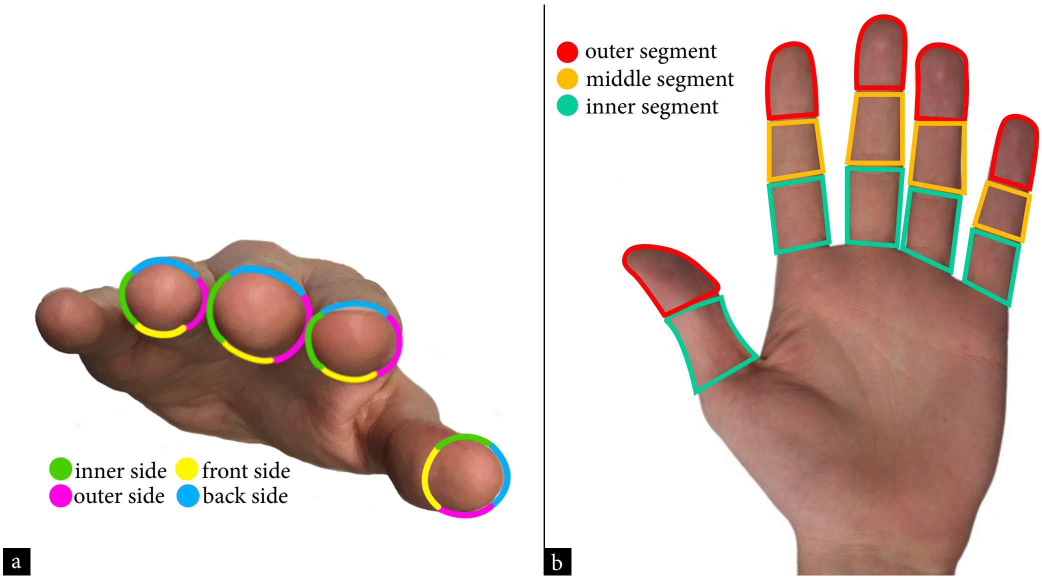

Figure 3.1: The surface of the fingers can be divided into (a) four sides and (b) two to three segments.

Apart from intended gestures the user is likely to also perform non-gesture movements. It is important that the gestures are distinctive from these natural or accidental movements.

3.2 Gesture Definition

To help keeping track over all the possible combinations in the gesture space while at the same time considering the constraints, we propose a structure for it. The structure is given by a simple gesture definition to describe finger-to-finger touch gestures which we detail in this section. Not only does it help to keep an overview it also facilitates the implementation of interactive systems using finger-to-finger touch gestures combined with the hand pose.

Analyzing the properties introduced in 3.1.1, we found that three primitives are sufficient to integrate all of them: touch initiator, touch action, and finger flexion. The first two primitives describe the finger-to-finger touches, like simple touches or moved touches. The third primitive grasps the pose of the fingers and their movement. Each primitive has two to three features (Figure 3.2).

1. The touch initiator. The thumb or one of the fingers can initiate a touch. We differentiate between features for the touch initiator thumb-to-finger touch and finger-to-thumb touch. For **thumb-to-finger touch** the thumb touches one of the fingers with its tip. **Finger-to-thumb touch** works similar but the other way

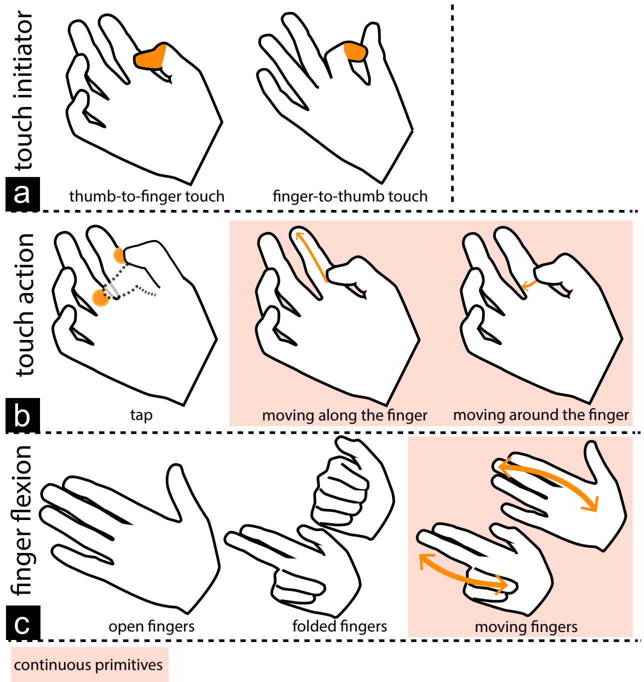

Figure 3.2: Gestures are defined by three primitives. (a) The finger initiating the touch, (b) the touch action, and (c) the finger flexion.

around. One of the fingers touches the thumb with its tip. Theoretically, fingers could also touch fingers. However, in most cases, this is quite unergonomic and furthermore prone to false activation, since fingers touch quite often during regular movement. As investigated by Huang et al. [13] there are certain comfort regions for thumb-to-finger touch input. Since the greatest part of the pinkie finger lays outside the comfort region we decided to exclude thumb-to-pinkie finger touch. Our user study 4.2 showed that this trend also appears for pinkie finger-to-thumb. We, therefore, exclude the pinkie finger from the possible touch initiators.

2. The touch action. By nature fingers are vertically divided into three segments and the thumb into two segments given by the bones and joints. To enable the description of gestures around all sides of the fingers we introduce a division of each segment into four sides; a front, back, inner and outer side (illustrated in Figure 3.1). There are three possible features. A **tap** is a touch of the tip of the thumb on an arbitrary finger segment or a touch of a fingertip on one of the thumb segments. This can be on any side of the finger, that is reachable for the touch initiator. Taps are a discrete primitive feature. Additional to discrete taps the touch action can also involve continuous movement. **Linear slides** can be characterized by a touch point moving along a finger, typically through multiple finger segments, crossing joints, and wrinkles. They can be on any side of the finger and can go from the tip down toward the palm or vice versa. **Rotational slides** are also continuous and the touch point moves around a single finger segment through multiple sides

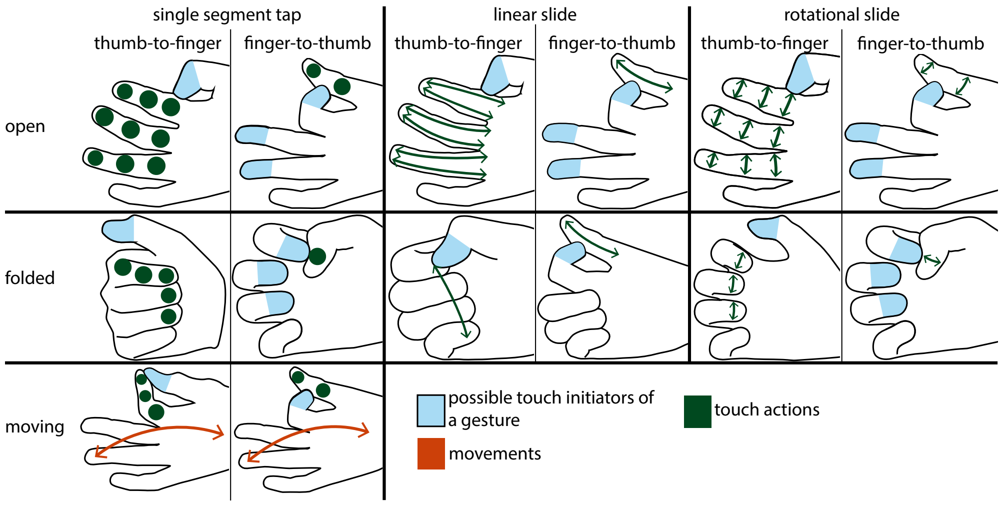

Figure 3.3: The gesture space can be divided into 14 classes, which are defined by the features of three primitives; finger flexion (rows) as well as touch action and touch initiator (columns). Each class is illustrated with some example gestures.

of the segment. They follow the curve of the finger segments. This movement can also be in both directions. Linear and rotational movement can be combined to more complex shapes.

3. The finger flexion. The features of finger flexion are **open, folded** or **moving**. We use the terms open and folded similar to Krupka et al. [17] who defined a finger or the thumb as folded when its tip resides in a certain area in front of the palm. A finger is moving when it changes from open to folded or vice versa during a gesture. This can be simple flapping with the fingers or other movements like drawing circles in the air. Moving fingers are a continuous feature. Sridhar et al. [25] reported that especially for bending the middle and ring finger there is a high co-activation of the neighboring fingers. We therefore only consider movements of the middle, ring and pinkie finger together.

Each gesture is a combination of features from all three primitives.These primitives can be combined to classes. Each class is defined by one out of two touch initiators, one out of three touch actions and one out of three finger flexion features. This results in 18 classes. A gesture belongs to a class if it has all of the class features.

As soon as one of the features of the class is continuous the gestures in it are continuous as well. However, combining two continuous features results in gestures with multiple independent movements, that are hard to coordinate and therefore difficult to perform. We excluded these classes from our gesture space. The remaining 14 classes are illustrated in Figure 3.3.

3.3 Gesture Set

We wanted to create a generic set of gestures that can be used in many different applications. For this, we chose seven representative gesture groups which are subsets of seven different classes. Within the groups, gestures are distinguished by the finger segments or sides involved in the touch action. Every feature of every primitive is covered by at least one of the gestures.

The choice of the gestures was an iterative process. We picked an initial set of gestures and reviewed it iteratively after conducting brainstorming sessions based on the following factors.

The set needs to cover gestures that map to the commands required for user interfaces. On a high level that are activations of one or multiple elements and changes to one or multiple values. Depending on the interface these commands can have different flavors. Activations are discrete commands ranging from picking an item, action or appearance to confirming actions. Changes are continuous commands, ranging from a change of the view to increasing or decreasing a value. In the final gesture set two gesture groups map to discrete activations, four gesture groups map to setting continuous values. We checked the completeness of the gesture set by doing an imaginary walk-through through several applications checking whether all functions were accessible with the gesture set.

The gestures have to follow previous findings on user preferences [1, 13, 32], finger individuation [25] and the advantages of tactile landmarks [9]. These facilitate the ease of use. Therefore, we used this information for refinements.

Within the set, the gestures should be distinctive. There are two motivations for differing gestures. On the one hand, there is the usability. Distinct gestures might reduce the risk of confounding one with the other. In a set of differing gestures, the chance to find a meaningful mapping to a virtual action is higher e.g. a linear finger movement maps to scrolling and image rotation is more similar to rotational finger movement. Meaningful mappings that make sense to the user are easier to learn and to remember. On the other hand, there is the technical feasibility. Distinct gestures tend to generate more distinct signals with a larger margin in between. It is, therefore, easier for the gesture detection system to distinguish them.

The iterative process resulted in a gesture set with two groups of discrete gestures, four groups of continuous gestures and one group integrating both (Figure 3.4).

3.3.1 Discrete Taps

The gesture set includes three groups of tap gestures. One for discrete activation of one out of multiple elements, one to confirm or trigger the main action and one combined with continuous movement e.g. for selecting a mode discretely while performing a continuous change.

Finger tap
The thumb is the touch initiator and performs a tap on the outer front side of a finger segment. Our gesture set includes taps on the inner and middle segments of the index, middle and ring finger, resulting in 6 taps. The position of the taps is rotated slightly towards the thumb to make them easier to perform. As validated during our user evaluation all six tap gestures are performable comfortably (4.1) and with a low workload (4.2). The option to tap on one out of six segments makes this gesture adequate for selecting one out of many options or discrete commands in an application. The positions of the finger segments provide a spatial layout which can be used for a spatial mapping to objects in an application.

Fist tap
The fingers are folded and the thumb performs a tap on the outer side of the index finger. By folding the fingers in, the outer side of the index finger forms a round surface, like a big button that affords to be pressed. Therefore this discrete gesture makes sense for triggering the most important action in an interface e.g. taking a foto in a camera application or playing and pausing a video or audio player. Since users have to roll in their fingers consciously this gesture is unlikely to be activated accidentally. Furthermore, it was rated to be the most comfortable gesture by all participants in our study (4.1). These are two more advantages of mapping this gesture to the most important discrete action in an application.

Tap-and-flap
For a tap-and-flap the thumb taps on the outer side of one of the segments of the index finger. While the thumb holds the tap, the other fingers flap from folded state to open or vice versa. In our pre-study (4.1) we found that this gesture was rated as comfortable as taps alone. It combines the discrete selection of a segment with a continuous movement of the fingers. This can be used for selecting a command or mode by holding a tap while setting a continuous value with the other fingers. Flapping maps to flipping pages conceptually, which also makes this gesture appropriate for browsing through a selection of photos or moving to a different screen in an application.

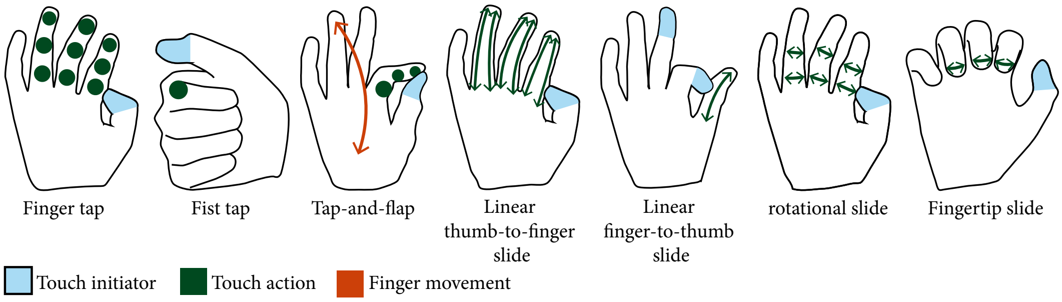

Figure 3.4: The gesture set contains 7 representative groups of gestures.

3.3.2 Continuous slides

The gesture set also contains four groups of linear and rotational slides. As continuous gestures, they are usable for setting one or multiple continuous values. Due to their different lengths, positions, and directions, they can be mapped to values of varying importance, spatial layouts or relation to each other.

Linear thumb-to-finger slide
The thumb is the touch initiator and performs a linear slide along the outer or inner side of the index, middle or ring finger. Linear slides are a good choice for important global continuous variables in an application because they are spatially the most extensive and go beyond individual segments. Having slides on two sides of the fingers does not only increase the total number of slides, but it is also useful for related actions. Two slides on the same finger could stand for similar values, e.g. when editing a photo the outer slide is used for the zoom level in the picture and the one on the inside is used for cropping.

Linear finger-to-thumb slide
For this linear slide finger and thumb flip roles. The index finger or middle finger performs a linear slide along the thumb. It can be performed without any lateral movement of the thumb or the finger by simply bending or unbending the finger. The result is a very quick and easy to perform gesture appropriate for continuous changes that need to be done often or in a repetitive fashion e.g. scrolling.

Rotational slide
The thumb performs a sliding gesture around one of the inner two segments of the index, middle or ring finger. Rotational movements like these have its own quality, bringing different affordances and therefore mappings. Such mappings could be for instance turning or rotating an object. A further property of the rotational slides is that they spatially map to the previously introduced finger taps. In our example applications, we use this for continuous changes related to selections made through finger taps.

Fingertip slide
Similar to the rotational slides the thumb performs a sliding gesture around a finger segment. However, for the fingertip slide, the finger bends inwards so that the thumb can touch the tip of the finger. The sliding motion thus goes around the fingertip following the curve of the nail. Fingertip slides benefit from the haptic path the nails provide for the thumb. Performed at the tip they are most comfortable [13]. The need to bend the finger only increases the workload slightly compared to rotational slides on the other side of the same segment (4.2). Because of the ease and the emphasis provided through the haptic guidance of the nail, they are appropriate for changing values that are more important than the ones mapped to the finger segments but less important than the ones mapped to the linear slides.

With this diverse gesture set, we are able to include a variety of properties that are beneficial for user interaction. All gestures use tactile cues due to the on skin touch, where the skin feels that it is touched. This helps to locate the touch. Additionally, some use tactile landmarks like wrinkles, nails, joints or the curve of the surface. Several analogies exist between real life shapes and actions and the gestures, e.g. the form of a button or the action of flipping through a book. Last but not least we exploit spatial mappings, e.g. the same layout of the segments is used for taps and rotational slides.

4 User Evaluation

4.1 Pre-Study

In a pre-study, we investigated how comfortable users considered the gestures from our gesture set.

4.1.1 Task and Procedure

In an interview, we introduced the participants to the gestures of the gesture set and let them perform the gestures with their right hand. After the participants were familiarized with the whole set we went through all gestures again and asked them to rate the comfort of each gesture on a five-point Likert scale [19], with five being very comfortable and one being very uncomfortable.

4.1.2 Participants

We recruited seven participants (three female). Ages ranged from 23 to 31 with an average age of 24.1.

4.1.3 Results

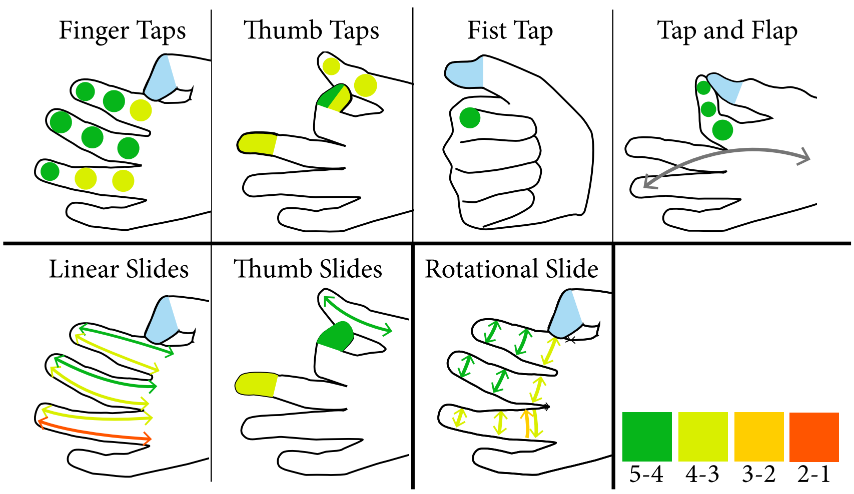

Figure 4.1: Average of the gesture ratings over all participants with 5 very comfortable and 1 very uncomfortable.

The most comfortable gesture is fist tap, it was rated 5 by all subjects. Participants tended to rate thumb-to-finger gestures higher that were closer to the fingertips and on fingers closer to the thumb. The two outer segments of the index and middle finger all received above 4.2. Overall, all of the finger taps were considered rather comfortable with average ratings of 3.5 and higher. Finger-to-thumb taps were rated higher for the index finger than for the middle finger (3.9 vs. 3.5) and for the outer segment of the thumb than for the inner segment of the thumb (3.9 vs. 3.5). Tap and flap gestures were rated 4.2 on average, which is similar to the taps on the index finger without flapping.

Averaged, taps received slightly higher scores than slides. While ratings of slides in the most comfortable regions were similarly high than for tap, the difference in rating between taps and slides increased towards the less preferred areas. Slides were perceived as less comfortable in these areas. Averages for rotational slides ranged from 2.9 to 4.7, for linear slides from 1.7 to 4.7 and for thumb slides from 3.7 to 4.6. Linear slides were rated higher on the outer side of the fingers (outer side: 3.6 - 4.7 vs. inner side: 1.7 - 3.6) and on fingers closer to the thumb (index: 4.1, middle: 3.8, ring: 2.6). The ratings for the thumb slides were similar to the thumb-to-finger slides on the corresponding finger (index: 4.4, middle: 3.7). Like for taps, rotational slides were considered most comfortable on segments close to the fingertips on fingers closer to the thumb.

4.1.4 Discussion

Participants indicated that fist tap was the most comfortable gesture. This matches to the fact that we suggest this gesture for the most important discrete action of an application.

Our results are in line with the results of a study on thumb-to-finger touch comfort regions in DigitSpace[13]. The positions of comfort regions of the two studies are comparable. The results of taps are similar and the results of our rotational slides are similar to the condition in DigitSpace where participants drew shapes on their finger segments. They found that stroke gestures were slightly less comfortable than simple taps, which our data agree on. However, our rotational slides did not score lower on the more comfortable finger segments than taps.

Linear slide gestures received higher ratings on the side of the finger facing towards the thumb. This might have to do with the distance of the finger side to the thumb but cannot be the only reason as the outer side of the next finger was always rated higher. One explanation could be, that users find it uncomfortable to reach around the finger. This is necessary in order to touch the finger side facing away from the thumb but not for touching the outer side of the next finger, which can be exposed by simply moving the neighboring finger away.

Finger-to-thumb touch was rated higher for the index finger and at the outer segment of the thumb than for the middle finger and the inner segment of the thumb. This indicates that the user preference for fingers closer to the thumb and for the tips transfers from thumb-to-finger input to finger-to-thumb input.

4.2 Main Study

The pre-study gave first insights and showed important trends. To broaden our understanding on the subjective view of users on our gesture space, we decided to expand the study based on the outcomes, which we describe in detail in the section above. We kept certain aspects of the study and refined others based on the insights we had gained.

The pre-study showed which fingers are the most and least comfortable to use. For one part of the study, where the number of testable gestures was limited we focused on these two extremes to limit the number of fingers to evaluate and instead expand in other directions of the gesture space like finger sides and finger-to-thumb touch.

During the pre-study, we used only the subset of the gestures that we expected to be usable. We observed, that participants tended to rate gestures relative to each other i.e. give the lowest possible rating to the least comfortable gesture in the set and the highest possible rating to the most comfortable gesture in the set, even though there might be more comfortable or less comfortable gestures outside of the set. For the other part of the study, where the number of gestures was less constrained we decided to include all gestures in the test set in order to normalize the scale, even though we expected some of the gestures to be very difficult.

We changed the measurement from rating comfort to NASA TLX standard questions in order to get more detailed information. NASA TLX is a commonly used multi-dimensional rating scale, designed to measure subjective workload [11]. It has six scales, each covering a different source contributing to the overall workload. Of these scales, we chose the three that we are most interested in: physical demand, mental demand, and performance. Using only some of the scales and analyzing them independently is a popular practice [10].

4.2.1 Research Questions

We are interested how users rate the extensions we made to the gesture design space compared to previous work. Particularly we want to investigate which fingers are preferable touch initiators and which differences between fingers sides exist. Previous work [13] and our pre-study found that touch on finger segments close on the fingertips and fingers close to the thumb is more comfortable. We would like to test whether there are similar relationships for finger-to-thumb interaction. Hence, we wonder whether the outer segment of the thumb is easier to touch than the inner one and whether it is easier to touch the thumb with fingers close to it. We also want to compare touch actions. Finally, we would like to find out whether finger-to-finger touch gestures are still convenient during walking since this is an important factor for mobile interaction.

4.2.2 Hypothesis

We have the following hypothesis based on our research questions.

H1: For finger-to-thumb touch the smaller the distance between the touch initiator and the thumb the lower is the workload.
H2: Touch gestures on the side of the finger facing towards the touch initiator are least demanding. In other words, the further the finger side faces away from the touch initiator the higher is the demand.
H3: Like for finger segments, the outer segment of the thumb is easier to touch than the inner one.
H4: Taps have a lower workload than slide gestures.
H5: If any, there is a small difference between performing the gestures while sitting and walking, with walking slightly more demanding.

4.2.3 Task and Procedure

Participants were situated in front of a screen which showed a sketch of a hand. On the sketch, for each trial, the fingertip of the touch initiator was highlighted in blue and the touch action was indicated with arrows or dots in red. The study was split into four parts; one for each of the three touch actions and a fourth one testing a representative subset of the gestures during walking (Figure 4.2).

Part one included thumb-to-finger taps on the frontside of all finger segments and finger-to-thumb taps by all fingers on the front of all thumb segments. **Part two** included linear thumb-to-finger on the inner, outer and front side of each finger and linear finger-to-thumb slides on four sides of the thumb initiated by all fingers. **Part three** included rotational slides of the thumb around the front of each finger segment and around the back of the outer segment. The fingers performed rotational slides on the front and back of all thumb segments in this study part. These three parts were counterbalanced between participants and the sequence of

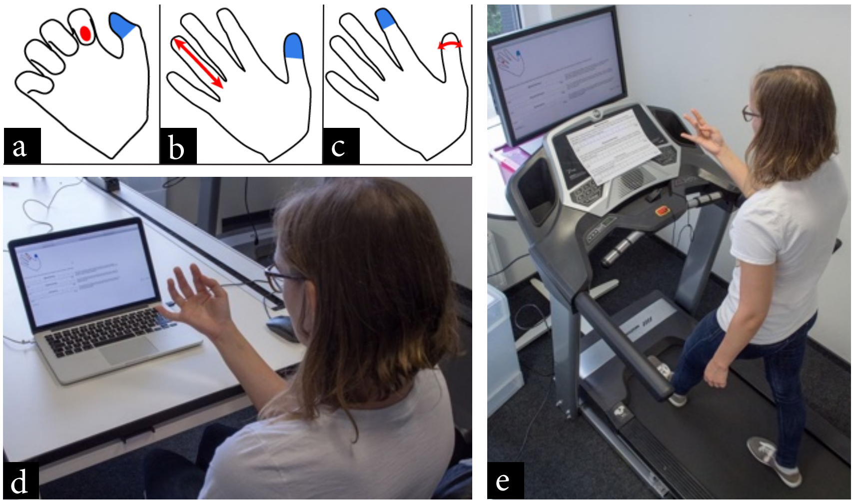

Figure 4.2: Participants were situated in front of a screen displaying an illustration of one gesture per trial with the touch initiator highlighted in blue und the touch action with a red dot to illustrate a tap (a) or an arrow to illustrate a linear (b) or rotational (c) slide. Participants performed and rated gestures sitting (d) as well as walking on a treadmill (e).

the gestures within the parts was randomized to avoid effects of order. The **fourth part** repeated a subset of each part while the participant walked on a treadmill (Horizon Fitness Paragon 6) with a speed of 4 km/h, which is a fast walking speed that has been used previously by Weigel and Steimle [31]. The subset included only the gestures involving the index finger and the ring finger which we expect to be the easiest and hardest gestures from out gesture set based on the results of our pre-study. Half of the participants did this part first, the other started with the first three parts. The order of the touch actions during the walking condition stayed the same for the participants.

Together all parts contained 138 trials. Before the start of a study part, the experimenter introduced the gestures and their visualization to the participant by four representative examples. After this introduction, the participants were asked to perform the indicated gestures with their dominant hand exactly three times. After performing a gesture they were asked to answer the three NASA TLX questions concerning mental demand, physical demand, and performance [11] about it. At the end of the study participants filled out a questionnaire asking for demographic information, whether they would use this style of gesture in public and additional comments. Participants had the option to have a break between study parts. On average the study took 40 minutes per participant.

4.2.4 Participants

For the main study, we recruited 12 participants (6 female). Ages ranged from 23 to 34 with an average age of 27. One participant was left handed. Hand size varied between 75 mm and 93 mm in width (average 82.8 mm) and 167 mm and 196 mm in length (average 183.3 mm).

4.2.5 Results

Before the analysis, we removed outliers that lay more than three times the interquartile range above the third quartile or below the first quartile.

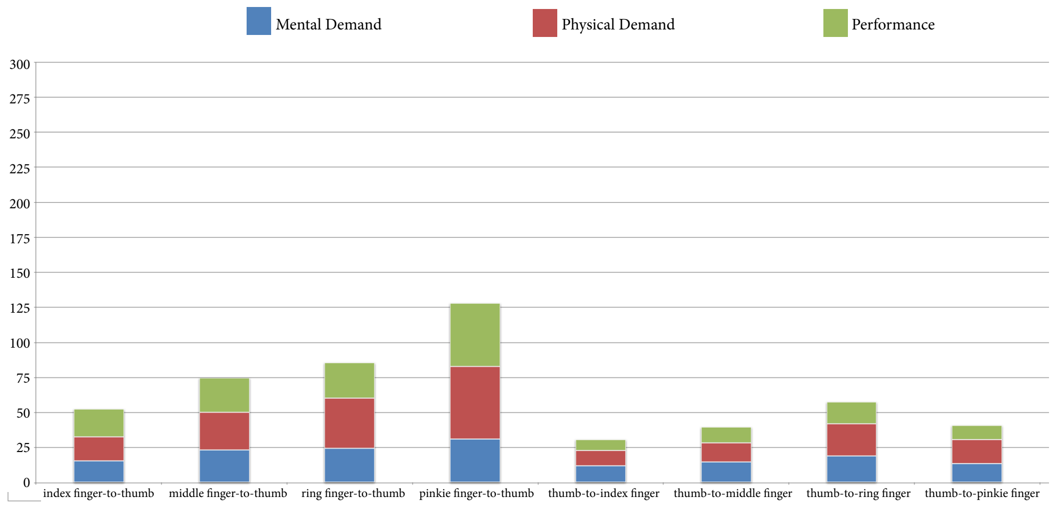

Figure 4.3: Mental demand, physical demand and performance for different touch initiators. Each metric is in the range between 5 and 100. Each bar illustrates the sum of the three metrics. Acting as touch initiators for linear slides on the front of the thumb the workload increases slightly with the distance of the finger from the thumb. Means for thumb-to-finger touch are typically lower than finger-to-thumb involving the same fingers. These differences are significant for the index and pinkie finger.

H1: We tested the differences between the four fingers (index, middle, ring, pinkie) acting as touch initiators for a linear slide on the front of the thumb. For all three scales, one-way ANOVAs revealed that there are differences between the four fingers (mental: F(3, 39)=5.02, p=0.00489, physical: F(3, 36)=17.54, p=0.0000003, performance: F(3, 39)=8.4, p=0.000198). There is a trend that the fingers closer to the thumb are considered less demanding than those further away. Using paired t-tests, significant differences could be found for the index finger (p=0.03) which was less mentally demanding and the ring finger and pinkie which were more physically demanding than their neighbors towards the thumb (ring finger: p=0.046, pinkie: p=0.0032). Except for the pinkie, which was significantly harder to use as touch initiator (p=0.012), there were no significant differences in performance. The direct comparison of each finger involved in thumb-to-finger and finger-to-thumb touch showed that on all scales thumb-to-finger touch is easier for the index finger and the pinkie finger. For the middle finger and ring finger, we did not measure significant differences. Average ratings can be seen in Figure 4.3

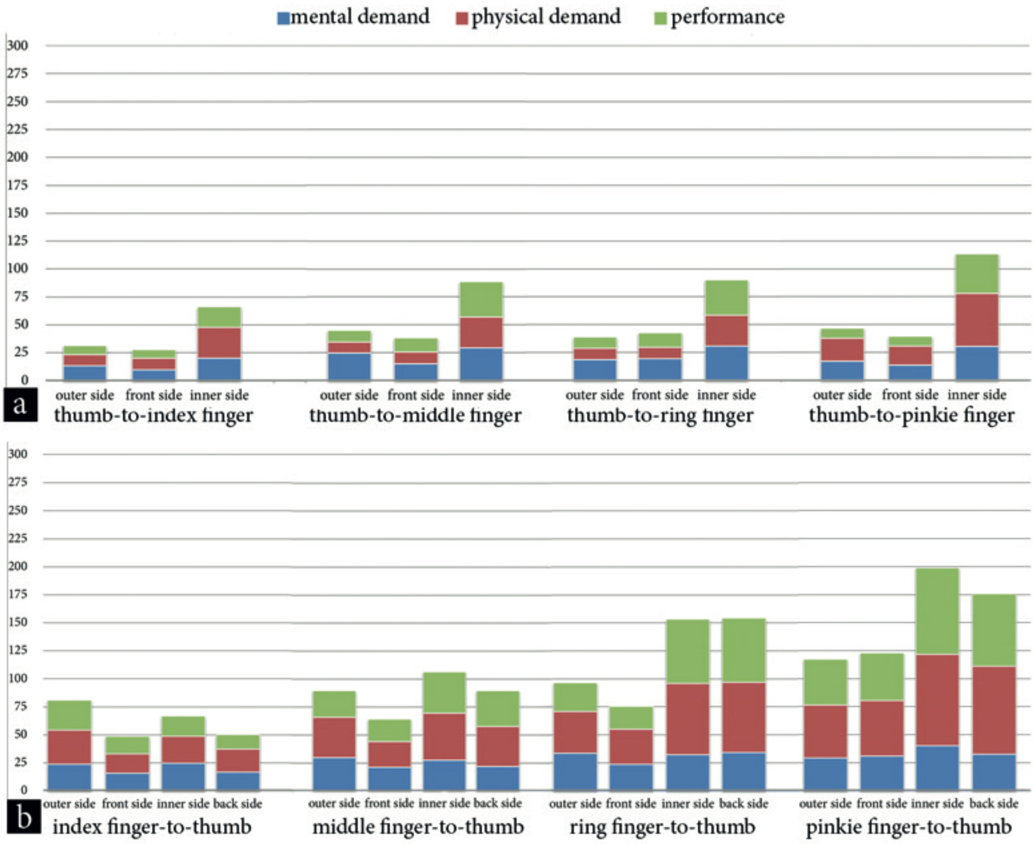

Figure 4.4: Mental demand, physical demand and performance for different finger sides. Each metric is in the range between 5 and 100. Each bar illustrates the sum of the three metrics. (a) Thumb-to-finger linear slides are little demanding on the outer and front side. On the inner side, they are significantly harder. (b) For finger-to-thumb touch the least demanding side to slide on varies between touch initiators. For the index finger, it is the front, inner, and back side, for the middle finger, it is the front, and for the ring and pinkie finger, it is the outer and front side.

H2: For thumb-to-finger linear slides there were no significant differences between the outer and front side in any of the three scales (p>0.11 for all t-tests). However, the slide on the inner side was rated significantly higher throughout all fingers for physical demand and performance (physical demand: p<0.0048, performance: p<0.038 for all t-tests). We observed the same effect for the mental demand on all fingers except for the index finger, where mental demand was not significantly higher for any of the sides (index finger: ANOVA F(358.33, 128.47)=2.78 p=0.081, outer vs. front side: p>0.13, front vs. inner side: p<0.011). Rotational slides on the nail and on the front of the outer finger segments tended to be rated easier on the front but showed only small differences (averages differed at most by 10) which mostly were not significant (nine of them not significant with p>0.08, three of them significant with p<0.04). The preference for sides is a little more complex when fingers touch on the thumb. Concerning physical demand and performance, we found the following significant differences. With the index finger it was harder to touch on the outer side than on the front (physical demand: p=0.018, performance: p=0.018) and back side (physical demand: p=0.043, performance: p=0.00089). The middle finger could best touch the front of the thumb (significantly lower physical demand than outer side p=0.0079, and performance than inner side p=0.033), while the ring and pinkie finger can touch the front and outer side better than the other two sides (p<0.041 for all t-tests). Figure 4.4 shows the average ratings of the finger sides for each touch initiator.

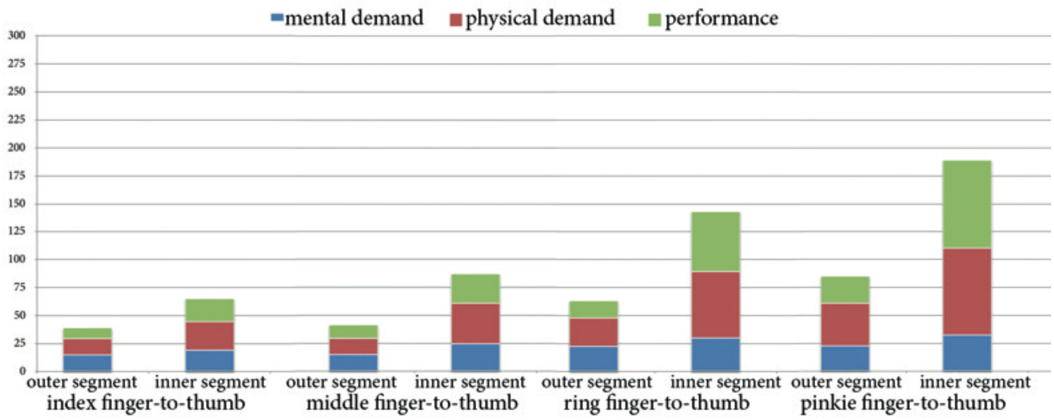

Figure 4.5: Mental demand, physical demand and performance for different thumb segments. Each metric is in the range between 5 and 100. Each bar illustrates the sum of the three metrics. Comparing the inner and outer segment on the front side of the thumb, we observe the trend that a touch on the inner segment is physically more demanding and harder to perform than on the outer segment. This effect becomes more significant towards the pinkie finger.

H3: We compared rotational slides on the outer and inner segment of the thumb performed by each finger. We observed that there is a tendency that the outer segment was easier to touch than the inner one (Figure 4.5). This effect was rather small and mostly not significant for mental demand. However, the further the distance between the touch initiator and the thumb, the bigger was the difference in physical demand, when comparing the outer and inner thumb segment (significant for all fingers, index finger: p=0.0059, middle finger: p=0.007, ring finger: p=0.00087, pinkie: p=0.00051) and performance (significant for middle, ring and pinkie finger, index finger: p=0.0648, middle finger: p=0.016, ring finger: 0.0011, pinkie: p=0.0001).

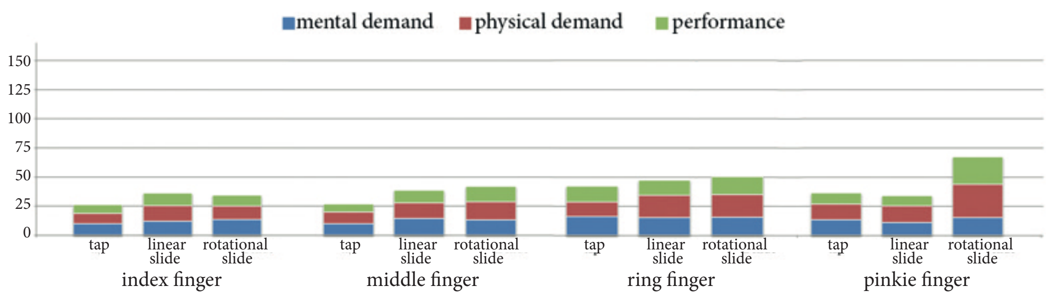

Figure 4.6: Mental demand, physical demand and performance for different touch actions. Each metric is in the range between 5 and 100. Each bar illustrates the sum of the three metrics. We compared taps, linear slides and rotational slides in the middle and front of the fingers. The demand of touch actions is similar. Only rotational slides on the pinkie finger are significantly harder concerning physical demand and performance than taps and linear slides on the pinkie.

H4: In order to investigate touch actions, we compared taps on the middle segment with linear slides on the front as well as rotational slides on the middle segments performed by the thumb on all four fingers. Except for rotational slides on the pinkie finger, there were only small differences in the means (of 6.6 or lower) which were not significant. On the pinkie finger, taps were rated significantly easier than rotational slides (13.33 vs. 28.33 physical demand (p=0.02), 9.58 vs. 23.75 performance (p=0.031)) (Figure 4.6).

H5: Paired t-tests revealed that there are no significant differences between sitting and walking. We tested this on each scale for each touch action (p>0.07 for all tests). The biggest difference between the means was 2.06 for the mental demand of linear slides (20.06 sitting, 18 walking), which is less than half a point on the rating scales.

On the questionnaire 11 out of 12 participants stated that they would use this type of gestures in public. Comments on this mostly referred to the discreetness of the gestures, e.g. "It's quite subtle compared to larger mid-air gestures or speech control" (P11), "They seem to be "little" enough to not look stupid while performing them" (P8) and "It's more or less standard and almost not noticeable. It won't arouse any questions" (P12). Some participants appreciated that most of the gestures were easy ("They were convenient and mostly easy to do.", P9), performable without holding a device ("You don't have to hold the device in your hand, in contrast to [a] touchscreen" (P11), matched to everyday behavior "some similarity to everyday hand gestures", P11), ("The gestures do not interfere that much with what my hand would be doing when walking or sitting.", P8) and the even distribution of workload between fingers ("I use all the fingers which might be better compared to a smartphone where I have a one-sided strain on the thumb", P11). Negative comments often referred to pinkie finger-to-thumb and ring finger-to-thumb touch e.g. "Some of the pinkie finger-to-thumb gestures were too hard to perform and felt uncomfortable." (P8). Participant 9 also commented that long nails make some sliding motions difficult.

4.2.6 Discussion

As expected by **H1**, there is a tendency of the workload to increase from the index finger towards the pinkie as touch initiators which is significant. An interesting side finding is, that this was not true for thumb-to-finger touch. Here values increased towards the ring finger and the pinkie received lower ratings. The difference between the perceived performance for linear slides on the ring finger and pinkie was significant (p=0.046). This is not in line with the results in DigitSpace [13]. It would be interesting to investigate in future work whether this can be replicated. It would be thinkable that it is for instance due to the differing measurement methods between DigitSpace and out work.

H2 can be considered partially true. For thumb-to-finger touch the outer and front side received similar ratings while the inner side was rated more difficult. Considering that the thumb is located slightly rotated towards the outer side this is in line with the hypothesis. Index finger-to-thumb and middle finger-to-thumb support the hypothesis: the outer side of the thumb faces away the most from the index finger and was rated significantly harder to touch than the other three sides. For the middle finger, the front side, which has the smallest angle to the finger, was rated easiest to touch. Therefore these touch initiators support the hypothesis. The ring and pinkie finger only partially support the hypothesis. The front side, which faces towards these two fingers, was rated among the easier sides to touch. However, the inner sides of the thumb were significantly more demanding than the outer side, even though both sides are rotated 90 degrees from the front side. The cause of the higher demand could be that the fingers have to turn inwards laterally in order to reach the inner side of the thumb. This is difficult because the finger joints do not have the necessary degree of freedom [7]. The differences in mental demand for touches on the thumb by these two fingers is not significant, which is a further hint that a physical factor causes this inconsistency with the otherwise approved hypothesis.

Concerning physical demand and performance we can accept **H3**, as the rotational slide was significantly harder on the inner segment of the thumb for most fingers (all for physical demand, all but the index finger for performance). However, the mental demand does not increase significantly towards the inner segment.

We have to reject **H4**, which said that taps are easier than slide gestures. The results rather hint that linear slides and taps are similarly easy. Rotational slides are significantly harder only on the pinkie finger. Since this effect is only limited to one finger, it is likely that the cause for the higher demand does not lie in the touch action but in the properties of the finger. The inner side of the pinkie finger was the side which was most demanding for linear slides. In contrast to the tap and linear slide on the front, the rotational slide touches this difficult side, which could cause the demand of the rotational slide to increase.

Concerning the mobile use of finger-to-finger touch we did not find a significant increase in the demand of the gestures during walking and can accept **H5**. This tells us that the physical activity of walking does not interfere negatively with performing finger-to-finger touch gestures and that it is not an obstacle for using them in a mobile setting. However, walking on a treadmill has a lower mental workload than walking in a real life setting because it does not require attention on where to go. Since participants rated the mental demand of the gestures lower than the other scales, we assume that they are likely to be performable together with the mental demand of walking. Still, it would be interesting for future work to investigate this further.

A few participants stated that some of the gestures were too hard to perform. With the NASA TLX scales it is difficult to find the border between gestures which are demanding and those which are too demanding because they do not implement such a border [10]. They can only tell which gestures are more or less demanding. However, the statements of the participants can provide some hints. Participants mentioned that the pinkie finger-to-thumb and ring finger-to-thumb gestures touching the inner segment of the thumb on the back and inner side were to hard. These gestures have in common that their average ratings for physical demand and performance both lay higher than 50.

4.2.7 Design Implications

From the study results, we derived a set of design implications. For thumb-to-finger touch the workload increases from the index finger to the pinkie finger. Therefore, common actions, that need to be performed often, should be performed with fingers closer to the thumb.

The finger sides facing away from the thumb have a higher workload and should, therefore, be considered with care. There might be good reasons to use them, e.g. to avoid accidental input or confusion with other finger sides, but the higher workload should be considered, especially for fingers further away from the thumb. For finger-to-thumb touch, the front of the thumb is a safe choice for all touch initiators. The back side is only easy to reach for the index and middle finger, while the outer side is a better choice for the ring and pinkie finger.

Thumb-to finger touch is preferable on the outer segment. For the index, middle and ring finger this segment is rather easy to touch, while the inner segment is more demanding.

Gestures with a physical demand and performance above 50 should be avoided, as participants stated that some gestures with such high ratings were too difficult. Among those gestures there are the tap on the nail of the pinkie finger, the tap on the back of the inner segment of the thumb with the pinkie and ring finger, rotational slides on the inner segment of the thumb with the pinkie and ring finger and linear slides on the inner and back side of the thumb with the pinkie and ring finger.

5 Interactive User Interfaces

In this section, we demonstrate two applications to showcase the use of our gesture set in a stationary as well as a mobile scenario with different devices. We describe how users interact with the applications and how we implemented them. As a technical enabler for the gesture recognition, we use a state-of-the-art hand tracker with enhanced functionalities, which we evaluated in three steps.

5.1 Technical Enabler

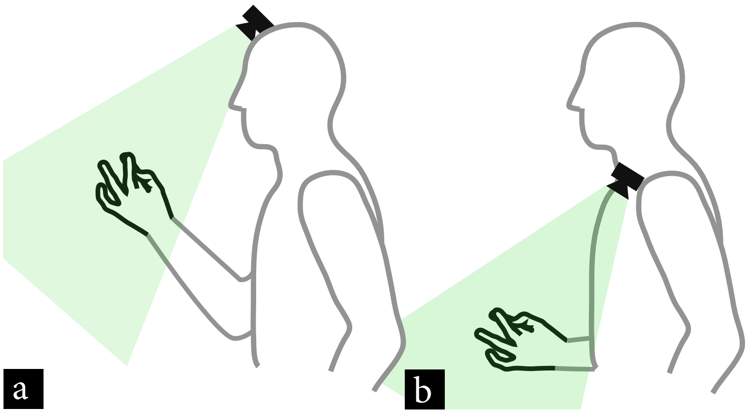

Figure 5.1: As a technical enabler we use a hand tracker. It facilitates the data of a depth camera to reconstruct the position and pose of the hand. The camera can be mounted on the (a) head or the (b) shoulder

For our implementation, we used an existing version of a hand tracker by Sridhar et. al. [26], with improvements implemented by Franziska Müller. It works with a depth camera and was optimized for the use from an egocentric view so that we could mount the camera on the shoulder or the head of the user (Figure 5.1). The tracker reconstructs the hand pose. Additionally, it detects finger-to-finger touch points. The visual approach, using a depth camera has the advantage to work without instrumentation of the hand, so that the movement of the hand is unrestricted and natural. The output of the tracker is a feature vector containing information related to the features of the gesture primitives (3.2) and the exact touch points. A gesture classifier uses this to predict which gesture was performed and sends the prediction to the applications.

We evaluated the improvements of the recognition pipeline, namely the accuracies of the hand pose estimation from an egocentric view, the touch detection, and the gesture classification.

5.1.1 3D Fingertip Position from an Egocentric Viewpoint

So far there exists no common public dataset with ground truth of depth information of a hand performing complex finger movements from an egocentric perspective. We, therefore, decided to record new egocentric sequences for the purpose of evaluation. We recruited two participants (1 female, aged 25 and 27 years). Their hands had a length of 201 mm and 192 mm and a width of 90 mm and 89 mm. The participants imitated a recorded video containing gestures from the gesture set, other gestures from the design space as well as non-gesture movements. The hand gesturing was recorded by the depth sensor that the participants wore on the shoulder. We recorded one sequence per participants. The resulting two sequences of similar length had a total number of 3,573 frames. The second participant rotated his hand more during recording. This caused more occlusion in the second sequence and thus made it more challenging for the tracker. We annotated the 3D fingertip positions in all frames resulting in ground truth similar to the Dexter data set [27] which is a commonly used data set of hand motion from the third person view.

This new egocentric data set allowed for testing the accuracy of the tracker by running the tracking algorithm on the recorded sequences and calculating the average fingertip localization error. The error was 13,92 mm for sequence one and 16,37 mm for sequence two, which is an improvement over the previous version of the tracker (22.5 mm and 38.0 mm on the same sequences).

5.1.2 Touch recognition accuracy

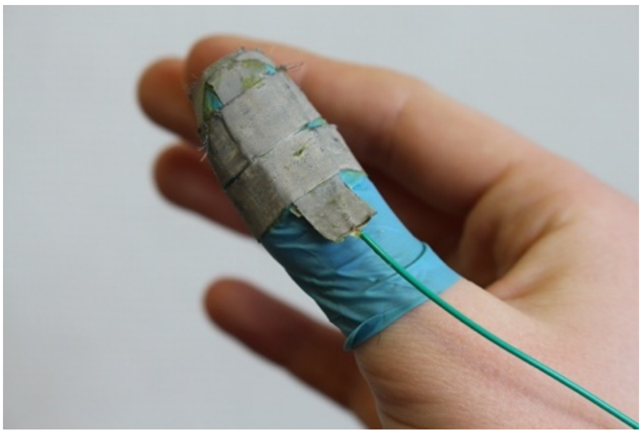

Figure 5.2: Capacitive thumb glove made of nitrile and conductive fabric for automatic ground truth annotation of finger-to-finger touch.

For the evaluation of the touch accuracy, we recruited two participants (1 female, aged 25 and 27). Their hands had a length of 201 mm and 192 mm and a width of

90 mm and 89 mm. The participants imitated a recorded video sequence containing a variety of touch gestures from our gesture space as well as hand movements without touch. With a depth camera mounted on the shoulder of the participants, we recorded their gesturing. To automatically annotate the depth data with ground truth for touch, participants wore a thumb glove out of nitrile. Conductive fabric was glued around the tip of the thumb in such a way that the shape and mobility of the thumb were preserved. Both materials did not distort the depth image of the camera. We connected this thumb glove to an Arduino Uno board[1] and used the CapacitiveSense Library[2] to detect the touch between thumb and fingers. Capacitive sensing can distinct touch from hover state. Hence it is more stable than an optical approach since the touch often times is occluded by the touching finger. The Arduino sent the current reading value to the recording software via a serial port every 10 ms, which is twice as fast as the maximum frame rate of the depth camera. The recording software synchronized the touch and depth data and annotated each frame with the touch measurement.

The tracking software was run on the two sequences and the result of the touch recognition was compared with the ground truth frame by frame and using different time windows, where we compared whether there was a touch in the tracking result and the ground truth within the window. Frame-by-frame comparison reached an accuracy of 87.5% and with a time window of 500ms, the accuracy rose to 95.8%.

5.1.3 Gesture classification accuracy

On top of the hand tracker, we used a gesture classifier to distinguish between gestures as well as non-gestures. We evaluated the whole gesture recognition pipeline with 10 participants (3 female, ages ranged from 23 to 31 with an average age of 25.2). The hand sizes varied from 155 to 212 mm in length (mean 180 mm) and 74 to 92 mm in width (86 mm). Participants performed 50 gestures from the gesture set which belonged to 7 different classes. While the participants performed the gestures we ran the hand tracking software on it in real time and recorded its output in form of a feature vector. We ran the classifier on the feature vectors and got an overall accuracy of 90.52% for recognition of the 7 gesture classes.

5.2 Implementation of Interactive User Interfaces

The hand tracker and gesture classifier provide an interface for applications that can be easily used in a modular fashion. Applications can subscribe to the gesture classifier via the zeroMQ protocol[3]. The classifier transfers identifiers

1 https://www.arduino.cc/en/main/arduinoBoardUno
2 http://playground.arduino.cc/Main/CapacitiveSensor?from=Main.CapSense
3 http://zeromq.org

for the gestures, which the applications receive and then react accordingly. We implemented two interactive user interfaces on top of this recognition pipeline. Thanks to the flexible architecture we could use different hardware as well as different programming languages.

In applications, the gesture set has some advantages over previous hand gestures. On the one hand, there are more direct mappings. For example, by providing rotational slides on multiple finger segments, the users do not have to first make a selection in a menu. Instead, they can make their changes directly. Compared to hand gestures with large arm movements, the control of our apps is less tiring. Thus, a linear slide along the finger requires less movement than a large waving gesture in the air, which is, for example how the volume is edited in Digits [16]. The various haptic cues provided by skin and bones and spacial layouts on the hand can be used for better mappings. We use, for example, linear slides for cropping an image and rotational slides for rotating it.

5.2.1 Photo App for Smartwatch

Our first application is a photo app for a smartwatch. Our finger-to-finger touch gestures enable the control over the app without the need to interact on the limited area of the touch screen. The photos on the screen stay unoccluded and the input on the fingers is more precise due to the larger area.

5.2.1.1 Application Walkthrough

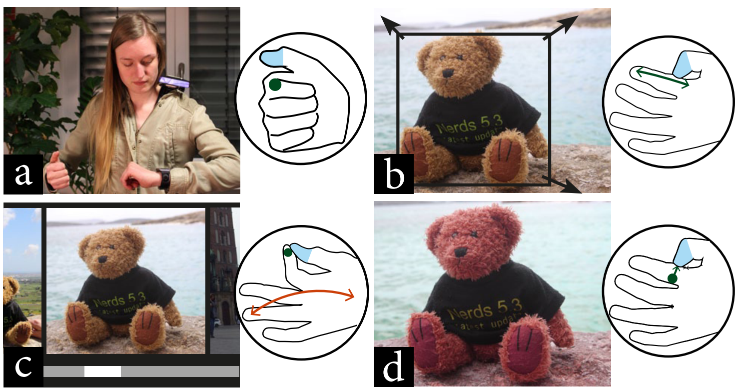

Figure 5.3: Users can control a photo application with (a) fist tap to capture a photo, (b) linear slide on the index finger to zoom in or out of a picture (c) tap and flap to browse through their gallery (d) tap on a finger segment to select an editing mode and alter its value with a rotational slide.

Taking a photo, the most important function of the application is mapped to fist tap. This is easy to perform but still unlikely to be performed accidentally, so no unintended pictures will be taken. After capturing a picture, users can view and edit it. They zoom in and out by performing linear thumb-to-finger slide on outer the index finger as this is a frequently used function, especially on small screens. Cropping the image, which is related to zooming can be done on the inner side. The lower comfort of inner side makes accidentally cropping the image instead of zooming unlikely. More editing functions, e.g. brightness, contrast, rotation or filters are mapped to the finger segments. Tapping and holding a segment with the thumb activates an editing mode. The value of this mode can then be adjusted with a rotational slide on the same segment. Users can browse through their photo gallery with tap and flap gestures.

5.2.1.2 Implementation Details

The prototype is programmed in Python 3[4] and runs on a Raspberry Pi[5] with an external display and the Raspberry Pi Camera Module V2 [6]. The program uses a state system which allows switching between the camera and the gallery mode. For previewing the camera image and taking photos the PiCamera API[7] is used. For displaying the gallery we use the native image viewer of the system. The image manipulation is done with the library Pillow[8].

5.2.2 Video Player

Our second application uses our gestures for a video player application on a smart television with context search. In contrast to a remote control, the hand cannot be misplaced. In addition, a remote control usually requires line of sight to the TV, which often leads the user to change his position in order to point to the television. With the tracking hardware for our system, mounted on the shoulder, the line of sight between the camera and the hand is independent of the user's position.

5.2.2.1 Application Walkthrough

Users adjust global, continuous values, like the volume and seek, with linear slides on the fingers. With a fist tap, they play and pause the video. When the video is paused, users can use finger taps, which are spatially mapped to the screen to select an area of the image for a context search. With rotational slides on the same segment, they can fine tune the selected area. With a tap and flap gesture inwards users flip to the search view and trigger an image search on the selection. They scroll through the results by performing thumb slides and return to the video with a tap and flap outwards.

4 https://www.python.org
5 https://www.raspberrypi.org
6 https://www.raspberrypi.org/products/camera-module-v2/
7 https://www.raspberrypi.org/documentation/usage/camera/python/README.md
8 https://python-pillow.org

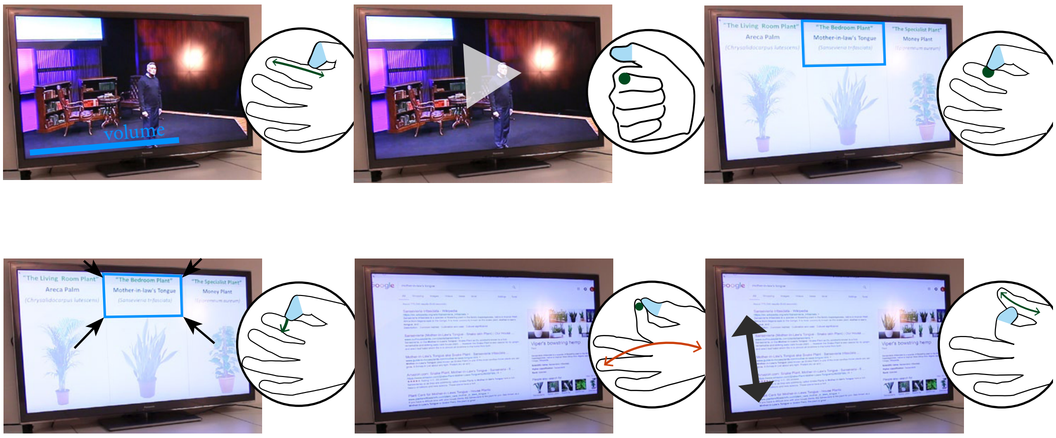

Figure 5.4: Users can interact with a smart television with (a) linear thumb-to-finger slides for important values like the volume, (b) fist tap to play/pause, (c) finger taps to select an area of the image, (d) rotational slides to alter the selection, (e) tap and flap to trigger a context search on the selection and (f) linear finger-to-thumb slide to scroll through the search results.

5.2.2.2 Implementation Details

The video player is implemented in Processing 3[9]. It runs on any desktop machine that supports the Java virtual machine. The image can be transferred to a television or projector. We switch between playback and search mode with a state system. The video is embedded and controlled with the Processing video library[10].

9 https://processing.org
10 https://processing.org:8443/reference/libraries/video/index.html

6 Limitations and Future Work

We decided to design subtle gestures, performable without any arm movements only within the hand. This is less tiring for the arm and usable in a discreet and private fashion. However, in-air movements and the orientation of the hand can communicate additional information. In certain scenarios, this could be interesting, e.g. in gaming or VR applications where engagement ranks higher than privacy. The in-air position and rotation of the hand are already given by the hand tracker. Future work could investigate how this could be used in combination with the hand pose.

This thesis focuses on finger-to-finger touch as well as combinations of it with the pose or movement of the fingers happening at the same time. A further step could be to also consider in-air movements of the fingers before, after, or in between the touch events. Future work could explore this addition. Especially use cases and mappings for these finger in-air movements could be interesting to investigate and how to prevent accidental input due to the natural movement of the fingers.

With the user studies, we presented findings on the subjective comfort and workload of the gestures. After this validation of the user perspective on the gestures, unbiased by a specific system or application, future work could test the usability of the gestures while interacting with a system. For instance, this could be a pointing task either with the tracker to evaluate the usability of the whole system or with an even more fine-granular tool to test the limits of the gestures. The spatial accuracy of tap buttons along the finger has been investigated by DigitSpace [13]. It remains to be tested how many rotational sliders and buttons distributed around the finger can be distinguished by users and how accurate fingers can touch on the thumb.

Concerning the implementation of the applications with the help of the hand tracker, there remain some limits in the recognition pipeline. So far the hand tracking requires so much computing power that it cannot be installed on small mobile computers. Besides optimization of the computation requirements, a further solution could be to connect the depth camera to a small mobile computer that runs the camera driver and sends the depth images wirelessly to a cloud for computation. However, this comes with the drawback of additional latency for transferring data, slowing down reaction time. A further improvement remaining for future work would be to augment the tracker with an automatically adapting hand model for new users. So far this is a step that has to be done manually for each user in order to get the best possible performance of the tracker.

7 Conclusion

We presented an extended design space for finger-to-finger hand gestures. It makes a fuller use of the dexterity of the human hand by including thumb-to-finger and finger-to-thumb touch while considering the pose of the hand. The touch can be performed on multiple sides of the fingers. The gestures are discrete, like a tap, or continuous, like a sliding motion. We investigated the properties of the gesture space. It can be described by three gesture primitives, the touch initiator, the touch action and the finger flexion. These primitives lead to a 14 classes structure of the gesture space.

From the presented gesture space we derived a concrete set of 7 diverse gesture groups. The selection was guided by several principles. Among these were the aim to support a variety of mappings, user preferences and capabilities from literature. These support the ease of use and learnability of the user interfaces. Two gesture groups map to activating one or multiple discrete elements, four groups map to setting one or multiple continuous values and one combines both. We proposed how they could be used in real applications.

In a pre-study, we evaluated the comfort of the gesture set. We found that most gestures were rated comfortable or very comfortable by users. Following up on the pre-study we conducted an in-depth study on the perceived physical demand, mental demand, and performance of the gesture space. The results show the following trends. It is easier for fingers close to the thumb to initiate a finger-to-thumb touch. Touching finger sides facing away from the touch initiator is more demanding. The inner segment of the thumb is more difficult to touch than the outer one. No significant differences could be found between the workload of the gesture during walking and sitting and between touch actions in the comfortable regions.

As a technical enabler for the gestures recognition, we used a hand tracker, which works with a depth camera mounted on the body. We evaluated the technical novelties of the hand tracker and used it to implement two interactive user interfaces: a photo application for a smartwatch and a video player. One demonstrated the usability in a mobile scenario, the other one in a stationary setting.

BIBLIOGRAPHY

[1] Edwin Chan, Teddy Seyed, Wolfgang Stuerzlinger, Xing-Dong Yang, and Frank Maurer. 2016. User Elicitation on Single-hand Microgestures. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 3403–3414. DOI:http://dx.doi.org/10.1145/2858036.2858589

[2] Liwei Chan, Yi-Ling Chen, Chi-Hao Hsieh, Rong-Hao Liang, and Bing-Yu Chen. 2015. CyclopsRing: Enabling Whole-Hand and Context-Aware Interactions Through a Fisheye Ring. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 549–556. DOI:http://dx.doi.org/10.1145/2807442.2807450

[3] Liwei Chan, Rong-Hao Liang, Ming-Chang Tsai, Kai-Yin Cheng, Chao-Huai Su, Mike Y. Chen, Wen-Huang Cheng, and Bing-Yu Chen. 2013. FingerPad: Private and Subtle Interaction Using Fingertips. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13)*. ACM, New York, NY, USA, 255–260. DOI:http://dx.doi.org/10.1145/2501988.2502016

[4] Andrea Colaço, Ahmed Kirmani, Hye Soo Yang, Nan-Wei Gong, Chris Schmandt, and Vivek K. Goyal. 2013. Mime: Compact, Low Power 3D Gesture Sensing for Interaction with Head Mounted Displays. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13)*. ACM, New York, NY, USA, 227–236. DOI:http://dx.doi.org/10.1145/2501988.2502042

[5] Artem Dementyev and Joseph A. Paradiso. 2014. WristFlex: Low-power Gesture Input with Wrist-worn Pressure Sensors. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 161–166. DOI:http://dx.doi.org/10.1145/2642918.2647396

[6] Niloofar Dezfuli, Mohammadreza Khalilbeigi, Jochen Huber, Florian Müller, and Max Mühlhäuser. 2012. PalmRC: Imaginary Palm-based Remote Control for Eyes-free Television Interaction. In *Proceedings of the 10th European Conference on Interactive TV and Video (EuroITV '12)*. ACM, New York, NY, USA, 27–34. DOI:http://dx.doi.org/10.1145/2325616.2325623

[7] George ElKoura and Karan Singh. 2003. Handrix: animating the human hand. In *Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation*. Eurographics Association, 110–119.

[8] Sean Gustafson, Christian Holz, and Patrick Baudisch. 2011. Imaginary Phone: Learning Imaginary Interfaces by Transferring Spatial Memory from a Familiar Device. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. ACM, New York, NY, USA, 283–292. DOI: http://dx.doi.org/10.1145/2047196.2047233

[9] Sean G. Gustafson, Bernhard Rabe, and Patrick M. Baudisch. 2013. Understanding Palm-based Imaginary Interfaces: The Role of Visual and Tactile Cues when Browsing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 889–898. DOI: http://dx.doi.org/10.1145/2470654.2466114

[10] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage Publications Sage CA: Los Angeles, CA, 904–908.

[11] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology* 52 (1988), 139–183.

[12] Yi-Ta Hsieh, Antti Jylhä, Valeria Orso, Luciano Gamberini, and Giulio Jacucci. 2016. Designing a Willing-to-Use-in-Public Hand Gestural Interaction Technique for Smart Glasses. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4203–4215. DOI:http://dx.doi.org/10.1145/2858036.2858436

[13] Da-Yuan Huang, Liwei Chan, Shuo Yang, Fan Wang, Rong-Hao Liang, De-Nian Yang, Yi-Ping Hung, and Bing-Yu Chen. 2016. DigitSpace: Designing Thumb-to-Fingers Touch Interfaces for One-Handed and Eyes-Free Interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1526–1537. DOI:http://dx.doi.org/10.1145/2858036.2858483

[14] Hsin-Liu (Cindy) Kao, Artem Dementyev, Joseph A. Paradiso, and Chris Schmandt. 2015. NailO: Fingernails As an Input Surface. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3015–3018. DOI:http://dx.doi.org/10.1145/2702123.2702572

[15] Maria Karam and others. 2005. A taxonomy of gestures in human computer interactions. (2005).

[16] David Kim, Otmar Hilliges, Shahram Izadi, Alex D. Butler, Jiawen Chen, Iason Oikonomidis, and Patrick Olivier. 2012. Digits: Freehand 3D Interactions Anywhere Using a Wrist-worn Gloveless Sensor. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. ACM, New York, NY, USA, 167–176. DOI:http://dx.doi.org/10.1145/2380116.2380139

[17] Eyal Krupka, Kfir Karmon, Noam Bloom, Daniel Freedman, Ilya Gurvich, Aviv Hurvitz, Ido Leichter, Yoni Smolin, Yuval Tzairi, Alon Vinnikov, and Aharon Bar-Hillel. 2017. Toward Realistic Hands Gesture Interface: Keeping It Simple for Developers and Machines. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1887–1898. DOI:http://dx.doi.org/10.1145/3025453.3025508

[18] Jaime Lien, Nicholas Gillian, M. Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. 2016. Soli: Ubiquitous Gesture Sensing with Millimeter Wave Radar. *ACM Trans. Graph.* 35, 4, Article 142 (July 2016), 19 pages. DOI:http://dx.doi.org/10.1145/2897824.2925953

[19] Rensis Likert. 1932. A technique for the measurement of attitudes. *Archives of psychology* (1932).

[20] Mingyu Liu, Mathieu Nancel, and Daniel Vogel. 2015. Gunslinger: Subtle Arms-down Mid-air Interaction. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 63–71. DOI:http://dx.doi.org/10.1145/2807442.2807489

[21] Christian Loclair, Sean Gustafson, and Patrick Baudisch. 2010. PinchWatch: a wearable device for one-handed microinteractions. In *Proc. MobileHCI*, Vol. 10.

[22] Manuel Prätorius, Dimitar Valkov, Ulrich Burgbacher, and Klaus Hinrichs. 2014. DigiTap: An Eyes-free VR/AR Symbolic Input Device. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology (VRST '14)*. ACM, New York, NY, USA, 9–18. DOI:http://dx.doi.org/10.1145/2671015.2671029

[23] T. Scott Saponas, Desney S. Tan, Dan Morris, Ravin Balakrishnan, Jim Turner, and James A. Landay. 2009. Enabling Always-available Input with Muscle-computer Interfaces. In *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology (UIST '09)*. ACM, New York, NY, USA, 167–176. DOI:http://dx.doi.org/10.1145/1622176.1622208

[24] Jie Song, Gábor Sörös, Fabrizio Pece, Sean Ryan Fanello, Shahram Izadi, Cem Keskin, and Otmar Hilliges. 2014. In-air Gestures Around Unmodified Mobile Devices. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 319–329. DOI:http://dx.doi.org/10.1145/2642918.2647373

[25] Srinath Sridhar, Anna Maria Feit, Christian Theobalt, and Antti Oulasvirta. 2015a. Investigating the Dexterity of Multi-Finger Input for Mid-Air Text Entry. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3643–3652. DOI:http://dx.doi.org/10.1145/2702123.2702136

[26] Srinath Sridhar, Franziska Mueller, Antti Oulasvirta, and Christian Theobalt. 2015b. Fast and robust hand tracking using detection-guided optimization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3213–3221.

[27] Srinath Sridhar, Antti Oulasvirta, and Christian Theobalt. 2013. Interactive Markerless Articulated Hand Motion Tracking using RGB and Depth Data. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 8. http://handtracker.mpi-inf.mpg.de/projects/handtracker_iccv2013/

[28] Emi Tamaki, Takashi Miyaki, and Jun Rekimoto. 2009. Brainy Hand: An Ear-worn Hand Gesture Interaction Device. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*. ACM, New York, NY, USA, 4255–4260. DOI:http://dx.doi.org/10.1145/1520340.1520649

[29] Koji Tsukada and Michiaki Yasumura. 2004. 2004), Ubi-Finger: a Simple Gesture Input Device for Mobile and Ubiquitous Environment, journal of Asian Information. *Science and Life (AISL* (2004), 111–120.

[30] Martin Weigel, Aditya Shekhar Nittala, Alex Olwal, and Jürgen Steimle. 2017. SkinMarks: Enabling Interactions on Body Landmarks Using Conformal Skin Electronics. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3095–3105. DOI: http://dx.doi.org/10.1145/3025453.3025704

[31] Martin Weigel and Jürgen Steimle. 2017. DeformWear: Deformation Input on Tiny Wearable Devices. 1, 2 (2017), 23.

[32] Katrin Wolf, Anja Naumann, Michael Rohs, and Jörg Müller. 2011. A taxonomy of microinteractions: Defining microgestures based on ergonomic and scenario-dependent requirements. In *IFIP Conference on Human-Computer Interaction*. Springer, 559–575.

[33] Sang Ho Yoon, Ke Huo, Vinh P. Nguyen, and Karthik Ramani. 2015. TIMMi: Finger-worn Textile Input Device with Multimodal Sensing in Mobile Interaction. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '15)*. ACM, New York, NY, USA, 269–272. DOI: http://dx.doi.org/10.1145/2677199.2680560