LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
Department „Institut für Informatik“
Lehr- und Forschungseinheit Mensch-Maschine-Interaktion
Prof. Dr. Andreas Butz

# Master Thesis

# Is Anyone Looking?
# Mitigating Shoulder Surfing on Public Displays through Awareness and Protection

Frederik Brudy
fb@fbrudy.net


03. September 2013 – 11. March 2014

Supervisor:
Prof. Dr. Saul Greenberg, University of Calgary

Reviewer:
Prof. Dr. Andreas Butz, University of Munich (LMU)

# Acknowledgments

Above all and most important I want to thank my family. Without you I could have never made it to where I am now. You supported me up to this point and always gave me good advice. Thank you for believing in me and supporting me!

First I want to thank my supervisor Saul Greenberg for being and outstanding supervisor, mentor and guide. The amount I learned from you is invaluable and goes beyond what is written in this thesis. Thank you for giving me freedom to find my research topic and guiding me through the past six months.

At the University of Munich I would like to thank Andreas Butz for getting me in touch with the Interactions Lab and reviewing my thesis.

I would also like to thank all members of the Interactions Lab. You are an outstanding group of people! You all make this place very unique with your openness, collaboration, discussion and conversations. Thank you for welcoming me with open arms and making my time at the University of Calgary a very special and unique one.

Everything is only half the fun without great friends to share memories. That's why I would like thank my friends who walked with me the past six years, especially Sebastian, Fabius, Max, Netti, Marlene, Iris, Simone, Felix, Franzi, Maija, Sonja, Jojo, Hanna, Sven, Laura. Thank you guys for being an outstanding *party group*!

# Abstract

When a person interacts with a display in open areas, information on the screen – which may be sensitive, personal and private – may be visible to anyone passing by. With small displays, a person's body suffices to shield that information from others, but this approach becomes increasingly ineffective as the display area increases (e.g., from large workstation screens, to multiple monitors, to wall-sized displays). This is especially problematic in public areas, as a person may be unaware that a stranger is present and looking over one's shoulder. To mitigate this problem, in this work proxemic relationships between the passerby, the person and the display are used to provide the person with both awareness of onlookers, and mechanisms to manually or automatically protect information by muting or covering sensitive content. Awareness of passerby's and onlookers is provided through visual cues, which range from flashing the borders of the screen, to a 3D model mirroring a passerby's position and gaze, to an indicator that illustrates the gaze direction of the onlooker on the screen the onlooker. The same information signals the onlookers that they may be intruding, where – if they obey social protocol – they then turn away. A person may act on that awareness information by performing a gesture that covers sensitive applications: by blacking out windows, or by moving them to one side. Alternately, the system can automatically darken screen regions that cover those display area visible by the onlooker, while leaving the display area shielded by the person's body unaltered (thus allowing the person to continue their actions). The person may also invite the person in for collaboration by explicitly turning off these protective mechanisms with a gesture.



# Kurzzusammenfassung

Währen der Interaktion mit Displays in öffentlichen Bereichen können sensible oder private Informationen möglicherweise von vorbeigehenden Personen eingesehen werden. Mit kleinen Displays kann dieses Problem oftmals umgangen werden, beispielsweise durch Verdecken von privaten Inhalten. Dieses Vorgehen wird durch die wachsende Anzahl großer Displays und dem Einsatz mehrerer Monitore erschwert. Besonders in öffentlichen Umgebungen kann dies zu Problemen führen da fremde Personen einem Nutzer über die Schulter schauen können und somit Einblick in dessen privaten Informationen erhalten können. Um dieses Problem abzumildern werden in dieser Arbeit die räumlichen Beziehungen zwischen Passant, Benutzer des Systems und dem Displays ausgenutzt um einerseits die Wahrnehmung zu fördern, die ein Nutzer über seine Umgebung hat, andererseits ihm Möglichkeiten zu bieten seine sensiblen Inhalte zu schützen. Diese Signale sind visuelle Indikatoren, wie beispielsweise ein blinkender Rahmen um das Display. Eine dreidimensionale Repräsentation des vorbeigehenden Passanten wird genutzt um dem Nutze Details über dessen Position, Orientierung und Blickrichtung zu geben. Wenn dieser auf das Display blickt kann der eingesehene Bereich hervorgehoben werden. Dies ermöglicht dann dem Nutzer seine Daten zu schützen, oder den Passanten zu Kollaboration einladen. Ein Nutzer kann durch implizite und explizite Techniken sicherstellen, dass seine Daten geschützt sind, beispielsweise durch automatisches Verschieben der Fenster oder Abdecken von Bildschirmbereichen. Auf der anderen Seite genügen diese Techniken auch um der vorbeigehenden Person zu signalisieren dass sie gerade persönliche Daten einsieht, welche nicht für sie bestimmt sind. Dies unterstützt weithin akzeptierte soziale Normen im Umgang mit anderen Menschen und deren persönlichem Bereich.

# Task Description

In personal workspaces, people follow a notion of territoriality, where they know what their territory is and what the territories of their coworkers are. This is especially true with personal items on a desk. Individuals usually know where the limits are when they are touching other people's belongings. When overlooking someone's work on a computer display, this border is easily overstepped. For example, when a person is looking at his banking account or personal emails, another person may be passing by and able to look at the screen. This problem also translates to interactions with a large wall or a public display.

However there are times when one person invites another person to enter his territory, in order to work together on the same desk or computer, either on his own or with the owner close by. It is also feasible that two people work on a large display together, but need their private spaces as well. In either case they need to (implicitly or explicitly) negotiate about privacy and their territory.

To solve this problem, we will explore whether we can use proxemic interaction (looking at people's relationships in terms of distance, orientation, etc.) to mediate territoriality and privacy while working on public or personal displays. We also will look at how people can collaborate, while still keeping their private zones.

Further we will look at how we can notify a person being overlooked about this situation (e.g. in form of a visual cue). We also want to explore how to limit the visible display area to a portion that can only be seen by the entitled person, while hiding it from a person passing by.

Herewith I declare that I have completed this work solely and only with the help of the mentioned references.

Calgary, 11th March 2014 _____

Frederik Brudy

# Table of Contents

# Table of Figures

# 1 Introduction

Large displays are increasingly appearing in public settings, many of which are not only for passive consumption of information and thus pose a threat of private information being overlooked. May it be in libraries, shopping malls, universities, museums, hotel lobbies or other urban spaces, they are becoming more and more interactive, allowing not only to read presented information but also use it in highly personalized ways. Furthermore the size of those screen increases, mainly because of the advancements in technology and sunken manufacturing cost. This opens up new interactions and with a larger area. Recent concepts enable to connect mobile devices to a monitor and keyboard allowing for a bigger work area [15], and commercially available solutions [23] wirelessly connect e.g. one's smartphone to a display, utilizing the bigger screen size. These advancements are favorable as in the future more digital data will be used on the go.

## *1.1 Privacy Violations*

On the contrary, people are voyeuristic when it comes to private data of other people, and especially with large displays this can become a serious threat of the integrity of personal information. This threat of personal data being overlooked is called *shoulder surfing* (shown in Figure 1.1) and has been observed in multiple variations and occasions (more about shoulder surfing in section 3.4).

Privacy intrusions can either be deliberately or inadvertent [7]. However, most people are rational [52]. Rational people respect other's privacy in a similar way that they want to protect their own data [11,52]. When following social protocol, inadvertent privacy violators, when becoming aware of the fact, will usually look away to not further intrude someone's privacy [11].

**Figure 1.1. A passerby is shoulder-surfing a user, working on a public display.**

### 1.1.1 Examples of Problematic Settings

There are multiple reasons why the use of large and public displays increases. For one, they offer more interaction space. With more information being used in a digital form, this space is needed. By using an existing display e.g. in a library environment the limited screen space of one's mobile device (e.g., laptop or smartphone) can be extended, allowing users to bring their applications and data and take it with them when they leave. Some other example problematic settings include:

- A person viewing and entering banking information at a *bright touch-sensitive ATM* when others may be in line behind them or passing by (Figure 1.2a).
- A person's *desk is located in an open office thoroughfare*, where his[1] works on multiple large desktop displays (Figure 1.2b and e).

[1] In this thesis the masculine version of the grammatical person will be used for an easier reading flow. It

- A person *preparing materials on a wall-sized interactive display* located in a shared and open break-out space (e.g. Figure 1.1).
- *Group work rooms in a library*, which have glass doors or glass walls often offer displays where people can connect their laptop to. The inside of the rooms are visible to anyone passing by, thus also allowing visual access to the data (Figure 1.2c).
- When *collaborating on a large display* a person often wants to bring in data, which he has prepared previously. He plugs in the USB flash drive to open a file, or log in to a cloud storage space, or his email account. The drive might also contain other, private, files or private emails which could be overseen by the collaborators or other people who they were not aware of.
- A *hotel with a public counter* with large desktop displays in its lobby, which guests use to do various personal tasks (for example, reading email, retrieving airline bookings, etc.).
- *Entering an email address on a publicly visible screen / terminal / display*. For example when entering a lottery or prize draw.
- A store that includes a large public display that lets people purchase items on it, and also allows people to pursue personal tasks (e.g., by offering a web browser) (e.g., Figure 1.1).
- In a research lab environment people roughly know who has access to the area. Usually people also do not want to hide their research from colleagues in the same lab. But sometimes there are *unexpected visitors* coming to the lab, who data needs to be protected from (e.g. data about participants from conducted studies).
- *Online dating* using mobile websites and applications. A study from early 2013 showed that every tenth person in the United States has used online dating sites on mobile applications [53].
- *Online banking* using mobile devices. A survey showed that 51% of adults in the United States do their banking business online. Almost one third of US adults check their banking account while on the go, using their mobile phone [21].
- *Public information directories*, e.g. in shopping malls, where e.g. a person can look up a store. The results of these searches not only contain the opening hours and the location, but also the store's logo and name in big letters. Not anyone might be comfortable publicly disclosing his searches. For example a husband might be looking for a lingerie store to find a gift for his wife, but he does not want anyone to oversee his current search. There is currently no way of protecting his privacy (Figure 1.2d).
- The shopping cart of online shopping websites. Even just a single search on some online retailer websites for a certain product suffices to have related products show up on the start page upon the next visit. With many online shops now selling not only books, but all sorts of products, including e.g. adult toys, this might not be preferable and people might consider that start page might to now contain sensitive information (see section 2.4 for the Proxemic Peddler, an example of this setting in public space).


**Figure 1.2. Private sensitive information in public settings, examples of possible shoulder surfing situations. a) Accessing banking information on an ATM. b) Open office environments. c) Workroom in a library. d) Information directory, e.g. in a shopping mall. e) Open office environment.**

must be noted that this is supposed to include the female version as well and does not endorse any gender discrimination whatsoever.

## 1.2 Shielding of Private Information

In the previous section a large number of different situations have been listed in which personal data could be overlooked on a public display. Almost everyone has a desire to protect their belongings and information about tasks they are performing. People have a natural sense of personal space and privacy. Although the conception of private data and the desired level of protection is to different degrees that depend on a person's personality and one's particular context [3,11], when it comes to private and sensitive information people want to be sure that their information is protected from unwanted views.

One example is how people use spatial features to shield how others in the surrounding space can see the details of what one is doing or what information they may be viewing [3]. Consider how a person limits how others can shoulder surf their computer display. When working on their own private computer or devices people can control who can see the content of their screens, for example by turning the computer screen out of sight. They can further close the door to protect their privacy [3]. These reactions vary from person to person, yet they always remain common and sometimes unconscious practices of everyday life. One of such practices is how people lower their voice when desiring a higher level of privacy and use their own body to occlude private data and avoid being overlooked [3]. To aid this privacy protection, desks are often positioned in a way that the display is not easily viewable by someone who approaches or enters the room. Therefore, a user will be naturally aware of when a person is about to enter the visible area of his display [3].

These dynamics of arranging displays away from passers-by's views or shielding one's work by using the own body, are a reasonable way to protect privacy-sensitive information on small computer displays. However, they are more difficult to control in an open environment: While still being a fairly easy task when the display region is small (mobile devices can be covered with a single hand), with increasing display sizes one's ability to reposition a computer display in an open office environment is somewhat limited and the display area easily outgrows the area someone can cover with his body. When working on multiple monitors or large displays it is almost impossible to protect sensitive information from passers-by, especially if they are situated in open shared areas. Thus shoulder surfing becomes easier and opportunities more frequent. As a result information becomes legible at larger distances. Because there is less control of who can access the area, passers-by can include strangers.

## 1.3 Awareness Leverages Protection

In order to protect private information, people must understand not only who has access to their data but also who can overlook their information [7]. More and more personal data is stored in the 'cloud' and access to it is available wherever people go. As a result information can be easily accessed and shared. In order to ensure their privacy, people need feedback about their environment and who might be overlooking their data when using it in a public environment. Even the most secure computing system fails in protecting data when an individual fails to protect personal information.

Most of these failures to protect this information are inadvertently and simply because of missing knowledge or false sense about what information is considered to be private. Thus providing awareness is a key step in protecting sensitive information [7,63]. However, the level between awareness and privacy has to be balanced in order to provide an ideal level of awareness versus intrusiveness. When too much awareness is provided, the user might feel overwhelmed by the amount of information and feels the system to intrude his autonomy. When too little awareness is provided, inadvertent privacy intrusions are possible, as a user cannot tell whether someone currently has (visual) access to his data [7].

A person typically relies on two factors to limit shoulder surfing. The first is the interplay between *awareness* and *social protocol*. Because text on a desktop computer is usually small, the passerby would have to be quite close to read content (at a distance, which is called the personal or intimate space in proxemics [26]). It thus becomes somewhat obvious to the person that the passerby is shoulder surfing, and both would normally consider this a rude behavior. Second, the person can

position his or her body to shield content from view from the approaching person.

The goal of this work is to mediate shoulder with a particular emphasis on public displays. To achieve this goal, notions of personal space, territoriality and proxemics are leveraged to provide participants with (a) *awareness of shoulder surfing moments* and (b*) implicit and explicit protection* of information when shoulder surfing is detected. These cues are provided to both, the user of a system, as well as also a passerby, as people follow social protocol to mediate their social interactions.

The rest of this thesis work is structured as follows:

Chapter 2 provides a background and summarizes previous work. Techniques to protect sensitive information on public displays and terminals will be explained. These techniques mostly focuse on a combination of special hard- and software to ensure privacy. The terms *territory* and *personal space* will be explained. Both of these theories are well explored in psychology research, and have recently led to *proxemic interaction*, utilizing physical relationships between people in interactive systems.

In chapter 1 several design challenges and considerations will be given, addressing different aspects of privacy, self-regulation and social interaction. The further course of this work draws heavily on thes concepts in order to mitigate shoulder surfing on public displays.

A number of solutions, each illustrating how participants can become aware of shoulder surfing episodes are then described in chapter 1, leveraging the previously mentioned design considerations. Chapter 1 explains how systems can afford some implicit and explicit protection in order to alleviate shoulder surfing.

Details about the implementation will be given in chapter 1, explaining how the system works in general and how a 3D motion tracking system is used to provide the techniques from chapter 1 and 1 in a prototyping lab environment.

Chapter 1 summarizes the work and discusses the solutions in a broader context. Chapter 1 provides an outlook for future work.

# 2 Related Work

This work is not the first to consider shoulder surfing over displays. Several somewhat specialized methods have already been disclosed, each offering some level of protection, as listed below. Most of the previous work focused on using special hardware or a combination of special hard- and software in order to ensure privacy of sensitive information.

## 2.1 Methods Tuned to Highly Specific Data

Security professionals are particularly concerned with shoulder surfing attacks of passwords, and have developed various password entry methods to protect against such attacks.

One such system is *VibraPass* [33], a system which provides secure authentication based on shared lies. When a user enters his authentication information at a public terminal, e.g. an ATM, the terminal communicates with the user's phone in his pocket, having it vibrate whenever he should enter a fake character to his PIN (a sequence of this interaction is shown in Figure 2.3). A user study showed that with a lie overhead of 30% (meaning almost every third digit was a fake digit) a trade-off between input speed and protection from shoulder surfers can be achieved. Nevertheless, this attempt adds additional overhead into the authentication process and the interaction with the public terminal.

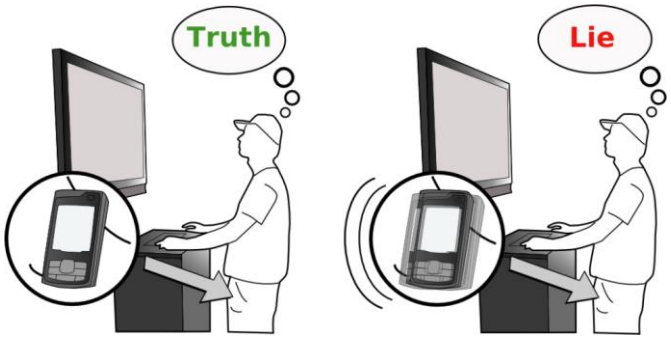


**Figure 2.3. Vibrapass, authentication based on shared lies between the user's personal device and the terminal. Figure from [33].**

By changing the way of how information is being entered into a public terminal, shoulder surfing can be mitigated or even eliminated as no information can be overseen. Especially with entering authentication information this has been studied frequently. One example is *EyePassword* [30], a gaze-based input method for sensitive data (such as PINs, passwords, etc.) on public terminals. This method adds a significant overhead on the input time, when compared to a normal keyboard: The average time for a password entry increased to 9.2 seconds to up to 12.1 seconds with their gaze based system, compared to 2.4 seconds for keyboard only entry. Yet, it has not been tested in terms of its resilience against shoulder surfing.

**Figure 2.4. The Spy-Resistant Keyboard. Picture from [59].**

Tan et al. [59] suggested using a *Spy-Resistant Keyboard* for input on public touchscreens. They managed to prevent data input from being overlooked by an inadvertent shoulder surfer, but for the cost of input time (almost 50 seconds for the Spy-Resistant Keyboard compared to ~22 seconds with a traditional soft keyboard), likely because of the unusual interaction technique introduced with this type of keyboard. Furthermore, the keyboard is not resistant against video camera supported shoulder surfing.

By using biometrical data to enhance traditional input, sensitive information can be protected against shoulder surfers by still allowing for fast input speed and high accuracy. One example of this is how De Luca et al. recorded how people moved their finger, the speed, touch size and pressure when performing a pattern unlock on a smartphone. They then compared it to other users and previous trials by the same users [32]. They found that security can be improved by still keeping usability at a usable level for valid users. The usability of biometric authentication has been tested by Coventry et al. [17]. After extended usage of an iris scanner when authenticating at an ATM 90% of users were satisfied and would elect this authentication method over the conventional PIN authentication. Compared to the previously mentioned systems, biometric approaches require enrolment of users into the process, as they have to provide the system with their biometrical features first.

The above stated techniques either change the way information is being entered into a system or alters the information itself by adding additional characters, thus creating an overhead for the user. The work in this thesis differs, as there is no knowledge in advance what information may be deemed sensitive and requires protection. Furthermore, the above techniques assume that an attack actually occurs and that someone deliberately tries to oversee someone else's information. As discussed in chapter 1, people usually respect other's territories and private data, and intentional privacy intrusions are rare.

## 2.2 Limiting what People See Based on Viewing Angles

Another approach physically limits what onlookers can see. As mentioned, this can be done by strategically locating displays in the environment to restrict how passers-by can view the display and its content. Further special display techniques can be used. For example Shoemaker proposed using a stereographic display in conjunction with shutter glasses, worn by the users, to show private content only to the user it belongs to, when working on a shared display [50,51]. The display shows a separate left and right eye view. Users had to wear special stereographic glasses, with either the left eye (user A) or the right eye (user B) covered. Private content for user A is shown on the screen refresh for the left eye, user B's private content is visible on the screen refresh for the right eye. Shared content is being shown on any screen refresh and thus visible to both users.

In a more recent approach Harrison et al. used an old LCD monitor to hide private information from passers-by [27]. They exploited an effect observed with (semi-old) LCD displays, resulting in color distortions when viewed from an oblique angle. The result is that color, hue, saturation and luminance can shift, depending on the viewing angle. This effect often is considered to be a problem and large viewing angles are usually a selling factor for displays. When it comes to hiding private content from passers-by this can be seen as an advantage, as this technique does not require the user to wear special glasses or the utilization of another device. By adapting the screen's content, two different images can be displayed, one only visible when viewed from an angle <30° and another one being visible when viewed from an angle >30° off the upright angle. An example application, simulating an ATM pin pad, can be seen in Figure 2.5. A combined image is also possible, which is visible from either angle. Their approach is easily distributed to various systems and can be used to hide personal information such as emails, password entry forms, etc. On the other side, this approach also requires the user to view his own content in a perpendicular angle onto the display in order. The personal content also becomes visible to a passerby when he changes his viewing angle and does not look from a slanted angle.

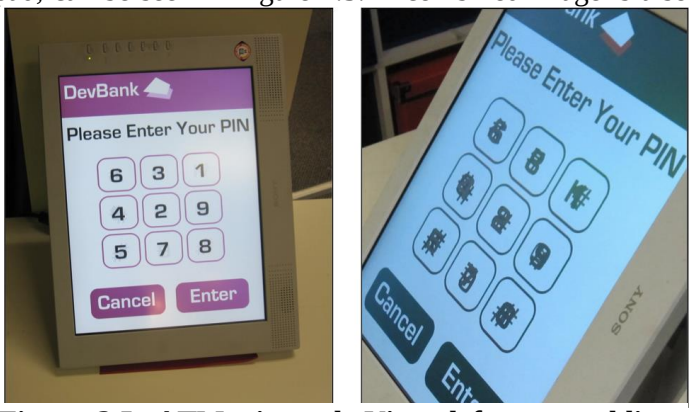


**Figure 2.5. ATM pin pad. Viewed from an oblique angle, the information cannot be seen on the display. Image from [27].**

More generally, commercially-available privacy filters – screens attached atop of displays such as [1] – cause the display to appear increasingly dark as the onlooker's viewing angle increases. Thus people looking at the screen from the side will not see anything. Because privacy filters do not stop a shoulder surfer from seeing the screen from a straight-on position, strategic positioning of the display and body shielding must still be done. These privacy filters are usually restricted to relatively small displays (e.g., tablet to desktop displays), likely because it would compromise how a person could look around a very large screen.

## 2.3 Offloading Private Information to a Trusted Handheld Device

Another approach considers how people can symbiotically use both a personal handheld device and a public display to help them perform a task more easily while still protecting privacy. Displaying and entering sensitive information only on a handheld mobile device rather than the public display, provides protection from shoulder surfers. This approach has been studied in multiple variations, of which some will be listed here.

### 2.3.1 Input of Personal or Sensitive Information

Sharp et al. [48] proposed a split-trust system to increase the security and privacy when browsing the web by using a trusted personal device. Sensitive information can be entered through a handheld mobile device in order to mitigate the threat of key loggers and other malicious soft- and hardware installed on the public device.

De Luca et al. also allowed a user to input sensitive data on a personal mobile device, using their *PocketPIN* system [31]. They state that using one's personal mobile device as an input device for public systems bears the advantage of being a trusted device: It is next to impossible for an attacker to manipulate the device, which is usually carried around by the user. Their system works by splitting up the user interface and input functionality between a public terminal and a user's personal device (shown in Figure 2.6). Whichever data the user deems to be private or sensitive can be entered on the mobile device. For convenience, non-sensitive data can be entered using e.g. a keyboard connected to the public display. A user study showed that data entry on a public terminal takes significantly longer when done on the mobile device. However they state that they observed a learning effect and expect users to be able to enter data more quickly over time. User satisfaction showed no significant difference between data entry using *PocketPIN* and a regular keyboard. The authors point this out as an advantage of their system.

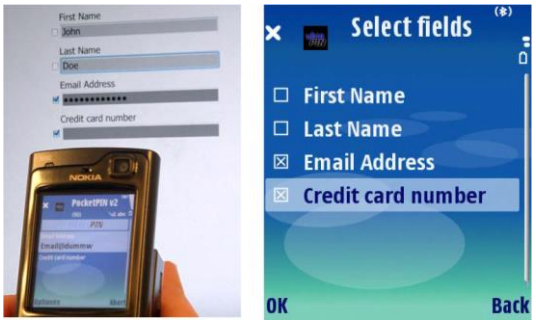


**Figure 2.6. Offloading input of sensitive information onto a personal device. The user can decide which data is deemed sensitive. Image from [31].**

*LuxPass* [9] is a systems, which enables users to input a PIN on their trusted personal device and transfer it to a public terminal, using light encoding. The system could be considered relatively secure against shoulder surfers: Because of temporal difference between entering the data into the personal device and usage at the public terminal and the spatial proximity during the transfer, an attacker, even when equipped with a video camera, could not shoulder surf the data, but this has not been tested.

### 2.3.2 Exchange and Display of Personal or Sensitive Information

The above mentioned systems are primarily targeted at the input of sensitive data on a public display. As soon as the user is past the step of authentication, their data on the display is visible to anyone passing by. Several techniques have been proposed to allow users to keep their sensitive information safe, through usage of a separate handheld device.

Greenberg et al. proposed *SharedNotes* [24], a system in which people can move from individual to group work when working with notes on a public display. They allowed users to create and edit personal notes on their handheld device. These notes can then be exchanged with a public display whenever they feel ready for it, thus moving it from their personal space to public space. They also enable collaboration on the public display, by allowing the group to create and edit public notes on the display, while private notes can be edited using their personal device.

With *Digifieds* several ways to exchange sensitive content with a public screen have been proposed [2]. Users can create data either on a public display directly, using offered input technology, such as an on-screen keyboard. When desiring for more private means of data entry, one can use his personal phone or even create the data at his home PC and transfer it to the display using a QR code, alphanumeric code or by bumping the phone against the display. That way, authentication information could be prepared when one is certain of not being shoulder surfed and then transferred to be used on the display. In a similar fashion content can be retrieved e.g. by scanning a QR code, touching the display with the phone or using an alphanumeric code. In a user study they found that there is no significant difference in time needed to enter data using the phone when compared to the display. This gives a hint that secure data entry on public displays does not necessarily need to come at a cost of being less user friendly. Their qualitative data showed that users are concerned when entering data, such as their email address, on a public display, as it can be overseen by passers-by. They favored offloading the entry of such personal data (e.g. to the phone or home PC) to better preserve their privacy, rather than using an on-screen keyboard.

Several systems have been proposed to allow users to protect their sensitive data, when shown on public displays. Sharp et al. [49] allow users to censor private data on a public screen either through blurring or blacking out data. The content surrounding the pointing device will be revealed on the user's trusted personal handheld device, as shown in Figure 2.7, whereas the entire content is shown fuzzy on the large display. While the public display cannot be used to read the private data, they claim that it still provides contextual information, such as where scrollbars are and the positions of windows. They also allow users to enter sensitive data, such as passwords, through their trusted personal device.

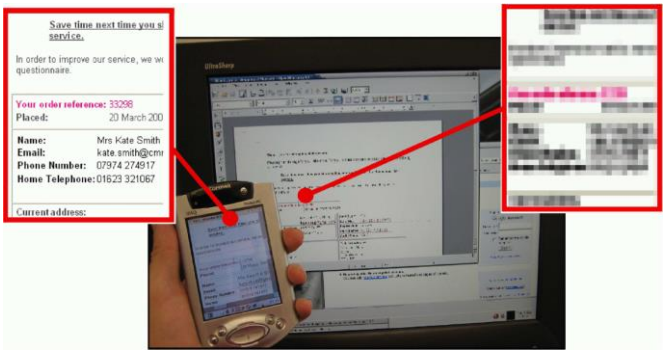


**Figure 2.7. The content is censored on the public display, while the area surrounding the pointing device is readable on the user's personal device. Image from** [49]**.**

Berger et al. [8] suggested to use *symbiotic displays* in order to black out sensitive words in an otherwise viewable document, being shown on a public display. The censored words can then be read on one's personal mobile device. A user could adjust his desired level of privacy, with the result of more or less data being readable only on the personal device. This approach relies heavily on text analysis. Their example used an email message, where the sensitive information are e.g. the sender, meeting locations which are mentioned in the email body and phone numbers. Again, the definition of sensitive information varies between different people (see section 3.3 for more details) and often fractions of an email message are enough to gain insights about the full content. When users want to protect larger portions of their work, this approach likely fails as either the entire display needs to be censored, making it obsolete, or a passerby can overlook more data.

Using the proximity between people and devices to interchange information and data with public displays has been formulated by Marquardt et al. in the *gradual engagement* design pattern [34]. They divided the interaction with the system into three stages: (1) awareness of the presence of a device and its connectivity, (2) reveal of exchangeable digital content and (3) transferring digital content between devices. Even though mainly about using proxemic information, they made some suggestions on how to ensure privacy in their settings. For example by leveraging a location's

context, as someone might be comfortable sharing his data in his own home but only wants to reveal its general availability when in his office. *Implicit* rules can define when someone wants to share information (e.g. when a device is in the user's pocket it stays private), whereas a user can always decide when he wants to stop sharing information by *explicitly* canceling a connection.

## 2.4 Territories, Personal Space and Proxemic Interactions.

The terms *territories* and *personal space* have first been used in zoological research, but can also be transferred to human behavior [3,55]. One's personal space can be described as the physical space surrounding him and the space that separates him from other people. Altman [3] and Sommer [56] describe the personal space as an *"invisible bubble"* (page 37 in [3]). They both note that one's personal space is different from one's territory: A person's territory usually refers to a fixed geographical location and rarely moves, whereas his personal space always surrounds him and stays with him wherever he currently is at. Subsequently one's personal space is always changing and dependent on the people around and the current context.

Edward Hall introduced *proxemics* as a theory describing how one's social distance is correlated to one's physical distance from another person [26] in everyday encounters. He described that people often use spatial relationships (e.g. orientation and distance) as a form to communicate their desired level of privacy. He observed that people associate certain social distances with a certain physical distance. As a result he defined four *proxemic zones* surrounding a person, which are shown in Figure 2.8:

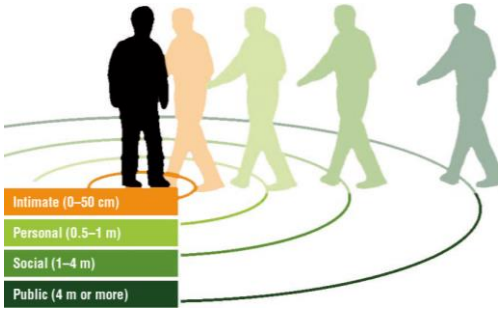


**Figure 2.8. Hall's proxemic zones, correlating physical distance with social distance. Figure from [36].**

- **The intimate zone** at the center. 0 cm – 50 cm surrounding a person. People can hear, feel and smell each other. People usually need a permission in order to enter that zone.
- **The personal zone.** 50 cm – 1.2 meters. At this distance physical interaction is possible, but not required. Conversations with close friends and family are often carried out at this distance.
- **The social zone.** 1.2 meters – 3.5 meters. For example conversations with co-workers. At this distance one has to speak louder and touching someone gets more difficult.
- **The public zone.** Every distance larger than 3.5 meters.

As the names imply, social engagement is expected to increase as one approaches the other. Hall's theory was exploited by Vogel et al. [61] to define four proxemic zones for large display interaction. These four zones are shown in Figure 2.9.

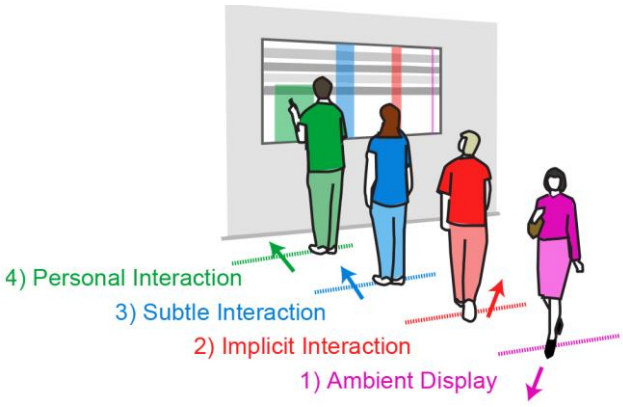


**Figure 2.9. Four phases of interacting with a public display. Figure from** [61]**.**

Vogel's system is an event calendar that shows both public and personal information. From afar, the display presents *ambient*, undetailed public information. Users are able to get a quick glance of the available information on the display. As one moves closer, the information presented and interaction allowed, become increasingly detailed and personal. In the *implicit* interaction phase the system recognizes the user's body position and orientation, inferring the user's openness to interaction. When the user engages in interaction, by getting closer to the display he enters the *subtle* interaction phase. The displayed information becomes more detailed. Previously displayed public

information, such as weather updates or publicly available calendar information, are augmented with private information, e.g. personal calendar events. In this phase users can explicitly interact with the display. This phase is only meant for short interaction and is followed by the *personal* interaction phase. Touch interaction with personal information is possible here. Their system shows details of people's personal calendar when getting closer. Some people might consider this information sensitive and would not want anyone else to overlook them (more on the meaning of what sensitive information is in section 3.3). If a second person enters the area, the display is split to provide each with an area to view their own personal information. As also observed by Altman [3], Vogel states that people can use body occlusion to shield sensitive content from eavesdroppers, thus being able to interact with more private content at a close distance. To further safeguard privacy, a person can perform certain gestures, or simply step back away from the display (into a more distant interaction zone) to hide or mute personal information.

Building on this work, Ballendat et al. describe the *Proxemic Media Player* [4], which incorporates people's position, orientation, movement and identity in order to control a media player. Their system detected the presence of people and to a certain degree their actions (answering a phone call, reading a magazine or talking to another person) and reacted accordingly: implicit interactions, based upon these actions are for example pausing a movie as soon as someone picks up his phone. In an explicit form, interaction is made possible by using physical objects, such as a pointing device, to interact with the system by e.g. selecting a movie and controlling its playback. They leverage people's identity for personalization (through personal media profiles), safeguarding (e.g. minors are not allowed the playback of horror movies) and history (the system automatically picks up the playback of a movie where a user last left). While they do not address privacy per se, they illustrate how the system can balance the needs of particular people in front of them. For example a person watches a movie when another person enters. To satisfy the entering person's desire to know which movie is currently playing the system could automatically display the title. When he gets closer to the display a short description will be shown, close to the position where he stands. All the while the first person can keep watching the movie. Therefore the system might take some automatic approaches of mediating collaboration on a large display. Figure 2.10 shows the three different stages of interaction with multiple users.

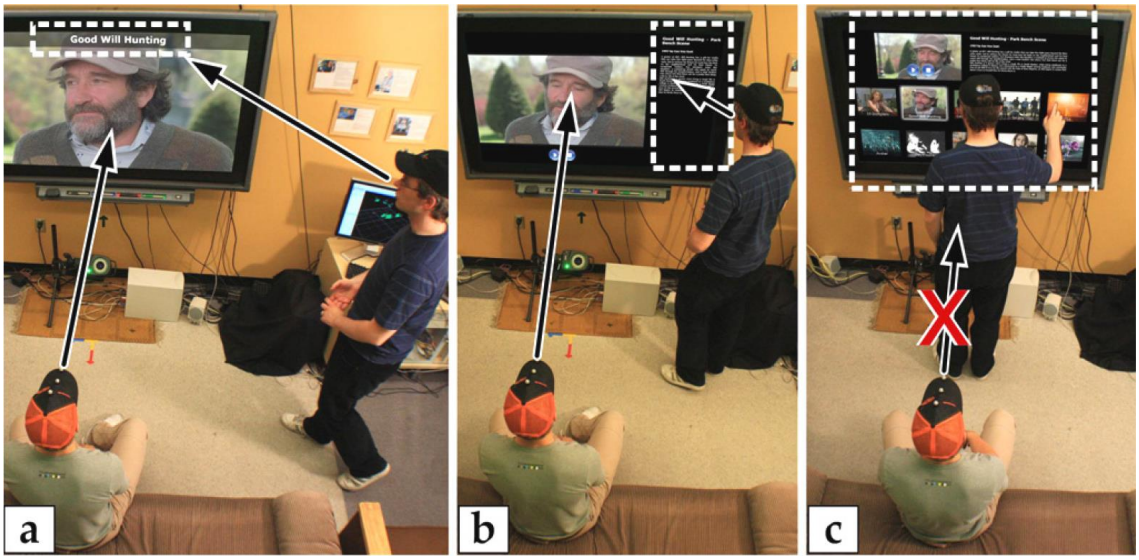


**Figure 2.10. Mediating between multiple people. a) While watching a movie, another person enters. The system can automatically displays the movie's title. b) As the person approaches the display, the system shows a description and allows for both users' interaction. c) The new user takes over the control, when he is in reach of the display. Image from [4].**

The *Proxemic Peddler* [62] by Wang et al. is a prototype of a public advertising display. It uses a person's identity, position and orientation in order to grab his attention, show personalized shopping tips from the Amazon.com website and allow interaction with the website as well as direct purchase. The suggestions are based on his shopping history and special offers the retail site is trying to sell. The various attention states, which are recognized, are shown in Figure 2.11). However, the definition of which data is considered to be of a sensitive nature varies from person to person [43] (a more detailed review on that can be found in section 3.3). Personalized shopping tips on a large, public display might be of interest for the merchant and in some cases convenient for the passerby, but it can cause a moment of embarrassment when certain shopping products are shown.

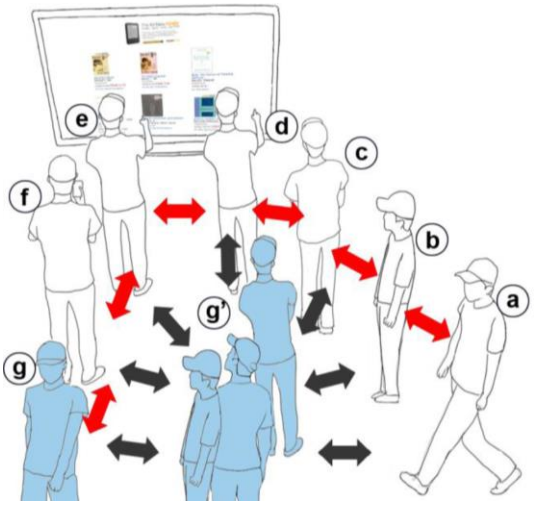


**Figure 2.11. The Proxemic Peddler displays personalized shopping information to a passerby. Figure from [62].**

## 2.5 Summary

Several techniques have been proposed to safeguard personal and sensitive information on public displays. These techniques range from special input technologies, as for example using a personal handheld device, to obfuscation through alternation of the actual data being entered, to using special display technology in order to limit what people can see. By splitting the content into a public and a personal view some systems try to combine the advantages of both: Having a large display to get a general overview by still ensuring privacy on a separate device.

All of these technologies assume to a certain extent that shoulder surfing is happening deliberately. Despite that, not only psychology research has shown that people usually respect other's territories and their personal space. Even when a privacy intrusion occurred, people are sensitive about it and step back as soon as they discovered that they are about to infringe someone else's space (more of that in the following chapter).

People's territories and their personal space have been exploited by various researchers in order to create interactive systems. Proxemic interactions have shown to help tackle the challenges in ubiquitous computing [25,36]. On the other hand, by using peoples' identity in an interactive system, privacy issues may arise as personal information becomes publicly visible. Many of these questions have been unsolved thus far.

# 3 Design Considerations

The above methods are based on simple assumptions of privacy. They treat all passers-by as suspected security threats. They have a strong notion that some information is clearly *private*, while others are clearly *public*. This section raises other considerations, which in turn provides a more nuanced design perspective of how one can mitigate shoulder surfing on public displays.

## *3.1 Privacy is a Boundary Regulation Process.*

In many cases, privacy is respected through a mutual negotiation of the parties involved [3,11]. People generally respect other's territories [3], and will do so if they are aware that they may be intruding across another person's privacy boundary [11]. For example, if they glance at a public display and see someone reading an email, they may self-regulate their behavior by looking away, or negotiate permission through social protocol [11]. Similarly, if a person realizes that another may be intruding, they will signal that.

Privacy regulation involves several different feedback systems, which are activated over time in order to achieve the desired level of privacy: from non-verbal cues to words to physical actions. Altman [3] lists the following behavioral mechanisms, which are used in order to communicate about privacy:

- **Verbal.** Through the content of speech (e.g. "keep out", "leave me alone", "join me").
- **Para-verbal.** Meaning the structure of speech (choice of words, grammatical structure, language style, etc.), and the way things are being said (voice quality, vocalizations (e.g. yawning, crying), pronunciation, etc.).
- **Non-verbal.** Such as the use of body language to communicate what one desires (e.g. blocking out the view of other people on one's work, or a head nod to invite people in).
- **Environmental behaviors.** Including clothing and adjusting one's personal space.
- **Cultural based privacy mechanisms.** As defined by the society an individual lives in (e.g. sharing a bedroom, or even a bed, is common in some societies).

Thus revealing *more* information rather than less can be a good privacy-preserving strategy: it enables mutual awareness of the situation, which in turn allows people to regulate their behaviors. Privacy abuses can either be deliberately or inadvertent as Bellotti notes [7]: If awareness is not sufficient, then privacy violations may occur inadvertently. People get defensive when someone breaks their personal boundaries, yet human territories and their personal space are usually respected by other people. The full range of reactions is not often needed [3]. Boyle et al. say that, when following social protocol / being rational, inadvertent privacy violators will look away to not further intrude someone else's privacy [11]. (For more details on human territories and personal space see section 2.4.)

## 3.2 Privacy as Social Distancing and as Territories.

People expect others to obey cultural expectations of social distance and proxemics [26]. In addition, people usually mark their territories through the use of symbols, objects and artifacts [3,46], which serve as further boundaries defining personal space. Upon infringement of either of them, people often react with resentment, anxiety or physical violence [3]. Altman [3] defines privacy on two different scales:

- **The desired level of privacy.** The level of ideal interaction a person wishes to have with another person or a group of other people.
- **The achieved level of privacy.** The amount of privacy an individual has achieved. This can be more or less than what was actually desired.

An optimum level has been reached when the desired level equals the achieved level. When there is a discrepancy, a person can either feel *crowded* or *lonely*. The first meaning that the achieved level is lower than the desired level (other terms are intrusion or privacy invasion). The latter meaning that the achieved level is higher than the desired level (other terms are boredom, isolation or intrusion). Figure 3.12 shows how the optimization process of privacy properties correlate.

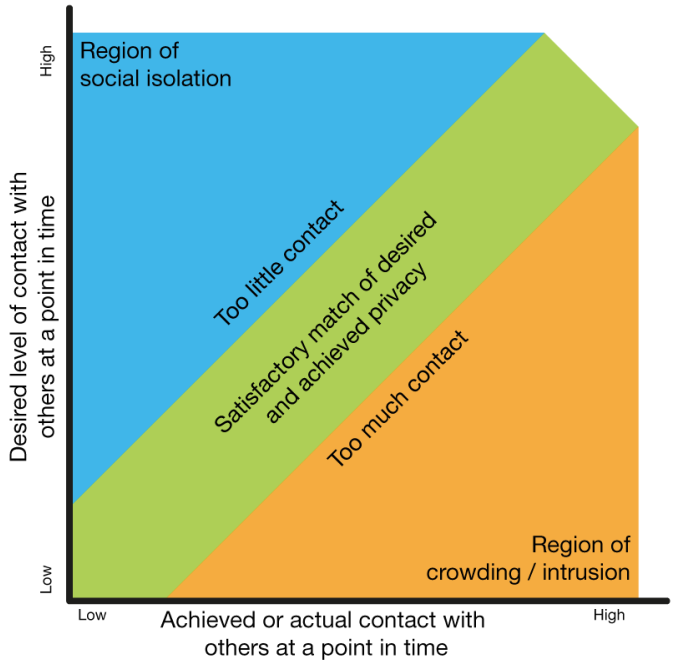


**Figure 3.12. Achieved *vs.* desired level of privacy. Figure based on page 26 [3].**

Similar to animal territorial behavior, marking of territories is also used by humans as a preventive function. Altman defines markers as symbols that help define the boundaries between the self and others [3]. Unlike in fauna, in the human world marking does not necessarily consist of marking places with body fluids. Barefoot et al. conducted a study [5] sitting at varying distances from a water fountain in a university. Their observations showed that a smaller percentage of passers-by used the fountain and they drank for a shorter period of time with the confederate sitting closer to it. The authors reason that people themselves act as a marker and that other people respect their presence, even in public spaces. In a similar study Sommer et al. investigated the effectiveness of various different markers. They found that personal markers, such as sweaters, jackets, etc. were more effective than less personal ones, such as a publicly available journal or brochure [54]. In a later study Becker confirmed these previous findings [6]. He further states that an actual person serves as the strongest signal of territory, compared to other physical objects, and an increasing quantity of markers also serves as a stronger signal to passers-by. Edney examined the use of markers to establish an even bigger territory. He found that home owner's tend to build fences, plant hedges or put up "no trespassing" signs with a longer lasting commitment to their place [18].

The territorial behavior has not only been observed in psychology research (e.g. by [3]) but also

in several more recent studies in the area of human computer interaction, some of which will be listed below.

When working on a whiteboard, people often take turns and wait for one another to finish their work before picking up or continuing with their own work [60]. They also found that people use the area in front of them as a personal space to store or collect information which is not shared. Kruger et al. [29] got similar results for work on a physical table. Scott et al. observed that people leverage territories when working on interactive tabletops [46] and organize their workspace into personal, group and storage territories (Figure 3.13) [47]. They observed that people understood and respected others personal territories without any communication needed, and people rarely performed any actions outside their own or the group's territory. Marshall et al. observed in an in-the-wild study that often social discomfort arises from intrusion of one's personal space [37]. They confirmed previous findings that people rely on body language and verbal cues to negotiate about their territories. In their study, as a last resort people left the public tabletop when they felt their personal space too much encroached.

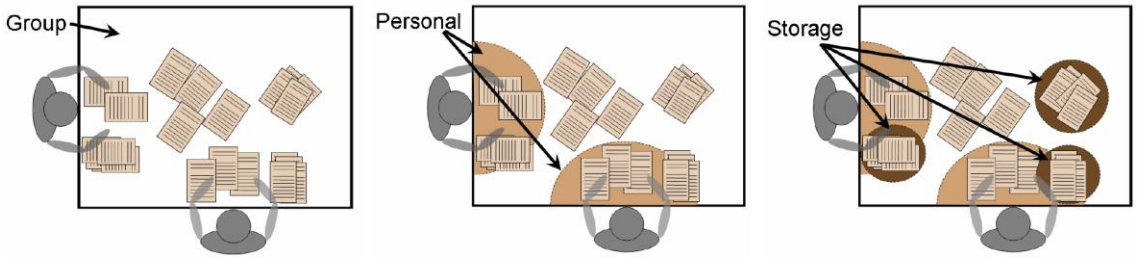


**Figure 3.13. Three different types of tabletop territories. Figure from [47].**

These studies show that humans respond to other people's markers and that they respect their territories. People use different means to acquire their desired level of privacy and physical means are usually not needed. Edney [19] stated in a later article that territories result in a more stable society as they serve as means for social organization and have a regulatory role to individuals, groups and communities. Social interactions, relationships between people and ownership of property are all negotiated by means of territories. Altman [3] carries this point to an extreme by saying that with perfectly set up and obeyed territories a social system will not fail.

People have a strong notion of personal space and territories. For one the mere presence of a person acts as a marking of a territory. Furthermore people utilize personal objects, such as books, jackets, bag packs, et cetera to occupy a territory. Social norms usually operate well and only in rare occasions people have to resort to physical actions in order to defend their territories. They respect social distance and territories of others. If someone breaks into another's personal territory or space, various protection mechanisms then come into play to negotiate about their territories, most of the times in different means than physical ones, for example by expressing their desire verbally, through eye contact or through their acting behavior (more details in section 3.1).

The problem is that a large display can change the dynamics of this process. Because the display and contents are visible at a distance, territorial boundaries and the size of proxemic zones can become ambiguous.

## *3.3 Private / Sensitive Information.*

The meaning of private and sensitive information varies between different people [43]. At one extreme, some people have little concern about their information (e.g., they may only be concerned about banking information). At another extreme some are highly sensitive to any information disclosure (e.g., routine purchasing behaviors). Of course, it also depends on context and what activity people are involved in. Therefore an automated system cannot successfully predict what data should be protected from shoulder surfing, as it is highly personal and context-dependent. Information can be categorized in three different categories, depending on the person it might have a sensitive meaning to:

- **The user of the information.** The most direct case is when someone wants his own personal information to stay private. For example this might be a person standing in front of a public information terminal, looking for directions to a clinic for sexual transmitted infections or someone reading his personal messages on a smartphone in public. This data can be easily overlooked and it is in the interest of the user himself that this does not happen.
- **The person represented by the data.** The user of the information and the person represented by the data do not need to be the same. There can be a divergence who might find this information be worth of any protection. An example would be a patient's record in a hospital or the financial information of a bank's customer. The doctor, looking at the records, or the banking clerk, checking for a credit score of a customer, might not consider this data private, as it does not affect him personally if someone else sees this information. The patient or customer himself very well wants this information to stay protected from third parties.
- **The passerby who oversees the information.** A passerby might not want to see private information or personal data of someone else. People are rational, meaning they want to protect someone else's information in a similar fashion as they desire protection of their own data [11]. From another perspective, they might not want to see the information as it might offend them. As an example serves the following: In the United States (because of the First Amendment, "freedom of speech") many public libraries and internet cafes allow the consumption of porn on the public computers and the usage of websites, listing personal ads including pornographic images. While the patrons, watching these, might not care whether they are being overlooked, worried parents might not be aware that their children, while at the library, can oversee graphical content. To tackle this situation, the San Francisco library installed plastic blinds to cover the screens [38], allowing their customers to consume whichever content they desire, while keeping it private from others.

On top of being private, information also has another property: It can either be *personal* or *public* as Greenberg et al. state [24]. Personal artefacts are ones that a user creates and keeps for himself, they *can* contain private information. Public artefacts on the other hand are owned by a group, usually of collaborator. In everyday life artefacts often change their property of being personal or public. For example, people prepare information before bringing it to a group meeting and exchanging them with others. Further, private information does not necessarily have to be private only to one person, it can also be private to a group of people [50].

Note, that these categories are not mutual exclusive though, as there can be overlaps. In all of these categories it is someone's interest that some information is kept from illegitimate eyes. For this someone first hast to know that a privacy intrusion occurs, meaning either he is infringing someone else's personal space or that someone is overlooking this data.

## 3.4 Public Displays: Shoulder Surfing and Honey-Pot Effect.

Shoulder surfing is the act of overlooking someone's data, via direct observation techniques (e.g. looking over someone's shoulder, using a camera, etc.). This is problematic when it comes to passwords, PINs or other sensitive information. Shoulder surfing often arises from curiosity rather than malicious intent, as people usually respect other's privacy and territory [3,11]. The easier it is to shoulder surf, the more likely it is that someone will do it. Tan et al. [58], for example, observed that people tend to be more voyeuristic with increasing display sizes. With a stem completion test they observed whether participants performed better when they subconsciously saw the priming words on a large

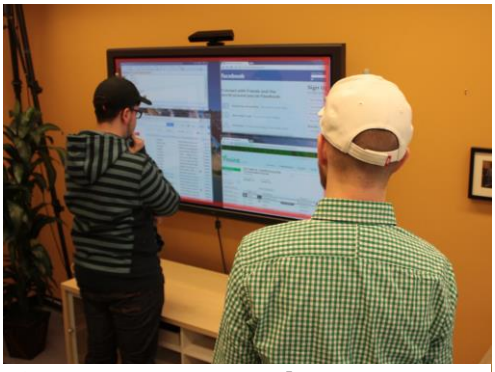

**Figure 3.14. People are more voyeuristic with increasing display sizes [58].**

display first. Their results show that participants named significantly more stems from the target words, when they have seen them on a large display instead of a small display, even if the contents of the screens was not supposed to be their primary concern.

Often people feel a barrier to interact with public display systems resulting from social embarrassment, and their conceptions of a system on a public display influences their feeling towards interacting with it, as Brignull et al. observed with their *Opinionizer* system [11]. On the other hand, they observed the *honey-pot effect* on public displays, showing the increased likelihood of being overlooked: People tend to interact with a display more likely when someone else is already working on it. Peltonen et al. also observed that people are more likely to interact with a display if someone else is already working on it and thus confirmed the honey-pot effect in their *CityWall* project [45]. On their display people could explore photographs, taken from an online photo platform. It was 2.5 meters wide, allowing for multiple users to interact with it at the same this. They also noted that the new person usually observes the existing user for a while before beginning to interact with the display and multi-user interaction was the most common type of interaction with the display. Even strangers worked together in order to achieve a goal, such as rotating a picture, although people usually respected others work and territory and tried to stay out of other users' way (for territorial behavior and social distancing also see sections 2.4, 3.1 and 3.2). When there was a conflict, passers-by negotiated about it, e.g. by taking turns when the space in front of the display was or when they wanted to work on a part of the display which someone else already occupied. Michelis et al. reported with their *Magical Mirrors* (passers-by's bodies are mirrored on large displays, and they could interact with the display, using gestures) that the honey-pot effect not only exists for multi-touch or keyboard interactions but also for gesture-based interactive public displays [39].

The above examples show that, when working on a large display, it becomes a honey-pot that attracts others to one's work. This is not necessarily a bad thing, for it could also encourage collaboration and engagement. Similar to Brignull [12], Müller et al. [40] state, this honey-pot effect is a very powerful cue to attract attention and increase engagement with public displays (see Figure 3.15).


**Figure 3.15. The honey-pot effect: Passers-by notice gesture interaction of a person in front of a public display. On the other hand they do not want to be seen by the system as they often try to hide from the camera's field of view. Image from [40].**

Koppel et al. [28] compared a combination of several screens. They setup their *Chained Displays* in public, with three different form factors: Flat, concave and hexagonal. Observations showed that the honey-pot effect is strongest on flat displays, because people in front of the display and their actions and effects can be observed by passers-by. Also the flat setup triggered the highest number of concurrent actors. Flat display setups are probably the most common formations found in public space nowadays.

## 3.5 Summary

Most shoulder surfing on large displays will not arise from malicious intent, as people are usually sensitive to other individuals content. Although the perception of sensitive information varies from person to person, people have a good feeling about when they are infringing someone else's personal space. Nevertheless there are several social behaviors, which might result in one's sensitive information being overlooked by other people. Behaviors such as voyeurism, the honey-pot effect, and territorial and spatial ambiguities have been observed. Inadvertent violations may result from the shoulder surfer not realizing that he or she is viewing sensitive (*vs.* public) information unless it is too late.

Consequently, systems that are supposed to mitigate shoulder surfing on public displays must meet two important criteria.

First, the system should make the passerby and the user of the display aware that shoulder surfing could occur or is occurring. If done well, both parties can regulate their behaviors via social protocol.

Second, the system should provide some degree of shoulder surfing protection over broad content (rather than about small units of information) until privacy is negotiated.

# 4 Awareness of Shoulder Surfers

Awareness can be provided to the user of a public display that someone is nearby and that his screen content might currently be shoulder surfed, thus having more information about his environment. On the other hand awareness provided also serves the passerby as an indicator that he might be infringing someone else's territory. When provided with awareness, a user can then decide whether his displayed data needs protection, and use social behaviors to regulate privacy. For example, he can ask the passerby (either explicitly or implicitly through body language) to respect his territory by not looking at the display. He can also hide his private data by either closing the application or covering sensitive information with his body. If need be people resort to physical means of defending their personal space. All of these techniques have been formalized, among others, by Altman [3] and Boyle [11] and are explained in more detail in section 3.1.

The general approach in this work uses visual indicators on the display to provide cues that another person is passing by or that someone actually is shoulder surfing, overlooking his screen content. While these indicators primarily inform the user that a passerby is present, they also provide the passerby, glancing at the display, with an indication that he may be intruding into the user's territory. When being rational, inadvertent privacy violators will then look away to not further intrude the user's privacy [11]. As the examples below illustrate, cues can range from abstract ones that provide only general awareness information, to literal and very precise cues that give fine-grained awareness of the passerby's whereabouts, movements and look direction. In all these cases the system does not take any automated means to provide protection, the mediation relies solely on the user.

Although many sketched figures are shown here, all of the proposed techniques have been implemented, details about the implementation can be found in chapter 1. Photographs of the working system are either shown with the description of the actual technique or can be found in Appendix A. A video, demonstrating some of the aspects has been accepted to the CHI'14 Video Showcase and will be presented end of April 2014 [14]. Some of these techniques have also been published as a Tech Report and are currently in submission [13].

## *4.1 Flashing Borders*

The simplest cue uses flashing borders. When a user is working on a large, public display ( a) he might not be aware of the presence of a passerby. When delved deep into his work, even audial cues of e.g. a passerby's footsteps might not have the power to make him aware of his presence. As soon as a passerby enters the visible area around the display, the system notifies him of the passerby's presence by flashing its borders. In its simplest case, the awareness provided only covers that someone is somewhere, but not the person's whereabouts and whether that person is actually looking towards the display. The cues can be enhanced with meaningful colors. The borders flash green when someone is nearby but not looking at the display ( b). The color transforms to red as the passerby turns his head towards the display ( c). To further emphasize the possibility of a territorial encroachment, the red borders flash more rapidly when someone is looking at the display, being less ambient and subtle. The relative distance of the passerby is coded into the transparency of the border: as the person approaches the display, the border color becomes increasingly opaque, adding another parameter of awareness. The position of the passerby can also be indicated by coloring only the sides and center / side border to roughly mirror

that person's location.

**Figure 4.16. a) No indication is being provided when there is no possibility of territorial encroachment (left). b) The borders of the display are flashing green when a passerby enters the visible area of the display (center). c) The color changes to red when the passerby is facing the display (right).**

### Discussion

These cues, while simple, can provide significant awareness information. With the green border, the user knows that someone has entered the scene but is not yet looking at the display. He can then take advanced action to mitigate the potential threat, such as by hiding privacy-sensitive information, or by signaling the other person that privacy is desired. When the user knows that someone is actually shoulder surfing (the red border), his actions can be even more decisive. The distance and location cues (transparency and border side), while approximate, provide the user with a sense of whether the passerby is moving through the area, has stopped, or is approaching the display. Because the border fades in and out (the flashing), the transparency does not have an absolute state indicating when someone is the closest or furthest away. Over time he might be able to tell a difference in the transparency despite the flashing. However, the abstract nature of these cues likely make it inappropriate in walk up and use settings, as neither the user nor the passerby will know what the flashing borders mean unless they are somehow taught it.

## *4.2 Mirroring the Passerby as a 3D-Model*

Awareness cues can be very precise, where they accurately portray the actual location and orientation of the passerby. This information is supplied via a mirror effect, where the passerby's relative location is portrayed as a 3D-model on the screen (Figure 4.17). When a passerby enters the display area, a 3D-model appears on-screen, where its position mirrors that of the tracked passerby relative to the display (Figure 4.18). As the person moves across the room, so does the model. The model's size changes with the distance of the passerby to the display, where the model increases in size as the passerby approaches the display. Additionally, the orientation of the model's head and torso are independently mapped to the tracked head and torso position of the person. For example, if the passerby turns his head (but not his body) towards the display for a quick glance, the model reflects that: the torso remains in its ~$90^0$ orientation from the display, while the head animates to turn towards the display (Figure 4.17). The model's transparency offers a further cue indicating how the passerby is attending the display, where the model becomes increasingly solid as a function of both distance and orientation of the passerby's head. It becomes more opaque when the passer-by is close to the display. When the passer-by turns his head away from the display the model becomes more transparent.

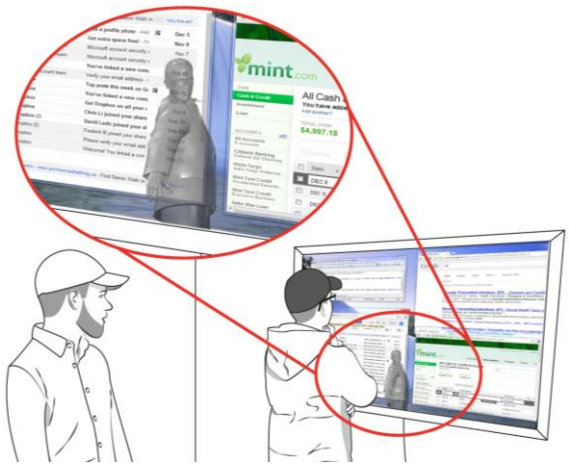

**Figure 4.17. Mirroring a passerby's position and orientation with a 3D-model. The model's head rotation correlates to the passerby's head rotation, as does the rotation of the torso.**

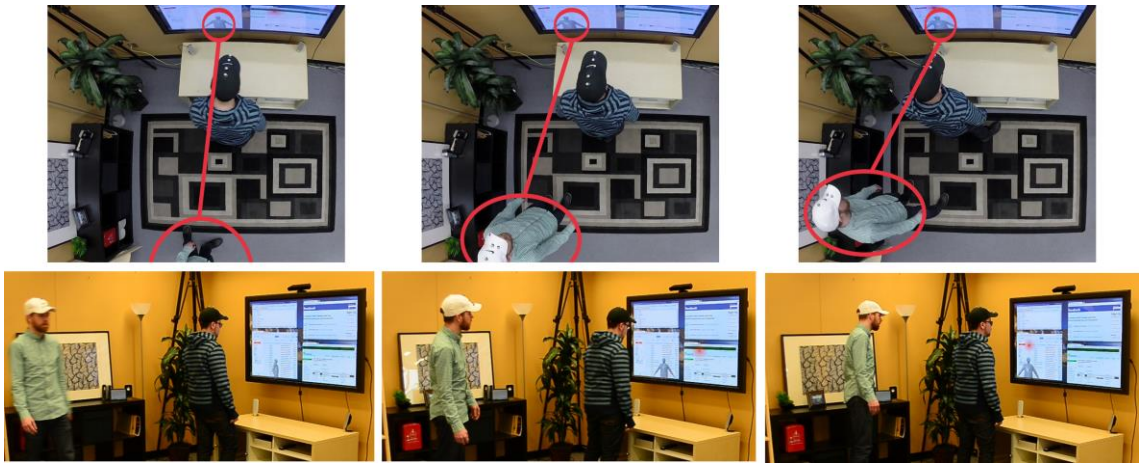

**Figure 4.18. The 3D-model follows the passerby's position in the room, tracked by a 3D motion capturing system.**

### Discussion

Unlike the abstract flashing borders, people can quickly comprehend that the model is mirroring the passerby, and understand its spatial relationship. The model informs the user not only of the passerby's presence, but also his position, distance and look direction in a natural manner. It gives a full indication of a passerby's current whereabouts and look direction. The information provided is not actual, e.g. no numbers are provided of how far a person is apart. Nevertheless a user understands distances without numbers, as they can be estimated from the size and opacity. Because the model is very responsive and animates in direct correspondence to the passerby's movements, the user can easily tell if someone is moving through the space, or has stopped, or is approaching, or is just giving a quick glance at the display, or is staring at it. Similarly, the passerby will see themselves on the display, and will understand that they have somehow intruded in the user's space by becoming part of it. Both parties can then act on this information as needed: The user can either reinforce his desired level of privacy or invite the passerby to work with him, whereas the passerby can look away or leave the area.

## *4.3 Gaze Awareness Indicator*

Another visual cue indicates where on the display the shoulder surfer is gazing, i.e., approximately what they are looking at. This cue is realized as a red fuzzy dot, which moves about in a manner somewhat similar to how eye-tracking systems portray eye-gaze direction. Because no eye-tracker is used in this system, the viewing direction is assumed from a person's head orientation.

In particular, the passerby's tracked head position and orientation are considered as a vector and

its intersection with the plane the display lies in, is calculated. Resulting in a point which marks the passerby's current viewing position on the screen. By displaying this point (Figure 4.19), the system gives the user, information about what part a passer-by is actually looking at. The size of the red dot is a function of the distance of the passerby to the display: The closer he is to the display, the smaller the red dot will be.

**Figure 4.19. The red dot indicates the gaze direction of the passerby.**

## Discussion

The gaze awareness indicator provides reasonably precise information about what screen region a passerby is likely looking at. Although previous research has shown that a person's head orientation somewhat correlates to his look direction [22,41,42,57], the current head-tracking implementation means that a shoulder surfer can trick the system by looking at the display from the corner of one's eye. This is why the gaze indicator should be seen to best be used in combination with other cues, such as the 3D-model, that gives additional information about what the passerby is doing.

Having the size of the red dot decrease with the passerby's decreasing distance might be somewhat confusing at first sight. The assumption here is to have the gaze indicator shrink when a passerby approaches the display, as with a smaller distance he is not able to oversee the entire display content without turning his head. From far away one does not have to turn his head in order to see the entirety of the display, therefore the size of the red dot is larger when the passerby is further away as he potentially oversees more content.

# 5 Providing Protection

Awareness is just the first step in helping the user protect his privacy, or in informing the passerby that he or she may be violating the user's territory. This may suffice for many situations. When a territorial violation appears imminent, people normally self-regulate their behaviors to resolve the issue (e.g., where the passerby simply turns away), or enter in some kind of signaling and direct communication to negotiate access [11] (for details see chapters 2 and 1). Yet there are times when further protection is needed. For example, when even a quick glimpse of the display by the passerby may compromise one's privacy. Or, the user may want to take explicit action to safeguard sensitive information, perhaps because the passerby is just too curious. Or, because the user does not wish to socially engage with the passerby, as for some people it can be distressing to socially engage with other people and they might prefer means of protecting their data, which involve less personal interaction.

In this section, it is shown how sensed information about people can be exploited to provide both *explicit protection* (a user can take quick action to gain protection when he or she becomes aware of a potential violation), and *implicit protection* (the system triggers protection when it senses a potential violation). He can either opt to have protection on a window based level or protection over broad content, until his desired level of privacy is negotiated. As described in section 3.3, private data has a different meaning to different people. Therefore a user of the system should have means of defining what he deems to be private.

## *5.1 Definition of Private Content*

Some of the systems described in this chapter act upon a user defined level, some provide protection over a broad range of information. For the first, a user first needs to define a privacy level for each of his applications. Three different levels are distinguished, as drawn from Altman's theory of desired levels of privacy [3]. A user can apply one of these levels, by dragging a selector to each application window (Figure 5.20 and Figure 5.21).

- **Green level, the public level.** The default level. Applications with the green level are visible for anyone. Examples of this are public always-on application on public displays, such as weather information or bus schedules.
- **Yellow level, the semi-private level.** Applications with this level contain sensitive information, but the protection of these windows will be left to the user. Only when the user turns away from the display or leaves the room the system will take over and protect these applications upon the presence of a passerby. Examples might be a private photo album, where a user would not like a passerby to snoop around when he is not around to regulate his actions.
- **Red level, the private level.** Applications with the red level will be protected from being visible to passers-by. When automatic protection is active, the system will try to hide content of those windows and not only provide awareness. These applications potentially contain privacy sensitive information to the user, which he does not want to be visible to other people. An example would be a user's banking information or his personal email account.

After the user is done with his selection, the covers disappear, to not further distract him from his work. If no selection is made for a particular window, the system falls back to either the green default level or, when desired, a keyword based approach: In particular, the application window's title is used to decide whether a window requires protection, e.g. a user deals with banking

information (e.g. keyword search for “bank”, “https”) or when he searches for health related information on the web (e.g. keyword search for “sti[2]”) which he does not want to be visible to someone passing by.

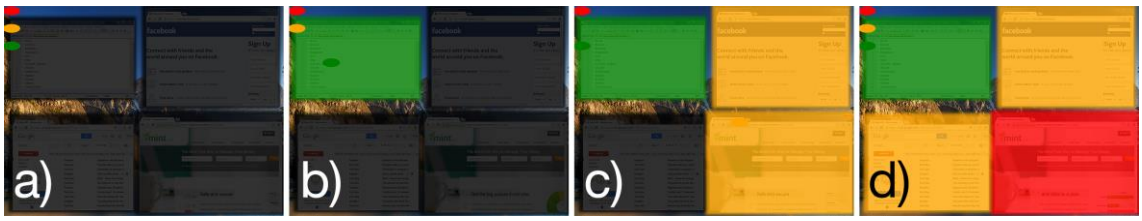


**Figure 5.20. A user can select his desired level of privacy for each window, by dragging a colored ellipse on the windows. a) All windows are in default / unknown state. The system falls back to keywords e.g. in the window’s title when deciding about a privacy status. b) The user explicitly set a public state for one window. c) Two windows have been marked as semi-private. d) The red color indicates that one window has been marked as highly sensitive.**

**Figure 5.21. A user defining the privacy level of his applications.**

## 5.2 Explicit: Moving or Hiding Content

When a person becomes aware of a shoulder surfing risk (e.g. through the awareness providing techniques presented in chapter 1), he may want to take action to mitigate that risk. Shielding sensitive data with one’s body is one such action in everyday live [3]. Yet shoulder surfing is a bigger risk, with increasing display sizes [58] (for more details see section 3.4). Because users typically spread application windows over the entire display area, shielding may be difficult or impractical in large display or multiple monitor settings. Alternately, the user may move, resize, hide, or even close windows containing sensitive information. However, conventional interface mechanisms require this to be performed one window at a time, which is a slow and tedious process.

Following the approach of Vogel et al. [61], in which a user can quickly invoke an action to safeguard privacy, the system’s particular safeguards allow the user to quickly move all windows to a portion of the screen directly in front of him. The first action is based on explicit gestures: the system recognizes a user’s hand wave as a command to gather all applications in front of him on the display. As a result it is possible for him to shield them with his body. He can also hide windows until privacy intrusion is no longer a concern. A sequence, illustrating this gesture is depicted in Figure 5.22. The second action is based on user orientation: the system recognizes when the user turns away from the display (for example, turning to face the passerby) and hides all windows by blacking out the screen (Figure 5.23). Both actions are quickly reversible, e.g., by the user waving his hand in the other direction to spread out the windows, or turning back towards the screen to reveal the windows.

[2] STI stands for Sexual Transmitted Infections.

**Figure 5.22. The user performs a gesture to gather his applications in front of him, thus being able to cover them with his body from a passerby's view.**

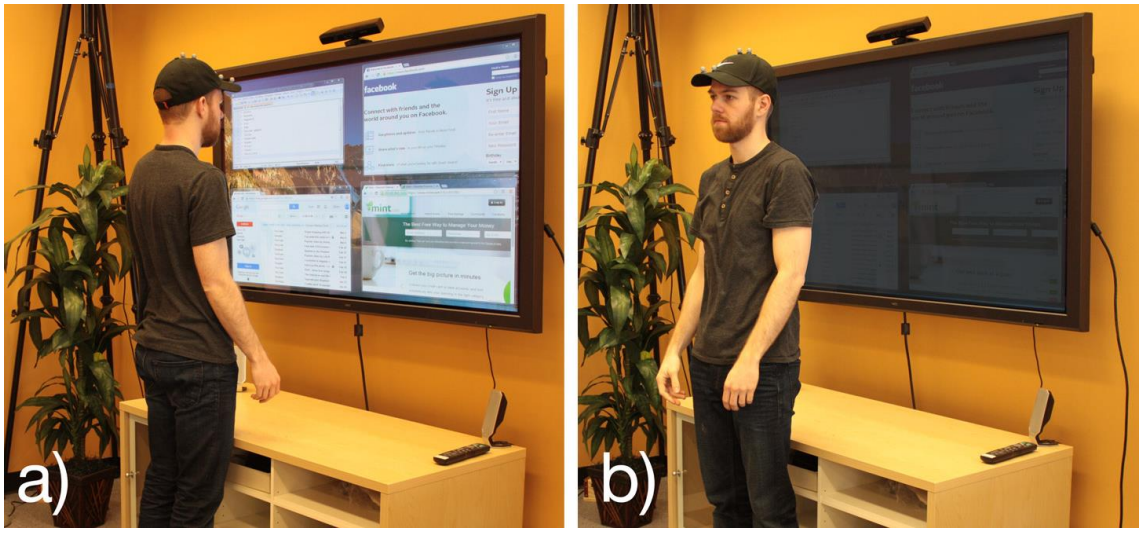


**Figure 5.23. The display content blacks out as soon when the user is not looking at it.**

### Discussion

These techniques not only protect information, but reinforce how the passerby understands a user's territoriality. The passerby sees information being moved or hidden as a result of a user's action, which feeds into self-regulation and further negotiation. The downside is that explicit action takes extra work, and that the resulting window re-organization (or hiding) can disrupt what one is doing.

Easy moving of windows serves a dual purpose, where it can not only protect sensitive information, but also encourage sharing and collaboration rather than protection. For example, if the passerby wishes to use the public display for his own purposes (assuming the current user invites the passerby to do so), moving windows to one side of the screen frees up space for both to work side by side. The quick reversing of the actions taken by the user also allows him to invite other people to collaborate or share information with them, when he feels comfortable.

## *5.3 Implicit: Blacking Out Sensitive Content*

Because the system can implicitly recognize potential shoulder surfing moments based on the passerby's relative position and look direction, it can take action to shield sensitive information from view. Ideally, the information will remain visible to the user but not to the passerby.

The implemented system does this on a window-level, where particular windows are tagged as public *vs.* personal. For example, the system may know what public windows it has provided (e.g., always-on public weather updates) *vs.* personal windows (e.g., ones the user has created, or has somehow marked as sensitive; details on how the user can teach the system what applications should be treated as private *vs.* public can be found in section 5.1). Or, the system may keep a list of applications that are privacy-sensitive, such as an email reader, or search for keywords that

identify sensitive content (e.g., “bank”, “mail”, “https”). The system then covers each of the windows, where it tries to strike a balance between masking the window’s contents from the passerby, while still making it legible to the user using transparency of the cover. Windows are fully visible when no passers-by are present. As a passerby enters the area at a distance, the transparency levels of private windows are set to make them hard to read from afar but easy to read by the user (who is close to the display). Figure 5.24 portrays this situation: Three of the four windows have been detected to be private by the system (either because the user has set their privacy level or because of a keyword analysis), thus overlaid with a black, semi-transparent cover. Opacity changes (and thus window legibility changes) as a function of the passerby’s distance and viewing direction: the closer the passerby gets to the display the more opaque the private windows become. Similarly, when the passerby turns his view away from the display, those windows become more transparent, as they cannot be overlooked as easily.

**Figure 5.24. Blacking out sensitive application. The opacity of the cover is set so that it is still readable when standing close, but difficult to read from a distant.**

Every implicit system’s action can be overridden on a user’s demand, e.g., by un-hiding windows using an explicit hand-wave gesture as described in the previous section 5.2. The user may want to do this for various reasons, such as inviting a colleague into collaboration. For example when going through his banking information, a user wants protection from his friends as he is not willing to fully disclose his financial status to them, whereas for his spouse it is perfectly fine to see everything. Thus the overall strategy is one where the system tries to automatically protect sensitive content (to mitigate privacy intrusions), but allows the user to easily override the system.

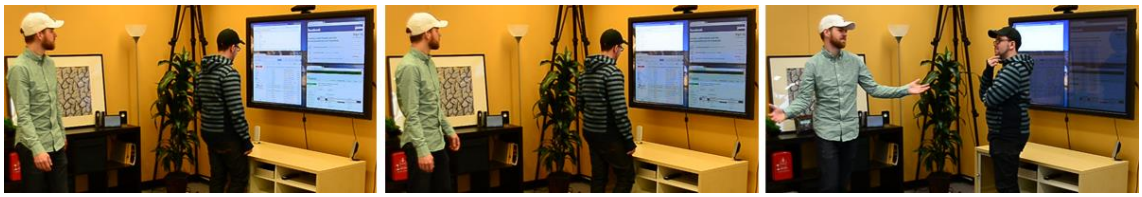

**Figure 5.25. Private applications are hidden when a passerby enters the visible area of the screen. The user becomes aware of it and can renegotiate his desired level of privacy.**

### Discussion

Blacking out of selected content based on inferences of privacy incursion is a somewhat radical approach. By using a semi-transparent cover the user is still able to see the content underneath allowing him to continue his work, while it is not as visible from a distant. Its advantage is that it not only offers protection, but it also clearly marks a potential privacy intrusion to both passerby and user. Another advantage is that public information remains available to the passerby (e.g., a public window showing the time or weather would remain visible). However, the particular implementation is not a sure-fire safeguard of privacy. First, it is difficult to balance occluding personal information from onlookers while still making it visible to the user. The strategy of occluding personal and sensitive information does not provide full security from intruders. However it can provide some protection [61]. Second, it requires that the system somehow 'knows' the difference between sensitive *vs.* public content. As mentioned in section 3.3, the definition of what data is "sensitive" varies between different people, and an automated system can never predict with 100% certainty whether data needs protection or not. For that reason the system allows to use both, automated as well as manual means of selecting privacy relevant applications (see section 5.1 for details).

Furthermore, blacking out the windows provides not only protection, but also strong awareness to the user that someone is currently overlooking his work (Figure 5.25). He instantly knows that a possible intruder is nearby and can take further actions if desired. On the other hand this technique provides awareness to the passerby that he might have just walked in on someone's sensitive information. He then can take action to mitigate that problem and react in a socially accepted way.

## *5.4 Implicit: Silhouette Protection*

Because the system recognizes the spatial relationship between the passerby, the user and the display, it can roughly calculate what part of the display is shielded from view by the user's body (Figure 5.26). It can then use that calculation to black out (again via appropriate transparency levels) the areas of the screen visible to the passerby, while leaving the area shielded from view (through the user's body) visible to the user. That is, if the passerby is considered as an inverse light source, the user working on the display casts a 'shadow of visibility' onto the screen (Figure 5.26b green line), which is called a *silhouette*. The rest of the screen becomes muted using a black cover with appropriate transparency levels, where it too tries to strike a balance between hiding the content from the passerby's view (Figure 5.26b red line) while keeping it somewhat accessible to the user. The opacity of the silhouette is a function of the passerby's distance and look direction. The silhouette disappears entirely when the user turns away from the display, leaving a black screen behind.

The animated silhouette moves when either the user or the passerby moves, reflecting the changes in the area that would otherwise be visible to the passerby. The size of the silhouette changes as a function of the passerby's distance to the user: With decreasing distance the size of the visible area decreases, reflecting the smaller inverse *shadow of visibility* cast by the user on the display (Figure 5.27).

The silhouette is calculated by creating a vector, based on the position of the user and a passerby (Figure 5.26b, green line). Extending the vector results in the intersection point with the display, which is being use as the center-point of the silhouette. The silhouette's width (Figure 5.26b, white area) is a function of the distance between the user and a passerby. The vertical position and height of the silhouette on the display is based on the sensed height of the user.

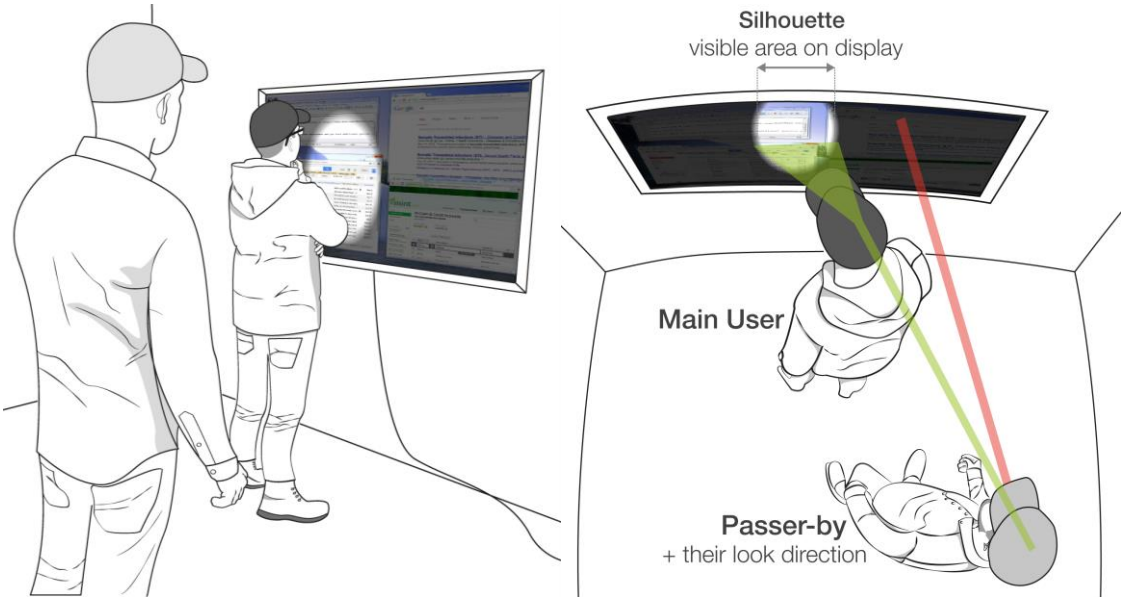


**Figure 5.26. a) The silhouette reveals only those parts of the display shielded from view by the user's body, allowing him to continue his work (left). b) The silhouette's area is calculated as a function of the vector between the passerby, the user, and the display, indicated by the green line. The red line indicates the passerby's viewing direction (right).**

**Figure 5.27. The silhouette moves and its size changes according to the position and distance of user and passerby. Here, the red line shows the passerby's look direction, whereas the green line shows which parts are covered by the user's body.**

## Discussion

Unlike the 'blacking out of sensitive content' approach, the system does not need to know what content is private *vs.* public. The silhouette acts on a physical metaphor, where it covers only those parts of the screen that can be overseen by a passerby. With the silhouette not all of the screen's content is visible to the user all the time. Because part of the screen's content is muted (especially if the passerby moves close to the display), it becomes more difficult for the user to employ the full display for his work. It tries to minimize interruption, as the user can continue to work on the visible area (which typically remains in front or close to one's body). Since the cover's transparency changes with the distance of the passer-by, the user, standing close to the display, is still able to see the covered contents as long as the passer-by is just walking by in a distant. The passer-by on the other hand, cannot tell the contents of the screen, as it is distinctively harder to read even just lightly covered data from afar. As with the animated 3D-model, the visuals are easy to understand by both user and passerby, making them both aware of possible intrusions. The silhouette also provides strong awareness of which parts of the screen a user currently shields with his body. It further provides awareness to the passer-by: He sees that he might be infringing someone's territory and can therefore refrain from closing in on the display.

# 6 Implementation

While in the previous chapters much of the implementation has been unmentioned, this chapter will give more details, explain how the system performs both, awareness and protection at a lower level.

The system is implemented in C#, using the .NET framework with Windows Presentation Foundation (WPF) for the graphical user interface. A Vicon motion tracking[3] system is used to track entities in a room. The Proximity Toolkit receives the tracking data from the Vicon system, encapsulates them to make them available via a TCP connection in a C# .NET program. The toolkit also allows to follow the proxemic relationships between two entities via asynchronous event notification.

## 6.1 General Tracking

The Vicon motion tracking system works by tracking, infrared-reflective markers, which are attached to various objects, using infrared cameras. In this particular system the objects are two baseball caps, and a vest. The baseball caps are being worn by the user, standing in front of the display and the passerby. They are used to track a person's position in the room and their head orientation. A view from one of the tracking camera's perspective is shown in Figure 6.28, highlighting the tracked information. The vest is being worn by the passerby, in order to track his torso's orientation.

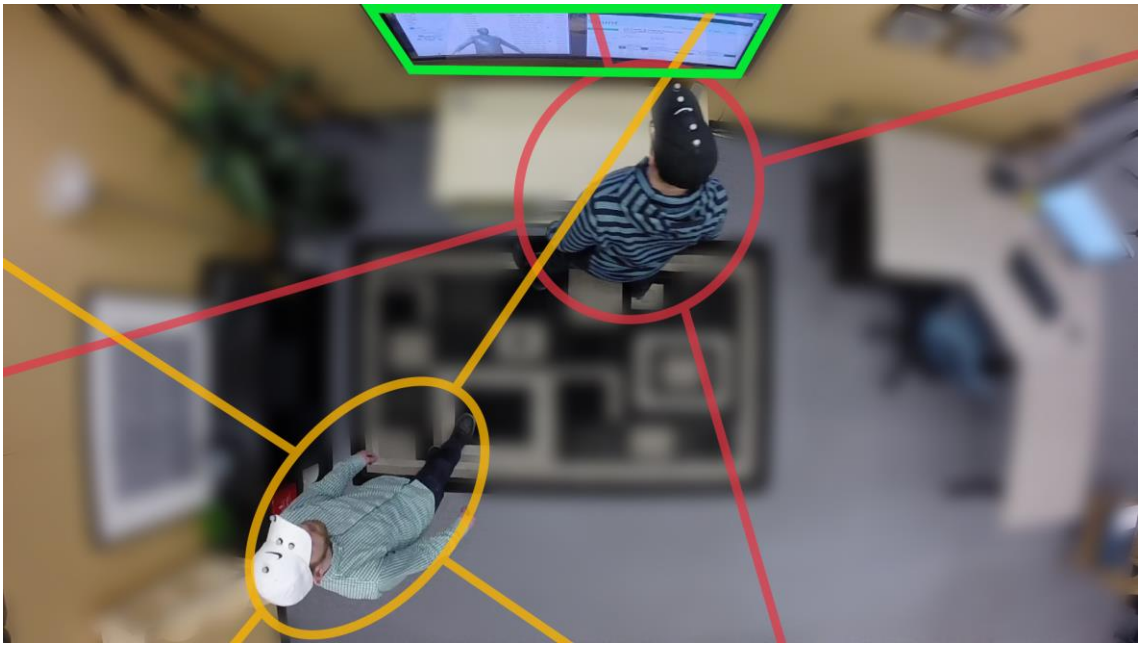

**Figure 6.28. The Proximity Toolkit gives precise information about each person's position and orientation relative to each other and the display.**

The level of detail of the Vicon system allows for sub-millimeter tracking accuracy. In this work 13 Vicon cameras have been used. Therefore two Vicon MX Ultranet servers had to be connected, each of them handling up to eight cameras. The following system versions have been used:

[3] www.vicon.com

Proximity Toolkit version 1.2.1, Vicon Nexus 1.5.

The Proximity Toolkit [35] consists of two main components:

- **The Proximity Toolkit server.** It allows multiple clients to connect via a TCP connection. It broadcasts the proxemic information to all connected clients. It can utilize various tracking plugins. In this work the Vicon tracking module has been used, using a Vicon motion capturing system for information about position, orientation and motion of tracked entities.
- **The Proximity Toolkit application programming interface (API).** The API is offered via an object-oriented C# .NET development library. It allows to easily connect to the Proximity Toolkit server via TCP networking. It enables for direct access to information available in the proximity toolkit using an object-driven approach. Further it offers an event-driven approach in a provider-subscriber manner where updates about an entity or a relationship between two entities can be received.

The general routine to use the Proximity Toolkit is as follows:

- **Setting up the Vicon Nexus.** This task has to be done only once. It consists of wiring the cameras, making marker sets and adding them to the system as *models*.
- **Calibrating the Vicon Nexus.** This task has to be done every few weeks, as the cameras tend to be very sensitive to even slight movements. The initial calibration after setting up the system is done by waving a special marker setup so that it can be seen by every camera at various distances and angles for several minutes. A re-calibration mode allows later for a quick calibration when there have only been minor changes in the setup or lighting conditions. Further noise can be masked at a camera level.
- **Setting up the Proximity Server.** This task has to be done only once. It consists of adding vectors to the models, available through the Vicon Nexus software. Also static entities can be added to a room, such as a display or other volumes, such as a couch or bookshelf.
- **Loading the Toolkit for a session.** Turning on the Vicon MX Ultranet servers which are controlling the cameras, launching the Vicon Nexus application and loading the previously created models of marker setups. When all the cameras are connected, the Proximity Toolkit can be launched. For a regular session these steps suffice and an application can connect to the server using the C# .NET library.

## 6.2 Coordinate system

Three different coordinate system had to be used, shown in Figure 6.29.

- **The regular WPF coordinate system for 2D graphics.** It originates in the upper left corner, the X values are on the horizontal (positive values proceeding to the right) and the Y values on the vertical axis (positive values proceeding downwards) (blue in in Figure 6.29).
- **The Proximity Toolkit coordinate system.** The origin is at a user defined point in the room. For this work the origin has been set at the floor about 50cm in front of the center of the display. The Z values are mapped to the horizontal axis (positive values proceeding to the left of the origin), Y values on the vertical axis (positive values proceeding upwards) and the X values on the axis being perpendicular to the other two (positive values pointing towards the back of the room) (red in in Figure 6.29). All data received from the Proximity Toolkit is accordingly to this system.
- **Helix 3D Toolkit coordinate system.** Used for the display of the 3D-model. The origin is in the center of the display, X values are on the horizontal (positive values proceeding to the right), Y values on the vertical and Z values perpendicular to X and Y (positive value proceeding upwards) (orange in in Figure 6.29). For details about the Helix 3D Toolkit see section 6.11.

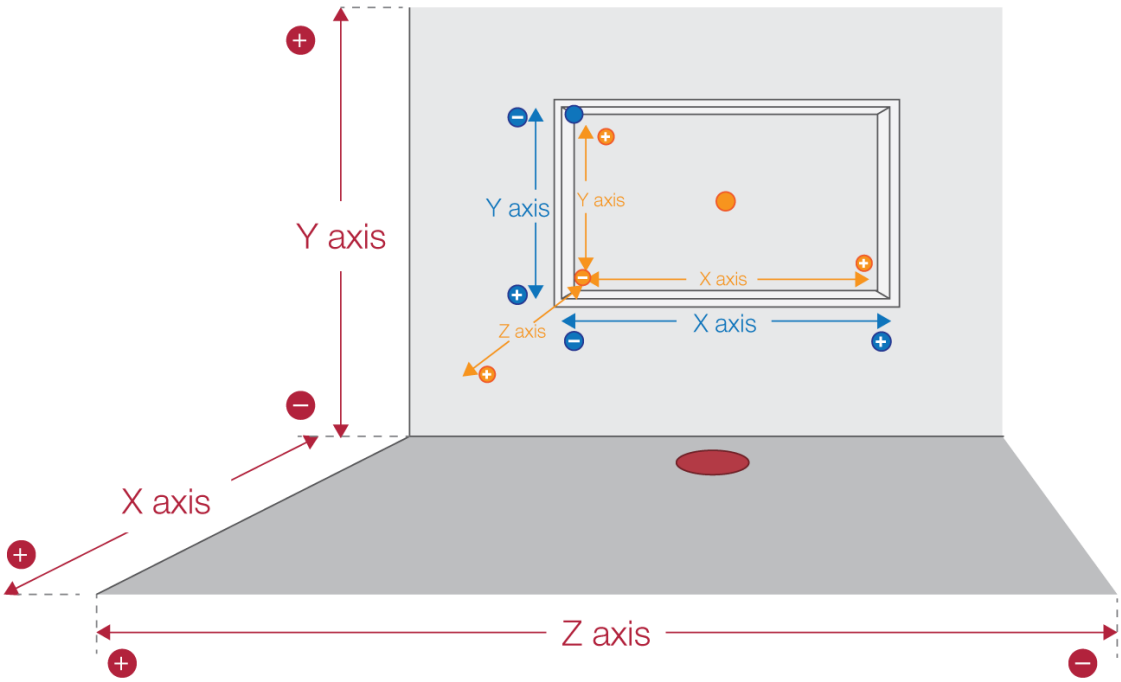


**Figure 6.29. The three different coordinate systems being used: The Proximity Toolkit's coordinate system (red), regular 2D coordinate system (blue) and Helix 3D Toolkit (orange). Figure based on a graphic by David Ledo. Used with permission.**

To keep the coordinates at a manageable level, all coordinates, received from the Proximity Toolkit are transformed to regular 3D coordinates (X-value on the horizontal axis, increasing values to the right; Y-values on the vertical axis, increasing values upwards; Z-values on the third, increasing values towards the center of the room). The center point was set to the top left point of the display, to allow easy remapping between display and world coordinates. By first transforming the Proximity Toolkit's coordinates, the usage of the different coordinate system was less confusing and more consistent.

## 6.3 Marker setup

The marker setup for each tracked entity has to be unique among the ones being used during one concurrent session and should not have any symmetry whatsoever (a picture of the baseball caps with its distinctive marker setup is shown in Figure 6.30). Setting up the markers can be a tedious task, and can easily result in symmetry, as one has to think about all the possible ways to create symmetry through rotation in 3D space. Further no reflective fabric should be worn by the user. Some materials, such as a stainless steel coffee mug or certain white shirts, can reflect the infrared light from the cameras, creating undesired noise in the tracking data. When there is a lot of noise, it causes tracked positions, and especially orientations of entities, to jump or vectors to flip. This results in unreliable tracking. Some other reasons for noisy tracking, observed during the development of the system, can be that the Vicon Nexus system needs to be restarted or even recalibrated. The software offers a recalibration of the entire system or just a single camera. From experience the full calibration should be used, as in the used version 1.5 of the Vicon Nexus software the single-camera-calibration can cause camera positions to jump. When a distinct marker setup is recognized by the Proximity Toolkit it is called an *entity*.

**Figure 6.30. The baseball caps being used to track the user and passerby with the Vicon markers.**

## 6.4 Generally available parameters

The Proximity Toolkit is the main information providers for data about the environment, the people and devices. Additionally a Microsoft Kinect sensor, separately connected to the system, is being used as information provider. In the following, parameters used from it will be explained.

The toolkit allows for two ways of retrieving information: either polling for any given entity's value at any time or subscribing to asynchronous update events. These events will be triggered whenever there is new information available from the Proximity Server. Observations showed that new events arrive approximately every 50-100 milliseconds. Any modifications in the UI which are a result of the proxemic information should therefore be made in a timed thread to achieve a fixed frame rate.

The following information is being received and used from the Proximity Toolkit:

- *Identity and position* of user and passerby's head in the room. For this both people are wearing a baseball cap with attached markers.
- Passerby's *torso rotation*, by tracking a vest which is being worn by the passerby.
- User's and passerby's *viewing direction* (forward vector defined in the Proximity Toolkit)
- *Distance* between passerby and user and passerby and display by subscribing to update events.

The vest is being worn by the passerby, in order to differentiate between the torso rotation and the head rotation. Each entity in the in the Proximity Toolkit offers the roll, azimuth and incline angle, therefore their full orientation in the room is captured. The rotation of the cap is being mapped to the azimuth rotation of the 3D-models head, the vest's azimuth angle is being mapped to the torso's rotation.

The prototyping area, called the *Homespace* is shown in Figure 6.31. A 61 inch screen is being used as a large public display. A total of 13 cameras are used for tracking of movements and orientations of various entities in the room (note that not all cameras are shown in the picture). A Microsoft Kinect, positioned on top of the display, is used to capture gestures performed by the user.

**Figure 6.31. The "Homespace". A prototyping area for proxemic interaction, using various tracking techniques.**

## 6.5 Calculated information

The information received from the Proximity Toolkit already contains already contains system information about e.g. the presence of user and passerby, their orientation and distances towards each other and the display. This information is used in various calculations. Two example procedures are explained below, as they are used extensively throughout the system.

### 6.5.1 Opacity

Some of the cues explained in chapter 1 and 1 encode a passerby's distance and his look direction in their transparency. The calculated opacity is not only being used for the 3D-model, but also in various other features of the program, e.g. to adjust the opacity of the silhouette and the cover of sensitive applications. It is made available as property `OverallOpacity` to allow easy event notification. On the one hand the distance of a person to the display is being mapped from its minimal to the maximum possible distance[4] (line 2-9 in Listing 6.1). On the other hand a person's look direction is considered in the transparency: The intersection point of the entity's forward vector with the display plane is being used as a metric in order to get a transparency value for the viewing direction. When the vector is centered on the display the overall opacity is increased by the maximum value set for the orientation setting. With the person's viewing direction moving further to the left or right of the display plane, the opacity decreases, as it is less likely that the passerby is seeing the content of the screen (line 10-18 in Listing 6.1). The minimum and maximum influence that distance and viewing direction have on the opacity depends on which option the user chooses to activate: When both options are activated, the distance has a maximum influence of 0.7 on the opacity and the viewing direction a maximum of 0.3. Listing 6.1 shows the pseudo code of how this mapping is done. The influence of the distance is much higher as a passerby can trick the system's viewing direction factor by gazing from the corner of his eye. An in depth explanation and system evaluation of the viewing point parameters can be found in section 6.5.2.

```
1  if (changeOpacity):
2     if (considerDistance):
3          if (considerViewingDirection):
4               min = 0.3; max = 0.7;
5          else:
6               min = 0.3; max = 1.0;
7          endif;
8          distanceFactor = Remap(currentDistance, minDistance, maxDistance, min, max);
9     endif;
10    if (considerViewingDirection):
11         if (considerDistance):
12              min = 0.0; max = 0.3;
13         else:
14              min = 0.3; max = 1.0;
15         endif;
16         viewingdirectionFactor = Remap(Math.Abs(intersectionWithPlaneX), centerX, maxOffsetX,
17                   min, max);
18    endif;
19    if(!considerDistance && !considerViewingDirection):
20         OverallOpacity = 0.8;
21    else:
22         OverallOpacity = distanceFactor + viewingdirectionFactor;
23    endif;
24 endif;
```

**Listing 6.1. Mapping of the passerby's distance and viewing direction to an opacity value.**

The minimum opacity cannot go below 30% as awareness about the presence of a passerby should still be available even when he is looking at a different direction from afar. When neither the distance nor the viewing direction should be encoded in the opacity, it is set to a default level of

[4] To get this number, the room's boundaries have to be known. A special system state has been implemented, allowing to measure various metrics, such as the full boundaries of the tracking in 3D space, and to set the position of the display. The control interface for the implemented prototype can be found in appendix A.

80%. This is done in order to allow for a good awareness to the user, while still being able to see the content underneath the model.

### 6.5.2 Person looking at display

With the flashing borders the color of the border changes when a passerby is looking at the display. In order to achieve this, the system must know where he is currently looking at. The Proximity Toolkit allows to follow a relationship between two tracked entities, such as the display and the passerby's hat. This relationship also includes the parameters `PointsAt` and `PointsTowards`. Unfortunately none of these parameters are suitable for this system. `PointsAt` defines whether an entity's forward vector directly intersect with the volume of the display. As the boundaries of the volume is limited, the ray does not intersect with it for the entire time a person is actually looking at the display. Especially with a larger distance this property is false most of the time. On the other hand `PointsTowards` tells whether the passerby's vector is pointing towards the volume of the display, meaning that it points towards the display at an angle between -90° and +90°. This means that a person can look at an angle parallel to the display plane and this property is still true. Therefore a custom implementation uses the passerby's position and the azimuth angle to calculate whether he is currently looking at the display. The assumption is that for each position in the room there is an angle to either side (left and right) where a person can just still see the content of the display and can tell what is shown. As soon as the passerby's angle is inside of this scope he is looking at the display. This angle varies, depending on the distance to the screen and position in the room.

#### Verifying parameters, system evaluation

In order to quantify it and to improve the calculation of when someone is looking at the display tracking data of eight different people has been collected[5]. They were asked to stand on six different positions, marked on the floor (circles in Figure 6.32). The middle column was positioned at the center of the display. The left and right column were at the edge of the tracked area, in order to get an angle value for the extremes of the room boundaries. Note, that the display is not centered in the tracked area, therefore the distance between the columns is not evenly distributed. The front row is at a distance which is slightly behind a comfortable working distance (110cm from the display), simulating a distance where the passerby would literally be looking over the user's shoulder. The middle row is at a distance of 170cm from the display, the back row at 270cm from the display, the furthest distance at which an entity can be tracked at a height of 180cm (approximately a person's height). The base angle was set at 0° when looking straight towards the display plane.

For each of the six positions, participants were asked to look at the display so they could see what is on the display. There was no need for them to be able to read the contents, but they should be able to comfortably tell what is currently shown. As there was a difference to be expected between whether a person is looking straight ahead or out of the corner of their eyes, data for two different conditions was collected. In condition A participants were asked to look straight ahead and try not to gaze to the side. In condition B they were asked to try to rotate their head as far away from the display as possible and to gaze as much out of the corner of their eye as they felt comfortable in doing so. In either condition they were allowed to rotate their head as this rotation was being tracked. When they felt comfortable to have found the maximum left / right angle the azimuth angle was recorded. The two conditions were counter balanced between the subjects. As for the positions, all participants started in the back row with the right position, first moving to the left then advancing to the next row, again beginning with the right position.

To evaluate the data and get a value, to be used with the system, the average of all participants was calculated and later used in the calculation of whether a passerby is looking at the display. To get a value for each position in the room angles are interpolated between two positions in order

[5] This system evaluation has not approved by an ethics board. It was conducted with volunteers from the lab environment.

to get a left and right angle. The combined visual results for both conditions for the left and right angle are shown in Figure 6.32b. The yellow lines indicate condition B, where participants were asked to look out of the corner of their eye (Figure 6.32a). The blue lines represent condition A, where participants were asked to look straight ahead (Figure 6.32c). The angles given are the absolute angles participants could cover by rotating their body. The detailed values for both conditions can be found in appendix A.

What should be noted is, that the angle for conditions B (blue lines) in the case of the bottom right position (furthest distance from the display, on the right side of the room) seems to be not correct, as its value for the looking direction to the right is greater than when participants were asked to gaze out of the corner of their eyes. This is most likely due to tracking issues, as that position is almost out of reach of the cameras field of view.

Some brief, informal feedback from the participants is worth to mention as this indicates room for future work:

- Some participants said they chose a certain significant object on the display in order to always have a fixed reference point to look at. In a follow up study it would need to be defined where participants should focus on and whether looking at the borders of the display is enough or if they should focus at the center.
- Participants, wearing glasses, said that if they tried to be able to tell what is on the display and actually would like to read it they would have to turn their head more so they could look through their lenses. Their gaze out of the corner of their eyes is limited by the frames of their glasses.
- One participant noted that, after looking out of the corner of his eyes for too long, he gets dizzy and would usually only do this for a short amount of time.

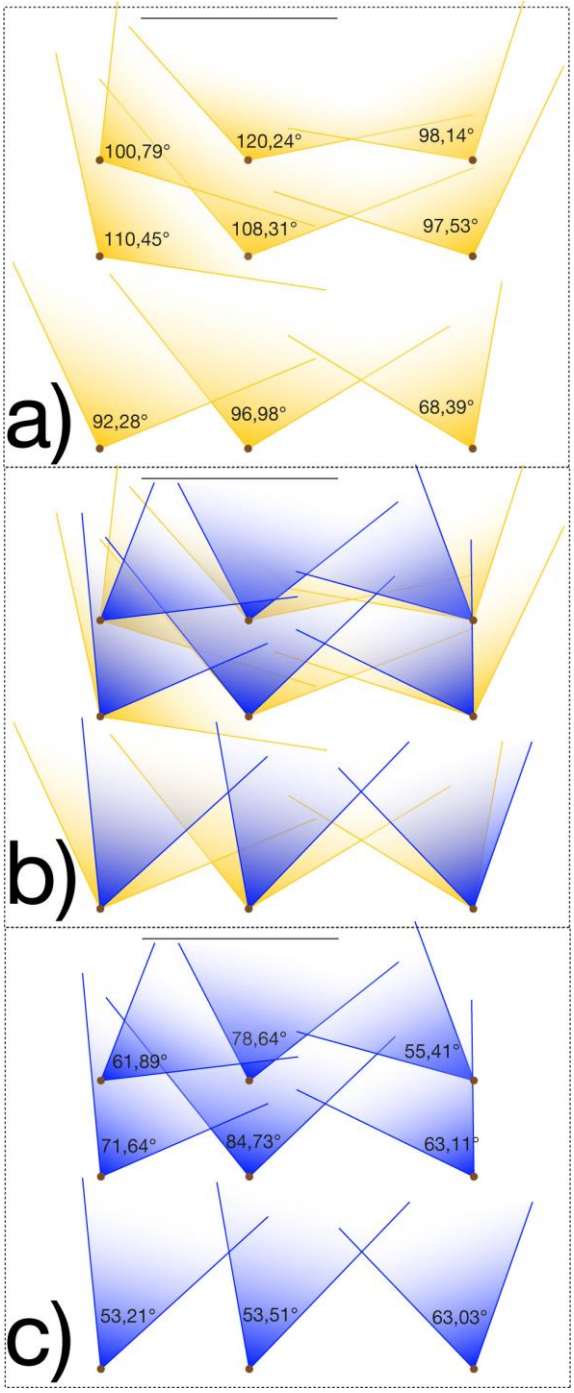


**Figure 6.32. Visual representation of the maximum left and right angle at which a passerby can comfortably tell what is shown on the display (black line). a) The yellow lines indicate when a passerby is trying to look out of the corner of his eye (condition B). c) The blue line represents the data when participants were asked to look straight ahead (condition A). b) A combination of both. Detailed values can be found in appendix A.**

## 6.6 Flashing borders

The Flashing Borders make use of the presence and look direction of the passerby. As soon as the passerby enters the tracked area the system's property `PasserbyPresent` is set to true. If the awareness about the passerby's presence is desired by the user, the display borders' fill is set to a linear gradient brush, with a green color. The opacity is at its fullest at the outer borders of the display, fading to full transparency to about 5cm towards the center of the screen. In a separate thread the opacity of the entire border increases / decreases in small steps, simulating the flashing. As soon as the passerby's look direction is directed towards the display the fill color of the borders changes to red, providing a less subtle awareness. To make it even less ambient the speed of the flashing is doubled, resulting in a duration for a full fade of 0.8 seconds instead of 1.6 seconds when the passerby is only present, but not looking. The distance of the passerby is further encoded into the maximum transparency of the border. The closer he gets to the display, the higher the maximum opacity. A minimum opacity of 30% is always kept, even with a large distance, in order to ensure visibility of the border. The position of a passerby is known, as are the room boundaries. Therefore it is known on which side of the room he currently is in. To allow for awareness about his presence, only the right border flashes when he is on the right hand side of the room and vice versa. All four borders flash, when he is not near the outer borders of the room.

## 6.7 Blacking out the display

The entire display can black out as soon as no permitted user is looking at the display anymore. This is done by using the forward vector of the user. When it intersects with the display plane the system assumes that he is looking at the display. As soon as the vector does not intersect with the plane anymore or the intersection with the plane is several meters away from the actual boundaries of the display the system sets the property state of `UserLookingAtDisplay` to false. The calculation of the user's intersection point is somewhat similar to the calculation of the passerby's viewing direction (see section 6.5 for details). The visibility of a rectangle, filled with a solid black color, is bound to that property, resulting in an entirely black screen when the user of the system is not looking at the display or not present at all.

## 6.8 Covering and moving windows

To cover applications, their positions have to be known first. This is done, through platform invoke (pInvoke) calls to the user32.dll files of the Windows operating system. As explaining the usage of this technique would be beyond the scope of this written thesis, the reader may be directed to the Microsoft Developer Network[6] for further information. Through calls to the native Windows API, it is possible to get and manipulate the information, such as the position of the window, of each running process on the system. To hide an application, simply a black rectangle is drawn on the position of that window. The rectangle's transparency is bound to the `OverallOpacity` property, mentioned before.

Through pInvoke also the position of the windows can be set. By doing this in a timed thread, the position can be smoothly animated. Several different easing methods have been implemented and can be selected at runtime. The windows' initial locations are saved before moving, so they can be moved back to their original positions.

Listing 6.2 shows how an easing method is being used in a time based animation in order to change the position of a point.

[6] http://msdn.microsoft.com/en-us/library/aa288468(v=vs.71).aspx [Last Accessed: 21-Feb-2014]

```
1  //initial setup
2  double xDiff = startPosition.X - destinationPos.X;
3  double yDiff = startPosition.Y - destinationPos.Y;
4  int lastTick = System.currentTime, timeElapsed = 0;
5
6  //the easing method returns the new position. Called in a threaded timer.
7  Point getNextMovementPointEasing(int currentTick){
8      timeElapsed = System.currentTime - lastTick;
9      while (timeElapsed > 30) {
10         //for every 30 milliseconds that have elapsed since the last movement
11         currentFrameCount += 1;
12         timeElapsed -= 30;
13     }
14     lastTick = currentTime;
15
16     //here, various easing methods can be used, as for the easing only a factor between
17     //0.0 (start) and 1.0 (end) is needed.
18     double factor = easeOutExpo(currentFrameCount, maxFrameCount);
19     return new Point(startPosition.X + xDiff * factor, startPosition.Y + yDiff * factor);
20 }
21
22 double easeOutExpo (int currentTime, int totalTime){
23     if (currentTime >= totalTime)
24         return 1; //easing is finished and animation should stop. Not listed.
25     return -Math.Pow(2, -10 * (double)currentTime / (double)totalTime) + 1;
26 }
```

**Listing 6.2. Calculation of the movement of windows, using an easing function.**

### 6.8.1 Gesture Recognition

The Proximity Toolkit could be used for gesture recognition by feeding the position data of a user's hand movement to a gesture recognition algorithm. After several attempts in doing so the decision was made to not use the 3D tracking data, as the computational power needed for continuous custom gesture recognition is fairly high (e.g. through usage of dynamic time warping). On the other hand resource friendly filters exist, such as the 1$ gesture recognizer [64]. The drawback of this filter is, that it needs a finite set of two-dimensional location data, it does not work for a continuous stream of data. Further it needs to be adapted to be used with locations in 3D space. The Microsoft Kinect sensor on the other hand offers easy to use gesture recognition. Simple gestures, such as a swipe gesture with one's arm can be recognized with no additional training of the algorithm and just about 20 lines of code.

**Figure 6.33. A user performs a swipe gesture. As seen by the Microsoft Kinect sensor.**

The Kinect sensor was positioned on top of the wall mounted display, facing the room. Swipe gestures from both, the left and right arm were recognized. The left arm's gesture caused the covering of sensitive applications upon user's request, the right arm caused the windows to move to that side of the display which the user waves them to (see section 5.2). Figure 6.33 shows a user, performing a swipe with his right arm, as seen by the Kinect's camera. Listing 6.3 shows a code snippet of how the Microsoft Kinect swipe gesture recognizer can be used to detect swipe gestures with either the left or right arm.

```
using Microsoft.Kinect;
using Microsoft.Samples.Kinect.SwipeGestureRecognizer;
[…]
void init_gesture(){
     // Look through all sensors and start the first connected one.
     // This requires that a Kinect is connected at the time of app startup.
     foreach (var potentialSensor in KinectSensor.KinectSensors){
        if (potentialSensor.Status == KinectStatus.Connected){
           this.sensor = potentialSensor;
           break;
        }
     }
     if (null != this.sensor){
        // Turn on the skeleton stream to receive skeleton frames
        this.sensor.SkeletonStream.Enable();
        // Use Seated Mode
        this.sensor.SkeletonStream.TrackingMode = SkeletonTrackingMode.Seated;
        // Add an event handler to be called whenever there is new color frame data
        this.sensor.SkeletonFrameReady += this.SensorSkeletonFrameReady;
        // Start the sensor!
        try{ this.sensor.Start(); }
        catch (IOException){ this.sensor = null; }
     }
     // Instantiate a recognizer
     var recognizer = new Recognizer();
     // Register for swipe gesture events from left or right
     recognizer.SwipeRightDetected += new
     EventHandler<KinectGestureEventArgs>(recognizer_SwipeRightDetected);
}
void SensorSkeletonFrameReady(object sender, SkeletonFrameReadyEventArgs e){
     Skeleton[] skeletons = new Skeleton[0];
     using (SkeletonFrame skeletonFrame = e.OpenSkeletonFrame()){
        if (skeletonFrame != null){
           skeletons = new Skeleton[skeletonFrame.SkeletonArrayLength];
           skeletonFrame.CopySkeletonDataTo(skeletons);
        }
        this.activeRecognizer.Recognize(sender, skeletonFrame, skeletons);
     }
}
void recognizer_SwipeRightDetected(object sender, KinectGestureEventArgs e){
     Console.WriteLine("swipe right");
}
```

**Listing 6.3. Using the gesture recognizer of the Microsoft Kinect sensor.**

The gesture recognizer uses the tracked skeleton to do its calculations, meaning that the user has to stand at a certain distance to be seen by the sensor. By activating the ‘seated mode’ of the Kinect, only the upper body of the user is being tracked. The user can then stand as close as 0.4 meters up to a maximum of 3.0 meters to the Kinect and his gestures will still be recognized (given that his arm does not leave the field of view, when performing the gesture).

## 6.9 Silhouette

For the silhouette cover (details in section 5.4), the area a user covers with his body from a passerby’s view needs to be known. This is done by assuming a vector from the position of the user and a passerby (Figure 5.26b, green line). By extending that line, an intersection with the display can be found (Figure 5.26b, white area). This intersection point is the center for the silhouette. For the silhouette itself a radial gradient is being used, with multiple gradient stops. To move the position along the axis, the center point and the gradient origin is being moved (lines 9-11 in Listing 6.4). The sensed user’s height influences vertical position and the silhouette’s height, with the top boundary being at the top of a user’s head. The passerby’s distance to the user influences the width of the silhouette (line 17-19 in Listing 6.4) and the extent of the blurriness (line 13-15 in Listing 6.4).

```
1  void InitSilhouette(){
2      silhouetteGradient = new RadialGradientBrush();
3      // […add gradient stops…]
4      silhouetteGradient.RadiusX = 0.2;
5      silhouetteGradient.RadiusY = 0.6;
6  }
7
8  void setSilhouetteXPosition(double PasserbyUserDistance) {
9      silhouetteGradient.Center = new Point(userPasserbyDisplayIntersection.X / displayWidth,
10       userDisplayIntersection.Y / displayHeight); //moves the entire gradient
11     silhouetteGradient.GradientOrigin = silhouetteGradient.Center;
12     //blurriness relies on the distance between the user and the passerby.
13     GradientStop gs1 = silhouetteGradient.GradientStops[1];
14     gs1.Offset = Helper.Remap(PasserbyUserDistance, proximityDistanceMin, proximityDistanceMax,
15       silhouetteGradient.GradientStops[0].Offset, silhouetteGradient.GradientStops[2].Offset);
16     //the width of the silhouette
17     var value = Helper.Remap(PasserbyUserDistance, proximityDistanceMin, proximityDistanceMax,
18       silhouetteWMin,silhouetteWMax);
19     silhouetteGradient.RadiusX = value / displayWidth * 0.75;
20     //[…similar remapping for the height of the silhouette, based on the user's height…]
   }
```

**Listing 6.4. For the silhouette a radial gradient is being used, with multiple gradient stops. To position it on screen the center point is shifted, according to the user-passerby vector's display intersection (x-axis; line 9) and the user's forward vector display intersection (y-axis; line 10).**

Figure 6.34 shows a visual representation of the three situations. The passerby is marked by a light-blue circle, the user by a green circle. The solid green and light-blue lines are the person's viewing direction. The dashed red line is the vector between passerby and user. The orange lines indicate the area shielded from view by the user's body. The width of the silhouette (orange line) changes with the user-passerby-distance. Figure 6.34a: User is looking at the display, the passerby cannot see what is shielded by the user's body. Figure 6.34b: The user is facing the display plane, but not actually looking at the display. The display is blacked out entirely, as there is no 'secure area' (orange) on the display. Figure 6.34c: The user is not facing the display. The display is entirely black.

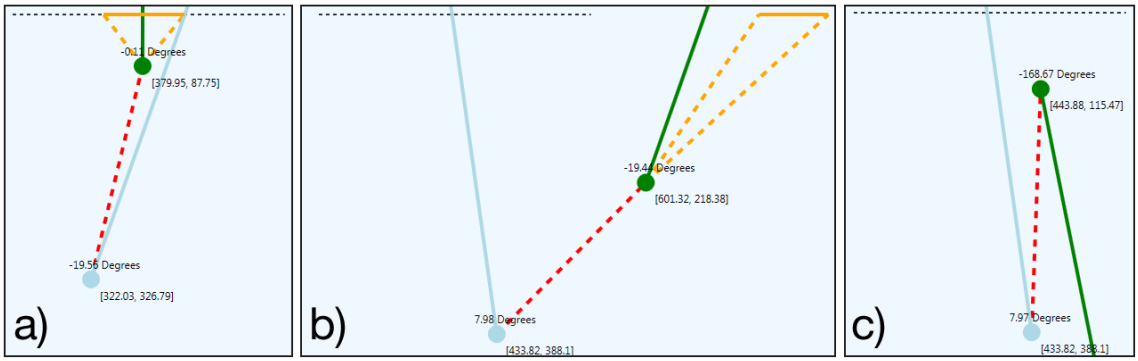


**Figure 6.34. Top-view of the sensed situation. a) The user is facing the display. b) The user is facing the display plane, but not looking at the display. No parts of the screen are covered by the user's body. c) The user is not facing the display, the entire screen area is blacked out in order to protect sensitive information.**

This representation has originally been implemented as a debug output, but has shown to be more helpful when tracking entities. It has also been used in the system evaluation, explained in section 6.5.2).

## 6.10 Gaze Awareness Indicator

The gaze awareness indicator shows the position on the display where a passerby is currently looking at (details in section 4.3). This is done by using the passerby's forward vector and intersecting it with the display plane. The position data, received from the Proximity Toolkit, is fairly stable and contains only little noise or jitter (unless tracking becomes unstable after a long

time of use or because of a many simultaneous markers being used). However, when working with the vectors, the data becomes increasingly noisy. The noise results in a jittering gaze indicator, and with increasing distances this noise intensifies which can be explained by an lever action.

A large wall display has been used, with a visible display area width of 135.1 cm and height of 76.0 cm. The resolution was set to 1600x1200 pixels, resulting in a pixel density of 30.1 ppi (horizontal) by 40.1 ppi (vertical). With this screen a jitter of 6 pixels (horizontal) / 8 pixels (vertical) correspond to ~5 millimeters of movement on the screen. Actual measurements for the jitter can be found in Table 1.1. A spatial jitter of that amount negatively influences how people perceive the system [44]. Therefore the data of the passerby's viewing position on the display was filtered, using the 1€ Filter [16]. The decision for this particular filter was made for two reasons: It does not add a big delay to the data and uses very little resources. Since the 1€ Filter only allows to filter a single decimal value, two filters had to be used, one for each of X- and Y-coordinate. The selected filter settings[7] accounted for smooth movements, adding a time latency (lag) of up to 500ms on its peak. A lag of this amount is easily noticeable when the person, controlling it, looks at the viewing indicator [20]. Considering the particular use case, this should not be a problem: The gaze awareness indicator is not meant to be viewed by the person controlling it (the passerby), but by the user. The intentions here were to have an ambient awareness indicator, showing an approximate location of the passerby's viewing position on the display without interrupting the user's work too much. Therefore a smooth appearance and gentle movements were desirable.

To quantify the spatial noise, introduced through the tracking, an entity (in this case the baseball cap usually worn by the passerby) was immobilized by securely mounting it on top of a tripod. The initial base-position of the intersection with the display was then set to the current position and for the following 15 seconds the maximum deviation from this base-point was measured, both along the X- and Y-axis. This was repeated ten times for each, the filtered and the unfiltered data. Unfiltered, the average horizontal jitter was 7.4 pixels, 12.1 pixels along the vertical axis. After applying the filter the jitter has been reduced by 62% (horizontal) and 66% (vertical), as shown in Table 1.1 and Figure 6.35.

The result of the filtered gaze indicator is a very smooth appearance of the red fuzzy dot, gliding across the display, as the passerby looks around.

| | Horizontal | | Vertical | |
|---|---|---|---|---|
| Unfiltered | 7.4 pixel | 6.24 mm | 12.1 pixel | 7.66 mm |
| Filtered | 2.82 pixel | 2.38 mm | 4.13 pixel | 2.61 mm |

**Table 1.1. Average jitter along horizontal and vertical axis. Ten independent measurements (15 seconds each) for each, horizontal and vertical spatial noise, before and after the 1€ Filter has been applied.**

[7] The following filter-parameters have been used: Minimum cutoff frequency 2.0, cutoff slope 0.

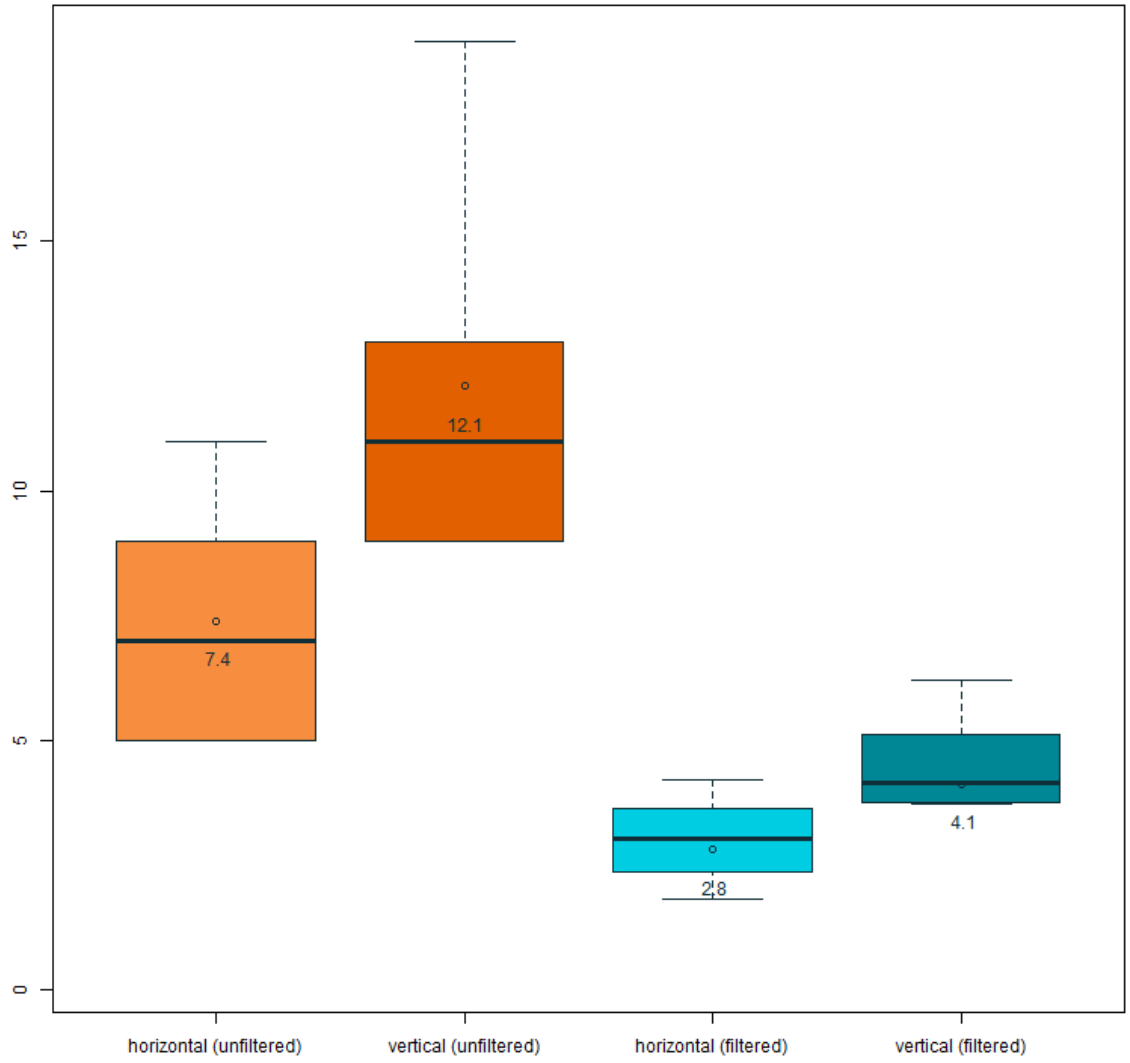


**Figure 6.35. The jitter of the intersection point on the display in millimeters. Circles are column means. The jitter n pixel can be found in Appendix**

## 6.11 3D-Model

For displaying the model the external Helix 3D Toolkit [10] was used. This framework was chosen over the 3D WPF framework because of its powerful and easy to use helper classes, e.g. for loading and displaying 3D content in an easy to use WPF-style custom control for 3D operations. Unfortunately the documentation of the Toolkit was very sparse at the beginning of this project (September 2013) and the API is not well documented. Therefore a lot of the functionalities had to be guessed from reading the source code and through trial and error. Some findings on how to properly use the toolkit will be listed here.

To use the Helix 3D Toolkit, the DLL file has to be referenced in the project and the required references be made in the header of each class. After that the 3D viewport needs to be initialized, and the camera and lights have to be added to the scene (Listing 6.5).

```
using HelixToolkit;
using HelixToolkit.Wpf;
[…]
private HelixViewport3D initHelixViewport(Windows.Controls.Grid parent){
        HelixViewport3D viewport = new HelixViewport3D();
        viewport.ShowViewCube = false; //turns off 3D orientation cube
        viewport.DefaultCamera = new PerspectiveCamera();
        viewport.DefaultCamera.Position = new Point3D(100, 0, 100);
        viewport.DefaultCamera.LookDirection = new Vector3D(-100, 0, -100);
        viewport.DefaultCamera.UpDirection = new Vector3D(0, 1, 0);

        viewport.Children.Add(new DefaultLights());
        viewport.ZoomExtents(5000);
        parent.Children.Add(viewport);
        return viewport;
}
```

**Listing 6.5. Initializing the Helix 3D Toolkit.**

After the primary initialization 3D objects can be added to the Children container of the `HelixViewport3D`. First a `ModelVisual3D` object has to be created, which will be used to display the model and added to the 3D viewport (lines 1-2 in Listing 6.6). Next the model needs to be loaded from the file system. The Helix 3D Toolkit supports various file formats, e.g. Autodesk 3DS, Lightwave LWO and 3D System's STereoLithography (STL). The Helix 3D Toolkit comes with a `ModelImporter` class which loads a model from a file system in a non-blocking manner (if desired).

```
ModelVisual3D loadedModel = new ModelVisual3D();
viewport.Children.Add(loadedModel);
[…]
var mi = new ModelImporter();
Model3D currentModel = mi.Load(@"models\man1\model.3ds", Dispatcher.CurrentDispatcher);
loadedModel.Content = this.currentModel
```

**Listing 6.6. Adding a container and loading a 3D-model, using the Helix 3D Toolkit.**

Finally it is set as the `Content` of the `ModelVisual3D` (lines 4-5 in Listing 6.6), which has been created earlier. The loaded model is then transformed multiple times to achieve a neutral, upright, front-facing startup position. The model is also being scaled, using a `Windows.Media.Media3D.ScaleTransform3D`. The initial scaling is done using the `ScaleTransform3D`. To change the size of the model in a later stage the zoom level of the entire `Grid` is being changed. The ScaleTransform3D and other `AffineTransform3D`'s are being grouped and applied to the loaded model. In order to easily manipulate these transformations later, a `Model` class is being used, encapsulating easy access to all the needed parameters of a 3D model. That way the affine transformations can be manipulated later, e.g. for rotating the model, when the passerby rotates his body. When the user decides he wants to display the passerby's head and body orientation separately, the 3D-model will be loaded twice. Both models are then cut off at the neck, one displaying only from the torso downwards, the other showing the neck and head, being displayed right above each other, giving the impression of one model. In order to change the position of the model on the horizontal axis the parent `Grid` is being moved on the X-axis. The passerby's location in the room, as reported by the Proximity Toolkit is then mapped between the room's minimum and maximum boundaries and the display area (the room's boundaries are known, as they were measured by using the provided system state 'measure room boundaries'). To make the movement along the horizontal axis, the entire parent `Canvas`, containing the `HelixViewport3D` is translated on the X-axis (Listing 6.7). That way only one transformation has to be applied, even when head and torso are displayed separate.

```
var x = Remap(Passerby.ProxemicLocation.Z, proximityZMin, proximityZMax, 0.0, displayWidth);
helixParent.SetValue(Canvas.LeftProperty, x);
```

**Listing 6.7. Remapping of the passerby's position to the display width.**

The same mapping function is being used for calculating the model's size. The distance between the passerby and the display is being mapped between the minimum distance those entities can have. The first assumption to use zero as the minimum distance has been discarded, as a person

usually never stands that close to the display. A comfortable reading distance of 30cm for the minimum has been chosen. The maximum distance is defined by the room boundaries and therefore the maximum distance a passerby can have from the display, while still being tracked, is approximately 4 meters. Another option would be to change the size of the model according to the distance between the user and the passerby. This feels very unnatural, as the model's size changes even when the user moves while the passerby stands still. This does not correspond to what the model represents, which is the passerby's position, orientation and distance. The model's minimum and maximum size can be changed in the controls interface, shown in Figure A.44 in appendix A.

# 7 Conclusion

The distribution of large displays in public increases. On the one hand this opens up new interactions with large interactive areas and digital information can be used while on the go. One the other hand, as the display area increases, it becomes more difficult to protect information using one's body. When it comes to private and sensitive information people want to be sure that their information is protected from unwanted views and shoulder surfing passers-by.

In this work, social theories have been analyzed in order to understand how people negotiate about territorial encroachments and their personal space. Previous research has shown that privacy intrusions rarely happen on purpose. Nevertheless they do happen, especially with increasing display sizes people tend to be more voyeuristic. The understanding of private information varies between different people. Many of previously proposed systems however rely on a single understanding of which data needs protection and assume that privacy intrusions occur on purpose.

This thesis explored how shoulder-surfing issues can be mitigated on public displays. All the methods are based on sensing the position, distance, and orientation between people and their environment, which in turn helps calculate and build upon social notions of proximity and territorial incursions. The techniques provide varying degrees of mutual awareness to allow user and passerby to engage in social protocol. Provided awareness helps them self-regulate their behaviors and/or negotiate their consequential actions, utilizing verbal or non-verbal cues. Human behavior and social protocol are well trusted entities. People usually act rational and are sensitive to someone else's information. People use many different means of renegotiating their desired level of privacy. This ranges from body language to verbal cues to physical means as a last resort. When being asked to respect someone else's privacy, people usually do so, even if their point of view of private / sensitive data might be different. Implicit *vs.* explicit actions allow for a negotiation of personal space and collaboration. Table D.2 (appendix A) shows an overview of the implemented systems in terms of their ability to provide awareness about certain aspects and the granularity of the surroundings.

The techniques in this work also provide some degree of protection of sensitive information to a broader range. By employing the physical relationship between people and devices, the system knows which data can be overlooked and in return protect it. Data can be hidden from passers-by, by still allowing the user, standing close at the display, to read it. This work does not claim that the protection mechanisms are entirely secure. Rather, they are useful to temporarily protect sensitive information from a passerby who happens to glance at the display, where again social protocol is expected to stop any serious attempt to breach one's privacy. A system evaluation has been conducted in order to refine several parameters, such as the silhouette's measurements and data has been collected to be used for the angle at which people can see content on a large display.

On many different occasions the system was demonstrated to various researchers from the HCI community. They were given the opportunity to test the system themselves and see how it performed and provided them with awareness about whether they are currently being overlooked or infringing someone's space. Informal feedback gave valuable insight that the shielding of information with one's body actually does work. They were aware of the fact that someone is looking and they liked the freedom the system gave them in order to negotiate about their personal space and allowed for collaboration and negotiation. This feedback was very valuable in the entire evolution of the system and exploration of how people use awareness and protection when protecting their personal information and negotiate their personal space. However some people noted that they see some of the system's implicit protection mechanisms fail for deliberate privacy intrusions.

# 8 Future Work

This thesis explored the design space of awareness and protection of sensitive information in a proxemic aware environment. Many social aspects of how people behave in everyday life have been exploited and this works relies on the fact that social norms are well known and accepted. In this section several aspects of how this work can be continued and built upon are described.

**Tracking technology.** The system has been implemented, using a Vicon motion tracking system. This system is very accurate, but also very expensive and tedious to set up. It is a good tool for prototyping but clearly not deployable in the wild as people are required to wear markers. However, alternate low-cost technologies can be used instead, such as the marker-less Microsoft Kinect. By distributing multiple Kinect sensors in an area, a similar level of detail can be achieved at much lower cost. With the soon available Kinect 2 all the required information, including body motion tracking (via skeleton tracking) is available. Even gaze orientation (via facial recognition) is possible and by analyzing skeletal features such as shoulder width even more data can be used for these purposes.

**Awareness indication.** The methods presented are suggestive of a broader range of other approaches. For example, the notion of indicating the presence and position of a passerby can be realized via many other cues. Examples are different visualizations, vibration or 3D sound. For each of these techniques it would be interesting to explore how the information can be encoded and whether a user can conclude a passerby's whereabouts and intentions from the system. Cues can be constructed to match the fidelity of the sensed information. This is already done by some of the techniques provided, as for example encoding the passerby's position in multiple ways in some cues, whereas others only provide a broad information about the presence of a passerby.

**Protection techniques.** Similarly, the idea of offering protection by masking information from view on the display can take many visual forms. Design trade-offs will include how understandable the cue is to all parties, the degree of awareness provided by these cues, the distraction caused by the cue or protection mechanism, the degree of security provided, and the amount of effort required by the parties to either explicitly control the system or override the implicit actions taken by the system.

**Collaboration.** This work is not only about privacy and securing one's personal information, but also about negotiation and collaboration. For example, blacking out certain parts of the display can be an easy way to focus someone's attention on the non-blacked-out parts. Because of their size, large displays can easily be used by multiple people at the same time. An aspect to study would be, how large displays can be used for collaboration, either among people who know each other but also among anonymous collaborators. An exploration can build upon a range of previous work. One interesting aspect would be how a seamless transition from single to multi-user operation could work.

**Small displays.** This work assumed the use of a large, public display. On the contrary, privacy violations can occur on small displays. How can these techniques be translated to small devices in a mobile setting? What kind of tracking technology could be used and what cues can provide awareness without cluttering the (small) screen?

**Observation of shoulder surfing.** During the collection of the viewing-direction data (section 6.5.2) people stated that they felt uncomfortable, viewing out of the corner of their eyes, especially when done for a longer time. Is this extreme case really practiced in real life? An observation of everyday shoulder surfing encounters can give further insight of how people actually overlook other people's information. In what situations and how likely are they to shoulder surf? Which data is of the most interest? All of this given, that many shoulder surfing incidents are inadvertent.

**Tricking the system and gaze direction.** While demonstrating the system to many people, it became clear, that some of techniques can be tricked by a deliberate privacy violator. It has been stated by several researchers, that a person's viewing direction can be assumed from his head orientation (see section 4.3). When using a 3D tracking system, how accurate are these information about the gaze direction? A passerby can easily trick the system, by looking out of the corner of his eye, while the tracked viewing direction will point away from the display. A more detailed survey cannot only help to improve rapid and inexpensive prototyping using gaze, but also allow to transfer these techniques to a real world setup. The University of Calgary Conjoint Faculties Research Ethics Board has approved the ethics application for a research study, investigating whether a person's viewing direction can be assumed from his head orientation. The details about this study application are listed in appendix A. Using inexpensive and easy-to-setup gaze tracking can spark new means of control of interactive systems.

**Multiple users and passers-by.** To this point, the explorations have considered only the case of a single passerby and a single user of the display. Thus they are likely appropriate for non-crowd situations where only occasional people pass by. Still, some of the approaches are somewhat scalable to include a few passers-by. For example, multiple 3D-models and gaze indicators (one for each person in the scene) can be included, or silhouette's size and position can be calculated as a function of multiple vectors representing each person. Again, there are tradeoffs. For example, the silhouette would shrink considerably or even disappear because there may be no display area that would be completely shielded from at least one person's view by the user's body (especially if passers-by are far apart). This can be remedied somewhat by weighting in the passers-by's viewing orientation, where those passers-by can be left out of the calculation, that are currently not looking at the display. On the other hand, when more than one user is working on the display, the area covered by their bodies is larger. Therefore the effectiveness of the silhouette increases with an increasing number of users.

# References


1. 3M. “3M Screen Privacy & Screen Protectors.” “3M Screen Privacy & Screen Protectors.” [Online]. Available at: http://solutions.3m.com/wps/portal/3M/en_US/3MScreens_NA/Protectors/. [Last Accessed: 03-Mar-2014].

2. Alt, Florian; Shirazi, Alireza Sahami; Kubitza, Thomas; and Schmidt, Albrecht. “Interaction techniques for creating and exchanging content with public displays.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’13*, ACM Press (2013), 1709–1718.

3. Altman, Irwin. *The environment and social behavior: privacy, personal space, territory, and crowding.* 1975.

4. Ballendat, Till; Marquardt, Nicolai; and Greenberg, Saul. “Proxemic interaction: designing for a proximity and orientation-aware environment.” *Proceedings ACM International Conference on Interactive Tabletops and Surfaces*, (2010), 121–130.

5. Barefoot, John C.; Hoople, Howard; and McClay, David. “Avoidance of an act which would violate personal space.” *Psychonomic Science 28*, 4 (1972), 205–206.

6. Becker, Franklin D. “Study of spatial markers.” *Journal of personality and social psychology 26*, 3 (1973), 439–45.

7. Bellotti, Victoria. “Design for privacy in multimedia computing and communications environments.” *Technology and privacy: The new landscape*, (1997), 63–98.

8. Berger, Stefan; Kjeldsen, Rick; Narayanaswami, Chandra; Pinhanez, Claudio; Podlaseck, Mark; and Raghunath, Mandayam. “Using symbiotic displays to view sensitive information in public.” *Third IEEE International Conference on Pervasive Computing and Communications, 2005. PerCom 2005.*, IEEE (2005), 139–148.

9. Bianchi, Andrea. “Authentication on public terminals with private devices.” *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, ACM Press (2011), 429–430.

10. Bjorke, Oystein. “Helix 3D Toolkit.” “Helix 3D Toolkit.”, *under the MIT License (MIT)*. [Online]. Available at: https://helixtoolkit.codeplex.com/. [Last Accessed: 23-Feb-2014].

11. Boyle, Michael; and Greenberg, Saul. “The language of privacy: Learning from video media space analysis and design.” *ACM Transactions on Computer-Human Interaction (TOCHI) 12*, 2 (2005), 328–370.

12. Brignull, Harry; and Rogers, Yvonne. “Enticing people to interact with large public displays in public spaces.” *Proceedings of INTERACT*, (2003), 17–24.

13. Brudy, Frederik; Ledo, David; Greenberg, Saul; and Butz, Andreas. “Is Anyone Looking? Mitigating Shoulder Surfing on Public Displays through Awareness and Protection.” *Research Report 2014-1056-07, Department of Computer Science, University of Calgary, Calgary, Alberta, Canada T2N 1N4.Department of Computer Science, University of Calgary, Calgary, Alberta, Canada T2N 1N4.*, (2014), 1–6.

14. Brudy, Frederik; Ledo, David; and Greenberg, Saul. “Is Anyone Looking? Mediating Shoulder Surfing on Public Displays (The Video).” *CHI ’14 Video Showcase, Proc. Extended Abstracts: ACM SIGCHI Conference on Human Factors in Computing Systems, 2014*, (2014), 1.

15. Canonical Ltd. “Ubuntu for Android.” “Ubuntu for Android.” [Online]. Available at: http://www.ubuntu.com/phone/ubuntu-for-android. [Last Accessed: 09-Mar-2014].

16. Casiez, Géry; Roussel, Nicolas; and Vogel, Daniel. “1€ filter: a simple speed-based low-pass filter for noisy input in interactive systems.” *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, (2012), 2527–2530.

17. Coventry, Lynne; Angeli, Antonella De; and Johnson, Graham. “Usability and biometric verification at the ATM interface.” *Proceedings of the SIGCHI conference on Human factors in computing systems*, (2003), 153–160.

18. Edney, Julian J. “Property, Possession and Permanence: A Field Study in Human Territoriality.” *Journal of Applied Social Psychology 2*, 3 (1972), 275–282.

19. Edney, Julian J. “Human Territories as Organizers: Some Social and Psychological Consequences of Attachment to Place.” *Environment and behavior 7*, 2 (1975).

20. Ellis, Stephen R.; Young, Mark J.; Adelstein, Bernard D.; and Ehrlich, Sheryl M. “Discrimination of Changes of Latency during Voluntary Hand Movement of Virtual Objects.” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, (1999), 1182–1186.

21. Fox, Susannah. *51 % of U . S . Adults Bank Online*. 2013.

22. Freedman, E.G.; and Sparks, D.L. “Coordination of the eyes and head: movement kinematics.” *Experimental brain research 131*, 1 (2000), 22–32.

23. Google Inc. “Chromecast.” “Chromecast.” [Online]. Available at: http://www.google.ca/intl/en/chrome/devices/chromecast/. [Last Accessed: 09-Mar-2014].

24. Greenberg, Saul; Boyle, Michael; and Laberge, Jason. “PDAs and shared public displays: Making personal information public, and public information personal.” *Personal Technologies 3*, 1 (1999), 54–64.

25. Greenberg, Saul; Marquardt, Nicolai; Ballendat, Till; Diaz-Marino, Rob; and Wang, Miaosen. “Proxemic Interactions: The New Ubicomp?” *interactions 18*, 1 (2011), 42–50.

26. Hall, Edward Twitchell. *The hidden dimension*. Anchor Books New York, 1969.

27. Harrison, Chris; and Hudson, Scott E. “A new angle on cheap LCDs: making positive use of optical distortion.” *Proceedings of the 24th Annual ACM Symposium on User interface Software and Technology*, (2011), 537–539.

28. Koppel, Maurice Ten; Bailly, Gilles; Müller, Jörg; and Walter, Robert. “Chained displays: configurations of public displays can be used to influence actor-, audience-, and passer-by behavior.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (2012), 317–326.

29. Kruger, Russell; Carpendale, Sheelagh; Scott, Stacey D.; and Greenberg, Saul. “How People Use Orientation on Tables: Comprehension, Coordination and Communication.” *Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work*, (2003), 369–378.

30. Kumar, Manu; Garfinkel, Tal; Boneh, Dan; and Winograd, Terry. “Reducing shoulder-surfing by using gaze-based password entry.” *Proceedings of the 3rd symposium on Usable privacy and security - SOUPS ’07*, ACM Press (2007), 13–19.

31. De Luca, Alexander; and Frauendienst, Bernhard. “A Privacy-respectful Input Method for Public Terminals.” *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*, ACM Press (2008), 455–458.

32. Luca, Alexander De; Hang, Alina; Brudy, Frederik; Lindner, Christian; and Hussmann, Heinrich. “Touch me once and i know it’s you!: implicit authentication based on touch screen patterns.” *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, (2012), 987–996.

33. De Luca, Alexander; Von Zezschwitz, Emanuel; and Hußmann, Heinrich. “Vibrapass: secure authentication based on shared lies.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (2009), 913–916.

34. Marquardt, Nicolai; Ballendat, Till; Boring, Sebastian; Greenberg, Saul; and Hinckley, Ken. “Gradual Engagement between Digital Devices as a Function of Proximity: From Awareness to Progressive Reveal to Information Transfer.” *Proceedings of Interactive Tabletops & Surfaces*, (2012).

35. Marquardt, Nicolai; Diaz-Marino, R.; Boring, Sebastian; and Greenberg, Saul. “The proximity toolkit: prototyping proxemic interactions in ubiquitous computing ecologies.” *Proceedings of the 24th annual ACM symposium on User interface software and technology*, (2011), 315–326.

36. Marquardt, Nicolai; and Greenberg, Saul. “Informing the Design of Proxemic Interactions.” *IEEE Pervasive Computing 11*, 2 (2012), 14–23.

37. Marshall, Paul; Morris, Richard; Rogers, Yvonne; Kreitmayer, Stefan; and Davies, Matt. “Rethinking ‘Multi-user’: an In-the-Wild Study of How Groups Approach a Walk-Up-and-Use Tabletop Interface.”*Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (2011), 3033–3042.

38. Matt Richtel. “Pornography in Public Causes Some to Gasp, Others to Shrug - NYTimes.com.” “Pornography in Public Causes Some to Gasp, Others to Shrug - NYTimes.com.”, *The New York Times*. [Online]. Available at: http://www.nytimes.com/2012/07/21/us/tablets-and-phones-lead-to-more-pornography-in-public.html. [Last Accessed: 02-Mar-2014].

39. Michelis, Daniel; and Müller, Jörg. “The audience funnel: Observations of gesture based interaction with multiple large displays in a city center.” *Intl. Journal of Human–Computer Interaction 27*, 6 (2011), 562–579.

40. Müller, Jörg; Walter, Robert; Bailly, Gilles; Nischt, Michael; and Alt, Florian. “Looking glass: a field study on noticing interactivity of a shop window.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (2012), 297–306.

41. Nancel, Mathieu; Chapuis, Olivier; Pietriga, Emmanuel; Yang, Xing-Dong; Irani, Pourang P.; and Beaudouin-Lafon, Michel. “High-precision pointing on large wall displays using small handheld devices.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM Press (2013), 831–840.

42. Nickel, Kai; and Stiefelhagen, Rainer. “Pointing Gesture Recognition based on 3D-Tracking of Face , Hands and Head Orientation Categories and Subject Descriptors.” *Proceedings of the 5th international conference on Multimodal interfaces*, (2003), 140–146.

43. Palen, Leysia. “Social, individual and technological issues for groupware calendar systems.” *Proceedings of the SIGCHI conference on Human factors in computing systems*, (1999), 17–24.

44. Pavlovych, Andriy; and Stuerzlinger, Wolfgang. “The tradeoff between spatial jitter and latency in pointing tasks.” *Proceedings of the 1st ACM SIGCHI symposium on Engineering interactive computing systems*, ACM Press (2009), 187–196.

45. Peltonen, Peter; Kurvinen, Esko; Salovaara, Antti; et al. “It’s Mine, Don't Touch!: interactions at a large multi-touch display in a city centre.” *Proceedings of the SIGCHI conference on human factors in computing systems*, (2008), 1285–1294.

46. Scott, Stacey D.; Carpendale, Sheelagh; and Inkpen, Kori M. “Territoriality in collaborative tabletop workspaces.” *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, (2004), 294–303.

47. Scott, Stacey D.; and Carpendale, Sheelagh. “Investigating Tabletop Territoriality in Digital Tabletop Workspaces.” *Technical Report 2006- 836-29, Department of Computer Science, University of Calgary., Calgary, AB, Canada*, (2006), 1–10.

48. Sharp, Richard; Madhavapeddy, Anil; Want, Roy; and Pering, Trevor. “Enhancing web browsing security on public terminals using mobile composition.” *Proceedings of the 6th international conference on Mobile systems, applications, and services*, (2008), 94–105.

49. Sharp, Richard; Scott, James; and Beresford, AR. “Secure mobile computing via public terminals.” *Pervasive Computing 2*, 2006 (2006), 238–253.

50. Shoemaker, Garth B.D.; and Inkpen, Kori M. “Single display privacyware: augmenting public displays with private information.” *Proceedings of the SIGCHI conference on Human factors in computing systems*, (2001), 522–529.

51. Shoemaker, Garth B.D. “Supporting Private Information on Public Displays.” *CHI ’00 Extended Abstracts on Human Factors in Computing Systems*, (2000), 349–350.

52. Simon, Herbert Alexander. *The sciences of the artificial*. 1996.

53. Smith, Aaron; and Duggan, Maeve. “Online Dating & Relationships | Pew Research Center’s Internet & American Life Project.” “Online Dating & Relationships | Pew Research Center’s Internet & American Life Project.”, *Pew Research Center Internet Project*. [Online]. Available at: http://www.pewinternet.org/2013/10/21/online-dating-relationships/. [Last Accessed: 02-Mar-2014].

54. Sommer, Robert; and Becker, Franklin D. “Territorial defense and the good neighbor.” *Journal of personality and social psychology 11*, 2 (1969), 85–92.

55. Sommer, Robert. *Studies in personal space*. 1967.

56. Sommer, Robert. *Personal Space. The Behavioral Basis of Design*. 1969.

57. Stahl, J.S. “Amplitude of human head movements associated with horizontal saccades.” *Experimental brain research 126*, 1 (1999), 41–54.

58. Tan, Desney S.; and Czerwinski, Mary. “Information voyeurism: Social impact of physically large displays on information privacy.” *CHI’03 Extended Abstracts on Human Factors in Computing Systems*, (2003), 748–749.

59. Tan, Desney S.; Keyani, Pedram; and Czerwinski, Mary. “Spy-resistant keyboard: more secure password entry on public touch screen displays.” *Proceedings of the 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future*, (2005), 1–10.

60. Tang, John C. “Findings from observational studies of collaborative work.” *International Journal of Man-machine studies 34*, 2 (1991), 143–160.

61. Vogel, Daniel; and Balakrishnan, Ravin. “Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users.” *Proceedings of the 17th annual ACM symposium on User interface software and technology*, (2004), 137–146.

62. Wang, Miaosen; Boring, Sebastian; and Greenberg, Saul. “Proxemic Peddler: A Public Advertising Display that Captures and Preserves the Attention of a Passerby.” *Proceedings of the 2012 International Symposium on Pervasive Displays*, (2012), 1–6.

63. Weisband, Suzanne P.; and Reinig, Bruce A. “Managing user perceptions of email privacy.” *Communications of the ACM 38*, 12 (1995), 40–47.

64. Wobbrock, Jacob O.; Wilson, Andrew D.; and Li, Yang. “Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes.” *Proceedings of the 20th annual ACM symposium on User interface software and technology*, (2007), 159–168.

# Contents of Enclosed CD

- This thesis as a Word and PDF file.
- A video, demonstrating some of the aspects [14].
- Tech Report, showing some of these techniques [13].
- C# Program of the implemented system. To be used in conjunction with the Proximity Toolkit and the Vicon Tracking system. Further the Microsoft Kinect SDK has to be installed.
- Tracked data for jitter measurements and viewing direction.
- R Scripts for evaluation of the tracked data (jitter and viewing direction).
- All the related work used in this thesis (if available electronically).

# Definition of Terms

**Sensitive information / private data:** Sensitive information has a different meaning to different people. Information can be sensitive to different parties: the owner of the data (e.g. the person represented by the data), the user of the data (the person working with the data) and the (inadvertent) viewer of the data (when catching a glimpse of data he does not want to see). A single person can engage in multiple roles. Private information does not mean it is private to only one person, it can also be private to a group of people [50]. A detailed review of what private information is can be found in chapter 3.3.

**Display:** A display can either be public, semi-public or private.

- Public displays: Can be overseen (and often used) by anyone. Examples are navigation terminals in a shopping mall, ATM terminals, train ticket vending machines, etc. For more example see section 1.1.1)
- Semi-public displays: Located in a shared space, but limited to a defined group, such as a display in an open office environment or research lab
- Private displays: Located in a very controlled environment. It is known (and usually controlled) who has access to the display. Usually only the owning individual himself and close member of his social circle can access it.

A display can be of various sizes. Not only wall-mounted displays should be considered, but also desktop, mobile and handheld devices.

**User of the system:** The person who is legitimately working on the display. Often it is their data being displayed on the screen. Two people can work together on a large display, without one of them necessarily being treated as a legitimate user and the other one as a passer-by / intruder to the other. They can have information up on the screen which is considered sensitive to both of them. Thus they are both considered to be a valid user of the screen and the contents on it.

**Passerby:** A person who is sufficiently close to a public display to be able to oversee the content [39]. The specific distance depends on the display itself and the area it is located in. With small displays, such as an ATM screen, someone passing by in 20 meters distance can most likely not see much of the content on the screen. Thus the threat of shoulder surfing is minimal (although still present), especially when it is located in a non-crowded place. With a large directory in a shopping mall the threat might be higher as the display area is larger and it often is located in a busy position with many people walking by.

**Awareness:** Understanding of a person, what other people are doing. Awareness can provided about information such as position and location, distance, movement, orientation, look direction, identity, intention, etc.

# A Appendix A

Some of the previously described techniques only included sketches. This appendix shows photographs of the actual system, if they have not been included in the previous chapters.

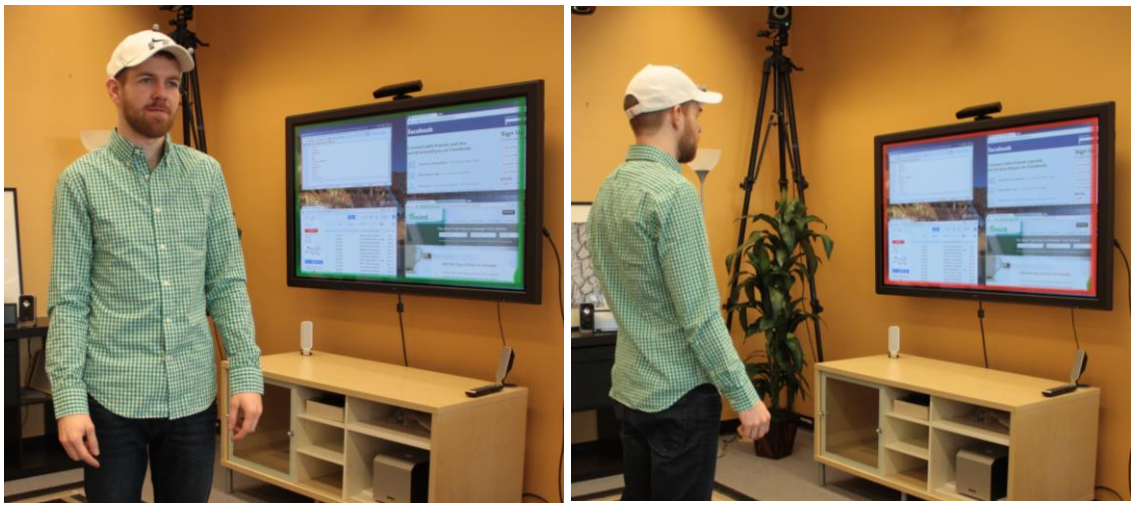

**Figure A.36. The display's borders flash green when a passerby is present. When he is looking at the display the color turns to red and the speed of the flashing increases.**

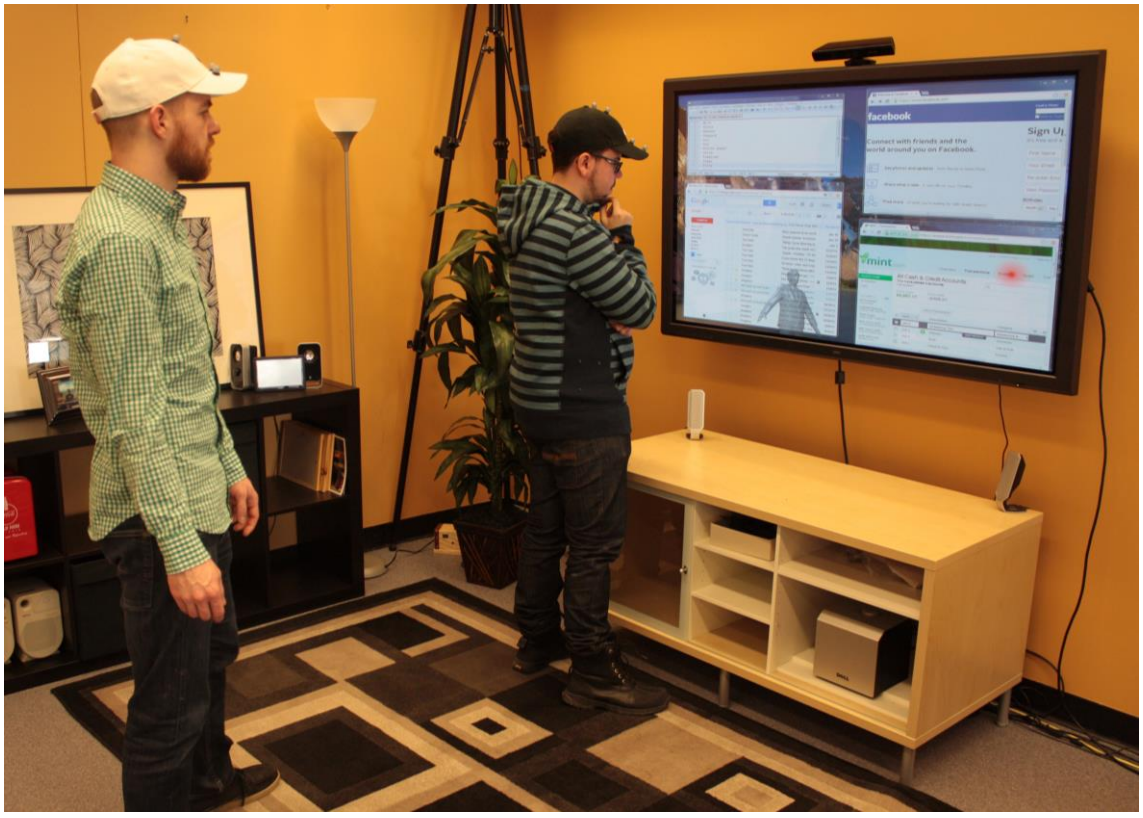


**Figure A.37. The 3D-model follows a passerby's position and orientation. The gaze awareness indicator (red dot) indicates the passerby's viewing position on the display.**

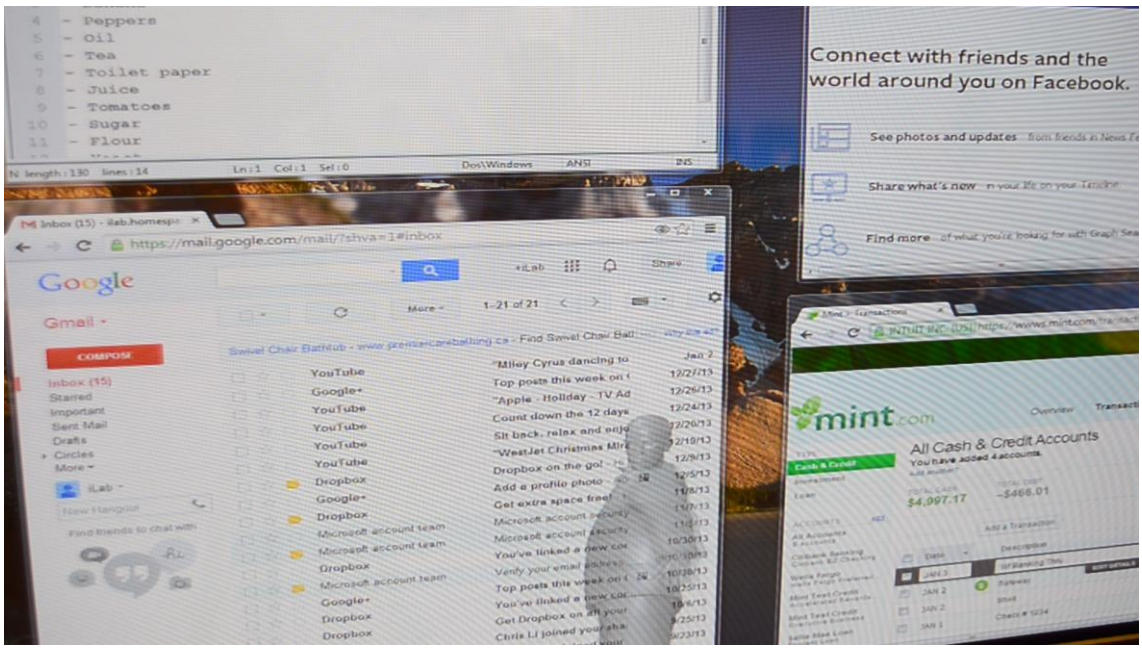


**Figure A.38. The passerby's head and torso are tracked separately and their rotation are mapped to the model's head and torso.**

**Figure A.39. Implicit: Blacking out sensitive content. The opacity of the cover is set so that it can still be read when standing close, but difficult to read from a distant**

**Figure A.40. The silhouette protection. Only those parts of the display are visible that are covered by the user's body.**

# B Appendix B

Detailed data about the measured jitter, when intersecting the forward vector of a tracked entity with the display. The millimeter measurements and further information can be found in section 6.10.

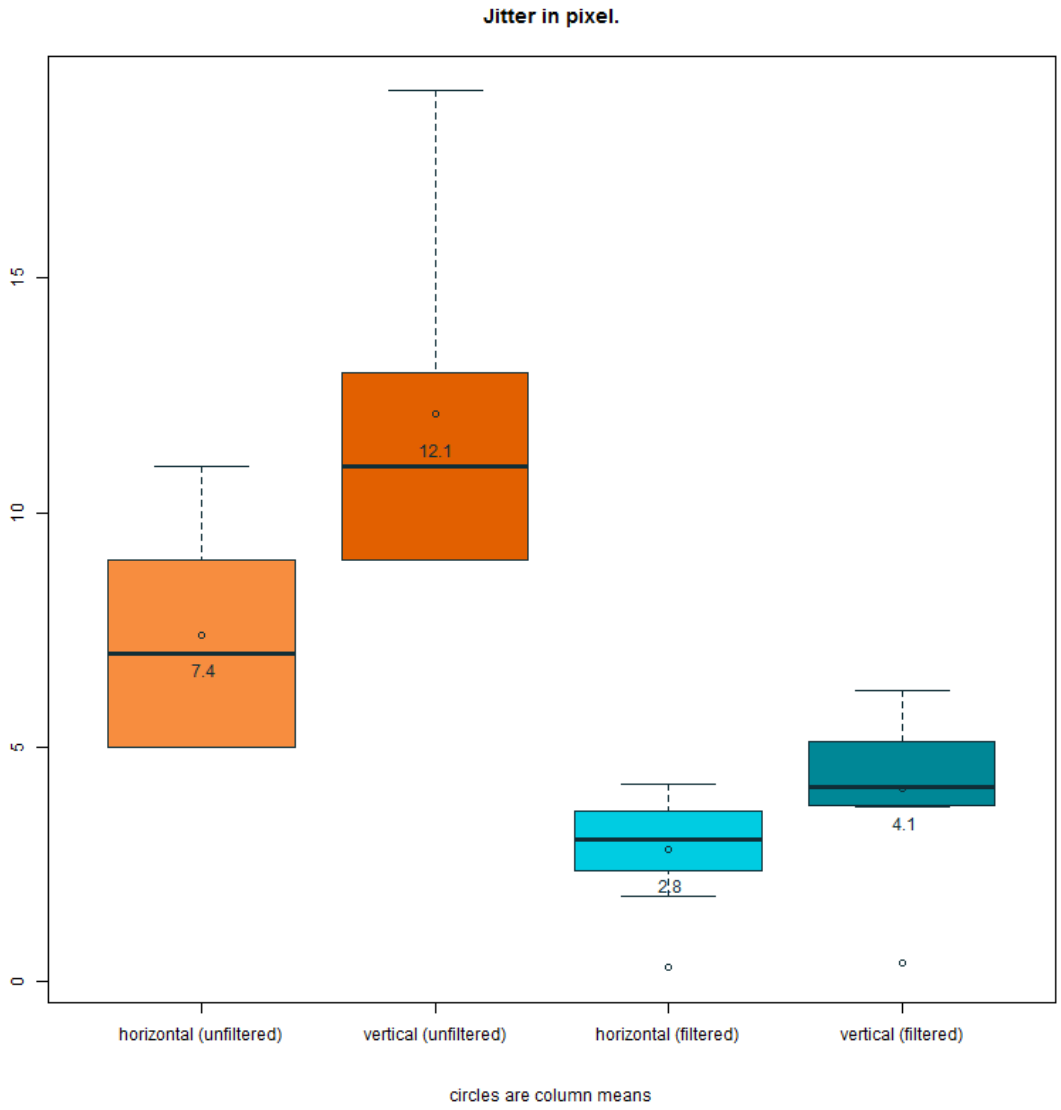


**Figure A.41. The measured jitter of the intersection point of the forward vector from a tracked entity with the display (in pixels). Measurements were taken during ten 15-second periods. The tracked entity was immobilized on a tripod in the middle of the room.**

# C Appendix C

At which maximum left / right angle are people still able to tell what is currently displayed on the screen. The data provided is the summary of 8 participants in a system evaluation. Details can be found in section 6.5.2.

**Condition A:** participants were asked look straight ahead.

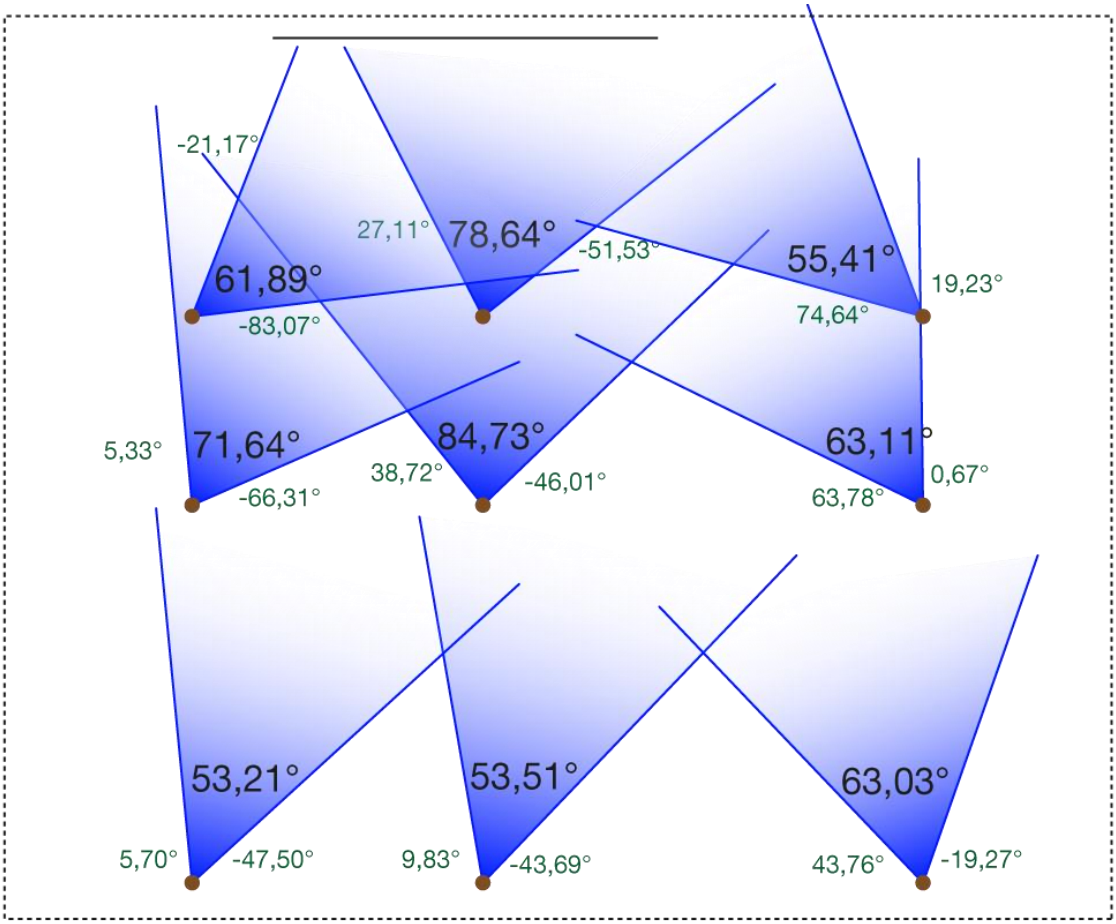


**Figure A.42. Looking straight ahead, being asked not to gaze out of the corner of their eye.**

| Distance | Position | Angle Left | Angle Right | Angle Absolute | average row |
|---|---|---|---|---|---|
| near | Left | -21,17° | -83,07° | 61,89° | |
| near | Center | 27,11° | -51,53° | 78,64° | 65,31° |
| near | Right | 74,64° | 19,23° | 55,41° | |
| middle | Left | 5,33° | -66,31° | 71,64° | |
| middle | Center | 38,72° | -46,01° | 84,73° | 73,16° |
| middle | Right | 63,78° | 0,67° | 63,11° | |
| far | Left | 5,70° | -47,50° | 53,21° | |
| far | Center | 9,83° | -43,69° | 53,51° | 56,58° |
| far | Right | 43,76° | -19,27° | 63,03° | |

**Table A.2. Average angles when people were asked to look straight ahead at the display.**

**Condition B:** participants were asked to gaze out of the corner of their eyes.

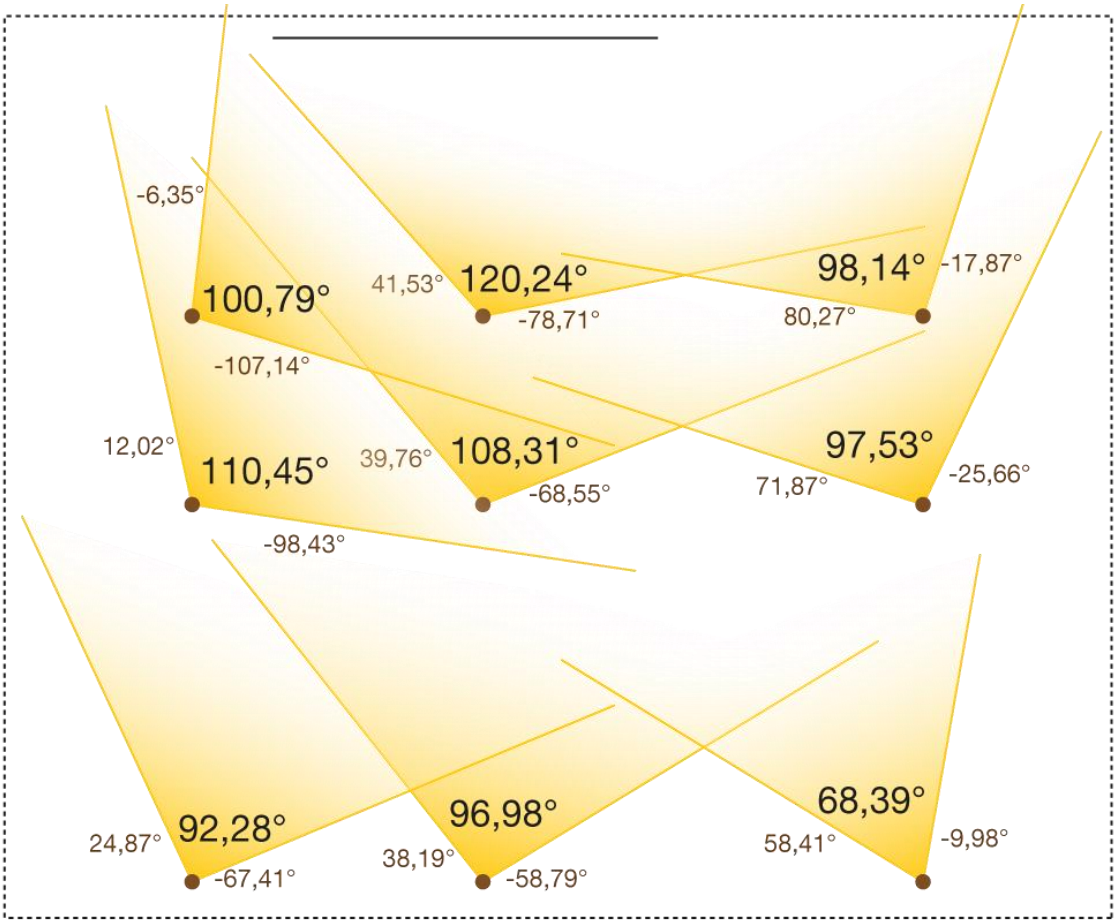


**Figure A.43. Looking out of the corner of their eyes.**

| Distance | Position | Angle Left | Angle Right | Angle Absolute | average row |
|---|---|---|---|---|---|
| near | Left | -6,35° | -107,14° | 100,79° | 106,39° |
| near | Center | 41,53° | -78,71° | 120,24° | |
| near | Right | 80,27° | -17,87° | 98,14° | |
| middle | Left | 12,02° | -98,43° | 110,45° | 105,43° |
| middle | Center | 39,76° | -68,55° | 108,31° | |
| middle | Right | 71,87° | -25,66° | 97,53° | |
| far | Left | 24,87° | -67,41° | 92,28° | 85,88° |
| far | Center | 38,19° | -58,79° | 96,98° | |
| far | Right | 58,41° | -9,98° | 68,39° | |

**Table 3. Average angles when people were asked to gaze out of the corner of their eyes.**

# D Appendix D

| | Awareness to User | Awareness to Passerby | Presence | Location | Movement | Orientation | Look Direction | Distance | Protection |
|---|---|---|---|---|---|---|---|---|---|
| Flashing Borders | ■ | ■ | ■ | | | | ■ | | |
| 3D-Model | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | |
| Gaze Awareness Indicator | ■ | | | | | | ■ | ■ | |
| Explicit: Moving / Hiding Content | | | | | | | | | ■ |
| Implicit: Blacking out Content | ■ | ■ | ■ | | | | | | ■ |
| Implicit: Silhouette | ■ | ■ | ■ | ■ | ■ | | | ■ | ■ |

**Table A.4. Comparison of the implemented systems on terms of their ability to provide awareness (orange), their level of detail (yellow) and whether they allow for protection of sensitive information.**

# E Appendix E

The entire system was built as single program. The controls window allows for customization at run time. The settings will roughly be explained here.

a) Selection of the desired means of awareness and / or protection.
b) Most parameters can be adjusted at runtime and will be saved for later use.
c) Setting whether to opacity should be used to represent information and adjustments to the 1€ Filter. Also the measurements for jitter are made here (see section 6.10 for details)
d) Setting which jitter should be measured.
e) Allows to use / demonstrate the system without people walking around.
f) Experimental setting which input source should be used for gesture recognition.
g) Various methods to either refine parameters or show verbose debug output.

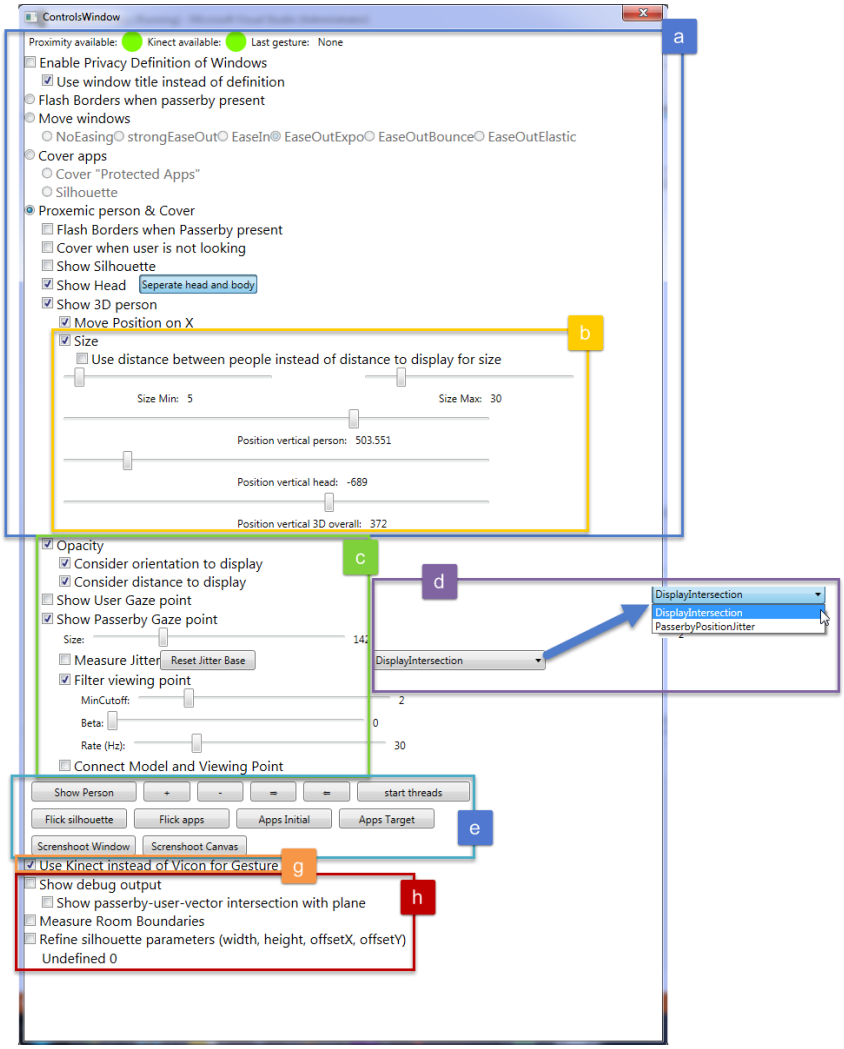


**Figure A.44. The main control window of the prototype, allowing for great customization.**

# F Appendix F

In this appendix, the verbal protocol, consent form and questionnaire is listed, to show an example of how a study could evaluate whether it is possible to assume a person's viewing direction from their head orientation when being tracked with a 3D tracking system. This study prosposal has been submitted in collaboration with David Ledo, Jiannian Li and Saul Greenberg and has been approved by the University of Calgary Conjoint Faculties Research Ethics Board.

## Verbal protocol

*The following description will be read to each participant at the beginning of the study to inform participants of the procedures prior to giving consent. Italicized text are instructions to the investigator.*

Hello and welcome to the iLab, my name is <*experimenter*>, and I will guide you through the experiment. Feel free to ask me any question at any time.

Before we start I need to let you know about your rights as a participant.

- If you feel uncomfortable you may quit at any time. The data that we have collected up to that point will be kept as long as you signed the consent forms.
- No data will be used without your explicit consent.

Now please read this consent form carefully, as it explains your rights as a participant and the conditions of the study, and sign it if you agree with these terms. <*hand form to the participant, go through consent form with participant and give them time to read on their own*>

Now, I will ask you to sign this sheet that indicates that you have received the $15 for your participation <*hand payment table*>.

If you find that you are struggling with the task, it is likely that there is a problem in the study itself that we need to address. Just let us know. Remember there is nothing you can do wrong in this study. If you have any questions at any point feel free to ask them.

Although I don't know of any reason for this to happen, if you should become uncomfortable or find this test objectionable in any way, you are free to quit at any time. Also if you would like to take a break just let me know and I will pause the system.

You may have a copy of the consent form for your own records.

Before we go on with the instructions of the experiment, I would like to let you know that we appreciate you helping us in this study.

First I'm going to introduce you to your task. I'd like to ask you to wear these special pair of glasses during the study, which enable us to track your head. During the study you will be asked to stand at four different positions, marked on the floor.

*Point to floor, showing the marks.*

Once you positioned yourself comfortably I would like to ask you not move your feet as little as possible, however you may move the rest of body as you like.

On the display you will see a circle appear. In the center of the circle a countdown from 3 to 1 will be shown, after which a letter will be shown inside the circle. Your task is it to read out loud the letter that is shown after the countdown is over. For reading the letter you will have two seconds time. The circle will then appear at a different position and the countdown starts from the beginning. Again, your task will be to read the letter, which is

presented to you at the end of the countdown, out loud.

*Start the demo application, which shows a countdown and a letter at the screen center.*

From time to time you will see written instructions on the screen. If you are presented with one of them just turn your attention to me for further instructions.

Do you understand what your task will be? Do you have any questions?

*If the participant consented to video recording, start the video recording now.*

Now that you are familiarized with your task, we will proceed with the actual study. I will kindly ask you to think aloud: this means, speak your mind about anything that you might think is related to our study. We will take notes of some of the comments you make, which will help us better understand if there are issues we need to address.

<*Start study application*>

*The application will guide the participant throughout the study. Make sure you ask the participant regularly (approximately every 5 minutes) if they are still feeling alright and if they would like to take a break. Make sure to give them the questionnaire after each distance, which belongs to the given distance (numbered from 1 to 4). Given that we have four different distances, this will happen four times. Note that to minimize learning effects, the distances will be shown in different order.*

*When the participant has finished doing all trials*

Next, I will ask you to please fill in this last form <*hand in final questionnaire sheet*>.

<*End study application*>

*After the last questionnaire is filled ask the following semi-structured interview questions –*

Now, I would like to take some time to ask you some questions

1. Did you find any issues while using the techniques, what were they?
2. What was your general approach for looking at a target? Did you have a particular strategy?
3. Did you feel you had to move your head a lot?
4. Do you think there is a discrepancy between the position your head is pointing at and your viewing position? How would you rate it?
5. Is there anything else you would like us to know?

Thank you very much for participating, I hope you have a good day!

<*End video recording, save tracking data*>

*Note: The questionnaires and verbal protocol included are indicative of what we will ask and say. Minor modifications may be made to smooth out our process. Additional questions may be asked depending upon particular comments and / or actions observed as the study progresses.*

# Consent Form


| Frederik Brudy | David Ledo | Jiannan Li | Saul Greenberg |
|---|---|---|---|
| M.Sc. Student | M.Sc. Student | M.Sc. Student | Professor |
| University of Munich | Department of Computer Science | Department of Computer Science | Department of Computer Science |
| E-mail: f.brudy@ucalgary.ca | E-mail: david.ledo@ucalgary.ca | E-mail: jiannan.li@ucalgary.ca | E-mail: saul.greenberg@ucalgary.ca |
| Phone: 587 968 5255 | Phone: 403 210 9499 | Phone: 403 399 8791 | Phone: 403 220 6087 |


**Title of Project: Correlation of head position to viewing direction**

This consent form, a copy of which has been given to you, is only part of the process of informed consent. If you want more details about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

The University of Calgary Conjoint Faculties Research Ethics Board has approved this research study.**1. Purpose of the Study:**

Current ubiquitous computing technologies make use of people's looking direction for two different purposes: first, they enable implicit actions, which computer systems can use for interaction; second, they allow people interacting with the system understand where they are looking at (gaze awareness). However, specifically tracking the gaze can be difficult: it requires sophisticated equipment with potentially intrusive technology, and also requires a thorough understanding of the meaning of the direction they are looking at. Conversely, we have the ability to track people's position in a room and we know their whereabouts, as well as the position and orientation of the head. We are interested in determining whether the head orientation and position is a good indicator of the looking direction for people.

**2. What Will I Be Asked To Do?**

We will first ask you to provide some basic demographic information about yourself. Next you will be given a pair of non-prescriptive glasses, which enable us to track your head position and orientation. You will then be asked you to look at various positions on a wall mounted TV screen. We will conduct this task at several different distances from the screen. After each distance you will be asked to answer a short questionnaire. If there are written instructions on the screen, just turn your attention to the interviewer for any further instructions.

If you find that you are struggling with the task, it is likely that there is a problem in the study itself that we need to address. Just let us know. Remember there is nothing you can do wrong in this study. If you have any questions at any point feel free to ask them.

If you feel you need a break at any time feel free to tell us and we will pause the study.

This experiment is expected to take about 45 minutes. We do not foresee any risks from participating in this study.

Keep in mind that your participation in this study is completely voluntary. You are free to withdraw from the experiment at any time without any kind of penalty. If you decide to withdraw, the experiment will be interrupted immediately. However, we will reserve the right to keep and use the data collected until the point of withdrawal. Participating in the experiment will grant you a total of $15. Should you choose to withdraw, you will still be allowed to keep the $15.

**3. What Type of Personal Information Will Be Collected?**

Should you agree to participate, you will be asked to provide basic demographic information. We will also be using a state-of-the-art tracking system. This will be used to track your head's position and orientation throughout the study.

We will take notes as you interact. Aspects that we might write down include: problems occurring during the interaction, such as when the application does not respond as expected, or some opinions that you may state. We may also record video and audio of this session with your explicit consent (see next page).

**4. Are there Risks or Benefits if I Participate?**There are no known harms or risks associated to the participation in this study. If you participate you will receive a compensation of $15 for your time. You will also have the opportunity of using a state-of-the-art interface facility.

**5. What Happens to the Information I Provide?**

The researchers will record your interaction with the computer, your responses to the questionnaires that you complete. Only the researchers will have access to the full recordings and the responses that you provide.

This information will be kept in a secure location (locked cabinets and password-protected drives). The information that we collect will not be associated to you personally. However, the researchers will publish the results of their analysis of your data in anonymized form in academic journals and conference papers.

The researchers might quote the responses in the questionnaires or any of your comments in anonymized form, and they may use still images taken during the interaction in research presentations and publications. Please check the boxes below to confirm that you understand this use of your data.

____I agree that the researchers may use any written or verbal comments and answers I may provide in research presentations and publications.

____ I consent for this session to be video and / or audio recorded.

____I agree that the researchers may use some of the video and / or audio content of my actions to illustrate their findings in research presentations and publications. I realize that once the video has been released through publications or presentations it is not under the control of the researches who has access to it.

All the collected data will be kept by the investigators for at least a year, where it will be destroyed after it is no longer required.

7. Signatures (written consent)

Your signature on this form indicates that you 1) understand to your satisfaction the information provided to you about your participation in this research project, and 2) agree to participate as a research subject.

In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from this research project at any time. You should feel free to ask for clarification or new information throughout your participation.

Participant's Name: (please print)
_____

Participant's Signature _____Date:
_____

Researcher's Name: (please print)
_____

Researcher's Signature: _____Date: _____8. Questions/Concerns

If you have any further questions or want clarification regarding this research and/or your participation, please contact:

Frederik Brudy
587 968 5255, fb@fbrudy.net
or
Jiannan Li
403 399 8791, jiannali@ucalgary.ca
or
David Ledo
403 210 9499, david.ledo@ucalgary.ca
or
Saul Greenberg
403 220 6087, saul.greenberg@ucalgary.ca

If you have any concerns about the way you've been treated as a participant, please contact the Senior Ethics Resource Officer, Research Services Office, University of Calgary at (403) 220-3782; email rburrows@ucalgary.ca.

A copy of this consent form has been given to you to keep for your records and reference. The investigator has kept a copy of the consent form

Questionnaire

*Note: The questionnaires and verbal protocol included are indicative of what we will ask and say. Minor modifications may be made to smooth out our process. Additional questions may be asked depending upon particular comments and / or actions observed as the study progresses.*

# Where are you looking at?

**Participant ID** _____

## Studying how people look at targets on a display

# Demographic Information

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Age** | | **Gender** | | **Corrected Vision?** | | **Dominant Hand** | | **Height** | |

## User assessment. For each, close, medium and far distance.

**1. Smoothness during operation was:**

Very rough ☐ ☐ ☐ ☐ ☐ Very smooth

**2. Arm Fatigue:**

None ☐ ☐ ☐ ☐ ☐ Very high

**3. Neck fatigue:**

None ☐ ☐ ☐ ☐ ☐ Very high

**4. Eye fatigue:**

None ☐ ☐ ☐ ☐ ☐ Very high

**5. General comfort:**

Very uncomfortable ☐ ☐ ☐ ☐ ☐ Very comfortable

**6. Overall, looking at the targets was:**

Very difficult ☐ ☐ ☐ ☐ ☐ Very easy

**7. How did you enjoy this distance?**

Really dislike ☐ ☐ ☐ ☐ ☐ Really Like

Please state your own comments (advantages / disadvantages) below:

*The following will be asked after all three trial rounds.*

## Please rank your preference on the different distances

far: _____

medium: _____

close: _____

## Additional general feedback (optional):