Institute for Visualization and Interactive Systems

University of Stuttgart
Universitätsstraße 38
D–70569 Stuttgart

Bachelorarbeit

Designing Gestures for Window Management on Large High-Resolution Displays

Marius Altmann

Course of Study: Informatik

Examiner: Prof. Dr. Albrecht Schmidt

Supervisor: Dipl.-Inf. Lars Lischke,
Pascal Knierim, M.Sc.,
Dr. Paweł W. Woźniak

Commenced: January 20, 2017

Completed: July 20, 2017

CR-Classification: H.5.m

Kurzfassung

Computer Monitore sind über die letzten Jahrzehnte ständig größer geworden und haben höhere Auflösungen erreicht. Große Bildschirme sind nun für den täglichen Gebrauch erschwinglich. Durch den technischen Fortschritt, werden in Zukunft Büros mit großen Bildschirmen und sogar Bildschirmwänden ausgestattet sein. Diese Bildschirme erzielen eine höhere Wirksamkeit bei der Arbeit mit großen Datensätzen oder der gleichzeitigen Verwendung von vielen Programmen. Standardeingabegeräte stufen die Vorteile zurück, die mit diesen Bildschirmen einhergehen. Um effizienter zu arbeiten, ist eine gewisse Distanz zu der Bildschirmwand erforderlich, welche durch Toucheingabe beeinträchtigt wird. Ein anderes Problem, das in dieser Arbeit angegangen wird, ist die Suche nach Mauszeiger, den die Benutzern auf den großen hochauflösenden Bildschirmen verlieren. Zusätzlich ist es anstrengend den Mauszeiger über die beträchtliche Ausdehnung des Bildschirme zu bewegen. Große Datenmengen und viele gleichzeitig geöffnete Programme müssen angeordnet werden, damit die Nutzer den Überblick behalten können. Um dies mit einer gewissen Distanz zum Bildschirm zu ermöglichen, erhebe und evaluiere ich in dieser Arbeit ein Freihandgestenset zur Verwaltung von Fenstern auf großen hochauflösenden Bildschirmen, mittels eines partizipativen Ansatzes. Die Erhebung, wie auch die Evaluierung des Gestensets, wird mit einer Benutzerstudie durchgeführt. Desweiteren werden für die Studie zur Erhebung der Gesten, Methoden verwendet, die den Einfluss von bereits bekannten Systemen eines Benutzers reduzieren. Diese Methoden werden innerhalb der Studie untersucht um ihren Effekt auf die Reduzierung der Befangenheit zu ermitteln. Die Ergebnisse der Evaluierung der Gesten zeigt, dass normale Endnutzer von Computersystemen in der Lage sind gute Freihandgesten zu entwickeln. Auch wenn das erhaltene Gestenset nicht mit der Maus konkurrieren kann, fanden es die Teilnehmer der Studie gut und die Angelegenheit des verlorenen Mauszeigers wurde gelöst.

Abstract

Computer monitors are permanently increasing in size and pixel density over the last decades. Large displays are affordable now, for daily purposes. Through this technical improvements offices will be provided with large displays and even display walls in the future. These displays improve the effectiveness when working with huge amounts of data or many applications simultaneously. Common input devices downgrade the improvements involved in these displays. To work more efficient a certain distance to the display wall is needed which is disrupted by touch input. Another problem that is tackled in this work, is the search for the mouse cursor which users lose on the large high-resolution displays. It also is exhausting to move the cursor over the display wall with its vast expanse. Huge data sets and many simultaneously opened applications need to be arranged so users can keep the overview. To make this possible with distance to the display wall, I elicit and evaluate a mid-air gesture set for window management on large high-resolution displays in this work through an participatory approach. Each, the elicitation and evaluation of the gesture set is performed by user studies. Further, for the elicitation study, methods are applied to reduce the influence of prior experiences of the users. These methods are investigated within the study to determine their effect on legacy bias reduction. The outcome of the evaluation of the gesture set shows that ordinary users of computer systems are good designers for mid-air gestures. Although the received gesture set cannot compete with the mouse, the study participants liked it and it solved the issue of the lost mouse cursor.

Contents

1 Introduction 15

2 Related Work 19
2.1 Multiple and Large Monitors 19
2.2 Gestures 21
2.3 Methodology 23

3 Designing and Testing Gestures 25
3.1 Gesture Elicitation Study 25
3.1.1 Participants 27
3.1.2 Apparatus 28
3.1.3 Procedure of the Study 29
3.1.4 Results 30
3.1.5 Discussion 37
3.2 Gesture Set Evaluation 38
3.2.1 Participants 38
3.2.2 Apparatus 39
3.2.3 Preparation of the Study 42
3.2.4 Procedure of the Study 44
3.2.5 Results 45
3.2.6 Discussion 50

4 Conclusion and Future Work 53

A Gesture Elicitation Study 57

B System Evaluation Study 61

Bibliography 69

List of Figures

3.1 Cards with the referents printed on them (referent cards). 26
3.2 A snapshot of the priming video for the gesture elicitation study. The user on the picture currently moves a window over the display wall. 27
3.3 The setup where the gesture elicitation study took place. 28
3.4 Shows the amount of participants that chose the respective referent. . . 31
3.5 The diagram shows the calculated priority for each referent dependent on the amount of participants that chose to not throw out the gesture and the according priority received by the card sorting method. Referents with lower values are rated more important. 32
3.6 Agreement Score: The diagram shows the agreement score of the chosen symbol for each referent. The agreement score over the whole gesture set is ≈0.104 34
3.7 The gesture set elicited from the user study. 35
3.8 Agreement score for the gesture set of the participants with priming and production. Overall agreement score: ≈0.145. 36
3.9 Agreement score for the gesture set of the participants without priming and production. Overall agreement score: ≈0.147. 36
3.10 Two pictures of a video that shows the gesture for maximize window. The hand has to move from fist to an open hand. 43
3.11 A participant moving a window with the gesture control during the study. 45
3.12 The average time the participants needed to solve the task lists. 46
3.13 The average NASA-TLX values for the single questions for the gesture control and the mouse. 47
3.14 The overall average NASA-TLX values for the gesture control and the mouse. 47
3.15 The SUS results for each task list with the according input methods. The y-axis shows the SUS points. 48
3.16 The diagram shows the average results of the SMEQ questionnaires for both input methods, for the single task lists and the combined outcome. The SMEQ value is displayed on the y-axis. 49

A.1 Gesture elicitation study consent form. 58
A.2 Gesture elicitation study background questionnaire. 59

A.3 Gesture elicitation study gestures list. 60

B.1 System evaluation study consent form. 62
B.2 System evaluation study background questionnaire. 63
B.3 System evaluation study task list 1. 64
B.4 System evaluation study task list 2. 65
B.5 System evaluation study SMEQ. 66
B.6 System evaluation study NASA-TLX. 67
B.7 System evaluation study SUS. 68

List of Tables

3.1 The results of the repeated measures anova for the referents on the SMEQ values and the elapsed task times. 50

List of Abbreviations

API application programming interface. 39

NASA-TLX NASA-Task Load Index. 43

OS operating system. 25

SDK software development kit. 39

SMEQ Subjective Mental Effort Question. 23

SUS System Usability Scale. 43

1 Introduction

Through the technical progress in the last decades, large high-resolution displays are now affordable for many applications [ABCD15; VB05]. In the beginnings of the personal home and office computers, there were nine inches large monochrome displays. Until in 1985 the first color display was released and the size increased to 14 inches. At that time they had resolutions like 560 px $\times$ 192 px. Since the end of the nineties flat screens became common and on 19 inches 1280 px $\times$ 1024 px can be displayed. It is so far not usual for home computers, but today, displays can easily outnumber 50 inches with resolutions of 7680 px $\times$ 4320 px [LMW+15b]. It is expected that large high-resolution displays will become commonplace in offices [ABCD15; LMW+16]. The size of the displays and their resolution still increase [LSF+16]. The advantages are the possibility to add contextual information and meaning through distributing information spatially [LHK+17; SSL+06]. However, interaction with such displays is still challenging [LSF+16]. It is still the standard that programs run in floating windows [WCH+16], although new systems often present themselves with so called apps which are only usable in full screen mode. So, what people see on monitors are mainly windows from active applications which are adjusted and arranged to their needs. Hence, users can have more application windows on their monitor at the same time, which are used to partition the available space on it [LMW+15a]. An emerging issue is how to arrange application windows and maintain clarity on the computer monitor. To make them easier and more intuitively accessible, I want users to take their hands for this purpose. Through this approach a further step towards gesture controls as common input device for computers is made. Further, this leads to a method that allows to interact with display walls from a distance, where a better overview is achieved, a problem that comes along with touch input. The commonly used tools therefor in office environments are the mouse and the keyboard [LSF+16; WCH+16]. Keyboard shortcuts only allow us to maximize and minimize the windows on the desktop and dock them to the sides and the corners. For more precise adjustments people have to use the mouse which can be unhandy when covering large distances. There are two possibilities to handle this problem: first, there is a lot of space on the desk to move the mouse over, and second, the mouse speed is increased. Both of these solutions restrict the user at their work, and results in increasing loss of time and effort for the user.

To approach the issue described above, this work is about finding a solution for gesture-based window management for the following options: selecting windows, moving windows, maximizing windows, minimizing windows, resizing windows, closing windows, docking windows to the left or right half and to a corner of the desktop, scrolling in a window up and down, showing the desktop, opening the task view and placing the mouse cursor to a certain point on the display wall. To receive a set of gestures that is likely to be applied by users, a gesture elicitation study has been executed. This study was conducted in an office environment, as displays in these settings will also increase. The setup consisted of six large high-resolution displays which represent the display wall, an arrangement as it could be standardized in some years. As the usage of windows for applications on computer monitors is the norm, users may be influenced by the current techniques and designs for window management. In this work the approach of Morris et al. [MDD+14] has been used to reduce legacy bias. In the concept of gesture elicitation from users, it cannot be ruled out that the results are influenced by previous acquainted technologies [MDD+14]. The participants use their experiences to design new gestures which are derived from known gestures of other interfaces, or they even try to transfer these gestures to the new technology. By the methods used in this work the participants is shown the consequence of an operation, what is called a referent, and afterwards they have to think of the operation that would cause the effect, called a symbol [MDD+14]. Morris et al. describe different ways to obtain these goals, there are three approaches: priming, production and partners.

In this work, the participants of the user study were divided into two groups to achieve this. On one part priming and production were used with a demo video to get to know how the technique works and what it is capable of. Additionally, these participants had to design five symbols for each referent. With this method, I aimed to go behind the gestures known from touch devices performed in the air, and lead the participants to invent new and innovative gestures. The other part neither uses priming nor production. From these findings, a gesture set was worked out. For recognizing the gestures, a program was implemented afterward the study which uses the Leap Motion controller to detect them. The implemented gesture set was then tested in another study and evaluated.

The results show, people are capable of inventing gesture sets that are learnable. And people can learn these, but as these techniques are new to many users, they need some time to achieve such a level that they can replace familiarized devices. Further, the gesture set is implemented and validated by a second user study.

This work presents the approach of inventing a gesture set, conducted through a user study, in which methods for legacy bias reduction, called priming and production, are applied and investigated.

In this work, the related terms screen, display and (computer) monitor are used interchangeably.

Structure

This thesis is structured as follows:

Chapter 2 – Related Work: This part describes earlier work on the subject of large displays, gesture controls and presents methods used in this work.

Chapter 3 – Designing and Testing Gestures: Here the conducted user studies are introduced along with their results and the implemented gesture control.

Chapter 4 – Conclusion and Future Work The final chapter summarizes the depicted work and includes a description of future prospects.

2 Related Work

This chapter gives an overview of previous work, associated with this work and its contents.

2.1 Multiple and Large Monitors

The monitor of the computer is the main output interface for interaction with the system. Usually, all information that comes from the computer to the user goes through the monitor. And mainly all the information users give to the computer through mouse and keyboard input is displayed on the monitor. This means that the monitor is probably the most important interface between the computer and the users.

There are many studies on multiple monitor setups which evaluate whether users are more productive with a design that uses more than one monitor and what the additional monitors are used for. These studies try to determine whether it is useful to have more than one monitor. For this purpose Grudin conducted a field study along with users of multiple monitor setups [Gru01]. As a result they perceive how people use additional monitors, namely for secondary things that do not need to be in the main focus at the moment. From this approach it is inferable how people use the six monitors of our setup, which ones they would use as main working area and which ones are seen as additional monitors.

Owens et al. probed different monitor configurations and how pleasant it is working on these for standard office tasks. They used four different monitor configurations and placed 60 participants in front of them to solve common office tasks. The configurations were: one 17 inch monitor, two 17 inch monitors, one 22 inch monitor and two 22 inch monitors. The results of their study show that the different configurations have no influence on the efficiency of solving the tasks. But the users prefer more display space. So the dual 22 inch monitors configuration was favored by most of the participants, followed by the dual 17 inch monitor configuration [OTN+12]. According to this I assume that more displays and larger displays are preferred by most people, which leads to the configuration used in this work.

Owens et al. did not show that there is a benefit by using two monitors compared to one monitor, apart from that users subjectively prefer two monitors [OTN+12]. The results of Kang and Stasko show exactly this fact. In "Lightweight task/application performance using single versus multiple monitors: a comparative study" [KS08] they present a trip-planning task executed by 28 participants who solved the task on a one monitor setup and a two monitors setup. Their results show that people on a two monitor setup experience less workload and finish the task faster. Additionally, they received the same feedback from the participants as Owens et al., that they subjectively prefer the two monitor setup [KS08; OTN+12].

In their work "Physically large displays improve performance on spatial tasks" [TGSP06] Tan et al. present their experiments to analyze the performance of users working on a regular desktop monitor compared to users working on a large wall display. They realized two experiments that cover the subject of mental rotation tasks and two experiments on 3D navigation and mental map formation and memory. The outcome of their studies show that a large display helps the participants on the tasks and offers them a higher level of immersion. Large displays make it easier for users to get a better imagination [TGSP06]. For this work, the results shown, are acknowledged and the setup is extended as six large high-resolution displays are used in one office workplace.

Another point is the arrangement of the monitors if people have several large high-resolution displays. Lischke et al. investigated how people order such monitors for office tasks, they received 19 different arrangements for four screens, where the so called screen band, was the most important one [LMW+16]. This arrangement shows the four monitors in portrait format arranged in a curve around the user and his desk. For this study a similar arrangement is used. The monitors are also in portrait format and arranged in a curve, next to each other, but instead of four, six monitors are used.

Large displays and multiple monitor workplaces are often used in control rooms. In "Understanding Work in Public Transport Management Control Rooms" Wozniak et al. provide insights in a control room of a public transport corporation, the presented setup has several workplaces with multiple monitors and a power wall [WLM+17]. They suggest touchscreens, so the employees could work more efficiently. These suggestions could be enhanced, one could probe if it would be practical to even handle the power wall with a gesture control.

The issue with the mouse on large high-resolution displays, already mentioned in the introduction, has a longer background [RCB+05]. In many offices it is common to use more than only one monitor at a workplace. So Benko and Feiner developed a technique which simulates one mouse cursor for every monitor by only using one physical mouse device [BF05]. The mouse cursor hereby is warped between the monitors which significantly reduces the distance the physical mouse has to be moved by the user. This technique was tested in a study with 8 participants where all of them suggested to rather use the Multi-Monitor Mouse approach compared to their regular mouse. It is shown that users are likely to use multiple monitors, controlled by only one input device. For this work I adopt these results, as a display wall consisting of six single displays is used, and the input for window management is provided by one controller.

It is assumed that Fitts' Law ([Fit54]) also holds on large high-resolution displays [KBSR07; VB05]. Fitts' Law returns a value that describes how long it takes to hit a specific target when your pointer has a certain distance to that point. Later the equation has been modified by MacKenzie [Mac92]. It is shown in Equation (2.1).

Fitts' Law:

$$MT = a + b * \log_2 \left(1 + \frac{D}{W}\right) \tag{2.1}$$

The elements in Equation (2.1) are the following:

- MT is the movement time a user needs to point on a specified aim.
- a, b are device-dependent variables.
- D is the distance the cursor has to be moved.
- W describes the width of the target.

The display wall used for this work is about four meters wide, so the distance one has to move the cursor to hit a certain position can be very large. With the approach of a gesture control for window management this work aims to make this easier and reduce the effort.

2.2 Gestures

There are already several approaches for mid-air gesture controls on large display walls [WJ16], although this technology is not commonly used in general office environments and not widespread these days.

A more common method is the interaction with large displays through touch gestures [JJBH15]. Jakobsen et al. investigate when users select which input method. Their results show, users interact more accurate and on average faster with touch, but they tend to use mid-air gestures when they need more overview [JJBH15]. The here presented approach aims for a mid-air gesture control that is easy to use. To achieve this, a gesture elicitation study is conducted, so a good overview of the data and easy to use mid-air gestures can be combined.

It is still challenging using mid-air gestures as input mechanism, as people do not use them regularly [ACTK15]. For this reason good mid-air gestures first must be searched, found and evaluated. The approach of Ackad et al. is to install a large interactive public display with that passers-by can interact. From this setup they elicit gestures that can be used in such environments. For their project it is important to find gestures that are quick to learn, easy to perform, reliably recognized and comfortable to be used by passers-by in a public area. These points are what constitutes good gestures [OF13], so, I adopt these conditions for the elicited gesture set.

Early prototypes often use gestures designed by the system developers [WMW09]. This is useful for first tests but for marketable products it is mandatory to adjust the input gestures to users' needs. It is time to let people who do not think about technical possibilities and restrictions (users) design the gestures. Wobbrock, Morris, and Wilson conducted a user study with 20 participants to elicit a gesture set for tabletop surfaces. The goal was to get the gesture set from regular users without experience in designing such technologies. In their study they showed the participants the effect on a tabletop surface and then asked them to invent a gesture that causes the shown effect. Each participant invented a 1-hand and a 2-hand gesture for each of the 27 effects and rated that gesture on a 7-point Likert scale for goodness and ease. So they received 1080 gestures from which they designed a gesture set. The outcome showed that users prefer 1-hand gestures and they do not care much about the number of fingers used. To calculate the agreement score of their gesture set they used a calculation rule from Wobbrock et al. [WARM05] which is shown in Equation (2.2).

$$A = \frac{\sum\limits_{r \in R} \sum\limits_{P_i \subseteq P_r} \left(\frac{|P_i|}{|P_r|} \right)^2}{|R|} \tag{2.2}$$

The elements in Equation (2.2) are the following:

- r represents a referent from the set over all referents R
- P_r represents the set of suggested gestures for r
- P_i represents the subset of identical gestures within P_r
- A is the agreement score which range is $[|P_r|^{-1}, 1]$

At first the score for each suggested symbol for a referent r is calculated, this happens in the inner sum. P_r stays the same for one referent, hence it represents the amount of suggested gestures for r. $|P_i|$ presents the amount of participants who recommend this gesture for the referent r. At this point $|P_i|$ is divided by $|P_r|$, this is the part where a higher score for gestures that are suggested more often is achieved. Then the square is calculated. This is done for every suggested symbol for the referent r and the results are added up. For every referent $r \in R$ the score needs to be calculated. They are also added up and then divided by the amount of referents $|R|$. This calculation method leads to the final result which describes the agreement score for a gesture set.

2.3 Methodology

Since people mostly only know devices controlled by mid-air gesture input from movies [FGVT15; MWW10; NSMG03], people are not used to it and have no familiar motions for such a control. Instead, they are used to input methods from touch interfaces, and also mouse and keyboard input. For this work, I am looking for a mid-air gesture based input method. There is a legitimate concern about getting many mid-air gestures derived from touch gestures, as there is an influence of earlier experiences [NDL+09]. This issue is tackled by the usage of methods that reduce the influence of the already famous movement patterns. Fortunately, Morris et al. [MDD+14] investigated three methods to decrement the influence of already known motions for gesture elicitation studies. These three methods aim to reduce legacy bias with regard to gesture elicitation, the techniques are called: priming, production and partner [MDD+14]. For the gesture elicitation study in this work two of the three methods were used, the third method, called partner, was not usable in the study setup because every participant had to take part in the study alone, to gain more gestures from different people and not have them influenced by each other. Nevertheless, the other two methods, priming and production, were utilized to prevent the participants from only giving the obvious gestures. The priming method works as follows: the participants perform physical exercises or watch another person performing some related tasks, it is used to improve the creativity of the participants [MDD+14]. To reduce legacy bias through production the participant has to consider several symbols for each referent. This can be achieved either by telling the participant a minimum number of gestures they have to make up or by provoking them to come up with another one until the necessary number is reached [MDD+14].

In the gesture elicitation study I instructed the participants to order the functions according to their importance. Basis for this, was the physical card sorting approach by Hudson, this work describes an approach to sort objects by categories [Hud13]. The participants have to group cards where the objects are printed on. Through this approach the order of the cards is not predetermined, because the cards are given to the participants in an unsorted order. For this work there was no need to have the objects sorted in groups, but to get a priority order of how important the participants think the referents are. So I picked up the idea of using a single card for each referent and pass all cards at once to the participant so they can get an overview of them and arrange them by their subjective feeling.

Sauro and Dumas present three different one-question questionnaires [SD09]. To be able to differentiate the effort between the usage of the gesture set and the mouse I provided a scale, the participants could fill in their demands. The three different questionnaires are the here used Subjective Mental Effort Question (SMEQ), a Likert scale and an Usability Magnitude Estimation judgment [SD09]. The idea for the system evaluation study was to have a questionnaire that is very easy to understand and can quickly be filled in, because every participant had to fill in 60 of these post-task ratings. I decided to use the SMEQ as it has a very fine stepping, which is negligible for the participant (see Figure B.5). The SMEQ has a scale from 0 to 150 determined in steps of ten on the left side of a 15 cm high bar. The participant simply has to draw a line on the right side of the bar to record their effort. The value can later easily be read off with a ruler.

More display space decreases workload and is preferred by the users [KS08; OTN+12]. Through the possibility to arrange data more spatial they get a higher level of immersion and better imagination [TGSP06]. The display wall is arranged in a curve, as a user study proposes [LMW+16]. This work tackles the effort using a mouse on multiple display setups for window management [RCB+05]. Touch gestures as input method for large displays have been investigated and compared to mid-air gestures ([JJBH15]), through further distance to display walls the overview is increased. By the research on mid-air gestures through an elicitation study, this input method shall be adapted in a way that people rather prefer it compared to touch gestures. For this reason, it is important to aim for gestures that are quick to learn, easy to perform and to remember [ACTK15]. To achieve this, I ask the participants in the study directly to come up spontaneously with a gesture for a given function without any restrictions.

3 Designing and Testing Gestures

This chapter describes the approach how I came from the idea of making a gesture control for large high-resolution displays to the working system. The following chapters will explain the study in which the gestures were received and designed, then analyzed, implemented and tested in another study.

The aim of this work is to design a gesture control for the window management on large high-resolution displays. To achieve this, the first step is to define the different window management functions and to collect the gestures for these functions. For this work, the main window functions are those used by a standard version of the Windows 10 operating system (OS).

These include: minimize window, maximize window, close window, select window, move window, resize window, dock window (left/right or to a desktop corner), task view, show desktop function, scrolling in a window and, additionally, setting the mouse cursor to a specific position. These are standard functions which I assumed should be clear to everyone using a computer for office or development tasks.

3.1 Gesture Elicitation Study

This chapter explains the setup, the range of participants, the proceeding and the results of the study in which the gestures for the system were elicited.

Already, Wobbrock, Morris, and Wilson [WMW09] recognized that even three experts in human computer interaction are unfeasible of doing a better job than 20 participants when it comes to gesture invention. To take up these suggestion, in this user study 40 participants were consulted. Collecting the gestures through the participatory approach, enhances the possibility to receive gestures with higher learnability and better memorability [MDD+14].

The setup used for this study (see Section 3.1.2) describes an office how it could look like, when the development on large display continues as before [LMW+15a; LMW+16; LSF+16]. These display walls allow employees to arrange their content more spatially and get a better overview. As the setup represents an office workplace, the participants of the study have to sit down during the conduct, like in a normal office nowadays.

As mentioned above, I aimed for an order of the importance of the single referents. Therefor the approach of Hudson [Hud13] was used in an adapted form. So each of the referents was printed on a single card (see Figure 3.1) which I handed to the participants in an unsorted order. The participants sorted the cards and threw out the ones they would not need or never use in this setup. Through this approach I received a ranking from each participant. Out of these rankings the average was calculated and I obtained a sight on the most and least important rated referents. Another advantage offered by the card sorting was that the participants thought about the gestures beforehand and could ask questions during the sorting. So in case they did not know any of the functions they obtained an explanation so everything was be clear to them during the gesture elicitation.

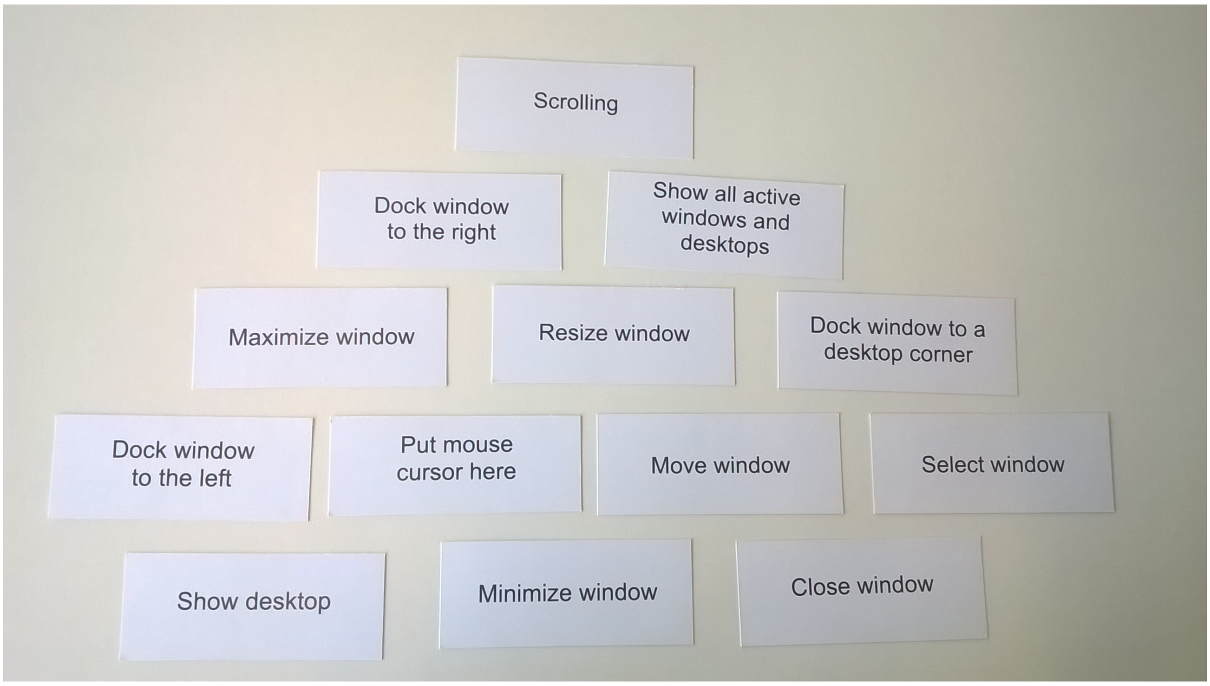


Figure 3.1: Cards with the referents printed on them (referent cards).

For those participants on whom the priming and production method was utilized I prepared a short video for the priming part. Priming can be applied in different ways. The participants can be shown gestures, either live by the experimenter of the study or on a beforehand recorded video. Additionally, the participants could have to imitate the shown gestures. Or they can have a look at gestures, that are more complex and developed by professionals [MDD+14]. I decided to record a video that shows a person using a gesture control on a display wall, it lasted 2:30 minutes. It was recorded at the same location as the study took place with a different orientation of the six displays (see Figure 3.2). As the participants should see the easy and standard gestures to become more creative for the actual study part the gestures on the video were taken from “Eliciting Mid-Air Gestures for Wall-Display Interaction” by Wittorf and Jakobsen [WJ16].

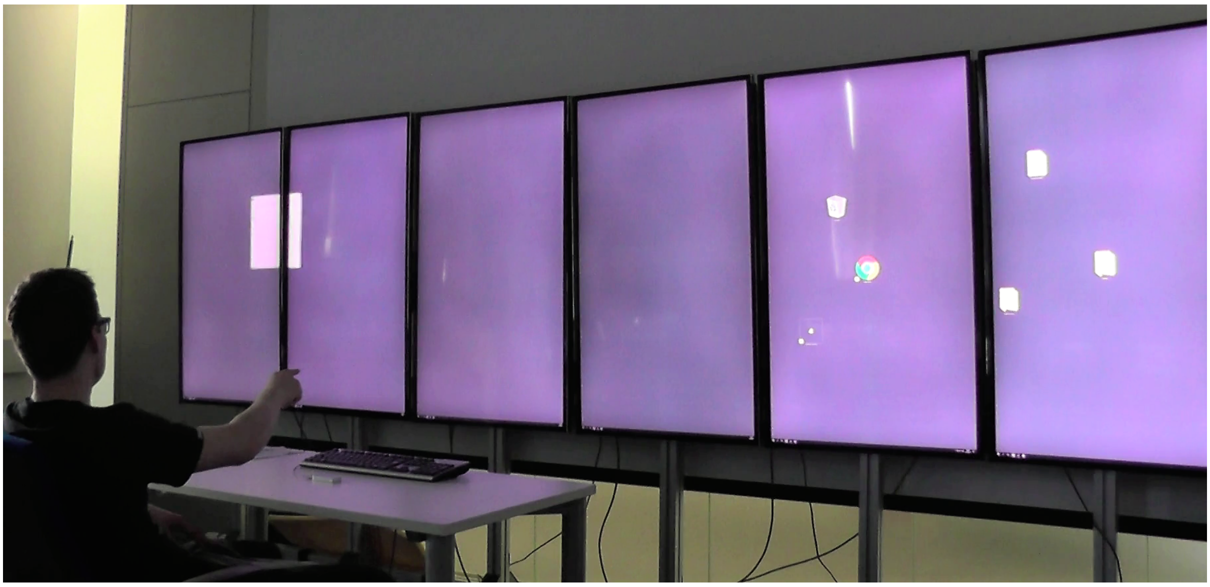


Figure 3.2: A snapshot of the priming video for the gesture elicitation study. The user on the picture currently moves a window over the display wall.

The method production is used to elicit more gestures per referent from each participant. Either the participant is told before the study that they have to create more gestures or they are prompted during the study to think up more. In this study, the participants that used this method, had to invent at least five different gestures per referent, as this was also the idea in the pilot study from Morris et al. [MDD+14]. The difference to their approach was, the participants in this study knew the minimum number before. The aim was not only to use these methods, but also to investigate them and to see the different outcomes in the two gesture sets and their agreement score. So if the methods would lead to benefits, they could be used on further studies.

3.1.1 Participants

The gesture elicitation study was conducted with 40 participants, seven female and 33 male. One participant worked as a research assistant, two were PhD students and one worked as a software developer. All other participants were students in the field of computer science or media informatics. They were aged between 19 and 31 years

$(M = 25.425, SD = 2.77)$. Among them, 34 were right-handed, five left-handed and one ambidextrous. The participants were mainly acquaintances of me or spontaneously met on the campus of the University of Stuttgart and asked to take part in the study.

3.1.2 Apparatus

Figure 3.3: The setup where the gesture elicitation study took place.

The apparatus used for the gesture elicitation study is displayed in Figure 3.3. The display wall consisted of six single 67.3 cm $\times$ 113.1 cm TX-50AXW804 high-resolution 50" 4K monitors from Panasonic. The monitors of the display wall were arranged in portrait format and each of them was adjusted to a resolution of 2160 px $\times$ 3840 px. During the study the participant was seated in the center of the display wall as it was arranged in shape of a semicircle around the user with the table in front of him. Between the participant and the display wall there was a distance of about 1.50 m. So from their position the participants had a very good sight on the whole display wall by simply turning their head. On the table there was only a mouse and a keyboard positioned as it would be in a normal computer workplace.

During the user studies I used two film cameras to record the audio and video data for subsequent analysis. One, the main camera, was positioned in the middle of the display wall directly beneath their lower edge. The second camera was positioned to the right of the table at which the participants sit during the study, to get two views for the case that a gesture could not be correctly detected from the front.

The following section presents the procedure of the user study.

3.1.3 Procedure of the Study

The following procedure was carried out for every participant who took part in the gesture elicitation study. When the participant came into the room where the study took place I welcomed and thanked the participant, introduced them to the setup and offered a seat at the table in front of the display wall. Before the actual study was initiated I handed the participant the consent form (see Figure A.1) to read and sign it. The document explained what data is collected during the study, what risks and benefits are related to it and the participant's rights, like that they have the possibility to withdraw their consent or discontinue the participation at any time without penalty. By signing the consent form the participant agreed that audio and video data is recorded during the study. Further information on the document were the approximated time the study would last and contact details to the researches involved in this project. The study lasted about 45 minutes per participant. Consecutively I asked the participant to fill out the background questionnaire (see Figure A.2). Through that form several data about the participant was collected, e.g. whether they are left- or right-handed, how many times per day they use a computer, how many monitors they typically use and what are their specifications and so on.

The next step was the study preparation where I explained the purpose of the study in detail to the participant, that it is about user-defined gestures for large high-resolution displays like the one introduced in the Section 3.1.2 (Figure 3.3) where the participant was during the study. After illustrating the setup and telling the participant how to use this computer and that they should envision this would be their workplace, we commenced with the introduction phase. The introduction phase was meant to familiarize the participant with the setup. To conclude this they had to start some standard applications, like Windows Media Player, Word and a browser and fulfill some prescribed assignments.

When the introduction phase was finished, I set out all the referent cards (Figure 3.1) at once on the table in front of the participant and exemplified the term window management by showing some useful standard options on a Windows Explorer window. The next task for the participant was to arrange the cards by their own opinion how important they think the described functions are for a user in this setup. Hereby the participant had the chance to ask further questions about the referents and get explanations if they did not know any of them.

I explained the participant that each referent needs a unique symbol and that it is important that they sort the referents. This is for the reason that they do not waste a symbol on a referent they rate less important which one they later on want to use for a referent they rate more important. The second reason why the participant had to sort the referents was to get a priority order of the functions. At this point in the study the participant also had the possibility to remove referents they do not need in an setup like the one presented in this study. So having all priority orders from the participants I could calculate the overall priority order of the referents. When the participant had sorted the list I transferred the order on a prepared sheet.

At this juncture it comes to the differentiation in the course of the study between the participants on whom priming and production was used and on whom it was not used. The next step in the study was the gesture elicitation. But before this was done the participants with priming and production had to watch the prepared short priming video to inspire their creativity. This was the priming part of the study.

When the participant started to perform the symbols for the single referents, the film cameras were switched on. Here the participant had to perform one gesture for each referent. I let some applications run on the computer so there were a few windows opened on the display wall but it was not possible to interact with them through gesture commands, so the participant had to envision the results of their actions. For those on whom priming and production was used had to perform at least five gestures for each referent and then choose one gesture they would favor. This is how production was accomplished for the study.

Finally, I started an open discussion with the participant. In that process the participant could give some feedback and propose improvements for the setup such as arranging the monitors in another way or adding further devices or tools to improve the workplace and make the handling with the gestures easier.

To conclude the study I thanked the participant for their attendance. The participation of the study was compensated with 5€.

3.1.4 Results

The main outcome of the study certainly is the received gesture set. To evaluate the gestures, I processed the data obtained in the user gesture elicitation study. Indeed every participant evinced their own gesture set. However, overall these single proposals one final result was formed. Overall from the 40 participants of the study I received 414 out of 520 possible gesture suggestions. So 106 gestures were not performed due to participants who removed referents they thought were not necessary for this setup (Figure 3.3). This means on average every participant decided to use 10.35 gestures and threw out 2.65 of them. Which referent has been chosen by how many participants is shown in Figure 3.4.

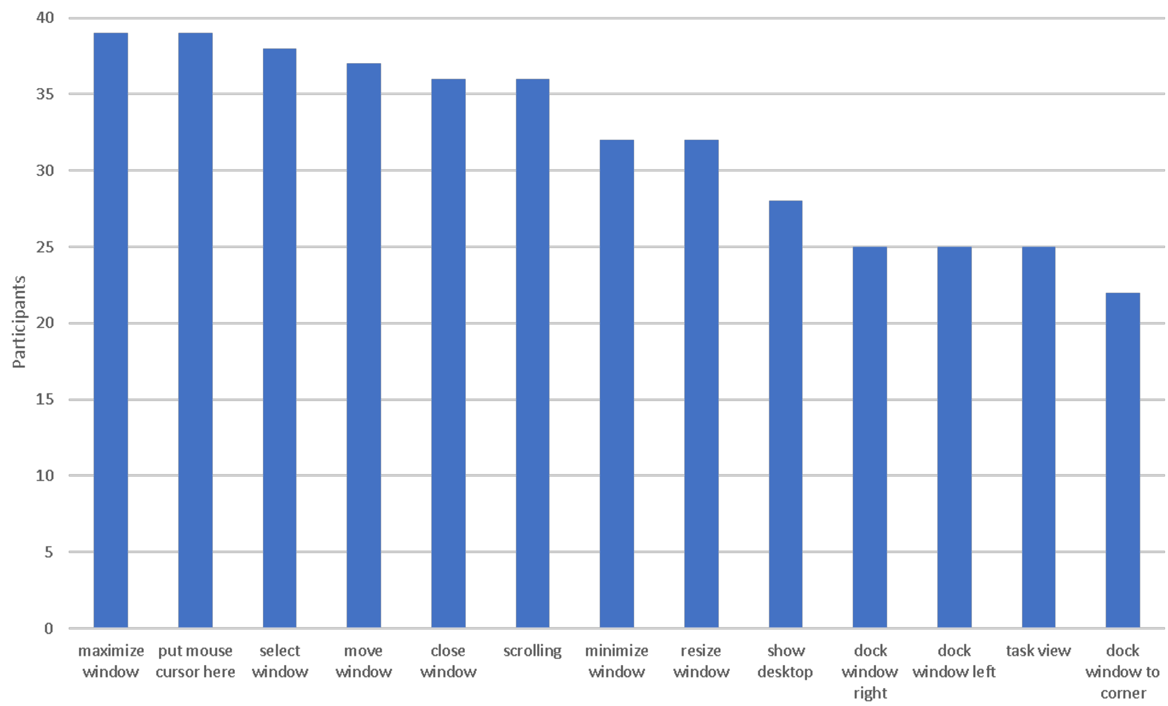


Figure 3.4: Shows the amount of participants that chose the respective referent.

The gesture set in the end anyway contains all the referents. But through the card sorting method and the possibility to leave referents out there is a priority on the importance of the gestures which is shown in Figure 3.5. As some referents were chosen by the same amount of participants, e.g. maximize window and put mouse cursor here, the card sorting method was important to receive a clear statement which of the referents is more important. So by every participant each referent had a priority from 1 (the most important referent) to the amount of referents they had chosen. When a participant threw out a referent, it was not ranked in their respective priority order. The consequence of this is that the referent is automatically rated worse. This is because for each referent, the single priorities of all participants were added and subsequently divided through the amount of participants that did not throw out this referent. So, the divisor was smaller when fewer participants used a referent, and so I received the mean value.

The chart in Figure 3.5 shows all referents with their according priority. The lower the value the more important was the priority rated for the introduced setup. So the function to set the mouse cursor to a certain position is the most important one, followed by selecting a window and scrolling. The least important function is docking a window

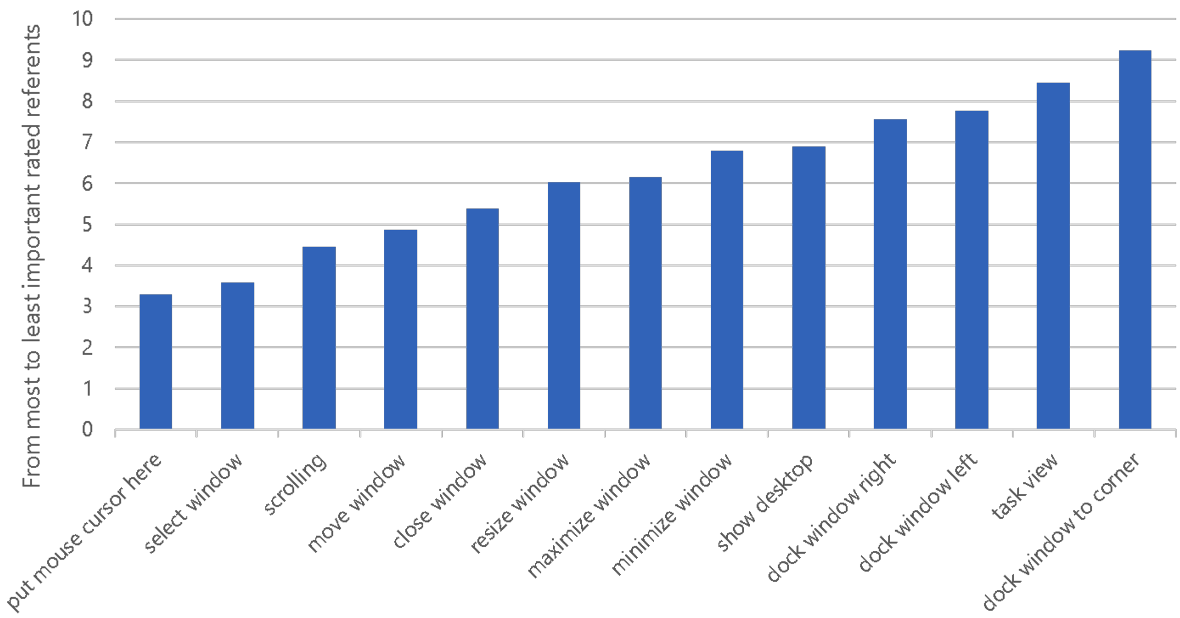


Figure 3.5: The diagram shows the calculated priority for each referent dependent on the amount of participants that chose to not throw out the gesture and the according priority received by the card sorting method. Referents with lower values are rated more important.

to a desktop corner. The minimum a referent could have reached was 1, this would have occurred if all the participants would have voted the same referent as the most important one. And the maximum a referent could have reached was 13, what would have occurred if all participants would have rated the same referent as the least important. This is not the case, so there is a rating for the referents between 3 and 10.

The result shows that the most important thing is setting the mouse to a new position ($\approx$3.28), this referent was chosen by 39 participants, followed by select window ($\approx$3.58), chosen by 38 participants, scrolling ($\approx$4.44), chosen by 36 and move window ($\approx$4.87), which 37 participants wanted to have in their gesture set. These are the functions most users usually perform with the mouse. The next function in the priority order, chosen by 36 participants, is close window ($\approx$5.389) what seems to be clear, as every window an user opened needs to be closed again. Presumably the most used functions to change the size of a window are resize window ($\approx$6.03), maximize window ($\approx$6.15) and minimize window ($\approx$6.78), but these functions are not used on every window, so it is understandable they also appear one after another in this ranking. Maximize window was under the top chosen referents, with 39 participants and resize and minimize window only were chosen by 32. The remaining ones are not used very often, at least not by the participants of the user study. However, 28 of the participants said they want to have a gesture for show desktop (≈6.89), even though some of them did not know the function but liked it when I explained it to them. For the functions dock window right (7.56) and dock window left (7.76) many of the participants said they would not need it in this setup, because the single monitors of the display wall are in portrait format and a high narrow window would not be comfortable. Still, 25 participants voted these referents to their gesture set. Also task view (8.44), chosen by 25 participants, only serves as a nice addition and is not a primary used function. Additionally, dock window to a desktop corner (≈9.23) was only chosen by 22 participants, with the argument that when someone has that much spatial freedom no one would need windows arranged in that way.

Also, I determined whether there is a significant difference between the two groups, the one I used priming and production on and the one without these methods for legacy bias reduction. As mentioned above 20 participants did use these methods and the other 20 did not. Until this point in the study the differences are not interesting, because they simply do not exist, hence to this point the study design was the same. But from here on the procedure of the study changed, the participants with priming and production had to watch a little video (the priming part) and make up more symbols for each referent (the production part). The video I presented to the participants showed a person in front of a display wall, operating it with a gesture control. Like Morris et al. mention, this was used to inspire the participants [MDD+14]. The same participants who watched the priming video, were encouraged to make up more than one gesture per referent. From them, at least five were expected. This method, called production, enhances the creativity of the participants and results in more variety within the gestures [MDD+14]. The other participants neither watched a video nor had to produce five gestures per referent, they were only allowed to suggest one gesture per referent.

The elicited gesture set for all 40 participants is shown in Figure 3.7. For all referents, except of one, the chosen symbol I could use was the one that had been chosen by most participants. Only for the referent select window I had to take the second most important symbol. This is because the top symbol for both, select window and put mouse cursor here, is the forefinger stab. 16 participants wanted to use this gesture for put mouse cursor here and only 12 wanted to use it for select window. The referent put mouse cursor here clearly won this symbol and for select window I had to choose the next most important one. This is forefinger and middle finger stab, which was only chosen by three participants. But it is the second most common answer and it is really similar to the first choice.

For each referent symbol connection I calculated an agreement score shown in Figure 3.6 with the rule from Wobbrock et al. [WARM05] and later the overall agreement score. The calculation rule is already pictured in Equation (2.2). The agreement score shows how well the chosen symbol goes with the referent. So if all participants would have chosen the same gesture for one referent it would fit perfectly and the agreement score for this referent symbol connection would be 1.0. Then one would say it fits 100%. The overall agreement score for the whole gesture set is simply the mean of all single agreement scores of the referents.

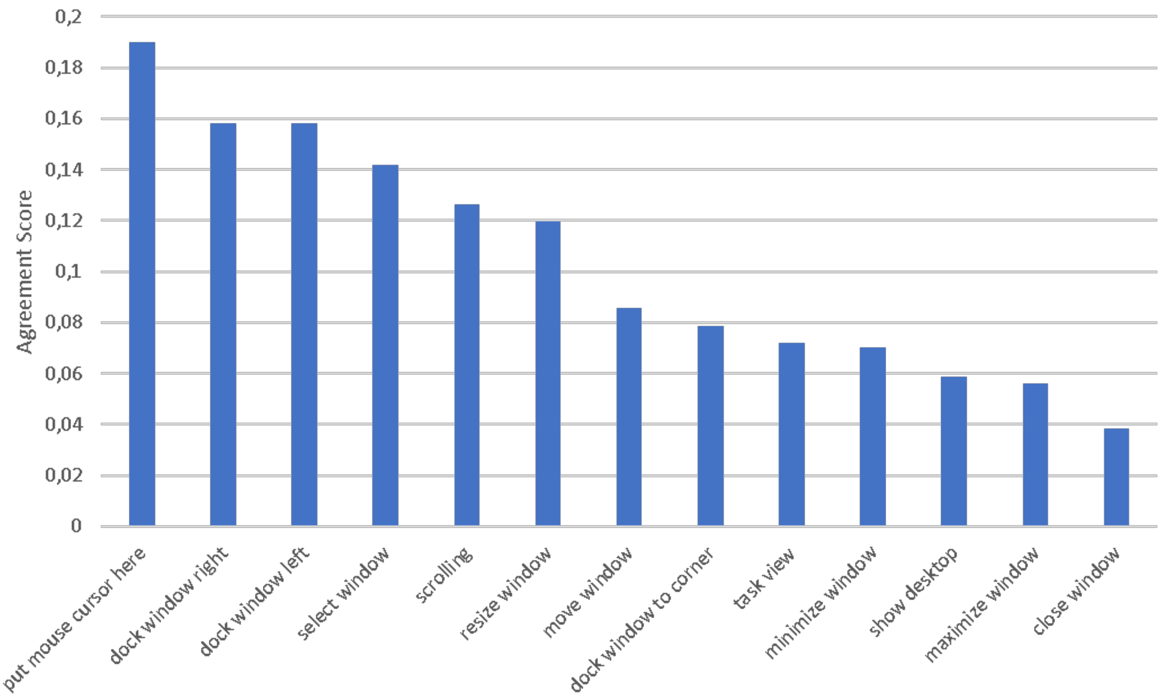


Figure 3.6: Agreement Score: The diagram shows the agreement score of the chosen symbol for each referent. The agreement score over the whole gesture set is $\approx$0.104

So the score for the gesture for put mouse cursor here is $\approx$0.190 while the one for select window is $\approx$0.142. To scroll in a window with this gesture set one has to point with the forefinger and the middle finger to the front and move the hand up or down, this gesture has an agreement score of $\approx$0.127. Moving a window is done by stabbing with the forefinger and the middle finger through a deeper imaginary layer and moving the hand to the new position in that layer, agreement score here is $\approx$0.086. A window can be closed by forming a horizontal fist and moving it to the right, agreement score: $\approx$0.039. One of the few referents in this setup that needs two hands to be executed is resize window. This is done by pointing with both forefingers on two corners of a window and then moving those corners with the appropriate forefinger. This gesture has an agreement score of $\approx$0.120. To maximize a window a user simply has to form a fist and then open the fist and spread all fingers, agreement score $\approx$0.056. The gesture for minimizing a window seems to be completely different, as therefor a user has to move and wave his hand down, the agreement score for this gesture is $\approx$0.070. To call the

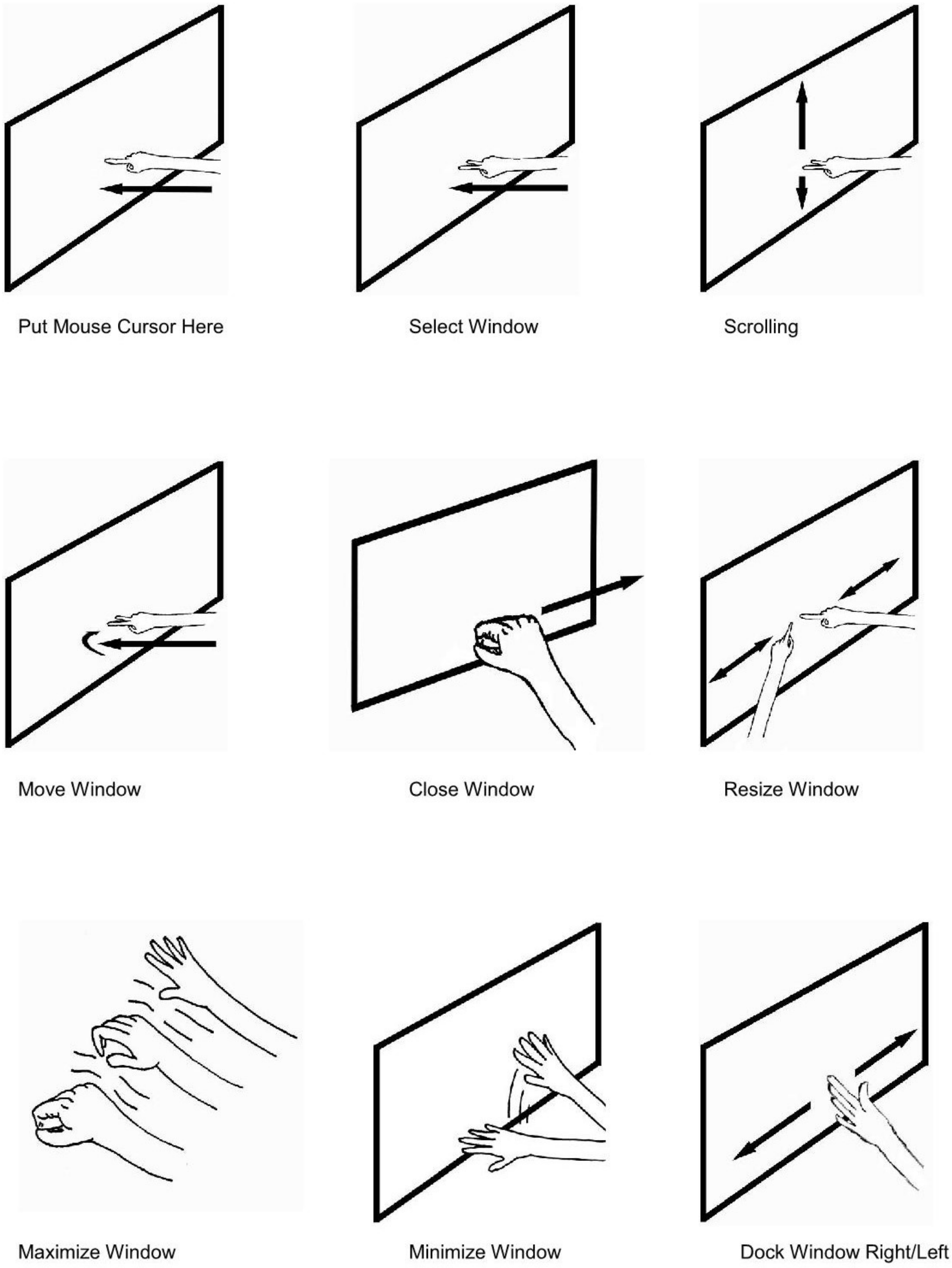


Figure 3.7: The gesture set elicited from the user study.

show desktop function both horizontal palms have to be moved down, agreement score: ≈0.059. To dock window to the right half of the desktop a vertical palm has to be moved to the right, agreement score: ≈0.158. The opposite action docks a window to the left desktop half, agreement score: ≈0.158. Opening the task view is done by moving both horizontal palms up and out, what means the right hand goes up and right and the left goes up and left, the agreement score here is: ≈0.072. To trigger the function to dock a

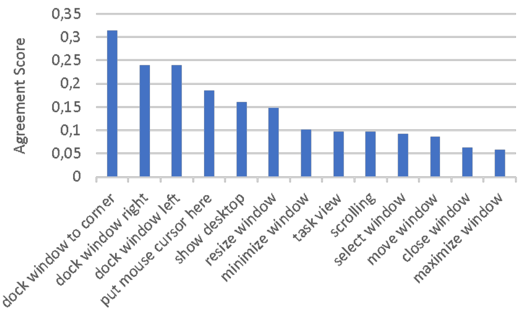


Figure 3.8: Agreement score for the gesture set of the participants with priming and production. Overall agreement score: ≈0.145.

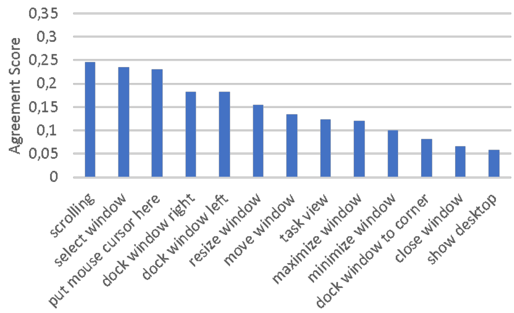


Figure 3.9: Agreement score for the gesture set of the participants without priming and production. Overall agreement score: ≈0.147.

window to a desktop corner a user has to move and wave one hand diagonally to the corner they want to dock the window, agreement score: ≈0.079. The overall agreement score for this gesture set is ≈0.104.

If we only take a look at the participants with priming and production we receive another result as outcome. Of course the same happens if we only look at the participants that did not use priming and production. In the following the different results of the two groups are demonstrated.

The single agreement scores for each referent are shown in Figure 3.8 for the priming and production group and in Figure 3.9 for the other group. The overall agreement scores hardly differ from each other as the participants with priming and production have a score of ≈0.145 and the others reached a score of ≈0.147. An result that could have been expected is that the agreement score of the priming and production group would be better. Fact is, that it is even slightly worse than the preserved result from the other group. As I have the agreement scores of both groups and compare them, which shows that they are equally good or bad, I come to the conclusion, that for this user study it did not make a meaningful difference, let alone profitable benefit. Although, it seems that the agreement scores for the referents that are higher rated on the priority order are higher in the group that used priming and production, it shows that the approach did not improve the result, because what in the end counts is the overall agreement score.

Additionally, on the background questionnaire which the participants filled out at the beginning of the study, they were asked about their habits on using their computers. So I received information about their monitor usage and arrangement and the time they spend on their computers per day. I also asked the participants about the specifications of their monitors, specially I asked for the size in inches and the resolution. Since most participants did not know the configuration of their monitors I cannot make a statement whether the users of large monitors with a high-resolution made better suggestions for the gestures because they felt more familiar with the setup.

3.1.5 Discussion

Some users removed the gestures to dock the windows on the left or right edge of the desktop because they were used to a tiling window managers and they said they would never use a Windows OS on this setup. As Wobbrock, Morris, and Wilson [WMW09] have already mentioned, users have a tendency to rather use one hand instead of two. In fact this can be confirmed by this user study. For some referents it seems to be more useful to work with widgets instead of mid-air gestures. Also Wobbrock, Morris, and Wilson [WMW09] already recognized this. In this study, this was only suggested by one participant (P19), but often participants mentioned gestures that are like touch gestures on their smartphones and pronounced this (e.g. P21).

I assume that the result of the priorities for the referents is strongly influenced by the habits of the participants for their standard office setup. As the mouse is mostly used more often than the window management functions, it reached the most important rank. Also scrolling is used more often than the adjustment of the windows, which most users only use when they open a new window and not many times afterwards. Scrolling instead is also often performed with the mouse wheel instead of the page up/page down or the arrow keys. To move or resize a window users also have to use the mouse, but these functions are not used that often in daily office tasks. Also the referent close window is placed between move and resize, it rather belongs to the group of maximize and minimize windows, hence these three options are on most windows displayed on the top right window corner in Windows OSs. And these three functions are often performed with the keyboard. On the first glance it might seem surprising that close window is rated on such a high rank, but actually it is pretty clear because usually users need to close every window they use.

All the other referents (show desktop, dock window right/left, task view and dock window to a desktop corner) are not often used as the participants told me during the study. Some of them did not even know these functions existed. This might be due to the fact that they are used to Linux systems. The cause that dock window to the right is ranked slightly before dock window to the left could concern to the fact that more participants were right-handed. In hindsight I assume that these two referents could have been counted as one and the difference between both of them can be neglected.

The overall agreement score of our gesture set seems to be a bit low with only 10.4%, but reflecting upon the study design this should be respectable. There were 40 participants, each with own ideas and a level of freedom of mid-air gestures which was not restricted by any guidelines. With this approach I ended up in 232 different gestures within an overall suggestion of 414 gestures for 13 referents.

In the course of this study, also the effect of methods to reduce legacy bias has been investigated. In "Reducing Legacy Bias in Gesture Elicitation Studies" Morris et al. propose to use their methods (production, priming and partner) to reduce the influence of familiar devices, within gesture elicitation studies [MDD+14]. The two methods (priming and production) used in this study, did not show any quantifiable enhancement for the received gesture set. The results show, that the gesture sets of the inspected groups are slightly different. And they show, that both groups, taken individually, reach a higher agreement score in their gesture than together, what shows, they had different ideas. Nevertheless, it has not been proved, that either using the methods nor not using the methods is better. Both of them reveal similar results in the agreement score. Hence, no statement can be made about the effect of the methods, so this has to be investigated in further studies.

3.2 Gesture Set Evaluation

This chapter describes how the gesture set I have received from the 40 participants of the gesture elicitation study was implemented, tested and evaluated. Therefore, the leap motion[1] controller was used. Of course the setup where the participants should try the control needed to be the same as the one where the gesture elicitation study took place, because the first study was designed in such a way that the gestures are suited for this setup. I invented a procedure for the study in which the participants had to pass through two similar sequences of tasks. Everyone of them should do some windows management tasks with the mouse and with the new invented gesture control. Within this study I tested whether the gestures that are invented by users are easy to remember and to perform and how hard is it to use these gestures compared to the same operations performed with a standard computer mouse.

3.2.1 Participants

In the system evaluation study 12 participants attended, one female and eleven male. One of them worked as an actuary, three of them were PhD students in the field of chemistry, one was a student of media informatics and seven were computer science

[1] www.leapmotion.com

students. The participants were aged between 23 and 28 years ($M = 25.5, SD = 1.58$). Nine of them were right-handed, two left-handed and one ambidextrous. Four of them said they had already used a leap motion controller before.

3.2.2 Apparatus

As the gestures were elicited in the setup described in Section 3.1.2 and they are actually intended for this setup, shown in Figure 3.3, the test phase also had to take place there. So, for this study I used the same assembly as for the gesture elicitation study. What has changed in the structure of the workplace of the participants is that on the table, additionally, to the mouse and the keyboard the leap motion controller is placed.

The next section describes the implementation of the gesture control what constitutes a very important part for this study.

Implementation

To implement the gesture control elicited from the first study, I chose the programming language C# because it is one of the languages supported by leap motion with a software development kit (SDK) and there is a very good way to access the Windows application programming interface (API). As the study should not prolong unnecessarily, because this could influence the participants in with regard to the level of frustration or temporal, physical and mental demand, I only implemented these 10 referents: put mouse cursor here, select window, move window, maximize window, minimize window, resize window, dock window left, dock window right, close window and scrolling.

Already mentioned above, the display wall is about four meters wide. The effective trackable field of the leap motion controller is between 25 and 600 millimeters. So I decided to match the leap motion coordinate system to the one of the display wall in a way that the boundaries and the origin of the systems are adapted to each other. Through this design approach it is not obvious for a user where on the monitor the current position of the pointer is. To tackle this issue, I created a little window which only shows a colored circle and is transparent apart from this and all operations fall through this window on the next control beneath it. This little window always moves it middle position, which also is the middle position of the colored circle, to the point the user's forefinger points to, relative to the coordinate system of the leap motion controller. Such a pointer was often suggested by the participants of the gesture elicitation study, especially if the coordinate system is only transferred and the displays do not show the actual position the finger is pointing to. The program does not differentiate whether an user was right- or left-handed, it responds to every forefinger. Also all the gestures which are performed with a single hand can be triggered by the left or the right hand. For the sake of completeness, it should be mentioned that the colored circle only shows up, when there is a hand recognized in the field of view of the leap motion controller. Furthermore, to trigger a function by a gesture, it must take place in the field of view of the controller, which means in this setup, above the controller.

The leap motion controller uses optical sensors along with infrared light to detect the hands. With the data received from these sensors it builds frames (`Controller.Frame`) which are forwarded to the `newFrameHandler`. In this function the frames can be further processed, additionally, the leap motion SDK holds the last 60 frames so they are accessible for gesture detection. So all the gesture recognition takes place here. If a gesture is detected the required frames are processed to the function that invokes the execution.

Put Mouse Cursor Here: The referent put mouse cursor here is performed by a single forefinger stab into the direction of the display wall. As a result of this the mouse cursor is set on the position of the center of the colored circle. The controller simply recognizes the movement of the forefinger tip to the front relatively and from every position. It then calls the function `SetMousePos` and passes it the position of the forefinger transformed to the coordinate system of the display wall by the function `GetRightIndexFingerCoords` or `GetLeftIndexFingerCoords`. `SetMousePos` gets a `System.Drawing.Point` and moves the mouse cursor to it. Hereby the program distinguishes whether only the forefinger moves or also the middle finger. When both fingers perform this movement select window is called.

Select Window: The referent select window is performed by a forefinger and middle finger stab into the direction of the display wall. As a result of this the window beneath the circle is selected and set as foreground window. The method works the same as for put mouse cursor here, but both, the forefinger and the middle finger tip, need to move to the front. The then called function gets the position of the forefinger and calls the window from that position through the `DllImport WindowFromPoint` which returns the windows `handle`. The imported function `SetForegroundWindow` uses this `handle` to set the focus on this window.

Scrolling: The referent scrolling is performed by pointing with the forefinger and the middle finger to the front and moving the whole hand up or down. As a result of this the window in which the mouse cursor currently is positioned scrolls up or down if it has a handle that is scrollable. This symbol is restricted to the area in front of the leap motion controller (`Pointable.TipPosition.z >= 0.0`), because instead it could come to conflicts with other gestures. In this case the controller simply polls the position of the tip positions of the forefinger and the middle finger and gets their pointing direction through the query `Pointable.Direction.z <= -0.9`. The `Vector` must be smaller, as it is pointing into the direction of the display wall, this is along the negative z-axis. If these conditions return true, the controller follows the vertical translation of the hand and activates the respective scroll function. This means if the `Hand.Translation().y` returns a positive value the respective document is scrolled upwards and if it returns a negative value it is scrolled downwards.

Move Window: The referent move window is performed by pointing with two fingers on a window and then moving the whole hand to move the window. The window that will be moved is always the currently foreground window, detected through the `DllImport GetForegroundWindow`. The function is activated by moving the tips of the forefinger and the middle finger behind the leap motion controller (`Pointable.TipPosition.z < -100.0`), then the active window's top left corner is moved to the center of the circle pointer, the window keeps its size and can be moved. It is released by moving the fingertips back to the front of the controller.

Close Window: The referent close window is performed by forming a horizontal fist and moving it to the right. The program then sends the Windows API close command (`SendMessage(handle, WM_SYSCOMMAND, SC_CLOSE, 0)`) to the active window handle which is the same as clicking on the 'X'-Button in the top right corner of the window or pressing `Alt+F4`. Of course the program has no chance to close itself. The leap motion SDK is not capable of recognizing a fist on its own. But there are several possibilities one can use to distinguish a fist from other hand forms. In this implementation I therefor used the property `Hand.SphereRadius <= 40.0`. If the controller recognizes a fist, the program checks the last ten frames whether the fist moved to the right (`Hand.Translation(Controller.Frame(10)).x >= 50.0`).

Resize Window: The referent resize window is performed by using both forefingers. If the leap motion controller recognizes the left hand and the right hand forefinger simultaneously it calls the function `ResizeWindow` and passes the Positions through `GetLeftIndexFingerCoords` and `GetRightIndexFingerCoords`. In this case the active window can be resized whereas the left hand's forefinger tip position transfers the position of the top left window corner and the right hand's forefinger tip position the bottom right corner position. Therefor another `DllImport` is used (`MoveWindow`).

Maximize Window: The referent maximize window is performed by first making a fist and subsequently spreading all fingers of the hand. For this gesture I used the `Hand.SphereRadius`-property again. If the program recognizes a widened hand (`Hand.SphereRadius > 100.0`) it checks the last 10 frames received from the leap motion controller and if it finds a fist in those frames it maximizes the active window on the respective monitor of the display wall with the `DllImport SendMessage(handle, WM_SYSCOMMAND, SC_MAXIMIZE, 0)`. Therefor the function from the Windows API is used to obtain the same effect as if one would click the appropriate button in the top right corner of the window.

Minimize Window: The referent minimize window is performed by holding a flat hand and moving it down towards the table. The leap motion controller hereby needs to recognize horizontal palm (`Hand.PalmNormal`) and a `SwipeGesture` with a main direction to the bottom. Therefor the function `GetSwipeDirection` analyzes the direction of the `SwipeGesture` (`SwipeGesture.Direction.y < -0.5`). If it does recognize such a gesture it calls the function `MinimizeWindow` which calls the Windows API as if one would click the according button in the top right corner (`DllImport SendMessage(handle, WM_SYSCOMMAND, SC_MINIMIZE, 0)`).

Dock Window Right/Left: The referents dock window right and dock window left are performed similarly as they cause a similar effect. They are activated by holding a vertical flat hand which is then moved to the according direction. To trigger the functions the leap motion controller hereby needs to recognize a vertical palm (`Hand.PalmNormal`) and a `SwipeGesture` to the right or to the left. Here again the function `GetSwipeDirection` analyzes the direction of the `SwipeGesture`. To dock the window to the right `SwipeGesture.Direction.x > 0.5` must be true, then the function `DockWindowRight` is called which triggers keyboard events, because there was no option available in the Windows API. So the program simulates keystrokes for the left Windows key (`keybd_event(VK_LWIN, 0, 0, 0)`) and the right arrow key (`keybd_event(VK_RIGHT, 0, 0, 0)`). To dock a window to the left the according events are triggered. The `SwipeGesture.Direction < -0.5` must be true, so the function `DockWindowLeft` is called. The keystrokes triggered here are the left Windows key (`keybd_event(VK_LWIN, 0, 0, 0)`) and the the left arrow key (`keybd_event(VK_LEFT, 0, 0, 0)`).

Show Desktop, Task View and Dock Window to Corner: The symbols for these referents were not implemented as their usage would have extended the evaluation study and the referents were not rated as very important by the participants of the gesture elicitation study.

The following section explains further necessary preparations for the study conduct.

3.2.3 Preparation of the Study

To show the participants how the gestures are used I prepared a single video for every referent. This was done so there is equality for everyone, and I do not introduce the gestures differently to every user. The videos showed a hand from the left in front of a white wall performing each gesture as displayed in Figure 3.10.

Also the tasks the participants had to solve needed to be prepared. I assumed that every referent should be used three times. As there is no method to open a window with the

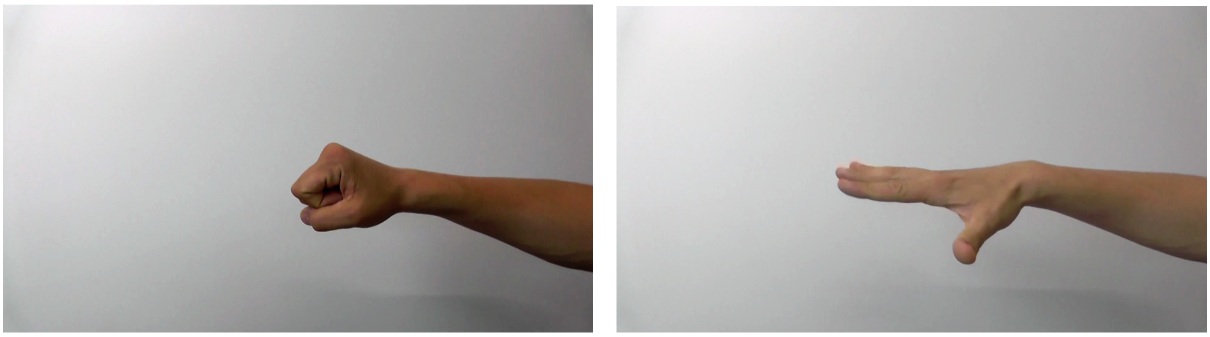


Figure 3.10: Two pictures of a video that shows the gesture for maximize window. The hand has to move from fist to an open hand.

gesture control and in order to prevent the participants to use another device during that part of the study, the referent select window had to be used six times. This is because the methods close window and minimize window were used three times, so six different windows were needed. These six windows needed to be set in the foreground to interact with them and this is done with the method select window. So I prepared two task lists, each of them contained 30 single tasks, because dock window left and dock window right were mixed up to one referent, as they only distinguish in the direction the hand needs to be moved. Every participant had to solve both of the task lists, one of them with the mouse and one with the leap motion and the implemented gesture control. The tasks the participants had to solve were designed for the options that are possible with the gesture control, so there were tasks like move a window from one monitor of the display wall to another one, maximize a window over one monitor or resize a window over two. The whole tasks are attached in Appendix B in Figures B.3 and B.4. To avoid influence of learning effects on the participants I used the latin square method for the order of the task lists and the order of mouse and gesture control input.

Furthermore, I needed methods to evaluate the different difficulties between the single tasks and to compare the mouse and gesture control input. From the participants I wanted to get a subjective assessment how challenging and difficult the single tasks are so I prepared a customized SMEQ [ZD85] for each task of every task list for all participants. To adjust the SMEQ, I omitted the issue of the mental effort and only asked the participants for the general effort (see Figure B.5). Along to that I gathered more general information about the used systems and how much effort the participants need to expend, so after finishing one task list with 30 tasks they had to fill in a NASA-Task Load Index (NASA-TLX) [HS88] and a System Usability Scale (SUS) [Bro86] questionnaire (see Figures B.6 and B.7).

I also clocked the time a participant needed to accomplish a given task and noted it in a table to the according task. These data should also provide insight in the ease of performance in later evaluations.

In case a participant forgot how a specific gesture was performed during the study they got the chance to review the video for this gesture. If this happened I made a note so this shows me later in the evaluation how good the users could remember which gestures and which were hard to remember.

3.2.4 Procedure of the Study

In this section I will provide an insight into how the study took place. The procedure of the study was uniform for every single participant.

In the very beginning of the study, when the participants entered the room I welcomed them and offered a chair to them in front of the display wall at the table. I thanked the participant for taking part in the study and told them the study is completely anonymous, that a video and voice recorder is on set and that the study will last around 60 minutes. Additionally, I told them they are allowed to withdraw their participation at any time or take a break as they wish.

Thereupon I handed the participant the consent form (see Figure B.1) to read and sign and the background questionnaire (see Figure B.2) to fill out. Subsequently I explained the purpose and the different stages of the study and introduced them to the setup. So they were told there was an earlier study in which a gesture set for window management on large high-resolution displays was elicited for the same setup which is used in this study. I explained the term window management and introduced them to the leap motion controller, as many of them had never seen this device before. Because most of the participants did not know the SMEQ, NASA-TLX and SUS questionnaires, I showed these to them and explained how they should fill them in.

The next part was the introduction phase. Here the participant watched the videos with the gestures for the referents and became acquainted with them by trying them until they felt good enough to use them. Only afterwards, when the participant said they are ready to start, I started the video camera and we began with the first task of the first task list. This works as follows, I told the participant what to do and they told me whether they still remembered the symbol for the required referent. If they did not remember the symbol they were allowed to watch the according video as many times as they wanted again, and I noted the times they watched it. When they were ready I gave them a sign to begin, simultaneously I started a stopwatch which I stopped when the task was finished. I noted the elapsed time and the participant filled in a SMEQ. This was repeated until the last task of the first task list was fulfilled after which I handed the participant the NASA-TLX and the SUS questionnaire. On that, the participant and I proceeded with the second task list and the second input method. Depending on whether the participant at the first time handled task list 1 or task list 2 with the gesture control or mouse input method I replaced the respective conditions. So every participant in the end had processed both task lists and used both input methods. In total three participants first used the gesture control for task list 1 then the mouse for task list 2, three used the mouse for task list 1 then the gesture control for task list 2, three used the gesture control for task list 2 then the mouse for task list 1 and three used the mouse for task list 2 then the gesture control for task list 1. In Figure 3.11 a participant performs the gesture to move a window, during the study.

Figure 3.11: A participant moving a window with the gesture control during the study.

After the actual study I conducted an open conversation with the participant about the gesture control. At this they had the possibility to give advices how it could be improved or changed.

3.2.5 Results

This section covers the outcome of the system evaluation study.

I compared the time the participants needed for the tasks. At this point it is worth mentioning that all participants were able to solve the complete task lists with both controls. The average time they needed to solve each task list with the respective input method is shown in Figure 3.12.

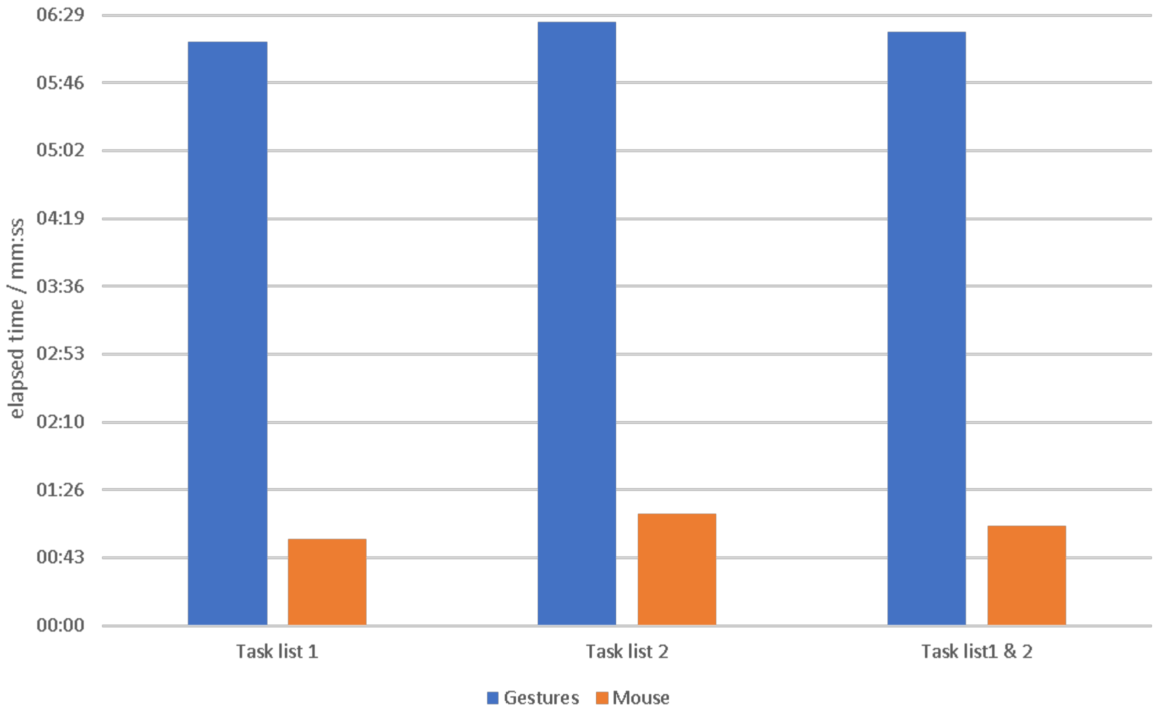


Figure 3.12: The average time the participants needed to solve the task lists.

Although the task lists included nearly the same exercises and the exact same amount of them, on average the participant needed slightly more time for task list 2 (see Figure B.4). The diagram clearly shows an indisputable difference between the time the participants needed to solve a task list with the mouse and the gesture control. On average the participants needed six times longer to fulfill the task lists with the gesture control than with the mouse. The time with the mouse amounts 01:03 min ($SD =$ 00:13 min) compared to 06:18 min ($SD =$ 02:01 min) with the gesture control. For task list 1 the average time with the mouse is 00:55 min ($SD =$ 00:09 min) and with the gesture control it is 06:12 min ($SD =$ 01.57 min). Processing task list 2 took the participants in general a bit longer, with the mouse the average time amounts 01:11 min ($SD =$ 00:11 min) and with the gesture control it amounts 06:24 min ($SD =$ 02:04 min)

From the NASA-TLX questionnaires, I received the results pictured in Figures 3.13 and 3.14. The diagram in Figure 3.13 shows the average values for the single points in the NASA-TLX questionnaire for the gesture control and the mouse. In Figure 3.14 the average values for the complete NASA-TLX questionnaires for the gesture control and the mouse are shown.

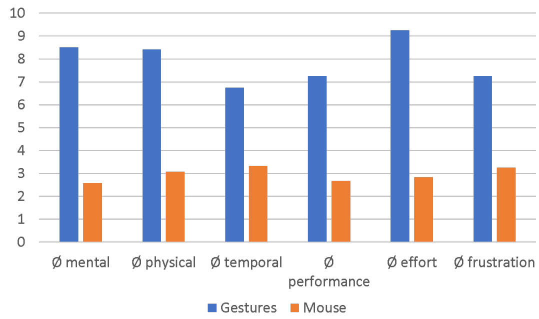


Figure 3.13: The average NASA-TLX values for the single questions for the gesture control and the mouse.

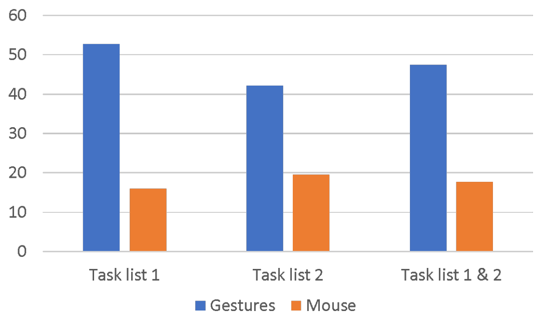


Figure 3.14: The overall average NASA-TLX values for the gesture control and the mouse.

Figure 3.14 determines that the demands for task list 1 are higher than for task list 2, although Figure 3.12 shows the participants needed more time for task list 2. The average mental demand of the gesture control for the two task lists is 8.5 ($SD = 4.07$), for the mouse it is only ≈2.58 ($SD = 0.86$). The average physical demand is with the gesture control is ≈8.42 ($SD = 5.00$), with the mouse it is ≈3.08 ($SD = 1.32$). For the temporal demand the average value for the gesture control is 6.75 ($SD = 3.35$), for the mouse it is ≈3.33 ($SD = 1.03$). The average for the performance with the gesture control is 7.25 ($SD = 2.65$), with the mouse it is ≈2.67 ($SD = 0.62$). The perceived effort the participants had is 9.25 ($SD = 4.30$) with the gesture control, whereas with the mouse they only had ≈2.83 ($SD = 0.90$). And the frustration with the gesture control was on average at 7.25 ($SD = 4.15$), and with the mouse at ≈3.25 ($SD = 1.30$).

The average overall NASA-TLX score for the gesture control is ≈52.67 ($SD = 20.27$) for task list 1 and ≈42.17 ($SD = 14.96$) for task list 2, and for both task lists together it is ≈47.42 ($SD = 18.57$). With the mouse the average overall NASA-TLX score for the tasks performed is 16 ($SD = 3.96$) for task list 1 and 19.5 ($SD = 3.86$) for task list 2. This results on average to 17.75 ($SD = 4.28$) for both task lists (see Figure 3.14).

Figure 3.15 shows the results of the SUS questionnaire for the mouse and the gesture control, each for task list 1, task list 2 and their average (task 1 and task 2).

The SUS value for the single task lists do not differentiate greatly within each input method (see Figure 3.15). For the gesture control, task list 1 has a value of 66.25 ($SD = 19.03$) whereas task list 2 has a value of 65 ($SD = 12.67$). Both task lists together result in a score of 65.625 ($SD = 16.17$). The mouse has better values, as for task list 1

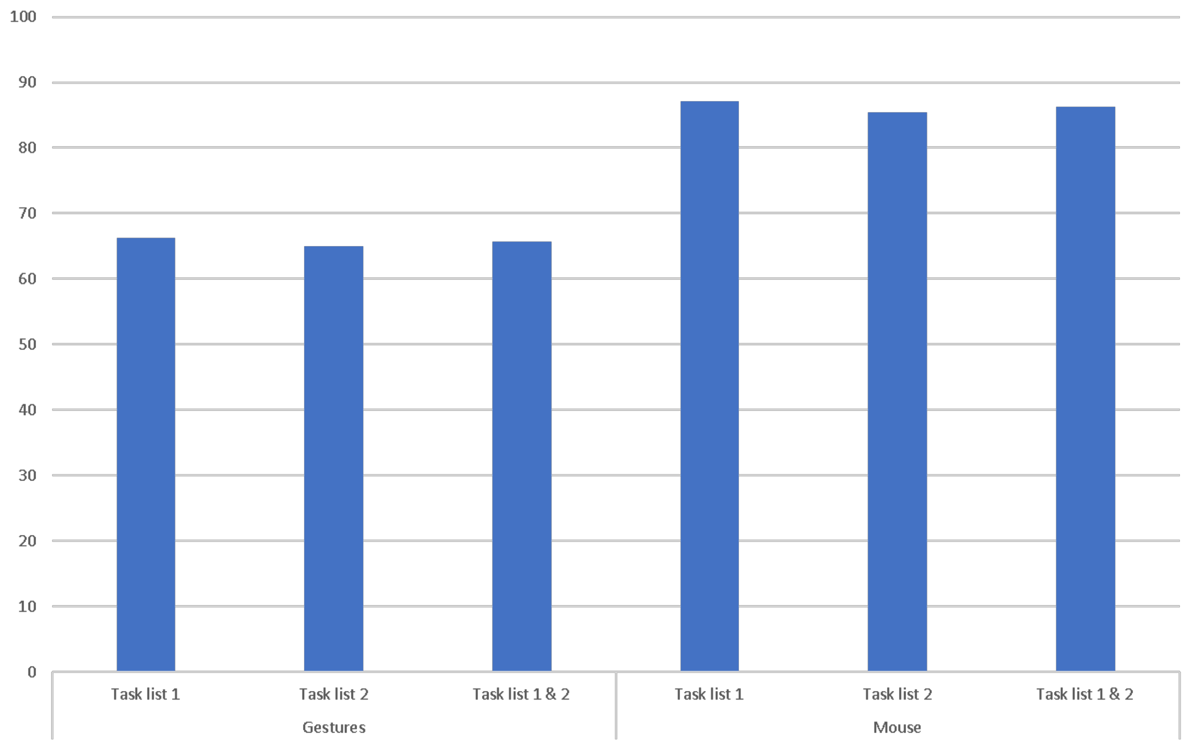


Figure 3.15: The SUS results for each task list with the according input methods. The y-axis shows the SUS points.

its result is ≈87.08 ($SD = 16.29$) and for task list 2 it is ≈85.42 ($SD = 12.45$). Together this results in a score of 86.25 ($SD = 14.52$) for the mouse.

The effort the participants experienced on every task was noted on the SMEQ questionnaires, the diagram in Figure 3.16 shows the calculated results.

The diagram shows that the effort is considerably higher with the gesture control. For task list 1 the average effort over all single tasks lies at 23.47 ($SD = 14.55$)for the gesture control. The same tasks performed with the mouse evince an effort of 2.11 ($SD = 1.81$). Task list 2 receives values of 18.48 ($SD = 12.13$) for the gesture control and 7.89 ($SD = 6.05$) for the mouse. On average over all tasks of both task lists the effort of the gesture control amounts 20.97 ($SD = 13.62$) whereas the effort of the mouse for the same tasks amounts 5 ($SD = 5.32$).

A repeated measures analysis of variance shows that there is a statistically significant effect of the input methods on the SMEQ value and on the time it took the participants to execute them, for all referents (see Table 3.1).

There is a statistically significant effect of the input method on the SMEQ value for the referent put mouse cursor here, $F(2, 12) = 12.734, p = .0044$, and for the time it takes

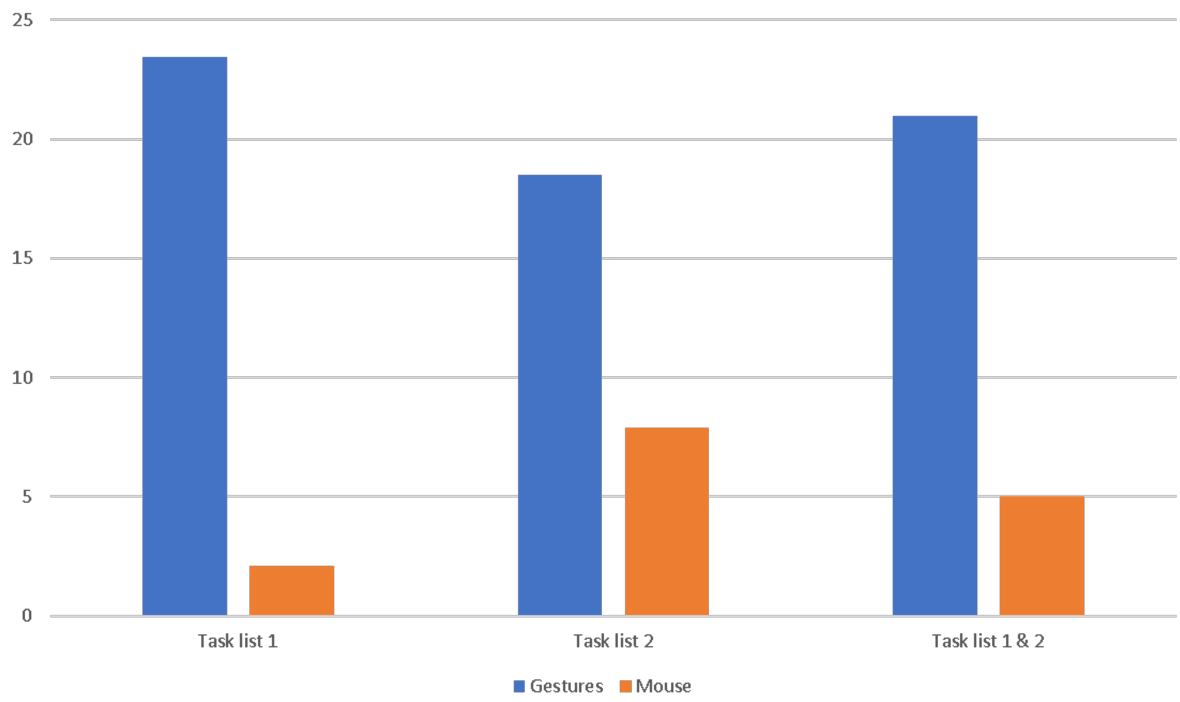


Figure 3.16: The diagram shows the average results of the SMEQ questionnaires for both input methods, for the single task lists and the combined outcome. The SMEQ value is displayed on the y-axis.

to execute it $F(2, 12) = 21.68, p = .0001$. There is a statistically significant effect of the input method on the SMEQ value for the referent select window, $F(2, 12) = 10.621, p = .0076$, and for the time it takes to execute it $F(2, 12) = 24.96, p = .00005$. There is a statistically significant effect of the input method on the SMEQ value for the referent move window, $F(2, 12) = 47.455, p = .00003$, and for the time it takes to execute it $F(2, 12) = 17.21, p = .0004$. There is a statistically significant effect of the input method on the SMEQ value for the referent resize window, $F(2, 12) = 10.241, p = .0085$, and for the time it takes to execute it $F(2, 12) = 8.834, p = .0070$. There is a statistically significant effect of the input method on the SMEQ value for the referent maximize window, $F(2, 12) = 50.33, p = .00002$, and for the time it takes to execute it $F(2, 12) = 12.99, p = .00157$. There is a statistically significant effect of the input method on the SMEQ value for the referent minimize window, $F(2, 12) = 34.933, p = .0001$, and for the time it takes to execute it $F(2, 12) = 8.586, p = .00775$. There is a statistically significant effect of the input method on the SMEQ value for the referent scrolling, $F(2, 12) = 47.283, p = .00003$, and for the time it takes to execute it $F(2, 12) = 28.89, p = .00002$. There is a statistically significant effect of the input method on the SMEQ value for the referent dock window left/right, $F(2, 12) = 24.022, p = .0005$, and for the time it takes to execute it $F(2, 12) = 18.11, p = .0003$. There is a statistically significant effect of the input

	SMEQ		Time	
Condition	$F(2,12)$	p	$F(2,12)$	p
put mouse cursor here	12.734	0.004407	21.68	0.000122
select window	10.621	0.0076118	24.96	0.0000532
move window	47.455	0.000026263	17.21	0.000419
resize window	10.241	0.0084514	8.834	0.00703
maximize window	50.33	0.000020078	12.99	0.00157
minimize window	34.933	0.00010158	8.586	0.00775
scrolling	47.283	0.000026698	28.89	0.0000214
dock window left/right	24.022	0.00047071	18.11	0.000323
close window	15.822	0.0021674	16.13	0.000579

Table 3.1: The results of the repeated measures anova for the referents on the SMEQ values and the elapsed task times.

method on the SMEQ value for the referent close window, $F(2,12) = 15.822, p = .0022$, and for the time it takes to execute it $F(2,12) = 16.13, p = .0006$.

In the following I will discuss the results gained during the conduction of the user study where the participants tested the implemented gesture control and compared it with a computer mouse.

3.2.6 Discussion

The results of the study show that the participants come along better with the mouse as input device than with the gesture control. This is because people are familiar with the mouse as input device on computers and gestures are usually only known from touch devices but not in mid-air to operate with a computer with a large display wall connected.

The time the participants needed to accomplish the specified tasks with the gesture control is many times over the time they needed with the mouse. This issue is in a certain way due to the leap motion controller. The controller needs short time to recognize a hand when it comes to its field of view. It is not a long time, but it is noticeable. During the study, the participants put up their hand when they started a new task. Often the colored circle did not react immediately and the participants started to move their hand fast to search the circle. The time the controller needs to recognize a hand was not measured but this constitutes a certain time period that should be subtracted from the time the participants needed to solve the tasks with the gesture control.

Furthermore, the responses on the SMEQ questionnaires characterizes the implemented control as exhausting compared to the mouse. But on the SMEQ scale, the received average value is located between "not very hard to do" and "a bit hard to do", so the gesture set is definitely not classified as strenuous.

Taking a look at the average of the NASA-TLX results, the mouse achieved the better results, here the values are less than the half of the ones from the gesture control. While working with the mouse none of the participants had to think about how to perform any referent, they just know it as they use a mouse every day. To complete the tasks with the gesture control the participants sometimes needed to think about how the demanded function is invoked. Also the hands need to be held above the table to perform the gestures, which causes a higher physical effort. This, along with the fact that in some cases the gestures did not trigger the according functions leads to more temporal demand and frustration.

The SUS result demonstrates that the participants consider the mouse as an above average input device. The average of the SUS score is 68. The gesture control reached 65.625 on average, what means it is below average in regards to the SUS result.

Nevertheless, the participants liked the gesture control and they said that they would use it in an office environment like the one set up for the study. They also mentioned they would need some more time for familiarization.

What the participants missed in the gesture control was a click function to be able to perform a right and a left click like with the mouse. This leads to the assumption that the participants could imagine to replace the mouse with the gesture control for the here presented office environment, as they wanted to add the mouse's main functions to the gesture control. Overall, the participants liked the office setup in which the study took place and the gesture control as input method.

4 Conclusion and Future Work

This work describes the approach of inventing a user-designed mid-air gesture set for window management on large high-resolution display walls and its evaluation. It presents a validated gesture set. Furthermore, for the elicitation of the gestures I applied two methods for the purpose of reducing legacy bias and investigated whether there are concrete benefits or not. These methods were applied to half of the participants of the gesture elicitation study. The referents were rated by the study participants according to their importance for the office environment where the study took place.

Although the two groups came up with different orders with regard to the importance of the referents and in few cases with different symbols, the calculated agreement score was nearly the same. So, the outcome of the investigation of the methods for legacy bias reduction shows that it did not lead to an advantage that justifies the additional expenditure. In fact, there was no difference between the agreement scores of the two groups. Thus, I merged and analyzed all the results and extracted a gesture set.

The received gesture set was subsequently implemented so it could be used with the leap motion controller. For evaluation, I conducted another user study in which the participants compared the gesture control with the standard input device for window management, the mouse. In this study, the participants had to process two similar task lists, one with the gesture control and one with the mouse. The results show that there is a statistically significant effect of the input method on SMEQ value and task completion time for the referents. Moreover, the effort the participants perceived was higher for the gesture control and it took them several times longer to solve the tasks. Yet, the feedback they provided was positive, all of them would like to use mid-air gestures for window management at such workplaces as the one I prepared for the study and some of them even in their office. Through the priority rating and the further feedback, I recognized that not all the referents I implemented must be provided by the gesture control. Then again, there were functions the participants missed, e.g. they proposed to add a left and a right mouse click option to the gesture control, so they would be able to start applications, open the start menu or documents without using the mouse at all.

Besides the benefits of large display walls there are some known problems that appear when working on them. While a user can overlook much more important information simultaneously, loosing track of the mouse cursor is a big issue on large high-resolution display walls, but this issue can be solved. With the gesture control invented and introduced in this work, it is easy to set the mouse cursor to the desired position. Through the colored circle that is significant larger and thereby more conspicuous than the mouse cursor, and the relative feeling of its position through the hand, users find it really straightforward to find it and set the mouse cursor to the right position.

The implementation introduced in this work and the study conducted with it, show that people are likely to use mid-air gestures in an office environment with large high-resolution display walls. It further shows, that it is possible to elicit gestures by user studies, that they are learnable and well thought out. But it also shows, that it is not efficiently usable yet and a satisfactory gesture set for everyone needs further research.

All participants came from a university environment. Maybe the results of the studies would be different if the participants were from other ethnic groups or if they were of another age. But these circumstances have to be investigated in further studies since this goes beyond the scope of this work.

Future Work

In the implementation of this work there are not all gestures that were elicited integrated. So, the next part would be to add and test them. In a further study the gestures should be shown to participants and they should try them again and provide suggestions for improvements for the gestures. With this approach it would be possible to find a gesture set that matches better and provides more consistency. I assume this helps to remember the gestures easier, e.g. the referents minimize window and maximize window are opposites, but their symbols are not related to each other in this implementation. The gesture set was compared with a computer mouse and even on the large high-resolution display wall where the mouse is more exhausting to use than on a normal computer monitor the study showed that the mouse requires less effort. With further investigations and the related improvements it should be possible to better the gesture control for a new comparison with the mouse.

The implementation of the gesture set is not always recognizing all the gestures well. To address this issue the next approach is to use machine learning algorithms on the leap motion controller and therewith improve the recognition of gestures which leads to fewer frustration for users.

Acknowledgements

Firstly, I would like to thank my examiner, Prof. Dr. Albrecht Schmidt, for giving me the chance to write my thesis about these futuristic and untrodden technologies and for the possibility to research in his laboratories.

I am most grateful for the support of my supervisors, Lars Lischke, Pascal Knierim and Dr. Paweł W. Woźniak, for their ideas and the fruitbearing and interesting discussions we held.

A Gesture Elicitation Study

University of Stuttgart
Germany

Human Computer Interaction Group (MCI), VIS

Prof. Dr. Albrecht Schmidt

Consent Form

DESCRIPTION: You are invited to participate in **a research study** on **designing gestures for window management on a large display-wall.**

TIME INVOLVEMENT: Your participation will take approximately **45 minutes.**

DATA COLLECTION: In this study, you will design gestures for window management on a display-wall. Video and Audio data will be recorded while you are participating the study. In addition, you will be asked to fill in several questionnaires.

RISKS AND BENEFITS: No risks are associated with this study. The collected data is securely stored. We do guarantee no data misuse and privacy is completely preserved. Your decision whether or not to participate in this study will not affect your grade in school. You can decide whether the recorded data can be published or not.

PARTICIPANT'S RIGHTS: If you have read this form and have decided to participate in this project, please understand your **participation is voluntary** and you have the **right to withdraw your consent or discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled**. **The alternative is not to participate.** You have the right to refuse to answer particular questions. The results of this research study may be presented at scientific or professional meetings or published in scientific journals. Your identity is not disclosed unless we directly inform and ask for your permission.

CONTACT INFORMATION: If you have any questions, concerns or complaints about this research, its procedures, risks and benefits, contact following persons:
Lars Lischke (lars.lischke@vis.uni-stuttgart.de)
Pascal Knierim (pascal.knierim@vis.uni-stuttgart.de)
Paweł Woźniak (pawel.wozniak@vis.uni-stuttgart.de)
Marius Altmann (altmanms@studi.informatik.uni-stuttgart.de)

VIDEO AND AUDIO DATA: (select one)

- ☐ Please **do not publish** the video and audio data recorded during my participation of study.
- ☐ I allow you to **publish** the video and audio data recorded during my participation of study.
- ☐ I allow you to **publish** the **anonymous** video and audio data recorded during my participation of study.

By signing this document, I confirm that I agree to the terms and conditions.

Name: ________________________ *Signature, Date:* ________________________

Figure A.1: Gesture elicitation study consent form.

Designing Gestures
Background Questionnaire

ID: ___

Date: ____________

How old are you? ______ years

What is your profession? ________________________

What is your gender? ____________________

Are you right- or left-handed? ____________________

How many hours per day do you use a computer? _______ hours

How many monitors do you use on your computer? _______

What specifications do they have?

Size: __________ inches

Resolution: _______ px x _______ px

How are the monitors adjusted?

Figure A.2: Gesture elicitation study background questionnaire.

Designing Gestures
Gestures list

ID: ___

Date: ____________

Priming and production: ☐yes ☐no

Please sort the gestures according to your personal feeling of relevance with numbers in the column at the front:

	Select window (click with mouse in window/choose via ALT+TAB)	
	Move window (take window on the top with the mouse by clicking and holding the left mouse button and use the mouse to move the window over the desktop)	
	Maximize window (WINDOWS BUTTON + ↑ or move with mouse to the top)	
	Minimize window (WINDOWS BUTTON + ↓)	
	Resize window/change window size	
	Dock window to the right (WINDOWS BUTTON + → or move with mouse to desktop right edge)	
	Dock window to the left (WINDOWS BUTTON + ← or move with mouse to desktop left edge)	
	Dock window to a desktop corner (WINDOWS BUTTON + ←/→ > WINDOWS BUTTON ↑/↓ or move with mouse into corner)	
	Show desktop (WINDOWS BUTTON + D)	
	Close window (ALT + F4)	
	Scrolling (mouse wheel up/mouse wheel down)	
	Show all active windows and desktops (WINDOWS BUTTON + TAB)	
	Put mouse cursor here / show mouse cursor	

You can discard any gestures which in your opinion are irrelevant.

If you have further ideas how to use gestures for window management on a computer, you can write them down here:

Figure A.3: Gesture elicitation study gestures list.

B System Evaluation Study

University of Stuttgart
Germany

Human Computer Interaction Group (MCI), VIS
Prof. Dr. Albrecht Schmidt

Consent Form

DESCRIPTION: You are invited to participate in **a research study** on **testing gestures for window management on a large display-wall.**

TIME INVOLVEMENT: Your participation will take approximately **75 minutes.**

DATA COLLECTION: In this study, you will design gestures for window management on a display-wall. Video and Audio data will be recorded while you are participating the study. In addition, you will be asked to fill in several questionnaires.

RISKS AND BENEFITS: No risks are associated with this study. The collected data is securely stored. We do guarantee no data misuse and privacy is completely preserved. Your decision whether or not to participate in this study will not affect your grade in school. You can decide whether the recorded data can be published or not.

PARTICIPANT'S RIGHTS: If you have read this form and have decided to participate in this project, please understand your **participation is voluntary** and you have the **right to withdraw your consent or discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled**. **The alternative is not to participate.** You have the right to refuse to answer particular questions. The results of this research study may be presented at scientific or professional meetings or published in scientific journals. Your identity is not disclosed unless we directly inform and ask for your permission.

CONTACT INFORMATION: If you have any questions, concerns or complaints about this research, its procedures, risks and benefits, contact following persons:
Lars Lischke (lars.lischke@vis.uni-stuttgart.de)
Pascal Knierim (pascal.knierim@vis.uni-stuttgart.de)
Paweł Woźniak (pawel.wozniak@vis.uni-stuttgart.de)
Marius Altmann (altmanms@studi.informatik.uni-stuttgart.de)

VIDEO AND AUDIO DATA: (select one)

- ☐ Please **do not publish** the video and audio data recorded during my participation of study.
- ☐ I allow you to **publish** the video and audio data recorded during my participation of study.
- ☐ I allow you to **publish** the **anonymous** video and audio data recorded during my participation of study.

By signing this document, I confirm that I agree to the terms and conditions.

Name: ________________________ *Signature, Date:* ________________________

Figure B.1: System evaluation study consent form.

Testing Gestures
Background Questionnaire

ID: ___

Date: ____________

How old are you? ______ years

What is your profession? _______________________

What is your gender? ___________________

Are you right- or left-handed? ___________________

How many hours per day do you use a computer? _______ hours

How many monitors do you use on your computer? _______

What specifications do they have?

Size: __________ inches

Resolution: _______ px x _______ px

How are the monitors adjusted?

Have you used a Leap Motion before? ☐ Yes ☐ No

Figure B.2: System evaluation study background questionnaire.

Testing Gestures
Task 1

ID: ___

Date: ____________

Order: 1.☐ 2.☐

Method: Gestures ☐ Mouse ☐

Task	Video	Time
1. **Select** the browser window on the third screen from right		
2. **Move** browser to the leftmost screen		
3. **Put the mouse cursor** to the browser		
4. **Maximize** the browser		
5. **Move** the browser to the second screen from right		
6. **Resize** the browser to the lower screen part		
7. **Minimize** the browser window		
8. **Select** the pdf reader window		
9. **Put the mouse cursor** to the pdf reader window		
10. **Scroll** to the third page of the pdf		
11. **Maximize** the pdf reader window		
12. **Scroll** back to the second page		
13. **Dock** the pdf reader window to the left of the screen		
14. **Move** the pdf reader window to the second screen from right		
15. **Scroll** down to the fifth page		
16. **Close** the pdf reader window		
17. **Select** the explorer window		
18. **Maximize** the explorer window		
19. **Dock** the explorer window to the right of the screen		
20. **Close** the explorer window		
21. **Select** the Windows Media Player window		
22. **Put the mouse cursor** to the Windows Media Player window		
23. **Dock** the Windows Media Player window to the left of the screen		
24. **Close** the Windows Media Player window		
25. **Select** the Word window		
26. **Resize** the Word window over the two screens in the middle		
27. **Minimize** the Word window		
28. **Select** the Outlook window		
29. **Resize** the Outlook window to the upper half of the screen		
30. **Minimize** the Outlook window		

Figure B.3: System evaluation study task list 1.

Testing Gestures

Task 2

ID: ___

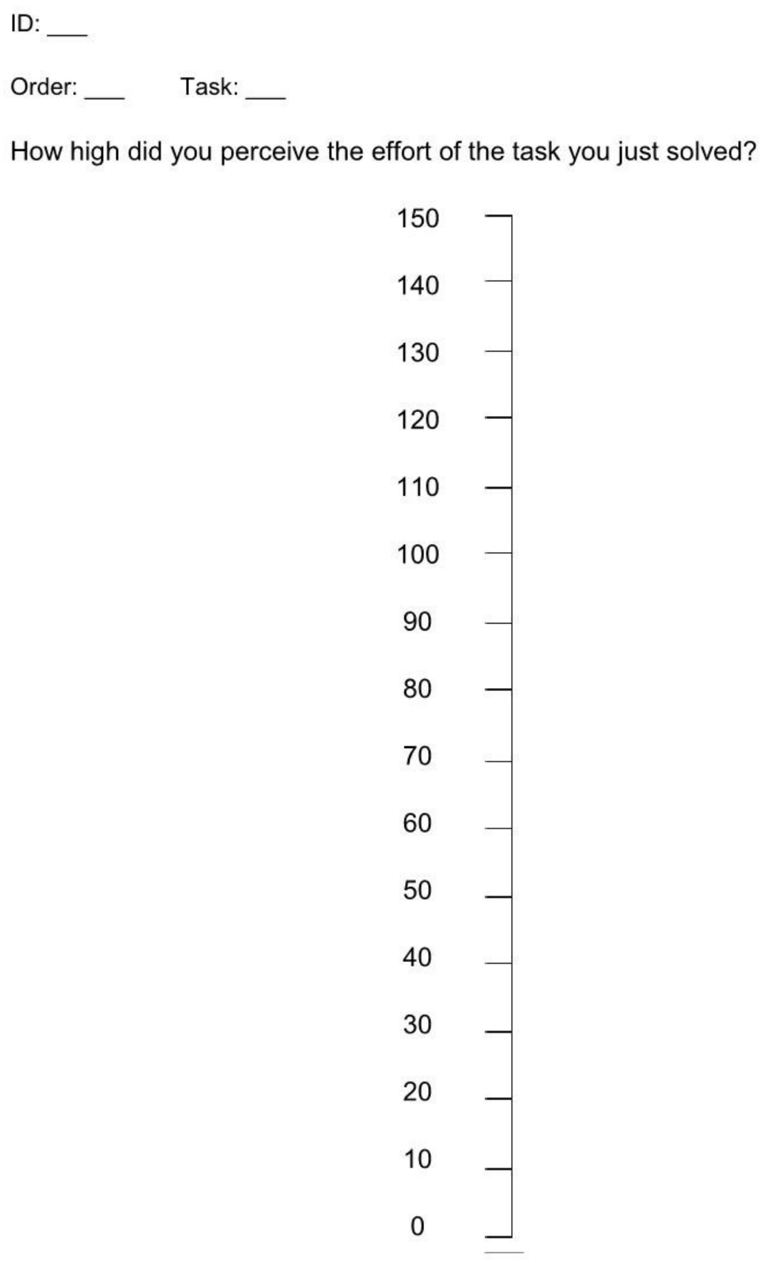
Date: ____________

Order: 1.☐ 2.☐

Method: Gestures ☐ Mouse ☐

Task	Video	Time
1. **Select** the pdf reader window on the third screen from left		
2. **Put the mouse cursor** to the browser window		
3. **Scroll** down to the next page in the pdf reader window		
4. **Maximize** the pdf reader window		
5. **Scroll** to the fourth page of the document in the pdf reader		
6. **Move** the pdf reader window to the leftmost screen		
7. **Dock** the pdf reader window to the left		
8. **Put the mouse cursor** to the pdf reader window		
9. **Scroll** to the second page of the document in the pdf reader		
10. **Close** the pdf reader window		
11. **Select** the browser window on the second screen from left		
12. **Move** the browser window to the second screen from right		
13. **Maximize** the browser window		
14. **Dock** the browser window to the right		
15. **Put the mouse cursor** to the browser window		
16. **Resize** the browser window to the upper half of the screen		
17. **Minimize** the browser window		
18. **Select** the Outlook window		
19. **Resize** the Outlook window over the two screens in the middle		
20. **Minimize** the Outlook window		
21. **Select** the explorer window		
22. **Maximize** the explorer window		
23. **Dock** the explorer window to the left of the screen		
24. **Close** the explorer window		
25. **Select** the Windows Media Player window		
26. **Move** the Windows Media Player window to the third screen from the left		
27. **Minimize** the Windows Media Player window		
28. **Select** the Word window		
29. **Resize** the Word window to the lower half of the screen		
30. **Close** the Word window		

Figure B.4: System evaluation study task list 2.


ID: ___

Order: ___ Task: ___

How high did you perceive the effort of the task you just solved?

150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

Figure B.5: System evaluation study SMEQ.

Testing Gestures
NASA TLX

ID: ___

Date: ____________

Mouse ☐ Gestures ☐

Mental Demand

How mentally demanding was the task?

Very Low — Very High

Physical Demand

How physically demanding was the task?

Very Low — Very High

Temporal Demand

How hurried or rushed was the pace of the task?

Very Low — Very High

Performance

How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort

How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration

How insecure, discouraged, irritated, stressed and annoyed were you?

Very Low — Very High

Figure B.6: System evaluation study NASA-TLX.

Testing Gestures

SUS

ID: ___

Date: ____________

Mouse ☐ Gestures ☐

	Strongly disagree				Strongly agree
1. I think I would like to use the system frequently.	1	2	3	4	5
2. I found the system unnecessarily complex.	1	2	3	4	5
3. I thought the system was easy to use.	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use this system.	1	2	3	4	5
5. I found the various functions in this system were well integrated.	1	2	3	4	5
6. I thought there was too much inconsistency in this system.	1	2	3	4	5
7. I would imagine that most people learn to use this system very quickly.	1	2	3	4	5
8. I found the system very cumbersome to use.	1	2	3	4	5
9. I felt very confident using the system.	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system.	1	2	3	4	5

Figure B.7: System evaluation study SUS.

Bibliography

[ABCD15] C. Ardito, P. Buono, M. F. Costabile, G. Desolda. "Interaction with Large Displays." In: *ACM Computing Surveys* 47.3 (Feb. 2015), pp. 1–38. ISSN: 03600300. DOI: 10.1145/2682623. URL: http://dl.acm.org/citation.cfm?doid=2737799.2682623 (cit. on p. 15).

[ACTK15] C. Ackad, A. Clayphan, M. Tomitsch, J. Kay. "An in-the-wild study of learning mid-air gestures to browse hierarchical information at a large interactive public display." In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*. New York, New York, USA: ACM Press, 2015, pp. 1227–1238. ISBN: 9781450335744. DOI: 10.1145/2750858.2807532. URL: http://dl.acm.org/citation.cfm?doid=2750858.2807532 (cit. on pp. 22, 24).

[BF05] H. Benko, S. Feiner. "Multi-monitor mouse." In: *CHI '05 extended abstracts on Human factors in computing systems - CHI '05*. New York, New York, USA: ACM Press, 2005, pp. 1208–1211. ISBN: 1595930027. DOI: 10.1145/1056808.1056878. URL: http://portal.acm.org/citation.cfm?doid=1056808.1056878 (cit. on p. 20).

[Bro86] J. Brooke. "System usability scale (SUS): a quick-and-dirty method of system evaluation user information." In: *Reading, UK: Digital Equipment Co Ltd* (1986) (cit. on p. 43).

[FGVT15] L. S. Figueiredo, M. G. Gonçalves Maciel Pinheiro, E. X. Vilar Neto, V. Teichrieb. "An Open Catalog of Hand Gestures from Sci-Fi Movies." In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15*. Vol. 18. New York, New York, USA: ACM Press, 2015, pp. 1319–1324. ISBN: 9781450331463. DOI: 10.1145/2702613.2732888. URL: http://dl.acm.org/citation.cfm?doid=2702613.2732888 (cit. on p. 23).

[Fit54] P. M. Fitts. "The information capacity of the human motor system in controlling the amplitude of movement." In: *Journal of experimental psychology* 47.6 (1954), pp. 381–391 (cit. on p. 21).

[Gru01] J. Grudin. "Partitioning DigitalWorlds: Focal and Peripheral Awareness in Multiple Monitor Use." In: *CHI '01*. 2001, pp. 458–465. ISBN: 1581133278. DOI: 10.1145/365024.365312. URL: http://dl.acm.org/citation.cfm?id=365312 (cit. on p. 19).

[HS88] S. G. Hart, L. E. Staveland. "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research." In: *Advances in psychology* 52 (1988), pp. 139–183 (cit. on p. 43).

[Hud13] W. Hudson. "Card sorting." In: *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* (2013). URL: https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/card-sorting (cit. on pp. 23, 26).

[JJBH15] M. R. Jakobsen, Y. Jansen, S. Boring, K. Hornbæk. "Should I Stay or Should I Go? Selecting Between Touch and mid-air gestures for large-display interaction." In: *Human-Computer Interaction*. Springer. 2015, pp. 455–473. DOI: 10.1007/978-3-319-22698-9_31. URL: https://link.springer.com/chapter/10.1007/978-3-319-22698-9_31 (cit. on pp. 21, 24).

[KBSR07] W. A. König, H.-J. Bieg, T. Schmidt, H. Reiterer. "Position-independent interaction for large high-resolution displays." In: 2007 (cit. on p. 21).

[KS08] Y.-a. Kang, J. Stasko. "Lightweight task/application performance using single versus multiple monitors: a comparative study." In: *Proceedings of Graphics Interface 2008*. Canadian Information Processing Society. 2008, pp. 17–24 (cit. on pp. 20, 24).

[LHK+17] L. Lischke, J. Hoffmann, R. Krüger, P. Wozniak, A. Schmidt. "Towards Interaction Techniques for Social Media Data Exploration on Large High-Resolution Displays." In: (2017) (cit. on p. 15).

[LMW+15a] 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, 2015, pp. 1845–1850. ISBN: 978-1-4503-3146-3. DOI: 10.1145/2702613.2732845. URL: http://doi.acm.org/10.1145/2702613.2732845 (cit. on pp. 15, 25).

[LMW+15b] L. Lischke, S. Mayer, K. Wolf, A. S. Shirazi, N. Henze. "Subjective and Objective Effects of Tablet's Pixel Density." In: *Proceedings of the ACM CHI'15 Conference on Human Factors in Computing Systems* 1 (2015), pp. 2769–2772. DOI: 10.1145/2702123.2702390. URL: http://dx.doi.org/10.1145/2702123.2702390 (cit. on p. 15).

[LMW+16] 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* (2016), pp. 228–234. DOI: 10.1145/2914920.2915027. URL: http://dl.acm.org/citation.cfm?doid=2914920.2915027 (cit. on pp. 15, 20, 24, 25).

[LSF+16] L. Lischke, V. Schwind, K. Friedrich, A. Schmidt, N. Henze. "MAGIC-Pointing on Large High-Resolution Displays." In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16*. New York, New York, USA: ACM Press, 2016, pp. 1706–1712. ISBN: 9781450340823. DOI: 10.1145/2851581.2892479. URL: http://dl.acm.org/citation.cfm?doid=2851581.2892479 (cit. on pp. 15, 25).

[Mac92] I. S. MacKenzie. "Fitts' law as a research and design tool in human-computer interaction." In: *Human-computer interaction* 7.1 (1992), pp. 91–139 (cit. on p. 21).

[MDD+14] M. R. Morris, A. Danielescu, S. Drucker, D. Fisher, B. Lee, C. Schraefel, J. O. Wobbrock. "Reducing Legacy Bias in Gesture Elicitation Studies." In: *Interactions* 21 (2014), pp. 40–45. ISSN: 10725520. DOI: 10.1145/2591689. arXiv: arXiv:1011.1669v3. URL: http://dl.acm.org/citation.cfm?doid=2608008.2591689 (cit. on pp. 16, 23, 25–27, 33, 38).

[MWW10] M. R. Morris, J. O. Wobbrock, A. D. Wilson. "Understanding users' preferences for surface gestures." In: *Proceedings of Graphics Interface 2010* (2010), pp. 261–268. ISSN: 13596454. DOI: 10.1016/j.actamat.2009.07.058. URL: http://dl.acm.org/citation.cfm?id=1839214.1839260 (cit. on p. 23).

[NDL+09] C. North, T. Dwyer, B. Lee, D. Fisher, P. Isenberg, G. Robertson, K. Inkpen. "Understanding multi-touch manipulation for surface computing." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 5727 LNCS.PART 2 (2009), pp. 236–249. ISSN: 03029743. DOI: 10.1007/978-3-642-03658-3_31 (cit. on p. 23).

[NSMG03] M. Nielsen, M. Störring, T. B. Moeslund, E. Granum. "A procedure for developing intuitive and ergonomic gesture interfaces for man-machine interaction." In: *Proceedings of the 5th International Gesture Workshop*. 2003, pp. 1–12 (cit. on p. 23).

[OF13] U. Oh, L. Findlater. “The challenges and potential of end-user gesture customization.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’13*. New York, New York, USA: ACM Press, 2013, p. 1129. ISBN: 9781450318990. DOI: 10.1145/2470654.2466145. URL: http://dl.acm.org/citation.cfm?doid=2470654.2466145 (cit. on p. 22).

[OTN+12] J. W. Owens, J. Teves, B. Nguyen, A. Smith, M. C. Phelps, B. S. Chaparro. “Examination of Dual vs. Single Monitor Use during Common Office Tasks.” In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 56.1 (2012), pp. 1506–1510. ISSN: 1071-1813. DOI: 10.1177/1071181312561299. URL: http://pro.sagepub.com/lookup/doi/10.1177/1071181312561299 (cit. on pp. 19, 20, 24).

[RCB+05] G. Robertson, M. Czerwinski, P. Baudisch, B. Meyers, D. Robbins, G. Smith, D. Tan. “The large-display user experience.” In: *IEEE Computer Graphics and Applications* 25.4 (2005), pp. 44–51. ISSN: 02721716. DOI: 10.1109/MCG.2005.88 (cit. on pp. 20, 24).

[SD09] J. Sauro, J. S. Dumas. “Comparison of three one-question, post-task usability questionnaires.” In: *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. New York, New York, USA: ACM Press, 2009, p. 1599. ISBN: 9781605582467. DOI: 10.1145/1518701.1518946. URL: http://dl.acm.org/citation.cfm?doid=1518701.1518946 (cit. on p. 24).

[SSL+06] G. Schmidt, O. Staadt, M. Livingston, R. Ball, T. Ni. “IEEE Xplore - A Survey of Large High-Resolution Display Technologies, Techniques, and Applications.” In: *Proceedings of IEEE Virtual Reality Conference 2006* (2006). URL: http://ieeexplore.ieee.org/xpls/abs%7B%5C_%7Dall.jsp?arnumber=1667648%7B%5C&%7Dtag=1 (cit. on p. 15).

[TGSP06] D. S. Tan, D. Gergle, P. Scupelli, R. Pausch. “Physically large displays improve performance on spatial tasks.” In: *ACM Transactions on Computer-Human Interaction* 13.1 (Mar. 2006), pp. 71–99. ISSN: 10730516. DOI: 10.1145/1143518.1143521. URL: http://portal.acm.org/citation.cfm?doid=1143518.1143521 (cit. on pp. 20, 24).

[VB05] D. Vogel, R. Balakrishnan. “Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays.” In: *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*. UIST ’05. Seattle, WA, USA: ACM, 2005, pp. 33–42. ISBN: 1-59593-271-2. DOI: 10.1145/1095034.1095041. URL: http://doi.acm.org/10.1145/1095034.1095041 (cit. on pp. 15, 21).

[WARM05] J. O. Wobbrock, H. H. Aung, B. Rothrock, B. A. Myers. "Maximizing the Guessability of Symbolic Input." In: *CHI '05 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '05. Portland, OR, USA: ACM, 2005, pp. 1869–1872. ISBN: 1-59593-002-7. DOI: 10.1145/1056808.1057043. URL: http://doi.acm.org/10.1145/1056808.1057043 (cit. on pp. 22, 33).

[WCH+16] A. Warr, E. H. Chi, H. Harris, A. Kuscher, J. Chen, R. Flack, N. Jitkoff. "Window Shopping: A Study of Desktop Window Switching." In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (2016), pp. 3335–3338. DOI: 10.1145/2858036.2858526. URL: http://doi.acm.org/10.1145/2858036.2858526 (cit. on p. 15).

[WJ16] M. L. Wittorf, M. R. Jakobsen. "Eliciting Mid-Air Gestures for Wall-Display Interaction." In: *NordiCHI '16* (2016), pp. 1–4. DOI: 10.1145/2971485.2971503. URL: http://dl.acm.org/citation.cfm?doid=2971485.2971503 (cit. on pp. 21, 27).

[WLM+17] P. W. Wozniak, 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* (2017), pp. 339–342. DOI: 10.1145/3022198.3026341. URL: http://dl.acm.org/citation.cfm?doid=3022198.3026341 (cit. on p. 20).

[WMW09] J. O. Wobbrock, M. R. Morris, A. D. Wilson. "User-defined Gestures for Surface Computing." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09. Boston, MA, USA: ACM, 2009, pp. 1083–1092. ISBN: 978-1-60558-246-7. DOI: 10.1145/1518701.1518866. URL: http://doi.acm.org/10.1145/1518701.1518866 (cit. on pp. 22, 25, 37).

[ZD85] F. R. H. Zijlstra, L. v. Doorn. "The construction of a scale to measure subjective effort." In: *Delft, Netherlands* (1985) (cit. on p. 43).

All links were last followed on July 19, 2017.

Declaration

I hereby declare that the work presented in this thesis is entirely my own and that I did not use any other sources and references than the listed ones. I have marked all direct or indirect statements from other sources contained therein as quotations. Neither this work nor significant parts of it were part of another examination procedure. I have not published this work in whole or in part before. The electronic copy is consistent with all submitted copies.

---

place, date, signature