

# Shopping Using Gesture Driven Interaction

Alexander Badju & David Lundberg

---

2015

Master's Thesis

Department of Design Sciences  
Lund University



# Acknowledgements

We would like to thank our supervisors Paul Cronholm at Crunchfish AB and Joakim Eriksson at Lund University for their help, guidance and support throughout the entire process. We would also like to thank Elinor Samuelsson, UX designer at Crunchfish AB for her invaluable input and feedback regarding our design of gestures and interaction as well as Carl Tönsgård and the rest of the Crunchfish AB team for their help and support and for making us feel at home at the office.

Last but not least we would like to thank all of the test participants that took the time to help us gain invaluable knowledge for this thesis.

Lund, May 2015

Alexander Badju & David Lundberg

# Abstract

Natural user interfaces are interfaces that attempt to feel natural to the user, eliminating the need for a traditional graphical user interface. Touchless technology with gestures is a way of achieving this.

With the fast-paced evolution of online shopping, via a computer or mobile device, it has become more important than ever for physical stores to regain customers lost to online dittos. The combination of touchless technology and shopping provides a new platform where the advantages of online shopping can be combined with those of physical store shopping.

This Master's thesis aims to explore this combination, creating prototypes for a gesture-based shopping window, where the line between physical store and mobile is erased. Using gestures to manipulate, search and browse items seen on a large screen outside a store, enables users to see if their size is in stock and whether there are any deals at the moment, before entering the store.

Two prototypes were created and tested, one for a realtor display window and one for a clothing store shopping window. Natural gestures were created and discovered, which can be used to interact with the prototypes, serving as a foundation for future implementations. The usability testing was performed on two groups of people, one group with no experience in touchless technology and another group with a lot of experience in the touchless technology field. Both groups performed well in the tests, however, the inexperienced group performed better than expected and the test participants were able to learn the gestures and understand the concepts faster than expected.

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	Introduction . . . . .	4
1.2	Background . . . . .	4
1.3	Purpose and Goal . . . . .	5
1.4	Scope . . . . .	6
<b>2</b>	<b>Technical Background</b>	<b>7</b>
2.1	Natural User Interfaces . . . . .	7
2.1.1	Voice Control . . . . .	7
2.1.2	Gestures . . . . .	8
2.1.3	Advantages of Natural User Interfaces . . . . .	10
2.1.4	Limitations of Natural User Interfaces . . . . .	10
2.2	Virtual Reality . . . . .	10
2.3	Augmented Reality . . . . .	11
2.4	Different Implementations of Natural User Interfaces in Commercial Products . . . . .	12
2.4.1	Kinect . . . . .	13
2.4.2	Microsoft HoloLens . . . . .	16
2.4.3	Oculus Rift . . . . .	18
2.4.4	Leap Motion . . . . .	19
2.4.5	Google Glass . . . . .	22
<b>3</b>	<b>Human behavior and shopping habits</b>	<b>23</b>
3.1	Shopping . . . . .	23
<b>4</b>	<b>Method</b>	<b>25</b>
<b>5</b>	<b>Investigation phase</b>	<b>27</b>
5.1	Brainstorming . . . . .	27
5.2	Gestures . . . . .	28
5.3	First mockup . . . . .	39
5.4	Interviews & Elicitation . . . . .	41
<b>6</b>	<b>Prototype Design phase</b>	<b>47</b>
6.1	Realtor use case . . . . .	47
6.2	Clothing use case . . . . .	62

<b>7</b>	<b>User testing</b>	<b>72</b>
7.1	Prototype evaluations . . . . .	72
7.2	Wizard of Oz . . . . .	73
7.3	Test participants . . . . .	73
7.3.1	Purpose and motivation . . . . .	74
7.3.2	Research questions . . . . .	74
7.3.3	Test descriptions . . . . .	75
7.3.4	Realtor use case . . . . .	76
7.3.5	Clothing use case . . . . .	78
7.4	Results . . . . .	79
7.4.1	Qualitative . . . . .	79
7.4.2	Quantitative Results . . . . .	83
<b>8</b>	<b>Discussion</b>	<b>86</b>
8.1	General discussion of the results . . . . .	86
8.1.1	Test participants . . . . .	86
8.1.2	Usability testing . . . . .	86
8.1.3	Prototype evaluations from a usability perspective . . . . .	88
8.1.4	Mid-Fi prototypes . . . . .	88
8.2	Industry relevance . . . . .	88
8.3	Continued work . . . . .	89
<b>9</b>	<b>Conclusions</b>	<b>90</b>
<b>10</b>	<b>Glossary</b>	<b>97</b>
<b>A</b>	<b>Personas and Scenarios</b>	<b>99</b>
A.1	Clothing use case . . . . .	99
A.2	Realtor use case . . . . .	100
<b>B</b>	<b>Test objectives</b>	<b>102</b>

# Chapter 1

## Introduction

This chapter will provide an introduction to the Master's thesis subject, goals and scope.

### 1.1 Introduction

A Natural user interface denotes an interface that is designed to feel natural, and easy to interact with. In order to achieve this feat, different approaches can be taken, for example touch screen or touchless technologies. Touchless technology allows users to perform actions, without pressing a physical button or screen, thus allowing those actions to be performed at longer distances than physical interaction allows. The touchless interaction can be achieved using e.g. gestures or voice commands.

The combination of touchless technology and shopping allows physical stores to offer the user the advantages of digital interaction e.g. easier searching for an item, with the advantages of a physical location where items can be tried on and visualized in 3D. Utilizing gestures, the digital visualization can be large, beyond the physical reach of the users, and allowing users to manipulate objects with their bare hands. A challenge with gesture-based interfaces is feedback. As there is no tactile feedback it would rely solely on visual or auditory feedback, e.g. in the form of a voice confirming an user's action.

The workload throughout this Master's thesis has been equally divided between the two authors as it was believed that working side-by-side enabled a constant flow of feedback and useful conversations to take place yielding better results than working alone would have.

### 1.2 Background

The company for which this thesis has been produced is called Crunchfish AB. They are a software and service development company operating in the mobile phone industry. Their headquarters is located in Malmö, Sweden but they also have an office in San Fransisco, USA. One of their largest customers is the Chinese mobile phone-developer Gionee. An example of their most successful applications is GoCam, an iOS application that allows users to take photos from a distance, using gestures (captured by the phone's or tablet's built in camera). However, as a result of today's technical limitations in the gestural interaction field, Crunchfish wanted to explore the possibilities with gesture-based interaction on the basis of not having any technical

limitations. They wanted to conceptualize what gesture interaction could become in the future, and how that interaction should be performed, once the technology has caught up.

This thesis aimed to explore gesture-based technologies and how touchless interaction can be used to improve customers' shopping experiences in physical stores. Furthermore, it will investigate how a touchless shopping interface should work and look like as well as what gestures should be used to interact with it and why.

### 1.3 Purpose and Goal

The purpose of this Master's thesis was to research already implemented touchless interaction methods as well as to explore the possibilities for touchless interaction combined with hologram/3-D interfaces.

One of the goals was to apply the purpose and create an end product that consisted of two conceptual gesture-based interaction prototypes. The other goal was to perform usability tests on the prototypes and to evaluate the tests according to scientific evaluation methods.

At the start of the Master's thesis, extensive research covering previous ways and methods used to interact with interfaces and how to design said interfaces was conducted. It was crucial for the thesis that an extensive literature study was performed before building any prototypes or making any design decisions, because it was of great importance to avoid gesture-based pitfalls others might have fallen into previously. It is better to learn from other people's mistakes than making the same mistake twice.

This Master's thesis provides new research and results leading to innovative user friendly and natural ways to interact with gesture-based interfaces, based on the results of the prototype evaluations.

#### **Problem statements**

In order to analyze the goals, problem statements were created that this thesis aims to answer.

- What are present day's Natural User Interface technologies strengths and limitations?
- What technical, as well as cognitive factors need to be taken into account when designing a gesture-based interface?
- Could gesture-based interactive shopping windows be a motivation for online shoppers to buy more items in physical stores?
- Are the built prototypes good representations of the use cases they implement and the final product they aspire to be the basis for?
- What kind of gestures are natural for users and how should they be applied in products that belong to the field of human machine interaction?

## 1.4 Scope

An important goal of the Master's thesis was to produce prototypes that supported gesture interaction. Due to the vast amount of technical challenges and limitations in today's gesture interaction field, the design of the prototypes was limited to a front-end interaction design, as opposed to a technical product implementation. As the aim was to explore the gesture interaction, the visual design was reduced in favor of increased focus on developing the interaction and gestures. With this scope in mind the end products lead to the appearance of the prototypes not representing the actual look of a finished interface, but rather a finished gesture-based interaction system laid on top of a mockup design.

As a consequence, the usability testing was aimed at answering questions about the interaction, design of the gestures and overall feel of the product, as opposed to questions about the designs of the objects and visualized feedback.

The area of application for the prototypes was shopping windows. The prototypes would explore how gesture interaction can be utilized in shopping windows to connect the customers with the store, via gestures and how it affects the shopping experience.



## Chapter 2

# Technical Background

This chapter aims at providing relevant background information about the theory described in the Master’s thesis, focusing on different touchless technologies and natural user interfaces.

### 2.1 Natural User Interfaces

A natural user interface (NUI) is a user interface (UI) designed to feel as natural as possible to the user (Christensson 2012). A normal user interface, e.g. Mac OS X, requires you to interact with the computer via a touchpad or a computer mouse. The Mac OS X interface is a graphical user interface (GUI) i.e. it is presented visually on the computer screen. A normal, non-graphical, user interface for example a remote control does not have a visual representation of the interface. Instead, the controls are presented as physical buttons.

In contrast to the examples above, a NUI is not visible (Christensson 2012). That it is not visible means that a user does not go through the UI in order to access a functionality, as e.g. pressing the volume up button on a remote control or clicking on the trash can in the Mac OS X GUI. To demonstrate this, consider the touch screen on an iPhone, an example of a NUI. If a user wishes to open the mail application on the iPhone, he or she simply presses on the mail icon. On the other hand, if the user was to carry out the same operation on a computer, the user would have to move the mouse to the mail icon and click or double-click on it (depending on the GUI). In other words, the interface on the iPhone is invisible, as the user interacts directly with the application, rather than through a secondary device (i.e. the mouse).

Christensson (2012) suggests that a touch screen feels natural as the objects on the screen are manipulated directly by touching them, much like physical objects in the real world. Another example of a NUI is voice control, e.g. Siri on the iOS (Christensson 2012). Siri responds to voice commands given by the user, e.g. saying “What is the weather going to be like tomorrow?” will return the weather forecast for tomorrow (Apple 2015).

#### 2.1.1 Voice Control

Speech is a natural way of communicating for humans, thus it is not difficult to see why voice control features would appeal to users. Hearst (2011) mentions that with the increased popularity of mobile phones with touch screens, the appeal of voice control has increased as touch

screens are not ideal for typing on. As previously mentioned, there are many products that offer voice control, albeit on different levels, e.g. Siri and Kinect. With different levels of voice control, command-based versus natural voice input is implied. According to Rubin (cited in Goth 2011), the best input device for gesture-based interaction is something you do not notice. Certainly this is true for voice input as well, if it requires a user to learn certain commands it will not be as natural as normal speech.

Speech recognition has been around for several decades, although the first examples could only recognize a handful of syllables (Juang and Rabiner 2005). Furthermore, Juang and Rabiner (2005) explain that speech recognition technology became popular in the 1960's and 70's, largely owing to movies such as "2001: A Space Odyssey" by Stanley Kubrick. In Kubrick's movie, an intelligent computer called "HAL" both understood and spoke perfect English, while being able to intelligently converse with humans (Juang and Rabiner 2005). In George Lucas's famous movie "Star Wars" existed an intelligent humanoid robot named C3PO that, just like HAL, could understand and speak perfectly. Fast forward to present day, where voice control is common in a lot of products publicly available, voice control is still not the norm for input devices. Hearst (2011) points out a major disadvantage of voice control, which might be reasons as to why voice control is not more popular than it is, namely that people can hear what others are saying.

Despite voice control products such as Siri in iOS being capable of understanding natural speech, in contrast to only basic commands, there might be other reasons as to why it is not the preferred interaction mode than privacy concerns. Negroponte (cited in Malizia and Bellucci 2012) said that the connection between humans and computers is "sensory deprived and physically limited". Applied to a computer mouse, which does not come near to a human hand in terms of physical freedom, it becomes clear what Negroponte is saying. Similarly, the same saying could be applied to voice input, as e.g. commands are not natural speech, hence the user is physically limited when it comes to interacting through voice input.

When voice control becomes capable of always interpreting the user correctly, voice control might become more popular. Given the limitations of voice control, it is in conjunction with gestures that it has the possibility to shine.

### 2.1.2 Gestures

Rempel et al. (2014) conveys the importance of the "selection and design of 3D modeled hand gestures for human-computer interaction should follow principles of natural language combined with the need to optimize gesture contrast and recognition".

Looking back at history, gestural interfaces have been subject for research for more or less 40 years (Garber 2013). The point and idea with gestural interfaces is that they strive for so called "natural interaction". In today's language and definitions, NUIs are mostly used as a method and way to distinguish between the "classical" computer interfaces. The classical computer interfaces use artificial control devices, a user has to learn these operations before being able to interact with the interface (Malizia and Bellucci 2012). However, today's gestural interfaces are not completely "natural" since a developer and designer designs a gesture to perform a specific operation and then the user must learn the designed gesture when interacting with the gestural interface. As conveyed by Malizia and Bellucci (2012) and Norman (2010), a completely natural interface has the capability properties so that a user interacts with the technology in the same

way he or she interacts with real objects in the real world, “as evolution and education has taught us” they continue. In spite of this fact, new interaction systems do not take into account/are not able to take into account a user’s natural spontaneity when communicating. On the contrary, new interaction systems restrain the user’s to think and act freely and spontaneously since the systems force the users to learn and adopt a series of commands that have already been pre-defined by someone else (Malizia and Bellucci 2012).

An important factor to take into account in this aspect is that not everybody on the planet has the same cultural background, meaning that the way people use gestures and body language differs from culture to culture around the globe. This particular parameter is important to take into account when designing gestural interfaces (Malizia and Bellucci 2012).

Malizia and Bellucci (2012) continues by saying that another of the most important factors that needs to be taken into account when designing gestural interfaces is to ask the question, is this gestural interface natural in such a way that it “offers a higher degree of freedom and expression power when compared to a mouse-and keyboard interface.”? Or is the point with gestural interfaces to empower users with communication means so they can communicate with computer systems in such a way that they feel more comfortable and familiar with them?

As mentioned earlier, for an interface to be natural the user should not have to learn a new “gestural language” depending on the product, device etc. Naturally, many companies today working with-and developing gestures as a mean to communicate and execute operations, develop and design their own gestures since they want their products to be unique and they believe that their way of doing things is the correct way. However, from a user’s standpoint it can be difficult and rather tiresome having to learn a new gestural language depending on the application, device or product. Furthermore, Malizia and Bellucci (2012) stresses the fact that imposing a standard gestural language for everyone to use might be a bit counter productive as well because of the cultural differences around the world. To exemplify this case, Malizia and Bellucci (2012) brings up the language Esperanto. This language did not succeed in being adopted and spread widely as a consequence to its artificiality. In this respect technology is also prone to fail from time to time as a result of an artificial way to communicate. However, the salvation of an artificial communication method is when the gesture or way of communication becomes natural thanks to a global widespread.

Certain gestures work in spite of not being natural gestures, e.g. the pinch-zoom that is used in iOS. Malizia and Bellucci (2012) stresses that pre-defined gestures inhibit users as they are forced to conform to an existing set of gestures, rather than the system conforming to the user’s natural gestures. Metaphors are often used in interface designs, like the “Save button”, which resembles a 3.5” floppy disk, exploits users’ prior knowledge of floppy disks. Today, USB devices are the standard, yet the save button is still used in many interfaces. UI metaphors are used to increase the familiarity when using interfaces by utilizing prior knowledge about other domains, and applying them to a UI e.g. a word processor using a typewriter metaphor (Helander et al. 1997; Szabó 1995). Much like the floppy disk, typewriters are not the standard device for text input anymore, yet the metaphors work even for people who have never used one. The floppy disk icon for the save button might, if you ask someone who has never seen a physical floppy disk, just be known as the “save button” rather than a floppy disk, without the person knowing what a floppy disk is. Similarly, the pinch gesture to zoom is not a natural gesture, although it does exploit sensorimotor skills, which combined with continuous feedback through dynamic resizing of e.g. a map, aids in making it feel more natural (Gillies and Kleinsmith 2014). Malizia and

Bellucci (2012) said “Have you ever seen a person make that gesture while speaking with another person?”. However, with widespread use of gestures, e.g. swiping, pinching or scrolling, they can become natural (Malizia and Bellucci 2012). Despite this, the pinch gesture works because it has been implemented into many applications, thus, like the floppy disk, pinching the screen has become synonymous with zooming.

### 2.1.3 Advantages of Natural User Interfaces

Except for the advantages and limitations that has been brought up in the previous sections there is promise and future hopes for NUIs because the ultimate goal for NUIs is for them to be completely natural in such a way that a user is able to interact with the systems spontaneously, in the same way humans interact with each other in everyday life situations (Malizia and Bellucci 2012).

### 2.1.4 Limitations of Natural User Interfaces

Norman (2010) states that the strength of GUIs are not their visual representations of a product’s functionality. Rather, it is the ease of remembering actions that is aided by the graphical representation that makes them powerful.

As mentioned in the article by Malizia and Bellucci (2012), a user is constrained and restricted when using a NUI if he or she is not able to interact with the NUI in a spontaneous manner but rather has to learn interaction patterns designed by a third party. They claim that today’s NUIs in reality are “artificial naturality” due to the previously stated reasons.

## 2.2 Virtual Reality

Virtual Reality (VR) is a computer-generated environment, which a person can observe and/or interact with, where physical sensory experiences can be recreated (Wikipedia 2015e). Bryson (2013) states that a VR device must be capable of head-tracking for it to be called VR. Head-tracking enables objects in the VR-environment to have spatial presence i.e. the feeling of real-life existence. The concept of spatial presence theoretically enables VR technology to create environments in which a person cannot tell it apart from the real world where the person is actually located (International Society for Presence Research 2001, cited in Wirth et al. 2007).

VR as we know it today started around the 1950’s, more of a vision than reality as the technology was not developed enough at the time (Schnipper 2014). Even though the PCs became popular and a mainstay during the 90’s, the VR business did not follow the success of the PC and was subsequently overshadowed by the increasingly popular Internet (Schnipper 2014). The resolution simply could not come near reality, thus the technology faded into the background. 2012 saw the arrival of Oculus Rift, a HMD with VR technology. Suddenly, VR was popular again (Schnipper 2014).

### **Applications and products**

The most famous VR product at the moment is most likely Oculus Rift, developed by Oculus VR. Oculus VR was founded in 2012 by Palmer Luckey (Oculus VR 2015; Wikipedia 2015d). As previously mentioned, the modern VR technology started to take form around 1950. Robertson and Zelenko (2014) state that what might be considered the first VR product ever was patented in 1962 by Mort Heilig. Heilig's machine was called "Sensorama" and was equipped with a 3D-display, surround sound, a vibrating seat and a scent producer. However, the Sensorama never had the breakthrough Heilig anticipated, in contrast to the success of Oculus Rift. Oculus Rift was originally funded through a Kickstarter campaign (Oculus VR 2015). The goal of the campaign was to raise \$250,000, which it exceeded by more than \$2,000,000 within a couple of hours (Kickstarter 2012; Rubin 2014). Palmer Luckey (cited in Schnipper 2014) said "the thing stopping people from making good VR and solving these problems was not technical. Someone could have built the Rift in mid-to-late 2007 for a few thousand dollars, and they could have built it in mid-2008 for about \$500. It's just nobody was paying attention to that." Facebook announced in March 2014 that they were buying Oculus VR for two billion dollars and the deal was closed in July the same year (Dredge 2014).

## **2.3 Augmented Reality**

Augmented reality (AR) is a view of the real physical world where elements have been augmented or changed by a computerized process, presented directly or indirectly live in real time on a device's screen (Singh and Singh 2013; Wikipedia 2015a). When comparing AR to VR the main and largest difference is that virtual reality's whole experience and environmental setup is computer generated whilst in AR the real world is portrayed and visualized but with computer generated elements (Bonsor 2001). The added elements can be e.g. sounds, video, graphics or even GPS data. As a result a user's view and perspective of the real world has been enhanced. When applying AR-technology to the view of the real world, information and actions become more interactive for a user (Bonsor 2001; Wikipedia 2015a).

### **Applications and products**

AR is implemented in many different types of products and applications. For example when a football game is on television and one of the teams has received a free kick, the broadcasting network might add a visual tool depicting e.g. the distance from the ball to the goal.

IKEA has an AR-application available for iOS and Android. The app is an AR-catalogue including some selected furniture from IKEA where the user can film a room or a setting to see if a piece of furniture might be a good fit, before purchasing said furniture (Evans 2014).

On today's market there is a piece of technology that uses the front camera of said device to film a human's face. The program is then capable of adding makeup to the face that is being filmed, hence letting a user see how a certain make-up product looks without needing to apply it or purchase it (Evans 2014).

There are a number of apps and products on the market today where a user is able to try and test certain possibilities in order to make the decision making process smoother and reduce possible

waste of effort, time, materials etc.

## **Future**

There are many different application opportunities for AR-technology. As mentioned in the previous subsection, AR-technology already exists today where a user is able to try a product before purchasing it. However, there will be a larger development and growth of AR-products in this “try before you buy”-field in the future (Total Immersion 2015).

In the future of digital marketing, AR will be an important tool and factor when attempting to reach potential consumers. According to the website Total Immersion (2015) the idea is for a consumer/user to interact and engage with a piece of technology that has implemented AR. The setting could for example be a store, at home, at a restaurant etc. A few important aspects of this type of digital marketing are; augmented packaging, on-street marketing, interactive products, so called “advergaming” (advertisement done via an interactive game) and geolocalized apps, to name a few. For example, AR applied to geo-location could provide a user with information on his or her GPS device regarding a restaurant’s location, where a bar is or where a particular grocery store is located (Total Immersion 2015).

Total Immersion (2015) and the article that has been written about the future of AR portrays the belief that the educational field will benefit from AR-technology in the future. Historical events can be recreated, structural views of the galaxy can be presented for different educational purposes and so on. As reported by Total Immersion (2015) it is believed that an educator, e.g a teacher or a professor, who uses AR in his or her teaching process will provide the students with a deeper and more tangible understanding of the subject that particular educator is teaching.

Industrial, medical and military fields can also benefit greatly from AR-technology. For instance in an industry setting, a 3D model can be used to validate a design someone has produced, the point is to turn and flip the model to see different views and angles of the model for validation purposes. On the battlefield, where soldiers need intelligence of the battlefield at hand, information regarding the surrounding terrain, enemy and ally positions is essential for a successful operation. If soldiers could have a live, accurate and detailed 3D model of the battlefield it could mean the difference between life and death (Total Immersion 2015).

## **2.4 Different Implementations of Natural User Interfaces in Commercial Products**

Natural user interfaces have become increasingly popular over the course of the last 5-6 years. In 2010, Steve Ballmer (the CEO of Microsoft at the time) called NUI the next major trend in technology, a few months later Microsoft released the Kinect for Xbox (Ballmer 2010). After Kinect, other products have followed including Google Glass, Leap Motion, Oculus Rift and Microsoft HoloLens to name a few.

Implementations of NUI can be achieved through gestures, voice input, touch or any combination of the three. Different products combine these three input functions differently, and to different extents. For example, the Kinect accepts certain voice commands, Google Glass a more natural talk interaction and Siri the most natural (Apple 2015; Google 2015; Microsoft 2015f). In other

words, the term is not set in stone, there are many ways to achieve a natural user interface.

### **2.4.1 Kinect**

The Kinect, or project Natal as it was referred to during the development phase, is a device for Xbox that is able to detect and understand human motions and gestures. Microsoft is the company that develops the Kinect and the first-generation of the device was released in November 2010 for the Xbox 360 and the second generation was released during the summer of 2014 for Xbox One. Project Natal was first announced in June 2009 and the same day a promotional and prototype video of the controller-less and motion-based gaming approach was released (Crawford 2010).

The Kinect is a horizontal bar-shaped object, which stands on a small platform. The platform-base is motorized so it is capable of following a user's movements and gestures in a fluent manner (Crawford 2010).

#### **Interaction approach**

The Kinect has a NUI, which utilizes face, body and voice recognition together with motion and depth sensors to allow the user to interact without the need for a handheld device, unlike its rival the Nintendo Wii, which requires a handheld remote controller (Crawford 2010). Furthermore, Kinect has support for more than one user at a time, enabling multiplayer features and collaborations (Crawford 2010).

Rule (cited in Crawford 2010) states that the first version of Kinect detects and tracks 48 points on a person's body, from which a digital model is then created. Kinect is not just a device for voice commands and gestures, it also tracks every movement a person makes, which is utilized in e.g. interactive games on Xbox (Kuchinskas, cited in Crawford 2010).

#### **Gestural commands**

Kinect uses a series of different gestures to accomplish the human machine interaction. There are a few differences between the first-and second generation Kinect regarding the gesture possibilities. These are a few examples from the possible navigation gestures for the second generation of Kinect for Xbox One. When a user is engaged in a game or in the menu-navigation phase and wants to return to the home menu, the user has to hold out both hands to the edge of the screen and then close them and then finally proceed to move them towards each other in front of the chest (Microsoft 2015a).

To make a selection the user is required to lift his or her hand with an open palm towards the Kinect sensor. The next step is to move the hand over the specific item or tile the user wants to select. The last step in the selection process is to push the hand towards the item to be selected and then move the hand backwards toward the body and chest again (Microsoft 2015a).

For scrolling purposes the user first needs to engage the Kinect sensor by facing it with an open palm and when a hand icon presents itself on the screen the user has to close its hand over the area where he/she wants to scroll. Finally, the user has to move the closed hand with a pulling motion to the left or to the right depending on which direction he/she wants to scroll. Moreover,

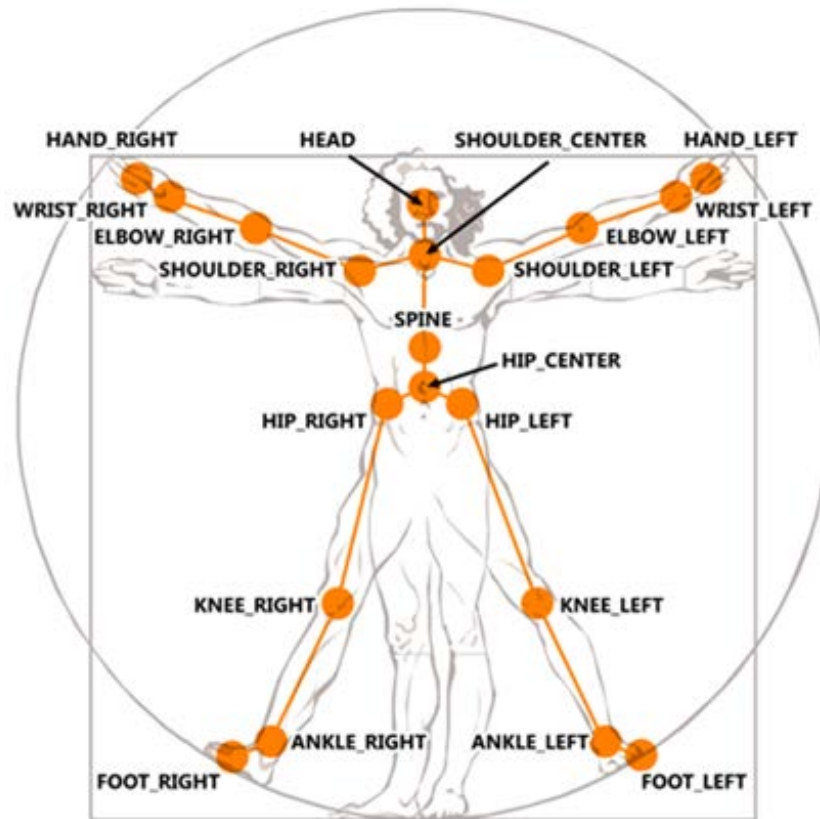


Figure 2.1: The figure illustrates all the points on the human body that the Kinect is able to recognize with its camera (Microsoft 2015c).

there are some apps that also allow vertical scrolling, meaning that a user can pull its closed hand upward or downward to scroll vertically (Microsoft 2015a).

Another important action a user might want to perform is to zoom in and out. First and foremost, the user must present an upward-facing open palm to the Kinect sensor to engage the system. Secondly, when the hand icon appears on the screen the user has to close its hand over the area he or she wants to zoom in-or out of. For the last phase the user pulls his or her closed hand towards the body to zoom in and pushes the closed hand away from the body to zoom out (Microsoft 2015a).

### Technology

As Crawford (2010) illustrates in her article, the user's body becomes the console's controller. The Kinect does not only rely on gestures and body movement for navigation-and interaction, it also supports voice control. The voice control feature can be used to interact with-and play games but it can also be used for menu-browsing purposes. Moreover, the Kinect uses the voice control feature as a mode of communication between users in the Xbox Live community, which



is live online gaming for Xbox-users. This video communication mode is called Video Chat and can support four people communicating at once (Stevens 2010).

The second generation Kinect, developed for Xbox One, is a better and more developed version of the first generation. The camera's peripheral view has improved and through the new voice-control system a user is able to turn on its Xbox One-console by simply uttering the words "Xbox on". Other commands such as "Watch TV" (to watch TV) or "snap Internet Explorer" (to open the Internet browser Internet Explorer) are a few examples of other voice commands that are available with the second generation of Kinect (O'Brien 2013).

The new Kinect is capable of reading and monitoring a player's heart beat as well as recognizing six different players, simultaneously, in the room. Furthermore, the Kinect is capable of communicating with the "regular" Xbox-controller and is able to use its movements as input for the active game, e.g. if a player lifts the controller, which in this example represents a shield, the Kinect registers the movement leading to the shield being lifted in the game (O'Brien 2013).

The Kinect is also available for Windows 7 and 8 as an interactive alternative. Apart from the interactivity, users are also able to create desktop applications developed in C++, C# and VisualBasic (Microsoft 2015b).

## **Limitations**

There is an interesting study made by Yao and Fu (2014) where they have looked into the limitations of the first-generation Kinect but also limitations of all systems that implement gestural interaction.

When the camera tries to distinguish a hand and later on understand which gesture the hand is trying to do achieve there can be many difficulties hindering the performance. Yao and Fu (2014) especially emphasizes the difficulties for the camera and technology to recognize a gesture when the hand is close to the background of its setting. When the system tries to segment the hand it can be difficult since the background features might be mistaken for the hand and vice versa. Furthermore, the shape a hand has in combination with the different motions a hand can perform makes it more difficult for the system to understand in comparison to the rest of the body.

According to Errol et.al (2007) cited in Yao and Fu (2014) there are two main challenges when it comes to developing hand gesture-based systems. The first problem is to figure out how to locate the naked hand and the second issue is how to reconstruct the hand pose from the raw data.

## **Progress**

Kuchera (2015) describes the promotional video for Kinect from 2009 as a bit misleading, in reality the functionalities of the first generation Kinect were not as good as the promotional video depicted. The Kinect had trouble tracking multiple users and the voice commands were often "hit and miss".

O'Brien (2013) continues by mentioning the camera's great improvement from the old VGA sensor to the new 1080p camera system. With the new camera system as well as new development breakthroughs, the Kinect is able to sense and recognize a larger number of reference points on

a human's body compared to the old model.

### **Marketshare**

The first generation Kinect sold 8 million units in the first 60 days from its release (Gates 2011). In an article written by Graziano (2012) where he referenced Yusof Mehdi, the CMO of Microsoft's interactive entertainment division, announced that the Kinect had sold more than 19 million units at the time and that Microsoft and Xbox held a 47% market share at the time.

### **Future**

In the previous section covering NUIs, the gesture interaction section in particular, Malizia and Bellucci (2012) and Norman (2010) stress the notion that a gesture based user interface that has been designed by a third party will not be a natural user interface according to their definition and is doomed to complicate user interaction since the user is forced to learn gesture patterns to be able to interact with the system.

As Carmack (2013) stated at the Quakecon convention in 2013, the Kinect still has some problems with the frame rate, latency and user interaction. However, he also continues by saying that the Kinect probably still will sell for a large consumer group but as mentioned earlier he believes there are some fundamental problems with the product.

To summarize, since a user has to read a manual to understand how to interact with the system and learn which gestural patterns are mapped to a certain operation, according to the definition of NUIs by Malizia and Bellucci (2012) and Norman (2010) the Kinect, in today's form, is not an NUI. Despite the technical problems Carmack (2013) stated and the lingering interaction issues with a learning curve of sorts, consumers still buy the product, which means that Microsoft apparently are doing something right with the Kinect.

## **2.4.2 Microsoft HoloLens**

HoloLens is an augmented reality headset developed by Microsoft, which is yet to be released (Yee 2015). Contrary to VR and other AR products such as Google Glass, HoloLens uses interactive holograms that are fused together with the physical world (Moorhead 2015). Microsoft themselves use the term "Mixed Reality" to distinguish HoloLens from other AR products, as HoloLens uses interactive holograms rather than just portraying non-interactive images in the physical world (Yee 2015). The actual device is bigger and clunkier than Google Glass and are not meant to be used continuously throughout the day, as Google Glass. HoloLens is not designed to be hidden and perfectly integrated, making users forget they are wearing it, instead they are focused at performing certain tasks for a few hours a day (Moorhead 2015).

Microsoft HoloLens uses Windows 10, which has support for holographic images (Microsoft 2015d). HoloLens uses normal applications, e.g. Skype, Netflix or weather forecast and integrates them into the real world, instead of immersing the user into a virtual world. For example, HoloLens is capable of projecting a screen on a wall, on which Netflix can be used, thus eliminating the need for a physical screen (Alderman 2015). The AR environment makes it possible to interact with both the HoloLens holographic objects, as well as physical objects at the same

time, in contrast to a VR environment in Oculus Rift.

Moorhead (2015) said that in a demo of HoloLens, Microsoft showcased some future applications for the HoloLens, including manipulating objects on Mars, installing a dimmer switch with the help of a person via a Skype call, playing Minecraft and creating 3D models in a program called Holo Studios.

In a promotion video for HoloLens, Microsoft show the aforementioned applications as well as a few others such as weather forecast and To-Do lists (Microsoft 2015e). Moreover, Skype calls presented in the field of view of the user, resizable “TV screens”, interactive 3D models that can be merged with a physical object, the surface of Mars portrayed on which the user can inspect objects on the surface more close up and To-do lists that can be attached to the fridge are shown in the video.

Based on the applications shown in the promotional video, one can make the assumption that Microsoft wants HoloLens to be able to do several things from entertainment, to productivity to social interactions. The difference between HoloLens, at least what is shown in the video, and existing AR e.g. Google Glass is that Google Glass presents the information on a tiny screen at the upper right corner of the user’s eyes, whereas HoloLens presents the information all around the user, mixed with the physical world (Microsoft 2015d; Swider 2015).

### **Interaction approach**

As mentioned, HoloLens uses holograms to create an augmented reality, in which a user interacts with both physical as well as holographic objects simultaneously. HoloLens combines gestures, sound and various sensors to track the user, the environment in which the user is located, and for output and input to and from the device (Microsoft 2015d).

Microsoft (2015d) states that HoloLens tracks a user’s eye movements, in order to navigate the user around a e.g. a room. This is a natural way of interacting with one’s environment, nevertheless it is a feature a lot of other VR/AR products lack. For example, Oculus Rift does feature head tracking, but not eye tracking, so a person has to turn the head to look around (LeClair 2014). A company called SensoMotoric Instruments (SMI) have developed eye-tracking functionality for the Oculus Rift, although it is not something that is included out of the box (LeClair 2014).

In addition to eye-tracking, HoloLens features spatial awareness i.e. it is aware of the surrounding environment (Gilbert 2015). These capabilities allows HoloLens to create a map of the area around a user and overlay it with e.g. the surface of Mars, as mentioned previously. Since holograms are the main feature of HoloLens, the other sensors support and improve the capabilities of the holograms. In addition to spatial awareness, HoloLens uses spatial sound for its holograms. This lets a user pinpoint the exact location of a hologram e.g. if it is behind, in front of or to the left of the user (Microsoft 2015d). This could be especially important for some of the applications, such as Skype, games or notifications.

## **Gestural commands**

Not much has been said about what type of gestures will be supported for HoloLens, except for a few visible in Microsoft’s promotional video (see further up). Kinect focused on pre-designed gestures that users have to learn, i.e. they are not natural Malizia and Bellucci (2012). Gestures that do not have an “evolutionary”, nor culturally, connection to the users are unnatural to use, hence detrimental to the success of NUIs such as Kinect (Malizia and Bellucci 2012).

HoloLens uses the “pinch zoom” gesture, for resizing holograms and finger gesture that resembles a mouse-click, with the same functionality as clicking the mouse on a computer (Microsoft 2015d). Otherwise, the specifics of which gestures will be available in HoloLens have not been disclosed at this point.

## **Future**

According to Moorhead (2015), Microsoft have, compared to other products on the market, chosen to take a vertical approach to AR with HoloLens. A vertical approach focuses on perfecting a few things, as opposed to trying to do everything, which often results in less than perfect results (Moorhead 2015).

As HoloLens has not been released yet, it is difficult to say how well it will sell, how useful it will be and how well the final product will live up to the expectations of the promotion video. Norman (2010) has criticized gesture based interfaces, saying that the gestures are often unnatural, hence requires the users to learn them and subsequently remember them.

If gestural interfaces are instead designed to utilize sensorimotor skills, rather than manipulation skills, the problem with unnatural gestures could be limited. The pinch zoom gesture in iOS is, as mentioned before, good at providing continuous feedback to the user as it dynamically zooms in and out, in contrast to gestures that perform a certain task after the gesture is finished (Gillies and Kleinsmith 2014).

### **2.4.3 Oculus Rift**

There is a term called “Binocular depth cues”, which is a term defining different types of clues the human eye uses when observing the observable world. These clues could be the depth or distance in a picture for example. Since our eyes are located approximately six centimeters from each other, an object that is closer than 30 meters from our eyes will appear differently for the right and left eye respectively. e.g. a person can hold an object in front of his or her eyes and then proceed in closing one eye at a time, then do the same action in quick succession, it seems as if the object shifts position, this is because the right and left eye perceive the object from different angles. Binocular depth cues defines stereopsis whilst another term called “Monocular depth cues” is the definition of a more accurate distance estimation (Pescarin et al. 2013).

The Oculus Rift device has a screen for each eye and the picture projected in each screen is projected in such a way that the phenomenon described earlier with the shifting object is eliminated. When looking through the Oculus Rift, the picture is calibrated in such a way that it looks real and accurate according to the settings and surroundings in the virtual reality (Pescarin et al.

2013).

### **Limitations**

Palmer Luckey (cited in Amato 2013) stated in 2013 at Boston Virtual Reality Meetup that the issues with Oculus Rift are not hardware or software related. Developers will keep designing and improving both hardware as well as software over the upcoming years so even if the technology is not perfect today it will keep improving with time. However, Luckey continued by saying that the largest hurdle are the users themselves, in the sense that the human's physiological and neurological factors are the real limits to Oculus Rift (Amato 2013).

The whole idea and purpose of the Oculus Rift is to immerse the user in a virtual world and for the user to feel as if he or she is transported into a new dimension that is realistic. Luckey (cited in Amato 2013) explains that this is the strength with VR and the Oculus Rift but at the same time it is the weakness of the technology. Luckey proceeds with an example, when the user runs towards a flight of stairs in the virtual world, the brain is prepared for the bouncing motion that comes with running up for a flight of stairs. However, in VR a flight of stairs is used as a ramp which the user glides upwards instead of bouncing, this leads to the brain expecting a movement that does not happen and thus the user can become motion sick, which is a problem with VR (Amato 2013).

### **Future**

As Luckey stated (cited in Amato 2013), hardware and software wise the Oculus Rift technology will continue to improve. The technological future of Oculus Rift is therefore not hanging in the balance nor is it surprising when advancements in that field is made. The future of Oculus Rift lays in the application possibilities. Oculus Rift will take gaming to the next level but the application field is larger than computer games. In the future surgeons or pilots for example can simulate realistic work environments. The movie entertainment business can change and the way people consume any type of entertainment, e.g. sports events, concerts etc. might change as well. Users could watch a basketball game with courtside seats from all over the world for example, there are many different application possibilities for Oculus Rift and VR in the future (Yarm 2014).

#### **2.4.4 Leap Motion**

Leap Motion is a piece of hardware the size of a candy bar that works as a touchless controller for HCI. The Leap Motion-device has sensors that are able to sense and distinguish a user's hand and finger motions and use them as an input for computer interaction. The technology is analogous to a regular mouse but does not require hand contact or physical touch (Wikipedia 2015c).

The technology for Leap Motion began development in 2008, however, the company Leap Motion Inc. was founded in 2010 by David Holz and Michael Buckwald. In 2012 the company released their first product called "The Leap". A few months after the release of The Leap they launched a software developer program. About 12,000 developers received a unit and the point was to have developers play around with and develop applications for The Leap. When developers have completed their apps for The Leap they are able to upload their creations to Leap Motions own app store called Airspace. There users can download and buy applications for their own Leap Motion-devices (Crunchbase 2014; Wikipedia 2015c).

## Technology and Interaction Approach

The Leap Motion is a USB-device with a height of 13 mm, a width of 13 mm, a depth of 76 mm and a weight of 45 grams. The Leap motion should be placed on a physical desktop, facing upward. To observe the settings around it and the hand movements being made above the surface, it has two monochromatic IR cameras and three infrared LEDs (Wikipedia 2015c).

The device is able to observe an area, which is more or less shaped in a hemispherical way. The distance it can observe objects to is more or less 1 meter. The Leap is able to detect a user's hands from a 150 degrees field of view with a Z-axis view as well for the 3D-perception of the hands. The Leap Motion controller tracks the hands all 10 fingers with the precision of 1/100 of a millimeter and it has the capability to track a user's movements with a frame rate of 200 frames per second (Leap Motion 2014).



Figure 2.2: The figure illustrates how a user's hands are realized on a screen when using the Leap Motion-technology (VJs Mag 2015).

An important aspect to keep in mind is that Leap Motion Inc. say that the Leap Motion-controller is not a substitute to the keyboard or mouse, it is simply a tool to compliment existing computer interaction technology (Leap Motion 2014).

The Leap Motion has generated enough traction and success that it will be directly incorporated in a few products Hewlett Packard produce. Some of the products are HP's new computers and new keyboards (Terdiman 2013; Hewlett-Packard 2015).

## Applications

Leap Motion (2014) describe many application possibilities for their Leap Motion controller. They say it will "do things you never thought was possible". Some examples they have are; a

user can browse the web, flip through photos, read articles and play music by “simply lifting a finger”. They continue to describe the modeling possibilities, a user can mold, bend, stretch and build 3D objects. With a fingertip it is possible to sketch, paint, draw and design whatever the user wants to.

In the game field a player can use its fingertip to slice fruit in a similar way to the popular game Fruit Ninja, the difference is that Leap Motion makes the game touch less. In other games a user’s hands can steer an airplane, a car or other vehicles (Leap Motion 2014).

For the music enthusiasts there are many musical possibilities, with the hands and fingers a user can play air guitar, air harp and “air everything” as Leap Motion (2014) describe it.

A company called Erghis, in collaboration with Lund University, has invented a spherical interaction method using the Leap Motion controller as hardware for user input. The product’s purpose is to simplify touchless computer interaction. A user will have an invisible sphere in his or her hands and use natural gestures such as twist, pull or tap to interact and give input to the sphere, which in its turn converts the commands in such a way that the computer understands what operation the user wants to perform. Each gesture and action the user wants to perform is mapped to a character or command. The feedback from the sphere interaction is visual and is shown on the computer screen in front of the user or through smart glasses worn by the user. The software for the spherical interaction is dependent on a third party’s hardware, in this case they have made the spherical application for the Leap Motion controller (Blénessy 2015; Erghis Technologies 2015).

### **Limitations**

In theory and in the Leap Motion’s promotional videos, the controller works effortlessly and looks more or less flawless. However, when Metz (2013) tried the Leap Motion she said “generally, I can’t get the controls to work as deliberately as I wanted.”. She had trouble using the music and paint apps, since she felt that her hand gestures and finger motions were not recorded and executed in an exact way, which she felt was needed for performing the actions she wanted to perform correctly. She also said that she got tired in her arm and shoulder after using the controller a while and that she the next day also felt a bit sore from the previous day’s interaction.

Hutchinson (2013) also describes his Leap Motion test and interaction in a similar way to Metz (2013), frustrating a part of the time and pleasant other parts of the time. However, even though Limer (2013) also felt frustrated in the beginning he felt that once he understood the technology and had used it for a while that it actually worked somewhat decently. He said that “Leap Motion-at its best the Leap Motion is a translator and at its worst a barrier-between the user and what is happening on the screen”.

### **Future**

According to Leap Motion themselves and their blog the future entails more game applications and especially for games in the VR-field (Colgan 2015).

According to tech journalists that have tested the Leap Motion, they believe that Leap Motion Inc. have to work-on and develop the technology further before the usage and user experience

will be as fluent and flawless as the promotional videos depict (Hutchinson 2013; Limer 2013; Metz 2013).

#### 2.4.5 Google Glass

Google Glass was thought to be Google’s revolutionizing wearable device. Google Glass is an Optical Head Mounted Display (OHMD) and is meant to be placed on the head and face like a regular pair of glasses. The information Google Glass wants the user to take part of is presented on a see-through small lens/prisma on the top right corner of the wearable device. The functionalities incorporated in the Google Glass are for example, to be able to read text messages as well as write and send messages using voice control, video chat, maps and turn-by-turn directions (Bilton 2015; Cass 2015; Goldman 2012; Wikipedia 2015b).

The Google Glass-project was developed in secrecy in an undercover lab called Google X. Google recruited top engineers, scientists and researchers from different companies and universities to develop and work on the project. In the development phase there were two factions amongst the developers, one half wanted the Google Glass to be worn throughout the day like a “fashionable device”. They considered the product to be a replacement for cell phones, computers and other similar devices for certain previously mentioned tasks. This focus on a wide set of applications is called a horizontal approach (Moorhead 2014). The other half wanted users to wear the glasses for specific tasks and functions, i.e. a vertical approach. Ultimately, the faction that wanted the Google Glass to be used with a horizontal approach won (Bilton 2015).

As a result of the marketing decision path chosen, Google marketed the glasses as a product that was supposed to be used continuously throughout the day. In hindsight this might have been a mistake since the Google Glass did not become the success Google hoped it would be. The whole project and the device became very hyped greatly thanks to the fact that people got the impression of the fact that they did not need to take them off. When the Google Glass-marketing strategy backfired “Google not only fanned the flames, but doused them with jet fuel” as Bilton (2015) described it, the backlash was even harder due to the fact that, as described in the section about HoloLens, when Microsoft launched the HoloLens they were clear about its purpose and usage, it was not supposed to be worn throughout the day, instead it should only be used for a limited amount of time when the user wants to perform certain actions and have access to certain information (Bilton 2015; Moorhead 2015).



## Chapter 3

# Human behavior and shopping habits

This chapter aims to give a brief introduction to shopping development and human behavior in order to help understand what drives sales. First, a brief introduction to shopping and shopping habits will be given.

### 3.1 Shopping

Online shopping has become increasingly popular over the course of the last decade. In the UK alone, e-commerce increased from 1.8 billion GBP year 2000 to 27.1 billion GBP year 2011 (Statista 2015b). The United States showed similar numbers, from 72 billion USD year 2002 to 322 billion USD year 2013 (Statista 2015a). Looking at this expansion, it is clear that online shopping is increasing. Analyzing how people shop is vital to understanding how to drive sales.

In the article Shopping Behavior and Satisfaction Outcomes, Jack and Powers (2013) have examined three primary shopping behavior types: price-conscious, recreational-conscious and impulsive-careless shopping behavior. Jack and Powers (2013) have examined and evaluated a few questions in their research, e.g.: Have the customers expectations been met? Has the expectations regarding the shopping experience in a particular store satisfied the customer? Does store satisfaction result in positive word-of-mouth communication about the store? Do these relationships differ between males and females?

According to Sproles and Kendall (1986, cited in Jack and Powers 2013) the price-conscious consumer represents a person who uses his or her decision making capabilities in such a way that they seek and sought the best value for the product they want to purchase. These type of customers and consumers shop carefully and systematically. This consumer group is very likely to compare what they want to buy with other similar products before making the purchase. They seek sale prices when they shop and low prices in general, like the name of this consumer group suggests they are aware of prices of products and want to get most value for their money. The shopping experience is confined to a purely functional purpose with the goal to achieve maximal economical value, opposed to enjoy the whole shopping experience and indulge. These two different approaches differ in a substantial way. The economic and utilitarian/functional-approach is applicable when consumers only shop and buy items out of necessity in order to gain the result

they want, best monetary value for the product they want to purchase, almost completely lacking the enjoyment of the shopping activity itself. Hedonic and recreational shopping activity is based on the appreciation of the shopping experience itself (Arnold and Reynolds 2003 ; Childers, Carr, Peck and Carson 2001, cited in Jack and Powers 2013).

Kaltcheva and Weitz (2006, cited in Jack and Powers 2013) proclaim that the recreational-conscious shopper's behavior pattern and satisfactory feelings come from the shopping activity itself. This type of shopping differs a bit from impulsive-careless shopping even though there is a part of this shopping behavior that consists of hedonic behavior that correlate to the satisfaction recreational-conscious shoppers feel when purchasing items. These feelings could possibly motivate consumers to shop more (Kukar-Kinney, Ridgway, and Monroe 2009, cited in Jack and Powers 2013). The recreational-conscious shopping behavior also includes efficient shopping and is considered task-oriented, meaning that the consumer focuses on efficiently completing the given shopping task without spending much energy on the activity. The task-oriented consumers feel satisfied when the outcome has been achieved, which is when the product has been acquired, not the shopping activity itself (Kaltcheva and Weitz 2006, cited in Jack and Powers 2013).

The last type of shopping behavior according to Jack and Powers (2013) is the impulsive-careless behavior. This type of consumer is characterized by his or her immediate shopping needs and purchases items without having pre-shopping intentions or planning ahead, he or she also ignores evaluating product alternatives (Bayley and Nancarrow 1998; Beatty and Ferrell 1998, cited in Jack and Powers 2013). However, after the shopping/purchase has been completed the consumer might reflect on the appropriateness of the purchase (O'Guinn and Faber 1989, cited in Jack and Powers 2013). The impulsive-careless behavior is related to a loss of self-control. The self-control loss is founded on to which extent a consumer is able to resist the temptation to purchase an item (Baumeister 2002, cited in Jack and Powers 2013). According to Mattila and Wirtz (2008, cited in Jack and Powers 2013) the level of stimulation a consumer feels in a store environment can affect the consumer's likelihood to do impulsive purchases. If a store has a high level of stimulation, the likelier a consumer is purchase unplanned products. Mattila and Wirtz (2008, cited in Jack and Powers 2013) describe how familiarity with a store or boutique also affects the unplanned purchases in such a way that shoppers have a tendency to shop more when they are in a familiar environment.

As E. K. Sproles and Sproles (1990, cited in Jack and Powers 2013) convey that each of the three shopping behaviors have their own learning process. The consumers that are included in the price-conscious behavior group, usually enforce a learning process that includes a methodical and precise planned approach using observation research capabilities to decide if a product is good for them or not.

# Chapter 4

## Method

This chapter aims at describing the methodology and work process of the practical part of the Master's thesis, i.e. the way the final products were created from idea to testing.

The design process of this thesis was iterative and incremental, the steps were not static but feedback from one part would be evaluated and most likely lead to a reiteration of an earlier step. The design process mainly consisted of three phases:

- Investigation phase
- Prototype design phase
- User testing

Each of the three phases contained several, iterative, steps. In addition, the phases themselves were iterative as the results from steps in the evaluation phase caused reiterations of steps in the creation phase (see image 4.1).

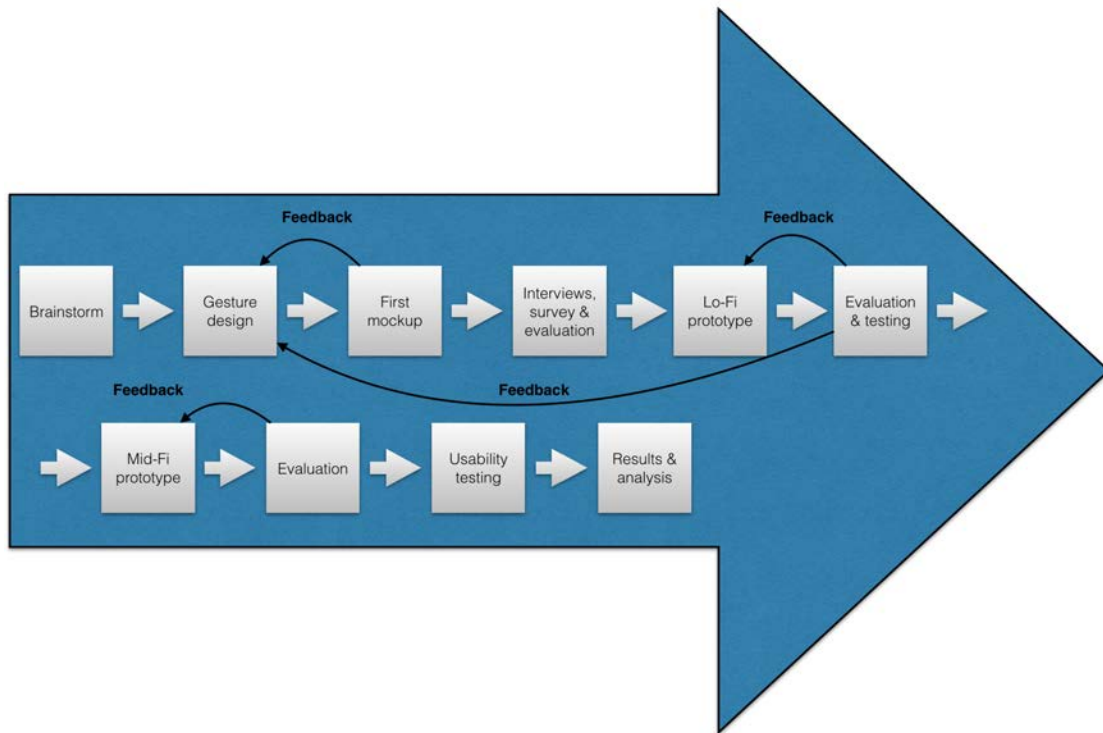


Figure 4.1: The figure illustrates the general steps of the design process.

During all three phases, brainstorming and discussions have been conducted as soon as an issue or question has come up. The discussions have been held between the thesis writers as well as colleagues at Crunchfish. For example, if there was a problem with a design feature, the UX designer at Crunchfish was consulted for ideas and input. When those ideas were realized in a prototype a quick evaluation was made with said UX designer so she could give her verdict on the end result. When both parties were happy with the results, that design feature was chosen and implemented. This quick and iterative approach was a key element that was kept throughout the whole work process.

# Chapter 5

## Investigation phase

This chapter describes the processes and results of the investigation phase.

The goal with the investigation phase was to gather the information and experiences from the literature studies done for the background studies, and together with new information and ideas gained while at the company explore potential applications, use cases and products. Furthermore, it consisted of creating generic mockups that served as a learning tool for the applications that would later be used. A big part of the investigation phase entailed interviewing stakeholders, and the general population about what they would expect, want and need from a gestural interaction product.

### 5.1 Brainstorming

The process began with a brainstorming phase in order to gather previous experiences and information from the literature studies at the beginning of the Master's Thesis and combine them with new ideas and product applications.

The team that did the brainstorming sat in a sound isolated room with a large table where quick sketches were performed and discussed. The room also had a large whiteboard where ideas and thoughts were written down along the process and some of the sketches were put on the whiteboard if they matched any ideas or thoughts. When brainstorming it is important not to criticize other participants ideas as it can harm the creative process, this aspect was taken carefully under consideration and avoided at all costs (Cook 2015).

The focus of the brainstorming phase was to discover potential sectors for the shopping application, which had been decided earlier in the thesis work.

Applications for existing online shopping actors were analyzed, Internet applications as well as iOS applications, in order to learn what could be improved, changed and removed from existing applications when designing the thesis product.

The goal for the process was to discover different use cases that would later form the prototypes. Assessing the use cases relevance was of uttermost importance, hence a decision to go out and ask people on the street about their thoughts on the two use cases, their usefulness, potential problems and potential upsides was made.

## Results

Two use cases were chosen for the thesis, a shopping window for a clothing store and a shopping window for a realtor firm. They are referred to as the Clothing use case and the Realtor use case respectively. The shopping window design was decided much on the part of them being large potential screens on which images could be displayed, but also because they were believed to have the potential of grabbing the attention of people walking by via the gestural interaction approach. The idea was to have a built-in camera in the shopping window that would recognize the users, their movements and gestures.

## 5.2 Gestures

When designing gestures for the thesis, the product ideas that sprung from the brainstorming session were the basis on which the gestures were built upon. A heuristic approach has been used as well when designing the gestures. The heuristic method however, was based on the extensive theoretical background studies. The theoretical background has provided a lot of information regarding different gesture applications, different application areas, which served as a guide when designing the gestures. The goal was to make the gestures as natural as possible, anchoring them to the physical world as well as possible, which was thought to shortening the learning curve compared to entirely made up gestures. The first gestures were included in the first mockup (see the next section), evaluated and the gestures were re-iterated until the final collection of gestures had been decided.

### Results

This project has lead to the development of 10 gestures, not all of which are presented and implemented in the two use cases. Two special, gestures, wave and grab, were created as well. Wave is used to start an interaction, grab is a basic gesture that forms the start of many other gestures but does nothing on its own. The use cases were designed to present and illustrate the most important gestures for two different scenarios to show how they can be used. In addition to the gestures presented in the scenarios, the other gestures are explained and illustrated below to give an understanding about where and how they are supposed to be used.

Each gesture consists of three states (except for the wave gesture), an uninitiated state, an initiated state as well as a finished state. Objects that can be manipulated through different gestures will hence appear differently in the uninitiated state than in the initiated state although the appearance varies depending on which gesture is performed on the object.

#### Wave

The wave gesture is a normal wave gesture, as used to greet people. The gesture was chosen because it symbolizes a greeting, it is a gesture that almost everyone ought to be familiar with and it is a non-manipulating gesture (i.e. it does not impact an object).

#### *Applications*

Used to start an interaction.

#### *Help*

Raise one hand, and wave like greeting a person.

## Grab

The grab gesture is a grabbing motion, that symbolizes the metaphor of grabbing an item e.g. a book, a bottle or a ball. When a grab is performed, visual representation is received in the form of the object shaking and detaching slightly from the background letting the user know that the object has been selected.

### *Applications*

Grab is used to start several other gestures, and to select an object to start manipulating

### *Help*

Raise one hand and hold it in front of the object to be manipulated and close the hand.

## Push to select

The push gesture mimics the push of a physical button. Push gestures are present in the physical world as well as on touch screen applications (e.g. an iPhone app). The gesture draws on the users' previous experiences with physical buttons and touch screen buttons to shorten the learning curve for the push gesture and make it feel natural to use.

### *Applications*

The push gesture is used for selections (e.g. choosing between two options), it is not used for selecting objects. The push gesture is used consequently throughout the UIs for both the Realtor and the Clothing use cases, even though the selections are not the same in the two use cases.

### *Help*

The gesture is carried out by pushing either with an open palm or the index finger towards the pushable object on the screen. The object that is being pushed will illustrate the action by being pressed in (see image 5.1 and 5.2).



Figure 5.1: The figure illustrates a pushable object in its uninitiated state.



Figure 5.2: The figure illustrates a pushable object in its finished state.

### **Grab and pull out**

The gesture is comprised of two physical gestures, namely the grab gesture and a pull out gesture. The pulling out gesture is then subsequently performed while still grabbing the illusioned “item”, by pulling the hand towards the body, causing the elbow to bend, much like picking up a phone from a table to interact with it.

#### *Applications*

Grab and pull out is used for selecting objects to view closer (see the example with picking up a phone above). In the Clothing use case, this is used to select an item of clothing from many items of clothing to view information about its price, sizes and colors etc. It is important to notice that this gesture does not simply enlarge an object, rather, it selects an object and presents information differently depending on what type of object it is (compare grabbing and pulling out a shirt with a house).

#### *Help*

The gesture is performed by first performing the grab gesture and then while still grabbing, pulling the arm towards the body. This is carried out by closing a fist with either hand and then pulling the closed fist towards the body in a straight line.

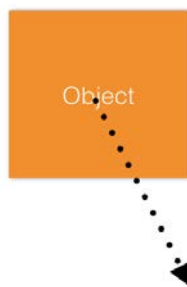


Figure 5.3: The figure illustrates an object that can be grabbed and pulled out in its uninitiated state.





Figure 5.4: The figure illustrates the same object as figure 5.3 in its finished state.

### **Grab and pull**

This gesture is based on the physical world metaphor of grabbing an item from a shelf e.g. at the supermarket, and putting it in a shopping cart. It relies on users being familiar physical shopping, as opposed to online shopping.

#### *Applications*

Similar to the grab and pull out gesture, this gesture is used for selecting items. However, instead of pulling out an object to view it, the object is dragged to e.g. a shopping cart (which in the Clothing use case is represented by the user's phone) in order to save it for some other purpose. What saving an object entails differs depending on the context, for example in the Realtor use case, saving an object both saves the object to be viewed later as well as informs the realtor that the user is interested in the property, whereas in the Clothing use case the same action adds the item of clothing to the user's shopping cart.

#### *Help*

The gesture is initiated by the grab gesture. The second step is dragging the object to the destination and letting go by opening the closed fist.

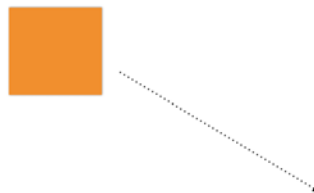


Figure 5.5: The figure illustrates a grab and pullable object in its uninitiated state.

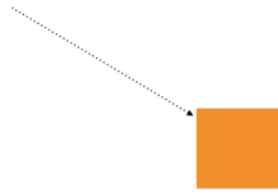


Figure 5.6: The figure illustrates the object in 5.5 in its finished state.

### **Grab and push**

The grab and push gesture is the opposite of the grab and pull out gesture. It is a metaphor for picking up an object and putting it back where it was first taken from. For example, removing a box of cereals from a shopping cart and putting it back on the shelf.

#### *Applications*

Congruent with the metaphor, the gesture is used to deselect objects that have been grabbed and pulled out. The gesture requires that an object has already been selected before it can be deselected.

#### *Help*

The gesture is performed by doing the grab gesture, before pushing the hand away from the body towards the screen and relinquishing the closed fist. The gesture requires that the arm is bent a little, otherwise the arm cannot be physically pushed away from the body.

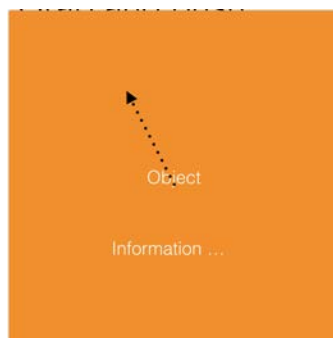


Figure 5.7: A previously selected item in the uninitiated state for the grab and push gesture.

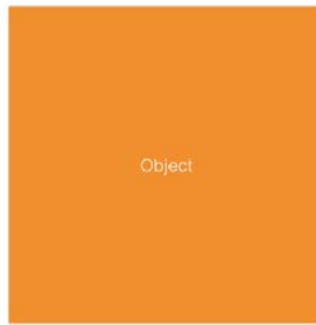


Figure 5.8: The object in 5.7 in the finished state.

### **Grab and replace**

Like the name implies, the gesture replaces an already selected object with another object. The design is meant to be a metaphor for placing an object such as a newspaper on top of another newspaper, thus rendering the first card blocked from the field of view.

#### *Applications*

Grab and replace is used to replace a selected object, which has been selected by performing the grab and pull out gesture, with another object. For example, if a shirt has been selected, another shirt from the suggestion bar can be put on top of that shirt, thus presenting information about the new shirt instead.

#### *Help*

The gesture is performed by doing a grab and pull gesture, where one object is dropped on top of another selected object. The gesture is only available if an object has already been selected.

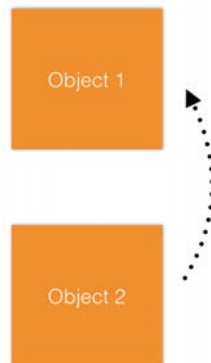


Figure 5.9: The figure illustrates a selected object, Object 1, and another object to replace Object 1, Object 2, In the gesture's uninitiated state.

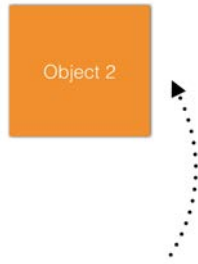


Figure 5.10: The figure illustrates the grab and replace gesture in the finished state, where Object 2 has replaced Object 1.

### Swipe to browse

The swipe gesture does not originate in a metaphor from the physical world, rather it builds on the swipe touch screen gesture that is present in e.g. iOS and Android. The swipe gesture relies on users' experiences with such touch screen devices in order to be natural even though it is not a natural gesture in the physical world.

#### *Applications*

Swipe is used to browse through objects and images e.g. changing the inspirational images in the Clothing use case. Swipe browses/changes through different objects depending on the context, however it is always used to browse or change objects no matter the scenario.

#### *Help*

The Swipe to browse gesture can be performed in two directions, left and right. To perform a swipe, the hand must be open (i.e. not a clenched fist) and moved from left to right, or right to left in a fluid motion. Swiping from right to left will browse to the next image/object to the right, similar to browsing photos on an iPhone.



Figure 5.11: Swipe to the right in the uninitiated state. The green arrow shows the direction in which the swipe is performed, the orange object symbolizes the hand's movement.



Figure 5.12: Swipe to the right in the finished state.



Figure 5.13: Swipe to the left in the uninitiated state. The green arrow shows the direction in which the swipe is performed, the orange object symbolizes the hand's movement.



Figure 5.14: Swipe to the left in the finished state.

### **Hand zoom**

The hand zoom is based on perspectives. When the hand is held close to the face, an object behind the hand appears small in comparison to the hand, when the hand is held close to the object, the object appears larger. That difference is utilized in the hand zoom gesture by enlarging items that are close to the hand i.e. the hand is held close to the object, and vice versa when the hand is held close to the body. The metaphor of having a pair of eyes in a hand can be used, if the hand moves towards an object, the object appears closer and moving the hand backwards will make the object appear further away.

#### *Applications*

The hand zoom gesture is used for zooming in and out on objects, images and text where the feature has been activated. Pushing the hand towards the object zooms in on the object, while pulling the hand towards the body zooms out.

#### *Help*

The gesture is performed by holding the hand up with the palm open flat against the screen, and then pulling or pushing the hand towards the body and screen respectively in order to zoom.



Figure 5.15: The figure illustrates an object where the hand zoom is in the uninitiated state i.e. not zoomed in nor out.



Figure 5.16: The figure shows the same object as in 5.15, where the hand has been pushed towards the screen, thus zooming in on the object.

### **Scroll**

Similar to the swipe to browse gesture, the scroll gesture is based upon scrolling on a touch screen, with no real physical context.

#### *Applications*

Scrolling can be performed e.g. when there is a lot of information that does not fit on the screen at once. Scrolling is always done up and down, never to the sides where it might have been confused with the swipe to browse gesture.

#### *Help*

Scrolling is carried out in the same way as swiping with one difference, scrolling is dynamic and can be precisely controlled by moving the hand up and down continuously as opposed to the swipe gesture, which is a single movement from one side to another.

The scrolling behaves like it does on a touch screen or a computer, which most readers will be aware of, thus no illustration is given below.

### **Grab and Flip**

The gesture is based on the metaphor of a book, which you turn over in order to read on the back. The information on the back is usually a summary of the book, i.e. more information about the book. This is the basis for the flip gesture, flipping the object over to gain more information.

#### *Applications*

The flip gesture is used on selectable objects in order to read more about them. What that “more” information is depends on the object at hand. Only objects that are selectable (i.e. objects on which the grab and pull out gesture can be used) are flippable due to the need to have a selected object to flip over.

#### *Help*

When an item has been selected by the grab and pull out gesture, the object can be flipped by performing the grab gesture and then turning the object over to the right or left depending on which hand is used. The ergonomical term for this joint motion in the wrist is called supination and pronation (Pheasant and Haslegrave 2006). The flip works without first grabbing the object,

just performing the flip motion. The two alternatives are rooted in the physical metaphor, where an object such as a book can be flipped over without grabbing it as well depending on how it is positioned. If attempted, it is evident that the only comfortable way to flip something over using the right hand is to the right, the other way around for the left hand.

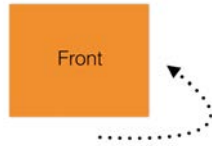


Figure 5.17: The flip gesture in the uninitiated state, showing the front of the object.



Figure 5.18: The flip gesture in the initiated state, the object has started turning over to the right.

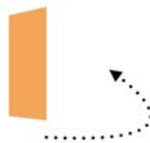


Figure 5.19: The flip gesture in the initiated state, the object is almost turned over.



Figure 5.20: The flip gesture in the finished state, showing the back of the object.

### **Card swipe**

This gesture mimics the action of swiping a credit card in a card reader e.g. when paying for groceries at the supermarket.

#### *Applications*

The gesture is intended for initializing payments, which can then be finished for example via a mobile phone.

*Help*

Performing the gesture is similar to swiping a credit card. The thumb and index finger should be pinched together, like holding a credit card. Then the hand should move down and up once in one rapid movement.



Figure 5.21: The figure illustrates the card swipe gesture in the uninitiated state.



Figure 5.22: The figure illustrates the card swipe gesture in the initiated state, where the hand is moved down.





Figure 5.23: The card swipe gesture in the finished state, where the hand has been brought back up to the starting position.

### 5.3 First mockup

The previous gesture and products analysis culminated in a first digital mockup, made in Keynote. The mockup aimed at highlighting potential problems early on and to learn how to make mockups in Keynote, which would be needed for the real prototype. The mockup was tested, filmed and evaluated and changed several times, the results leading to the design of the gestures.

#### Results

The first mockup described a shopping scenario, where there were a couple of items on deal at the moment, which could be added to a cart. The cart could then be viewed and altered, before proceeding to pay and using a gesture to perform the actual payment. Certain design aspects from the mockup were kept, and some were immediately discarded as they seemed unnatural when visualized on screen.

The main screen of the mockup (see 5.24), contained a search bar, four deals on different items and a cart to add the items to.

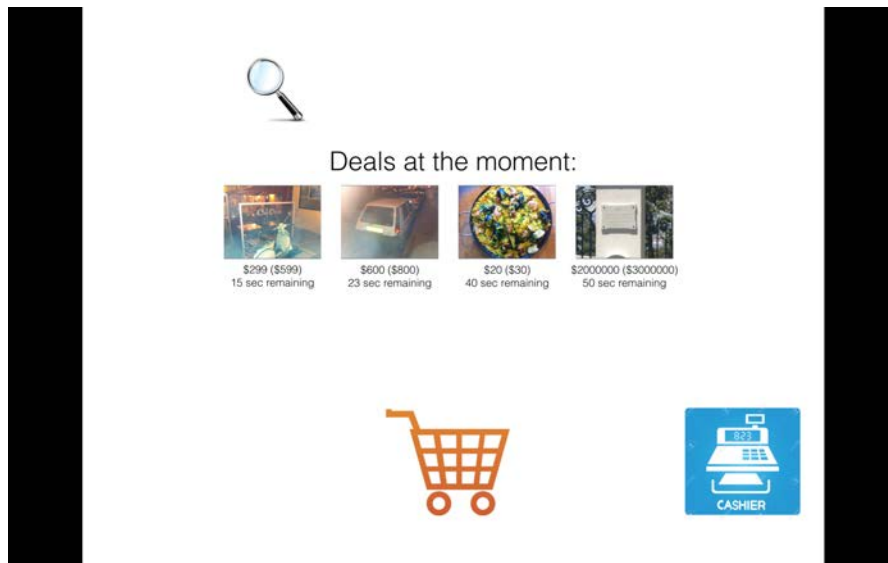


Figure 5.24: The figure illustrates the first mockup main screen.

The items could be added to the cart by grabbing the item, then pulling it down to the cart icon at the bottom of the screen. In order to view the cart, the user needed to push on the cart and subsequently push it towards the cashier icon. Upon completing the action, the user would be taken to a new screen, where all the items added to the cart could be viewed (see 5.25).

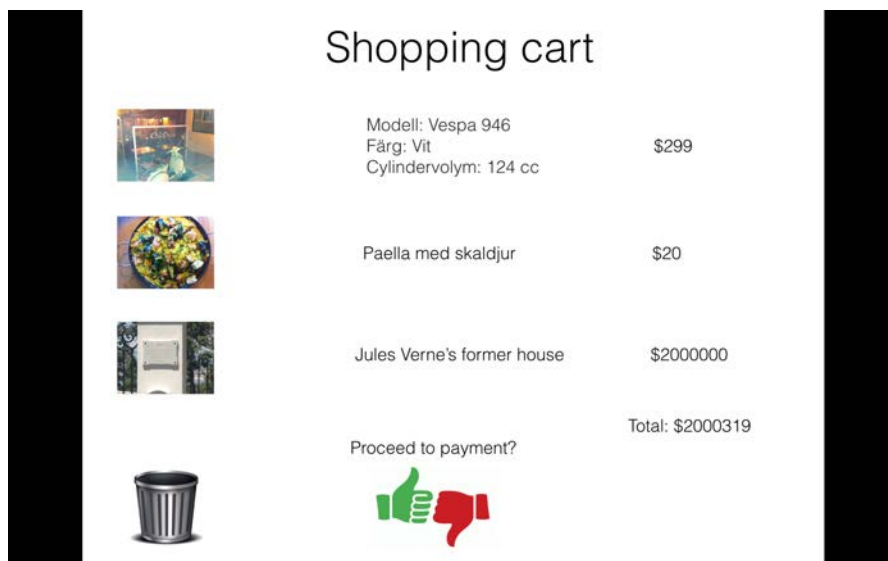


Figure 5.25: The figure illustrates the shopping cart view of the mockup.

On the main screen, only the price of the items was shown, in the shopping-cart view a description of the item was also shown. Grabbing and pulling an item down to the trash can icon would remove it from the cart, performing a “thumbs up” gesture would proceed to the payment view.

The mockup was projected on a large screen, using a projector, and different actions were tested, e.g. adding items to the cart, removing them and proceeding to payment. It was learned that creating this type of mockup and projecting it onto a large screen worked well to test gestures. Moreover, the testing gave an early insight to which gestures could work, and which would not. The grab and pull was thought to be natural and easy to learn, whereas the “thumbs up” to proceed to payment felt unnatural.

## 5.4 Interviews & Elicitation

As previously mentioned, it was deemed crucial to the success of the final prototypes that the use cases were sound and relevant to the general public (i.e. the potential users). It was decided that both a quantitative and a qualitative collection of data would be the most complete and provide the best outcome, thus a survey and personal interviews would be performed on regular people and potential future users. The survey aimed at testing people’s views on online shopping compared to physical shopping. The outcome was believed to help understanding how physical stores could attract customers that have abandoned physical stores in favor of online shopping. The information would be taken into account when designing the use case interfaces, e.g. what information they should display and what type of functionalities they need to have in order to be attractive to customers.

The survey consisted of seven to nine questions, depending on what answers were given. People who claimed that they never shop online and/or in physical stores were given extra questions about why they did not shop at respective places.

The questions regarding shopping frequency, i.e. how often something is shop9 online or in physical stores were for non-food items as that was thought to shift the numbers in favor of physical shopping due to the thesis writers’ experiences with grocery shopping online and at supermarkets in their surrounding areas.

Using the results from the survey, the goal was to be able to answer the questions:

1. How often people buy items online
2. How often people buy items at physical stores
3. What they buy online
4. Why they buy certain items online
5. Why they buy certain items in physical stores
6. What would make people choose to go to a physical store instead of buying the same item online

The interviews were directed at store owners/employees and shoppers, with different questions depending on whether the interviewee was a shopper or a store owner/employee. The interviews would be conducted at the shopping mall Emporia, located in Malmö, Sweden. The location was chosen because a shopping mall has a lot of different customers, from different backgrounds, and stores, which would provide different people to interview as well. Another aspect was considered as well: People who are at the shopping mall have decided to go to the physical stores, which

might impact the answers given. On the other hand, it was concluded that since the survey would be distributed to friends and family, many of whom are roughly the same age as the authors of this thesis, that the survey demography might impact the answers given in another way.

The interviewees chosen were from a mix of different stores, e.g. clothing stores, book stores and accessory stores. The shoppers were chosen in a similar fashion, asking people in different stores, of different ages and gender.

Five questions were asked to store owners and employees:

1. Have you noticed any difference in sales for the past few years as a result of the expansion of online shopping and its increasing popularity?
2. Do you suspect that customers try items in the store but then proceed to buy the same item online with the knowledge gained from the physical visit to the store?
3. If the answer to the previous questions are yes, why do you think people buy items online instead?
4. Do you think online-and boutique sales differ between sectors? E.g between the clothing sector and the make-up sector.
5. What are your thoughts on an interactive shopping window? (To be answered after a brief introduction describing what an interactive shopping window is)

The questions aimed at getting the store's perspective on why people shop online contra in physical stores and if they have noticed a difference the last couple of years towards more online shopping. The interview ended with a question about interactive shopping windows, where our use cases and ideas were presented and discussed informally together with the interviewee.

The questions for the shoppers, were similar, but from a shoppers perspective:

1. Have your shopping habits changed the past years in favor of online shopping?
2. Do you try and test items in a physical store with the intention of buying the same product online later on?
3. When do you usually shop in physical stores?
4. What are your thoughts on an interactive shopping window? (To be answered after a brief introduction describing what an interactive shopping window is)

As with the previous questions, the interview ended with a discussion about interactive shopping windows, more focused on what the customers would want out of it rather than what the stores thought they would need.

## Results

The survey was comprised of 89 responses, where the demography of the participants included people from a diverse age range with different genders, various occupations, and varying shopping habits. Using the data from the surveys, results were collected that attempted to answer the six questions formulated in the design of the survey (see 5.4).

The results showed that roughly 70% of the respondents buy something online at least once a month (see 5.26), while only 4 out of 89 never shop online.



Figure 5.26: The figure illustrates the results from the survey regarding online shopping.

Similar data was received for physical store shopping frequencies, were none of the respondents never bought anything at a physical store (see 5.27).

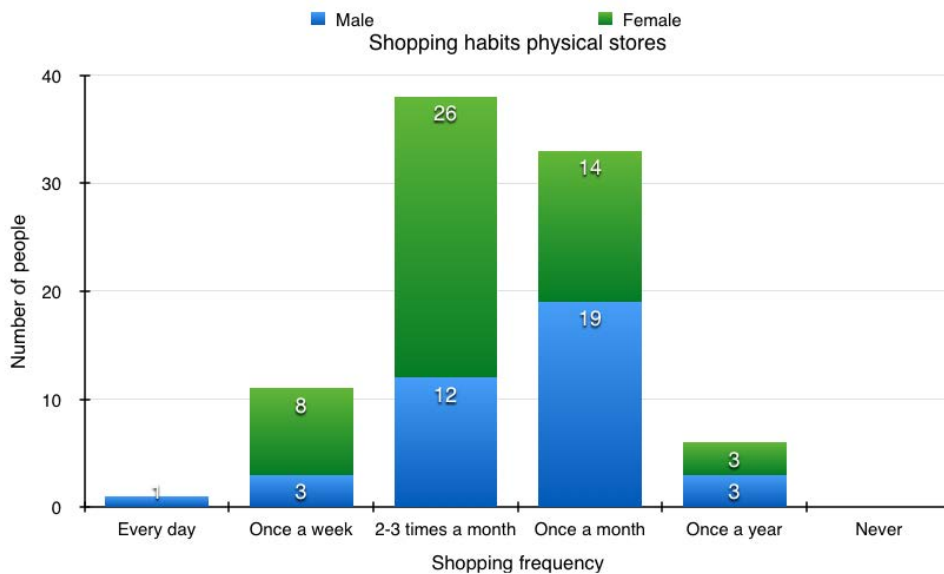


Figure 5.27: The figure illustrates the results from the survey regarding shopping at physical stores.

The biggest driving factors for shopping online versus in a physical store were discovered to be better prices, easier browsing, searching and that was thought of as more convenient (see 5.28). For physical stores, the same results were to be able to feel and try on items, assistance with shopping and to be able to take something home immediately compared to waiting for an item to be delivered (see 5.29).

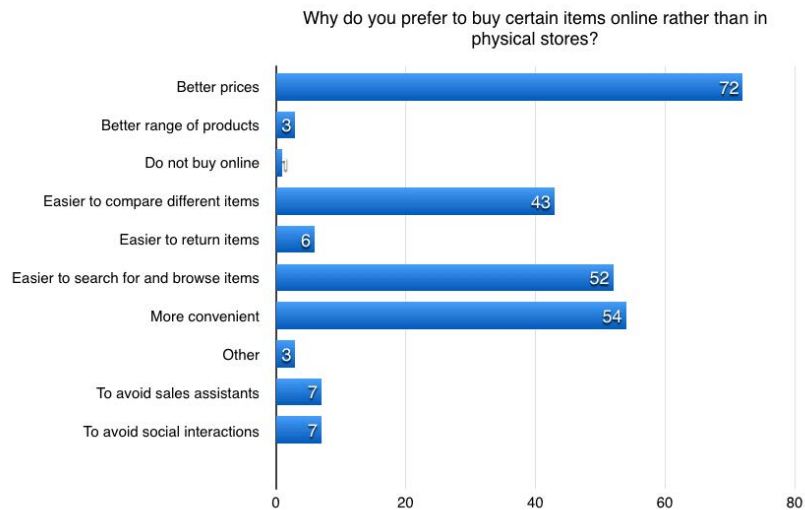


Figure 5.28: The figure illustrates the results from the reasons why some items are preferred to be bought online.

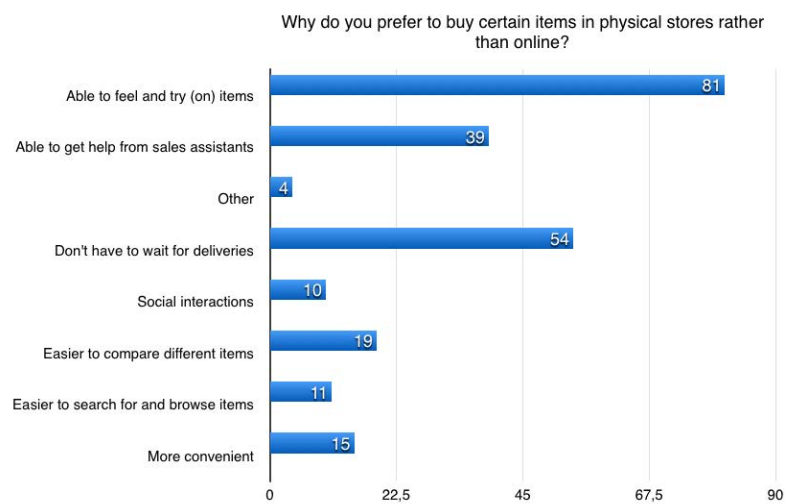


Figure 5.29: The figure illustrates the results regarding why certain items are bought at physical stores.

Studying what kind of items that are bought online showed that the three biggest areas are books, electronics and clothes (see 5.30). That information gives an insight into other potential application areas than the two use cases for this Master's thesis.

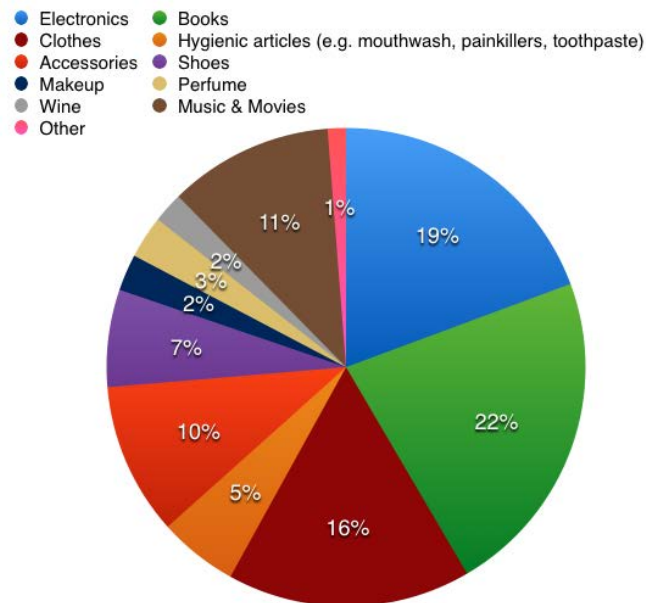


Figure 5.30: The figure illustrates the distribution of what kind of items are bought online.

Finally, since the hope was to find a way of attracting customers lost to online shopping, back to the physical stores, data was analyzed about what customers (i.e. the respondents to the survey) expected, demanded and wanted from physical stores in order to consider taking the time to go to a physical store to buy an item instead of buying that same item online. The results showed that the biggest factor are prices, followed by better browsing, service and deals similar to online coupon codes rather than the ordinary praxis of launching seasonal sales (see 5.31).

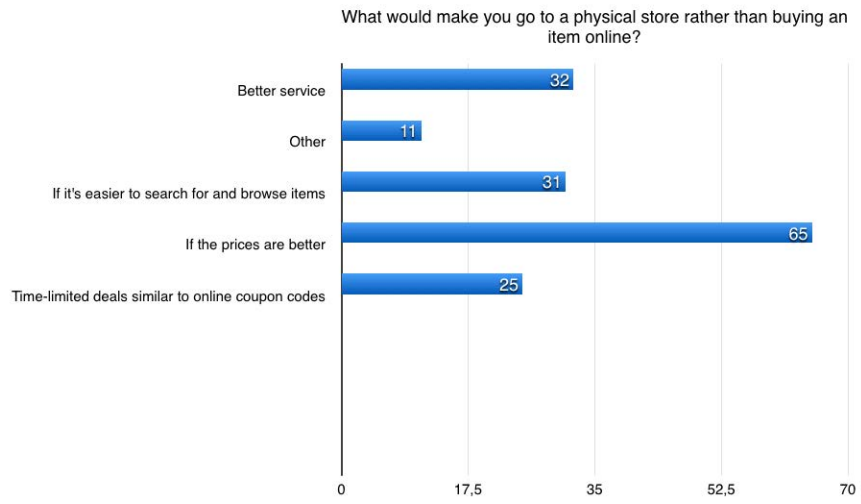


Figure 5.31: The figure illustrates what customers believe would make them go to physical stores instead of buying an item online.



## Chapter 6

# Prototype Design phase

This chapter will provide the processes and results for the prototype design phase.

The investigation phase was followed by the actual implementation of the ideas, the prototype design phase, where the first prototypes of the actual product was created. The process began by examining the results gathered via the surveys and interviews, and an elicitation process for the Realtor use case was carried out. Lo-Fi prototypes were created, evaluated and then made into Mid-Fi prototypes.

Once the use cases had been established, personas and scenarios for each use case were created to get a better feel of the scenarios and their relevance in the real world if they would have been implemented as finished products. Feedback about the Clothing use case had been received through the interviews and surveys, however, it was not believed that they would be a good source of information for the Realtor use case due to the various levels of experience with real estate of the general public. Instead, the focus would be on researching the existing real estate firms in Sweden.

All of the gestures that had been created were documented and animated in a separate digital mockup, which served as a guide when testing the prototypes and when designing the animations in the Mid-Fi prototypes.

### 6.1 Realtor use case

The design of the Realtor use case began with researching information, after which a Lo-Fi prototype was created. The Lo-Fi prototype was translated into wireframes before being finalized as a digital Mid-Fi prototype.

#### **Research**

In order to gather information about important information for the Realtor use case, several background investigations were performed. As a first measure to understand the whole real estate industry the search began with browsing the Internet, examining different real estate agencies' websites to get an idea of what information they believe to be the most important for potential customers. Naturally, each web site and real estate agency had a somewhat different approach

to the way they portray and market their products to potential customers but the essentials were clear: beautiful pictures, clear and visible information about the price, size, location and an impressionistic description of the object and its surrounding environment.

The next step entailed visiting a large, typical Swedish real estate agency and find someone to discuss the concept of a gesture based shopping window with. Before entering their office, the outside display windows were studied in order to see how they depicted their real estate objects at their physical location, using physical media (i.e. paper printouts with information and images).

The person interviewed at the real estate agency was a realtor who worked there. He was asked what aspects he believed were the most important regarding the objects in their display window. The realtor explained that their main focus were the pictures of the apartment or house in question, aesthetically as well as quality wise. Another important aspect of the display was the address of the object, the size in square meters and the number of rooms. The interviewee continued by saying that the setup of how objects are displayed physically in their display window, and how objects appear on their website is similar, the major difference being that the website contains more detailed information and more pictures since customers can scroll and navigate on the website. At the display window, a person is only able to see the most essential information due to the restricted space. When the thesis project and the Realtor use case were described, the interviewee was positive as he believed such a product would be able to present the aforementioned information in a good way visually due to the large “screen” and interactive interface.

### **Lo-Fi prototype**

Based on the elicited information, Lo-Fi prototypes in the form of paper mockups were created to test the scenario and the gestures. As was discovered during the research, the images of the property objects are crucial to a positive perception by potential customers. The paper mockups were created with that information in mind, and tested to ensure that the visual aspects were correct before proceeding to the next step of the design.

When the Lo-Fi prototype was finished, wireframes were created to show more detail and exact representation of information and gesture interaction between user and interface. The wireframes served as the foundation for the design of the Mid-Fi prototype and the creation of objectives for the usability tests (see Appendix B).

## **Results**

The persona for the Realtor use case was a man in his mid- to late twenties, who already owns an apartment and wishes to upgrade to something bigger (see Appendix A.2). The complementary scenario describes how he interacts with the gestural interface, exploring different functionalities and options of the interface as a fluid story (see Appendix A.2). The scenario could later be used when creating the paper mockups as a guide to what functionalities should be presented and how they should appear, visually.

The paper mockups were created using A3 papers and post-it notes, which could be moved to replicate actions performed by a gesture. The first mockup, the main screen of the Realtor use

case, contained four property objects and a phone zone, to which objects could be dragged to save them to a cell phone (see 6.1).

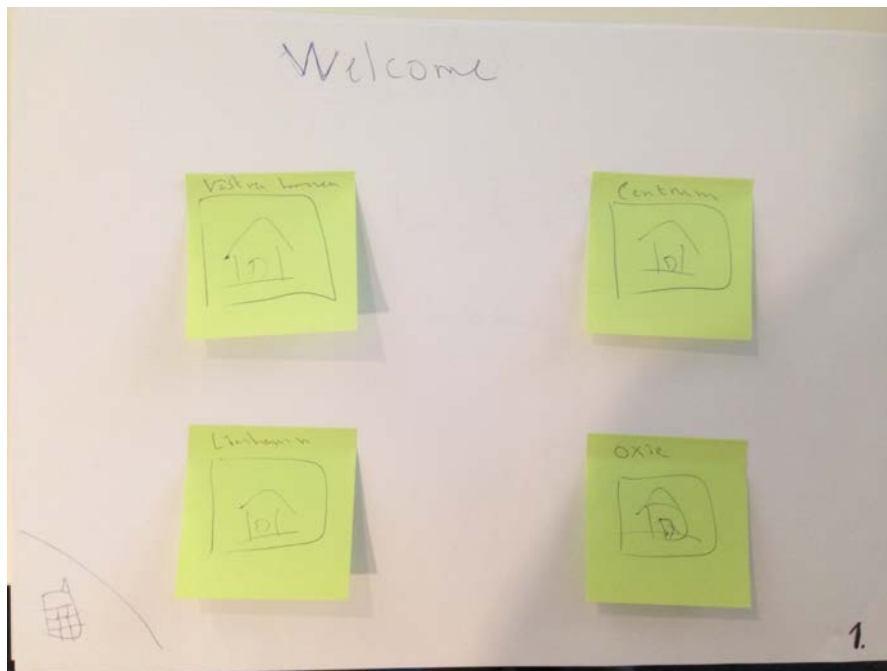


Figure 6.1: The figure illustrates the main screen of the Realtor use case as it looked in the Lo-Fi prototype.

The phone zone represents the user's cell phone, which is connected via aBubbl. The idea is that the shopping window has this feature built in with facial recognition. Meaning that when a person stops in front of the shopping window and his/her phone has the aBubbl technology implemented, the window will connect itself to the person's phone and that person is then able to transfer information from the screen to the phone, using aBubbl. The user's phone acts as a bookmark for the properties that he or she is interested in. Saving a property object to the phone lets the user view the property at home and also notifies the realtor that the user is interested in that particular property.

Owing to the information that was received during the research, the decision was made to focus on the visual representation of property objects. Four property objects shown simultaneously on the screen were decided to be ultimate as the property objects only have a few seconds to catch the attention of someone passing by. The property objects change, through a slideshow, thus enabling one window to display a lot of different objects.

Once a user stops and faces the screen, the user's face is recognized by the interface and the interaction begins. Two options were created, that the user would choose between after the interaction started: The Home matcher option and a browse option (see 6.2).

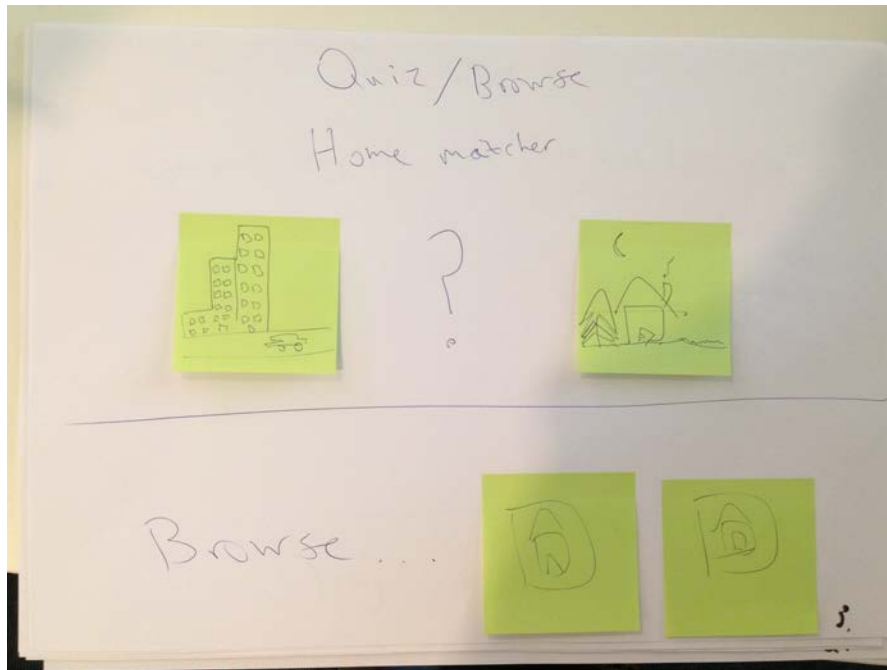


Figure 6.2: The figure illustrates the second screen of the Lo-Fi prototype, with the Home matcher and browse options.

The Home matcher is a quiz game where the user answers a few questions before four matched property objects are presented, based on what answers the user gave. Each questions presents the user with two options, e.g. a bottle of wine or a bottle of beer (see 6.3). The user then has to choose between the two options using different gestures for each question. The idea was to make the quiz fun, while at the same time, use it as a learning period for the gestures that the user would later have to use to interact with the interface.

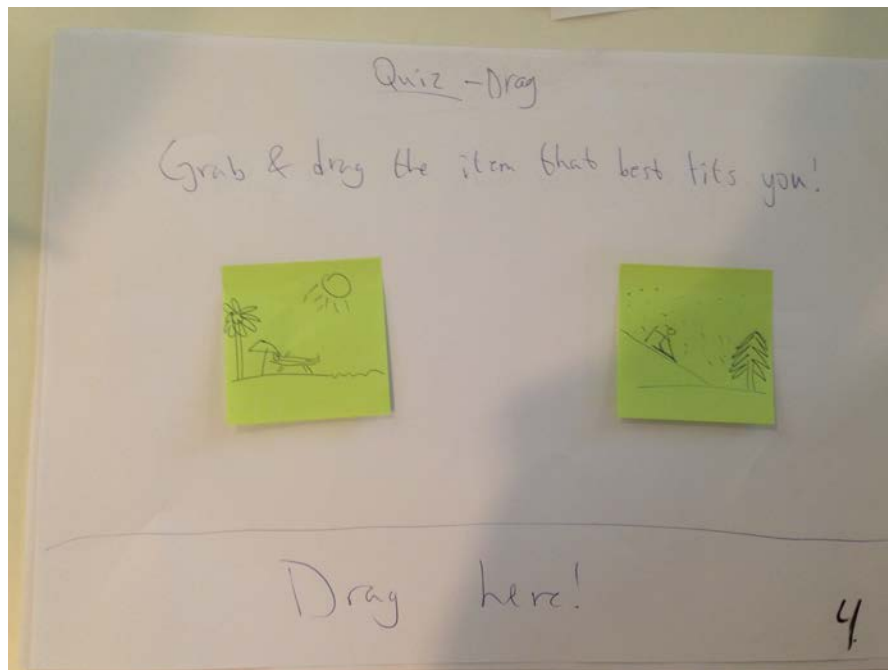


Figure 6.3: The figure illustrates the first quiz question for the Home matcher option.

The design was continued and completed using wireframes. Wireframes were created for each stage of the interface.

The first wireframe describes the passive start screen, where a slideshow of property objects are displayed, the start of the interaction and the Home matcher and browse options (see 6.4). A welcome message was added to the screen, which is presented after a user has been facially recognized. The message encourages the user to wave his or her hand in order to start interacting with the properties. A message was elected to be shown to alert the user that the interface is:

- (a) An interactive interface
- (b) The user has been recognized

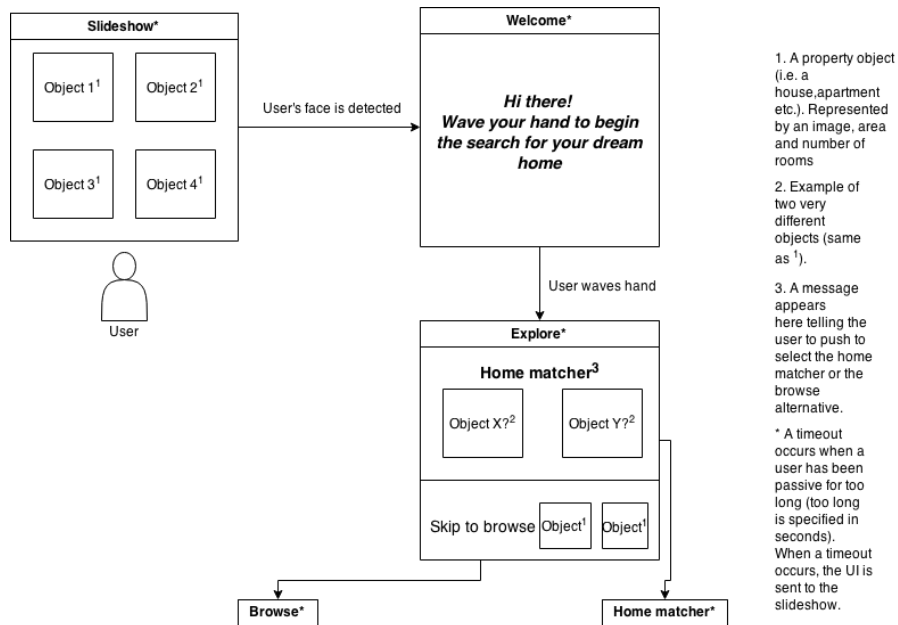


Figure 6.4: The figure illustrates the wireframe for the slideshow, welcome screen and the Home matcher/browse options.

Performing a wave gesture to start the interaction provides the user with the control, instead of having the interface start automatically upon recognizing a face. Furthermore, it was thought that by performing a gesture to begin the interaction, the user would be prepared for the coming gestural interaction.

The Home matcher and browse options were designed with the Home matcher option being at the top of the screen, and being larger in size than the browse option (see 6.4). The goal was to portray the Home matcher as the “main” alternative, to hopefully make users opt to select it over the browse option. As previously mentioned, the Home matcher is designed to provide the users with a fun experience, teach them the gestures as well as providing property objects that match the user’s personality. By having most first time users go through the Home matcher quiz, it was believed that the error rate for the gestures would go down, due to the element of learning.

The second wirescreen (see 6.5) depicts the three questions in the Home matcher quiz. The number three was chosen because it makes the quiz fast to go through, which is important as not to bore the user, and because it provides the opportunity to teach the user three important gestures. The three questions use the gestures grab and pull (see 5.2), Flip (see 5.2) and Swipe (see 5.2). The decision was made to show an animation for each gesture at the start of each question in order to show the user how the gesture is supposed to be performed.

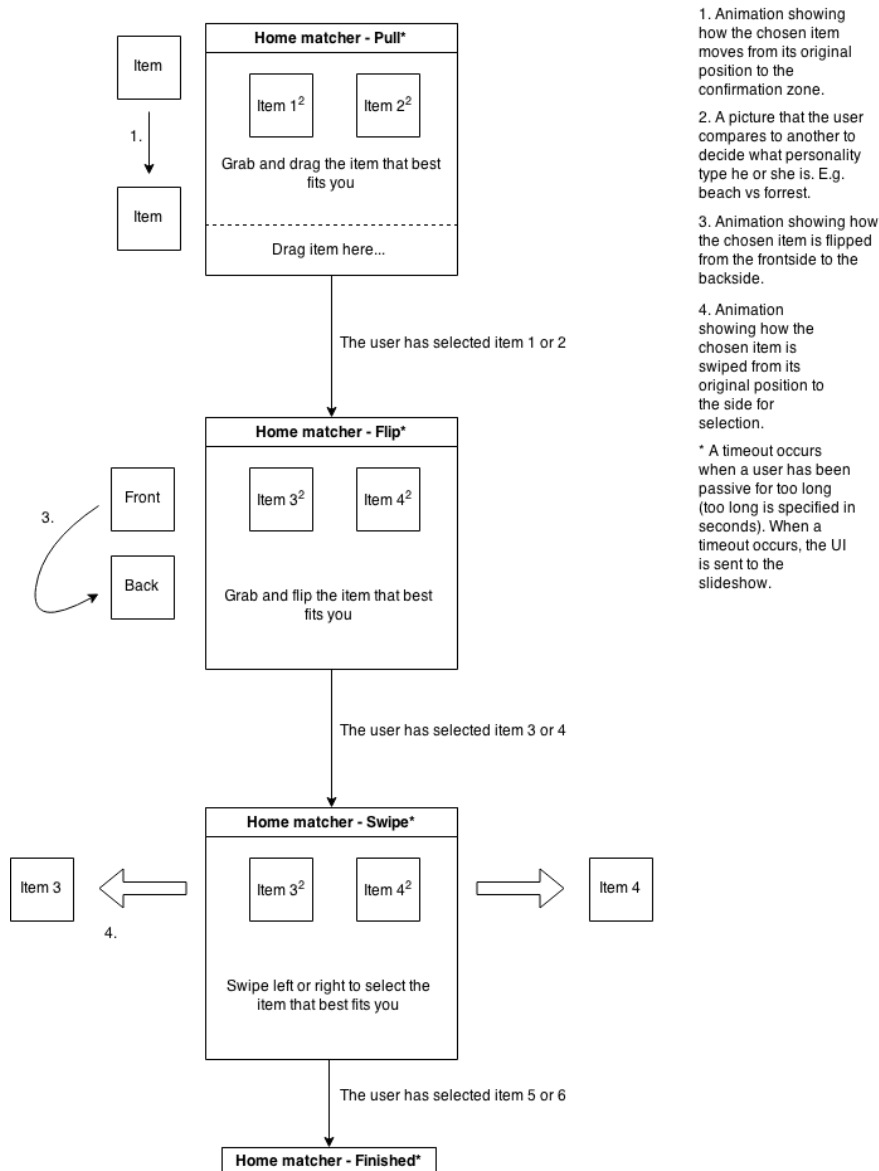


Figure 6.5: The figure illustrates the wireframe for three questions of the Home matcher quiz.

After having finished the quiz, the user is taken to the “matched browse” (see 6.6), where property objects are shown the same way as in the slideshow, except the properties shown have been selected due to the user’s quiz answers. The wireframe also illustrates what it looks like when an object is viewed, using the grab and pull out gesture as mentioned in the section about the Lo-Fi prototype. Relying on the learned information, it was decided that instead of displaying a lot of information on the front of the object, the object was given a backside where more information could be presented. The backside could be viewed by doing a flip gesture, using the metaphor of a card or a book as described in (5.2). Moreover, the action of saving an item to the phone is illustrated (see 6.6), the feature described for the Lo-Fi prototype.

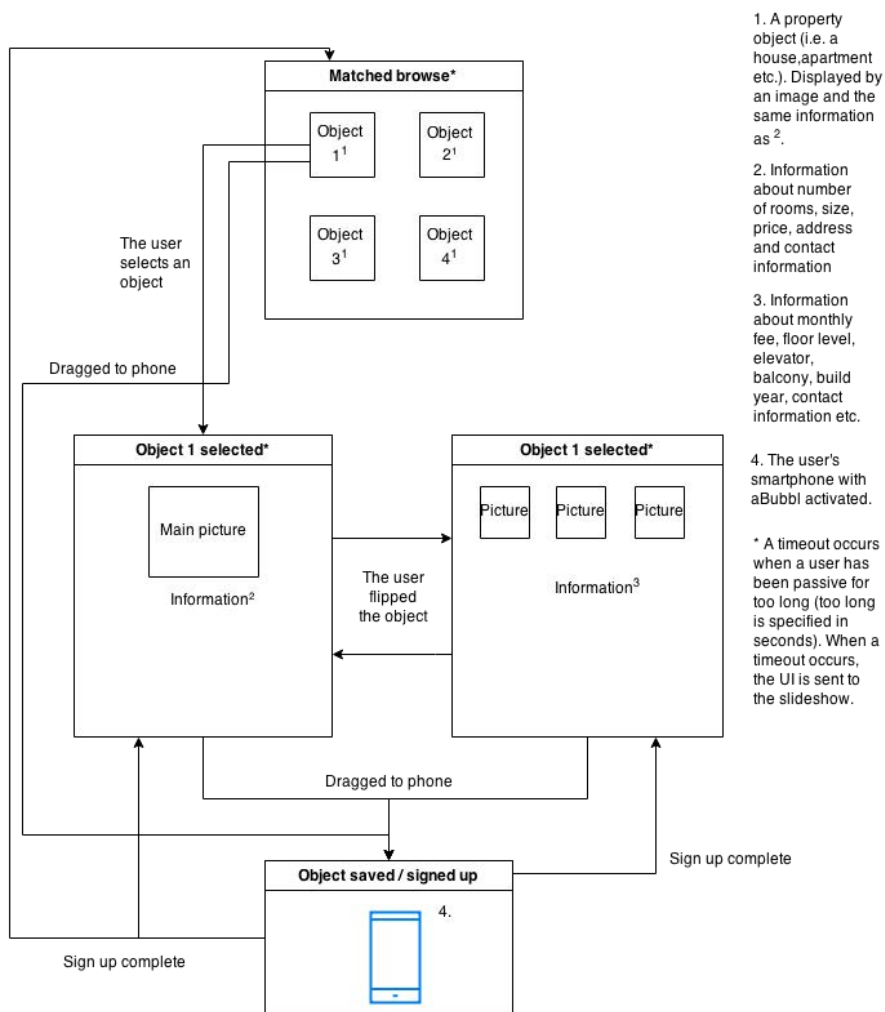


Figure 6.6: The figure illustrates the wireframe for the matched browse mode, and how an object can be manipulated as well as saved to the phone.

### Mid-Fi prototype

Once the wireframes were complete, the design of the Mid-Fi prototype began. The Mid-Fi prototype was created as a Keynote mockup, onto which animations were added for the objects and gestures corresponding to the objectives that the users would perform. Focus was put on the interaction, rather than the design, leading to the bulk of the effort being put into the animation speeds, movement and delay.

### Results

For the creation of the Mid-Fi prototype, a Keynote mockup was created with images and animations. It was decided to forgo the gesture animations that were included in the wireframes before each quiz question. The reason behind the decision was to study how users would perform the gestures, without seeing the animation of how to do them properly. Due to the goal of



making the gestures as natural as possible, it was interesting to study how the users perceived the gestures. However, instructions were added to the Mid-Fi prototype, giving the users some guiding about how to interact with the objects on the screen.

Other changes made to the wireframes when creating the Mid-Fi prototype was the information displayed on the property objects. Other than the image, area and number of rooms of a property, the address, price and size in square meters were added as well. The extra information was added because the prize and size were regarded to be vital to a user evaluating a potential home.

The first screen (see 6.7) displays the passive slideshow, before a user has been facially recognized.

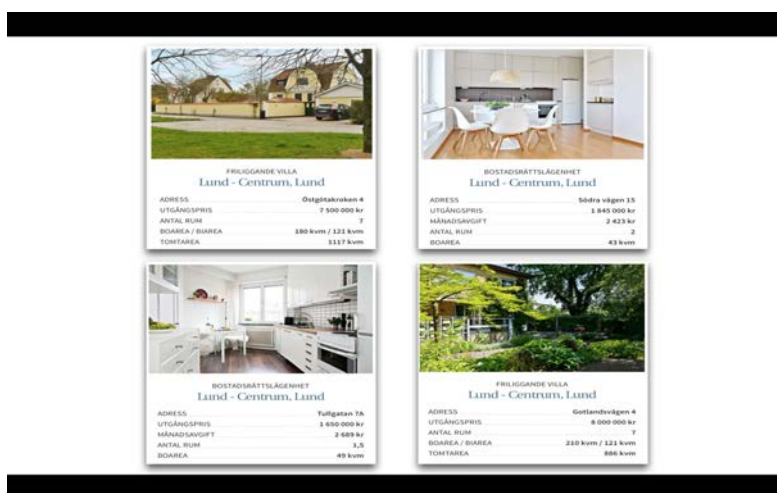


Figure 6.7: The figure illustrates the main screen, where a slideshow of objects is displayed.

After a user has been recognized, the welcome message appears. The decision was made to dim the objects in the background in order to make the message easier to read (see 6.8).



Figure 6.8: The figure illustrates the welcome message that appears when a user has been recognized.

The Home matcher and browse options screen were implemented as described in the wireframes, although with an added instruction that tells the user to use the push gesture (see 5.2) in order to select one of the options. The instruction was added in order to reduce the number of errors made by users (see 6.9). The push gesture was chosen due to the design of the two options as two distinctive, separate, options that were supposed to evoke the experiences of touch screen buttons from the user (see 6.10). Based on those experiences, it was thought to be the most natural to implement the push gesture for this action.



Figure 6.9: The figure illustrates the interface instructions shown at the beginning of the Home matcher and browse options.



Figure 6.10: The figure illustrates the Home matcher and browse options after the instruction has disappeared.

As decided before, the Home matcher would be implemented in the usability tests, thus it was implemented in the Mid-Fi prototype. The first question was chosen as an image of expensive red wine and one image of a bottle and glass of beer (see 6.11). The choices were believed to be fun, and a clear choice for the user, which was determined important for the success of the Home matcher. In addition, the gesture for the question is the grab and pull gesture, the metaphor enhanced by the physical scenario of grabbing a bottle of wine or beer.



Figure 6.11: The figure illustrates the first question of the Home matcher.

The second question features images of a person skiing and a beach respectively (see 6.12). The images play on the choice between winter and summer, cold and warm. The gestural command for the question is the grab and flip gesture, aided by the metaphor of two postcards from a ski- and a summer vacation respectively.

---

*Grab and flip the item that best fits you*



---

Figure 6.12: The figure illustrates the second question of the Home matcher.

The third and final question featured a hamburger and fries versus a plate of sushi, two popular kinds of take-out or fast foods (see 6.13). The swipe gesture used to select one of the options plays on the notion of the food being fast or quick, and browsing through images on e.g. a cell phone's photo library or social media application.

---

*Swipe left or right to select the item that best fits you*



---

Figure 6.13: The figure illustrates the third question of the Home matcher.

Once the quiz has been completed, the user is taken to the matched browse screen and greeted by an instruction encouraging the user to use the grab and pull out gesture (see 5.2) in order to view the different objects (see 6.14). Another instruction lets the user know that property objects can be dragged to the phone in order to be saved (see 6.15).



Figure 6.14: The figure illustrates the instruction for viewing property objects at the matched browse screen.

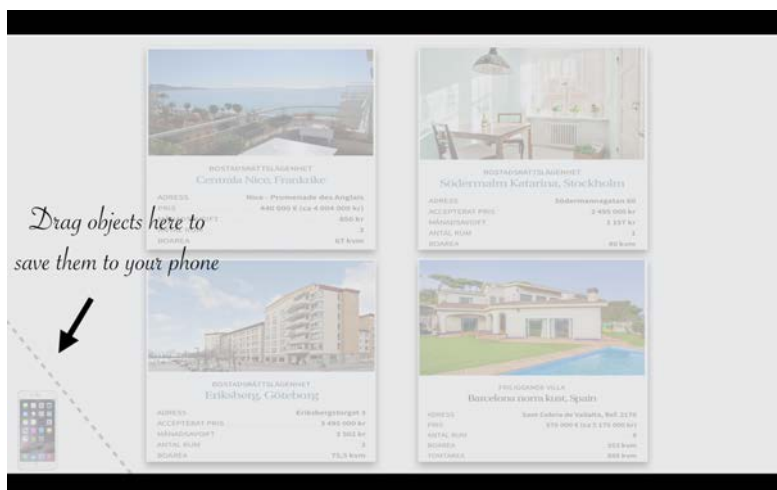


Figure 6.15: The figure illustrates the instruction for saving property objects to the phone at the matched browse screen.

The matched browse mode features four properties, as described in the Lo-Fi prototype and wireframes with the addition of the "phone zone" at the bottom left corner of the screen (see 6.16).

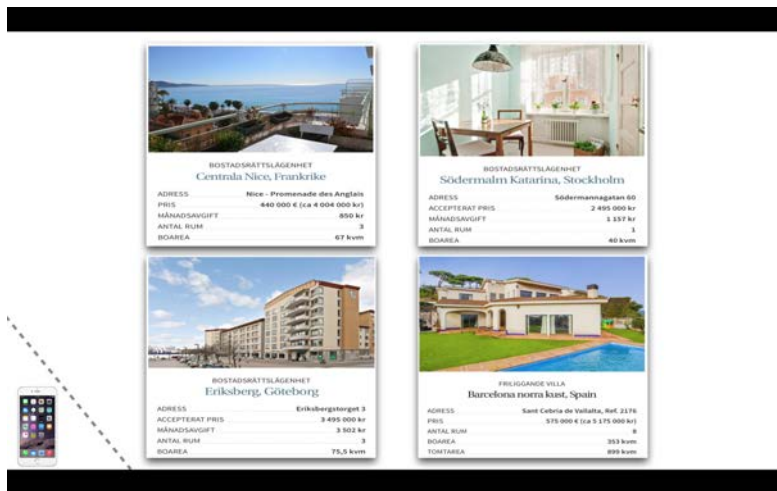


Figure 6.16: The figure illustrates the matched browse view after the instructions have disappeared, displaying four matched properties.

The image below (6.17) depicts the result of a user grabbing and pulling out a property object, causing the object to enlarge, thus allowing the user to examine the property more closely.

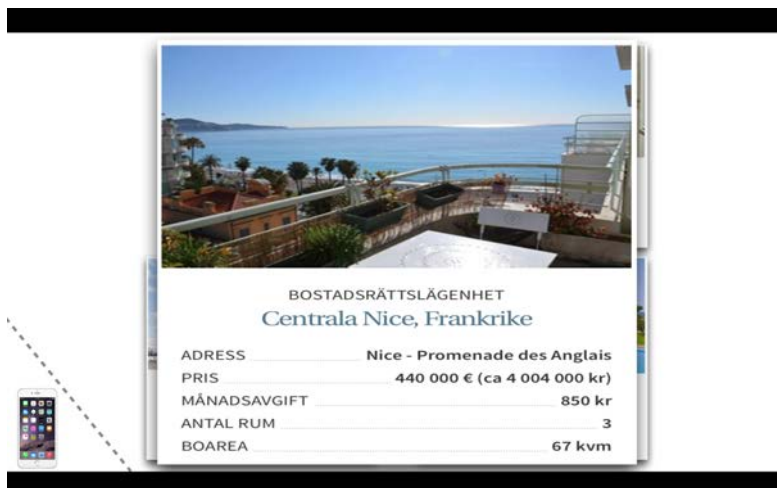


Figure 6.17: The figure illustrates a property object after it has been grabbed and pulled out, thus becoming enlarged.

Grabbing and flipping a property was implemented after the wireframe design, adding a photo of the realtor to give a more personal feeling to the interface (see 6.18). The information displayed on the back changes from property to property, depending on what type of property it is etc. The decision was made to show the price, size and number of rooms on the backside as well as the front side, because when a user decides a property is interesting, it is unnecessary to flip the object back in order to see check the price and size again before dragging the property to the phone.



Figure 6.18: The figure illustrates a property that has been turned over using the grab and flip gesture.

Feedback is important to notify the user of what has happened since the user might feel in less control using a gesture-based interface than a conventional touch screen interface. When an item has been saved to a user's phone, a message is shown (see 6.19) that notifies the user that the item has been saved and that a realtor will contact the user about e.g. open houses. Additionally, it added an extra layer of personality to the interface due to the nature of the message, assuring the user that someone will follow up on the property. After the message disappears, the user is taken back to the matched browse, where another property will take the place of the property saved to the phone. Saved objects are removed because if the user wishes to, the object can be viewed on the phone, and to remove the chances of a user trying to save the same property twice.

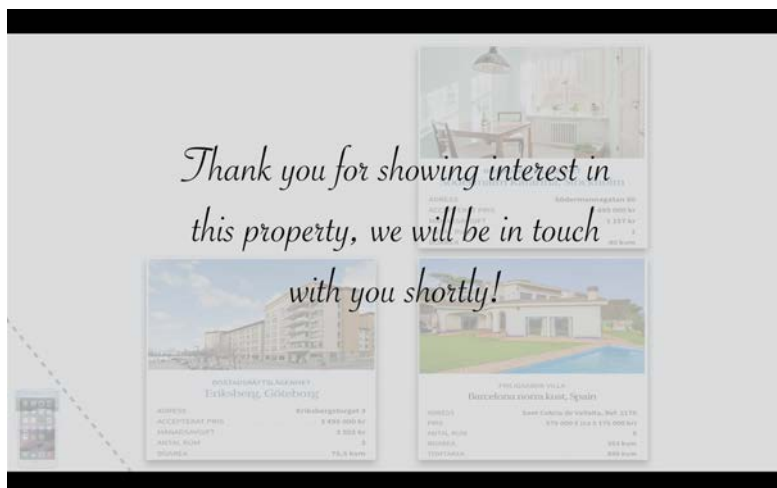


Figure 6.19: The figure illustrates the message displayed when an item has been saved to a user's phone.

## 6.2 Clothing use case

Before the Lo-Fi and Mid-Fi prototyping of the Clothing use case could begin it was crucial to conclude and analyze the results of the shopping survey as well as interviews conducted regarding people's shopping habits.

The results from the survey question "What would make you go to a physical store rather than buying items online?" a majority of the people who answered said that they would shop more often in physical stores if the prices were better and more competitive. This fact indicated that the gesture based shopping window had to have competitive prices and probably a deal system. A sizable amount of the people asked the same previous question also stated that they would shop more frequently in a physical store if browsing and find items was made easier.

### Lo-Fi prototype

The persona for the Clothing use case follows Eric a 23 year old single software developer (See Appendix A.2). In the scenario corresponding to Eric's persona he goes to the mall to shop for clothes and sees that a store has a gesture based shopping window and begins the interaction. The persona and scenario was helpful when developing the paper mock-up as well as the wireframes and therefore also instrumental in the Mid-Fi prototype.

When designing the Lo-Fi prototype, the final Mid-Fi usability tests were kept in mind. A storyboard was created and the parts that were sketched were the parts that a user in the usability tests would encounter. The storyboard was sketched on an A3-sized paper and it was basic in nature, to get an overview of the system.

The next step was to create digital wireframes for the Clothing use case based on the sketched storyboard to enhance the understanding of the system and make further development easier and more tangible. The wireframes depicted how the user would interact with the system and depicted the different phases and transitions of the system.

### Results

The first part of the Lo-Fi that was made was a storyboard sketch that depicted the inspirational picture slides. An inspirational picture is a picture where models are wearing clothing objects that are intended to be for sale in the store that has an interactive shopping window installed. The inspirational pictures portray the models in a certain environment doing certain activities to make a potential buyer want to interact with the shopping window and possibly purchase a clothing item. The next sketch in the storyboard was of an instructional window telling the user to wave his/her hand to begin the interaction. After the wave, there is a picture with feedback saying that the user had been facially recognized and a phone zone appears, as mentioned in the Realtor use case. The fourth sketch in the storyboard portrayed two choices, the men's department and the women's department. In the next sketch there was an inspirational picture with models wearing clothes on a time limited sale. The storyboard's second last sketch was an enhancement of a shirt one of the models was wearing, showing the sizes left as well as the price of the shirt. The last sketch was of a user grabbing and pulling down the clothing object to the phone zone.



The wireframes were built upon the storyboard sketches and became a digital extension to the storyboard that was made on paper. The wireframes show how the different stages of the system look like. They also illustrate what happens when a user performs a certain gesture at a given moment in time and what is supposed to happen then. In the wireframes there are also descriptions as to what the different parts of the wireframes are and what purpose they fulfill.

The first picture of the wireframe shows an inspirational picture of three models wearing clothes that are for sale in the physical store (see 6.20). In reality it is meant to be a slideshow, that's what the arrows in the picture indicate, more pictures are available and will appear after a certain amount of time.

The arrow between the first and second wireframe indicates that a facial recognition has been performed by the system and next a welcome screen is displayed with an instruction written on it (see 6.20).

The third wireframe picture shows the result of the instruction that the previous wireframe has presented, which is two choices, one being the men's department and the other being the women's department (see 6.20).

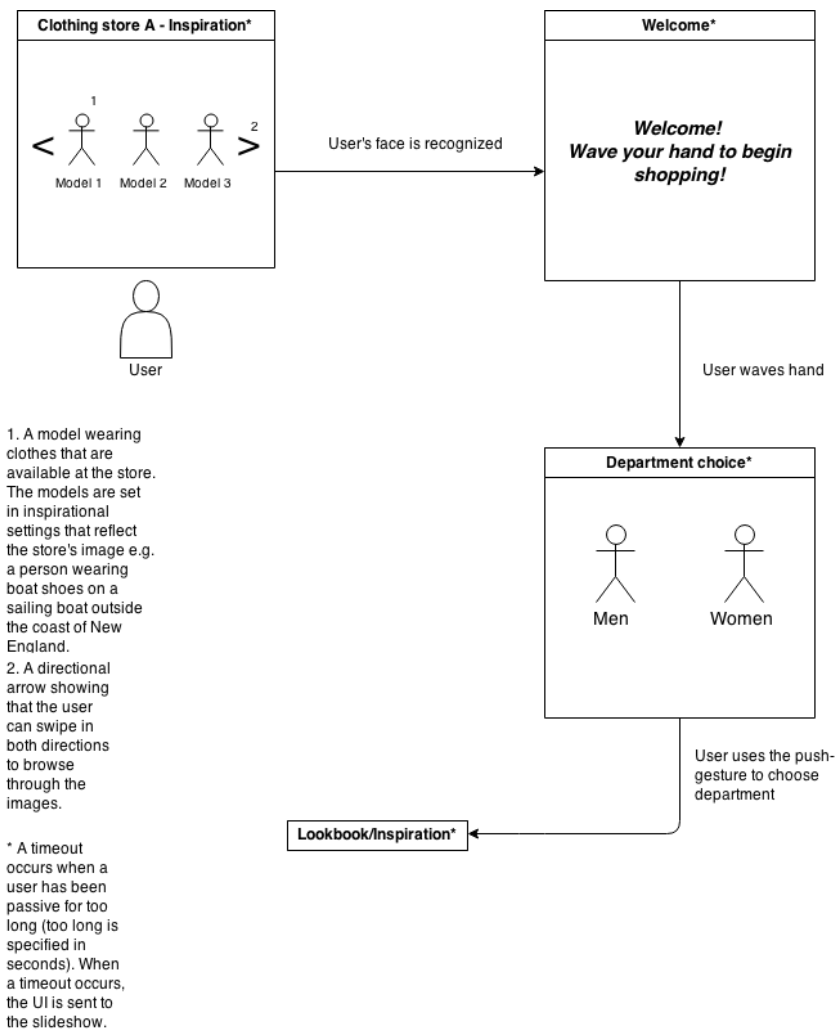


Figure 6.20: The figure illustrates the first part of the shopping window's wireframes.

When the instructed gesture has been performed then the transition is made to either the male or female clothing department. At the particular clothing department there are more inspirational pictures in the slideshow, this time only showing models of the specific gender correlating to the chosen department.

After an item has been selected through the instructed gesture, that item takes up the main view of the shopping window and will in the bottom of the picture concise of clothing suggestions based on the chosen clothing item in the main view. The wireframe (6.21) also shows what happens if a user chooses to do a grab and flip gesture.

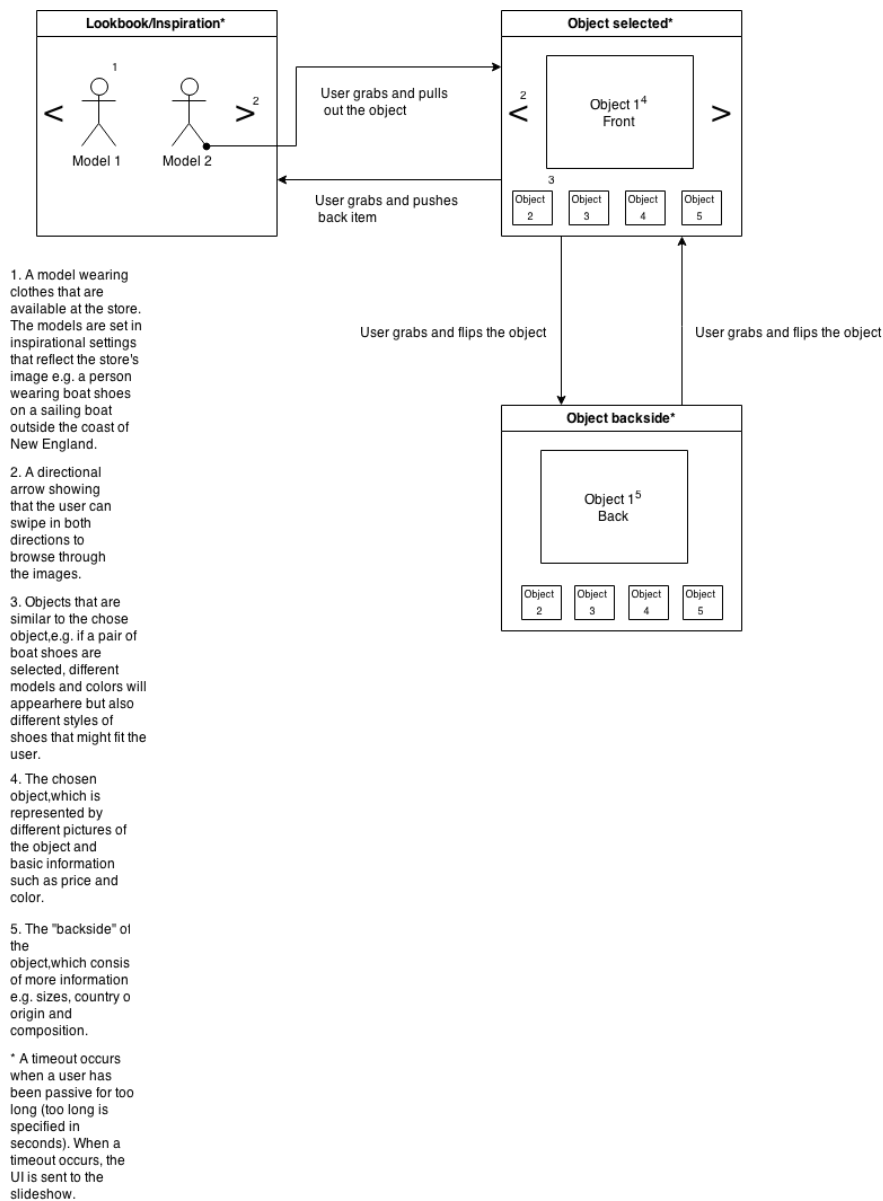


Figure 6.21: The figure illustrates the second part of the shopping window's wireframes.

### Mid-Fi prototype

The Clothing use case's Mid-Fi prototype was developed in Keynote and was completely based on the Lo-Fi prototype as well as the wireframes. The graphics and the design were not the main focus of the Mid-Fi prototype, the most important aspects were the interaction possibilities and to make the user experience comfortable in the regard of transitions between screens, movement speed, delay of objects etc.

## Results

The Mid-Fi prototype is a more developed version of the Lo-Fi version and the wireframes but it is largely based upon and built on both the Lo-Fi prototype as well as the wireframes.

As a result of copyright rules regarding pictures from regular retail stores, it was decided to use private pictures.

The picture below (6.22) depicts an inspirational picture of clothing models standing on top of the Chinese wall. This particular picture was chosen since everyone in the picture is happy and radiating a good time to make a potential buyer want to experience the same level of joy if he or she buys clothes from the store that the interactive shopping window belongs to.



Figure 6.22: The figure illustrates an inspirational picture before facial recognition has been performed.

The second picture (6.23) is the result of the facial recognition when a person has stood in front of the shopping window. The screen also shows a gesture instruction to begin the shopping. The picture in the background is yet another inspirational picture, the reason why it is not the same as the previous inspirational picture is because the slideshow has moved one step before performing the facial recognition.



Figure 6.23: The figure illustrates the welcome screen and the commencing wave gesture.

After a wave gesture has been performed a screen, which has the same background as the previous screen, with two options appears (see 6.24). The instruction in this case is to “push the female or male icon to select department”. If the female icon is pushed the next step is a transferal to the female department and if the male icon is pushed, the transferal is performed but directs to the male clothing department instead.



Figure 6.24: The figure illustrates the choice between the male and female clothing departments and a select instruction.

After the male department has been selected, the next two pictures depict gestural instructions to be used in a later stage of the shopping process. The grab and pull out gesture is used to find out more information about a clothing item (see 6.25). The phone zone has the same function as in the Realtor use case, grab and pull down objects to save them to phone (see 6.26). The picture in the background is an inspirational picture of a group of guys standing in Shanghai , displaying the clothes that are for sale. As the aforementioned inspirational picture, this one has

also been chosen to portray a certain lifestyle connected to the store's image.

This phase was not part of either the wireframes nor was it part of the original storyboard sketches. The reason why it was added was because the system needed to give more interactional instructions so a buyer is able to understand what to do when he or she wants to look at and save clothing items.



Figure 6.25: The figure illustrates how to interact with clothing objects



Figure 6.26: The figure illustrates the instructions on how to save clothing objects to a cell phone

When the instructions have faded there are more inspirational pictures on the slideshow presenting clothes that are for sale in the store (see 6.27), as seen in the second part of the wireframes (see 6.21).



Figure 6.27: The figure illustrates an inspirational picture with clothing objects that are for sale

In the inspirational picture below (6.28), there are five objects that are manipulable with gestures, they are represented with circles. The grey circle indicates that the item in question is not on sale whilst the red circle with the 30% mark indicates that the shirt is on a 30% sale at the moment. The reason behind the circle as an indication that an object is manipulable is that circles symbolize unity and a round object is formed as a grabable ball, but in 2D rather than 3D.



Figure 6.28: The figure illustrates an inspirational picture with five buyable and interactable objects.

When a grab and pull out gesture has been performed on the clothing object that is on sale, the object fills the main view with suggested clothing objects located in the suggestion bar below (see 6.29).



Figure 6.29: The figure illustrates a selected clothing object.

The arrows to the left and right of the shirt object indicate that there are more pictures available of the shirt in question and the point is to swipe left or right to see more pictures of the shirt (see 6.30).



Figure 6.30: The figure illustrates a new picture of the selected clothing object.

When wanting to switch to a similar clothing object that is located in the suggestion bar, a grab and pull out gesture is required. The old shirt is then placed in the red shirt's old position in the suggestion bar (see 6.31). The reason behind the grab and pull out gesture is to resemble a physical clothing store when a person grabs a shirt on hanger from the rack and holds it up in front of him/herself to inspect it. This piece of interaction is not described in either the wireframes or the storyboard sketches.





Figure 6.31: The figure illustrates a switch of shirts and the main view shows the red shirt.

When a clothing item has been grabbed and pulled down to the phone zone (see the Realtor use case), a feedback screen appears portraying the message that the clothing item in question has been saved to the phone (see 6.32).



Figure 6.32: The figure illustrates the feedback a shopper receives after pulling an object to their shopping cart.

To deselect and put back an item a reverse gesture to the grab and pull out needs to be done, in other words a grab and push or simply push gesture will suffice. The reason why this gesture is chosen for this purpose is that is similar to what a person would do in a physical store after he or she has looked at a clothing item up close. This option is also made possible if a potential buyer wants to browse for more clothing items and needs to clear the view to swipe between inspirational pictures to find other clothing objects.

# Chapter 7

## User testing

This chapter provides the execution and results of the usability tests carried out on the two Mid-Fi prototypes.

### 7.1 Prototype evaluations

The third phase of the design process was to perform user tests on the prototypes. In order to make sure that the prototypes were ready for the tests, they were evaluated.

The first step of the evaluation was to analyze the Mid-Fi prototypes by testing and trying different time values and calibrating the timing of the actions. This was done by running through the tasks and objectives and evaluating how well the animations reflected the gestures that were performed. As the desire was for the mockups to appear real and responding to gestures, functionality for remote controlling the animations was added, which would let the gesture animations be timed to the users' gestures without having to stand by the computer.

The evaluation of the prototypes provided benchmark values for animation times of the different gestures, setting up speeds that reflected the natural speed of the physical gesture carried out by a user. The resulting animation times for the different gestures can be seen in table 7.1.

Table 7.1: The gestures used in the Realtor and Clothing use cases and their respective animation time as implemented in the Mid-Fi prototypes

Gesture	Animation time (seconds)
Push to select	0.80
Swipe to browse	0.90
Grab and pull out	1.25
Grab and replace	1.0
Grab and pull	1.0
Grab and push	1.25
Grab and flip	1.50

## 7.2 Wizard of Oz

A Wizard of Oz test is sometimes used in the field of HMI and testing, it is a research experiment where the test participants believe, or partly believe that the piece of technology they are testing is fully functional, whilst the reality is that a person is controlling and doing the actions on a not fully implemented system.

In the prototypes and usability tests, when a test participant was given a certain task, he or she then performed the designated gesture and the Wizard of Oz pressed a button making the animation correlating to the performed gesture visible on the screen, hence making the interaction look functional. Only the tasks that were given to the test participants were implemented and functional. The other objects that were visible were there for a holistic experience.

### Benefits

The benefit with the Wizard of Oz method for usability testing is that product developers are able to perform simple, quick and efficient usability tests to find problems with e.g. a system's usability or find nuggets in the usability field that the developers have overseen when doing the initial brainstorming and design of the system.

## 7.3 Test participants

The usability tests were done on two different groups of people. One of the groups was comprised of five (5) people without any prior knowledge to gesture based interaction technology and these test participants were young adults/students, which belong to generation Y (people born between the late 1980s and the mid 1990s). These specific test participants were chosen since the conducted interviews and surveys covering shopping habits online and in physical stores indicated that people belonging to generation Y could be the group of people that are most prone

to use these types of interactive shopping windows if they are ever realized.

The other group of test participants consisted of six (6) people working at Crunchfish, between the ages of 23 and 40. The reason behind choosing this group as test participants was the fact that they work with gestural technology on a daily basis and are a rare group of people due to this fact. It is otherwise hard to find experienced people in the field of gestural technology. The idea was to compare the different groups performances to each other to see if previous knowledge of gestural technology had a significant effect on the performance results.

### 7.3.1 Purpose and motivation

An imperative part of the tests was to time the test participants when they executed the tests, so a comparison of the test participants times could be made. If the tests went smoothly and rather quickly, the conclusion could be drawn that the interaction, objectives and instructions were clear and easy to understand.

The purpose of testing both use cases on the two groups of people with a large knowledge gap in regard to gestural technology was to see how they both performed when given the exact same objectives and instructions.

The test participants did the tests in different order, if one test subject started with the Realtor case, then the next began with the Clothing use case and so forth. The reason why the tests were carried in this way was to avoid the transfer of learning effect. By using this "Within subject design"-approach the transfer of learning effect was supposed to be balanced out.

The extracted data from the tests that was desired was both qualitative and quantitative in nature. The ambition was to use semi-structured interviews, conversations, filming and observations to make out the qualitative data and utilize surveys to collect quantitative data.

### 7.3.2 Research questions

This section provides different questions that were formed before performing the user tests, they were clear goals to strive for and used to investigate extra important elements of the tests.

- Will the experienced group perform better than the less experienced group, time wise as well as the amount of failed gesture performing attempts?
- As a result of one group being experienced in the gestural field, would their mind set be restricted in any way due to them knowing the limit of today's gestural technology?
- How will the inexperienced group act, will they have an open mind and see other possibilities and approaches that the experienced group will not see?
- Do the test participants need graphical tips to understand how the gestures work or is it enough with written instructions?
- Each quiz-question in the Realtor use case has a new gesture for each quiz question, will the test subject understand the use cases better? Subconsciously or not.
- How will the test participants perform the grab and flip gesture in the Realtor use case?

- In the Clothing use case, how will the test participants change shirt from the suggestion bar to the main view, which gestures will they use?
- Is it clear for the test participants how to put back an object after interacting with it in the Clothing use case?
- In the Clothing use case, how do the test participants want to put back/remove an object from the main view?

### 7.3.3 Test descriptions

Each participant tested both the Realtor use case as well as the clothing store use case. However, the test participants did the tests in different order, e.g. if one test subject began with the Realtor use then the next test subject began his/her tests with the Clothing use case.

Before the tests began briefing with the test participants took place where the Master's thesis was described and some details about the test process was discussed. The test participants were also given a consent form to sign where they agreed to have their performance filmed, for qualitative data analysis only.

After the test participants had been debriefed on the project and the concept of the project, they were given a written scenario story to read, the two test cases each had their own scenario, to put the test participants into context before they performed the tests. The test participants read the scenarios one at a time, right before each test.

When the test participants had read the scenario belonging to the test they were about to perform, the test began. During the tests the test participants were encouraged to talk freely and describe what they were thinking, since the tests were filmed the test participants thoughts and expressions were supposed to be used for qualitative data analysis in the evaluation stage. After the tests were completed, the test participants received a questionnaire for each one of the two tests, this would provide the quantitative data.

#### Roles

Alexander had the role as the test leader and he gave the test participants the background information about both use cases, he also gave the test participants their tasks throughout the tests. If a test subject asked any question, Alexander tried to help out in such a way that he or/she did not receive any clues that were too descriptive and that would compromise the results of the usability test.

David's role was to be the Wizard of Oz as well as the observer. He was also the person filming the tests, for the collection of qualitative data.

#### Successful completion criteria

If a test subject has completed the tasks and objectives given to him or her good enough is decided with a subjective view. As a result of the tests nature where the main goal is to observe how a test subject spontaneously performs and completes a given task, each test subject may perform differently than other, which as previously mentioned is the the point of the tests. It is therefore up to the designers to decide if a test subject has performed acceptably.

## **Test environment**

The tests were performed in a room with a television screen, which was where both the use cases were displayed and it was with the television the test participants interacted with. However, since the test was a Wizard of Oz test, the computer that ran the prototypes was controlled from a distance through a remote control by the Wizard of Oz, David. In conclusion, the test participants were able to see the computer and the television in the room but they were only facing the television screen while the computer was placed in such a way that the screen was not visible for the test participants.

### **7.3.4 Realtor use case**

#### **Objectives walk-through**

This section describes how each part of the test was built up. Part meaning in what stage the test subject is at in the test. Objective means the task the test leader has given the test subject orally at that particular part of the test. Instruction from interface signifies what instruction the test subject receives from the interface at that particular part of the test.

Part 1. (See figure 6.7)

Description: The test subject encounters the realtor shopping window with a slideshow which depicts four different property objects.

Objective: Stand in front of the realtor window.

Instruction from interface: -

Wanted result: Facial recognition.

Part 2. (See figure 6.8)

Description: Facial recognition is performed (through the Wizard of Oz) and the test subject is presented with a welcome message with an interaction command, which is to wave his or her hand to the screen in order to begin the interaction.

Objective: Read the screen's instructions.

Instruction from interface: Wave hand to begin interaction.

Wanted result: Test subject waves hand.

Part 3. (See figure 6.9)

Description: The test subject is transferred to a window with an instruction, either to take the Home matcher quiz or to browse properties.

Objective: Select the Home matcher quiz.

Instruction from interface: Push to select the Home matcher quiz or the browse option.

Wanted result: Test subject uses palm or finger to push and select the Home matcher quiz.

Part 4. (See figure 6.10).

Description: When the test subject has performed the previous task, he or she is transferred to the first quiz question. The first question consists of two choices, a wine option and a beer option.

Objective: Select the wine option.

Instruction from interface: Grab and drag object to the confirmation zone.

Wanted result: The test subject grabs and drags the wine option to the confirmation zone.

Part 5.(See figure 6.11)

Description: The test subject is transported to the second quiz question, which has the same set-up as the first question. The options in this case is a skiing option and a beach option.

Objective: Select the beach option.

Instruction from interface: Grab and flip the object that fits you best.

Wanted result: The test subject grabs and flips the beach option.

Part 6. (See figure 6.12).

Description: The test subject is moved to the third and final question. In this question a burger and a plate of sushi is pitted against each other.

Objective: Select the burger option.

Instruction from interface: Swipe left to select the burger option or right to select the sushi option.

Wanted result: The test subject swipes the burger option to the left for selection.

Part 7. (See figure 6.14), See figure 6.15

Description: After the quiz, the Realtor use case gives the appearance that it is processing the answers from the previous questions, a parameter within the Wizard of Oz-test. Two instruction screens displaying how to interact with property objects are presented in succession at this point.

Objective: Read the instructions.

Instruction from interface: Grab and pull out a real estate object to find out information about it. Drag items to the phone zone, this acts as your shopping cart.

Wanted result: The test subject has understood the instructions and remembered them.

Part 8. (See figure 6.17)

Description: Four property objects are presented as home matches, based on the test subject's choices during the quiz.

Objective: Take a closer look at the property located in Nice, France.

Instruction from interface: -

Wanted result: The test subject has remembered the interface instructions from the previous part and uses the grab and pull gesture to interact with the property object.

Part 9. (See figure 6.17)

Description: Find more detailed information about the property object.

Objective: Identify more detailed information about the property, the information is located on the back of the property object. Use a gesture you have learned from the quiz and pick the gesture you thinks is the most suitable.

Instruction from interface: -

Wanted result: The test subject performs the grab and flip gesture or an equivalent flip gesture.

Part 10. (See figure 6.18)

Description: Switch the view of the property.

Objective: Flip the property object back to the front.

Instruction from interface: -

Wanted result: The test subject performs the grab and flip gesture or an equivalent flip gesture.

Part 11. (See figure 6.19)

Description: Sign up for information from the realtor in charge of the property in question.

Objective: Save property object to phone.

Instruction from interface: -

Wanted result: The test subject performs the grab and pull gesture and drags the property object down to the phone zone.

### 7.3.5 Clothing use case

#### Objectives walk-through

This section describes how each part of the test was built up. Part meaning in what stage the test subject is at in the test. Objective means the task the test leader has given the test subject orally at that particular part of the test. Instruction from interface signifies what instruction the test subject receives from the interface at that particular part of the test.

Part 1. (See figure 6.22)

Description: The test subject encounters the Clothing use case's shopping window with a slideshow that depicts inspirational pictures of models wearing clothing items, which are for sale in the fictional physical store.

Objective: Stand in front of the shopping window.

Instruction from interface: -

Wanted result: Facial recognition.

Part 2. (See figure 6.23)

Description: The test subject is presented with a welcome message and an interaction command.

Objective: Follow the screen's instruction.

Instruction from interface: Wave to start.

Wanted result: Test subject waves his/her hand.

Part 3. (See figure 6.24)

Description: The test subject is then transported to a window with two choices, men's department and women's department. The men's department is chosen using the push-gesture.

Objective: Choose the men's department.

Instruction from interface: Push to choose clothing department.

Wanted result: The test subject pushes his/her palm/finger on the men's department.

Part 4. (See figure 6.25, See figure 6.26)

Description: When the test subject arrives at the men's department there are two instructional pictures presented. One of which describes how to interact with an object and another describing how to save the object to the test subject's phone.

Objective: Read the screen's instructions.

Instruction from interface: Grab and pull out object to find out information about it. Drag items to the phone zone, this acts as your shopping cart.

Wanted result: The test subject remembers the instruction for later application.

Part 5. (See figure 6.27)

Description: The next section displays inspirational pictures of male models wearing clothes that are for sale in the physical store. All clothing objects that are manipulable have a circle placed on top of them. A grey circle indicates normal price and a red circle with a number representation indicates that the clothing object in question is on sale with a certain amount of discount.

Objective: View the next inspirational image to the left, twice.

Instruction from interface: -



Wanted result: The test subject swipes his/her hand twice in succession.

Part 6. (See figure 6.28, See figure 6.29)

Description: In the third inspirational picture, the test subject encounters a green shirt with a red circle on it indicating that the shirt is on sale.

Objective: Find out more information about the item on sale.

Instruction from interface: -

Wanted result: The test subject has remembered the interface instruction from part 4 so he/she grabs and pulls out the object in question.

Part 7. (See figure 6.30)

Description: The green shirt is now displayed in the main view.

Objective: Find a new picture of the shirt.

Instruction from interface: -

Wanted result: The test subject uses the swipe gesture to change picture.

Part 8. (See figure 6.31)

Description: The test subject is now told to switch the shirt in focus with the red shirt located in the suggestion bar.

Objective: View the red shirt from the suggestion bar.

Instruction from interface: -

Wanted result: The test subject grabs the red shirt and pulls it up to the main view.

Part 9. (See figure 6.32)

Description: Saving the clothing object to phone.

Objective: Save the shirt to your phone/shopping cart.

Instruction from interface: -

Wanted result: The test subject uses the grab and pull gesture to save the item to the phone.

Part 10. (See figure 6.28)

Description: Removing shirt from main view.

Objective: Put shirt back into the inspirational picture where you found it.

Instruction from interface: -

Wanted result: The test subject uses a grab and push/or just push-gesture to put the shirt back.

## 7.4 Results

The results for the Realtor and Clothing use cases are divided into two parts, first the qualitative results will be presented, secondly the quantitative results will be presented.

### 7.4.1 Qualitative

The qualitative results are based on observations regarding the test participants performances, which aims to present the reader with easily understandable data about how the test participants fared in the different parts of the tests.

## Realtor use case

Table 7.2: The table presents the qualitative results for the Realtor use case.

	<b>Group A (Inexperienced)</b>	<b>Group B (Experienced)</b>
Part 1. (Start)	All test participants performed correctly.	All test participants performed correctly.
Part 2. (Facial recognition)	All test participants performed correctly.	All test participants performed correctly.
Part 3. (Quiz choice)	All test participants performed correctly.	All test participants performed correctly.
Part 4. (First quiz question)	All test participants performed correctly.	All test participants performed correctly.
Part 5. (Second quiz question)	Only two out of the five inexperienced test participants performed the correct gesture, the other three ignored/missed the new instructions and used the previous question's gesture.	Four out of the six in the experienced group performed the correct gesture, the other two performed the previous question's gesture.
Part 6. (Third quiz question)	Four out of five test participants swiped left to select the burger option, one had misunderstood and swiped the burger option to the right instead.	Five out of the six test participants swiped left to select the burger option, one had misunderstood and swiped the burger option to the right instead.
Part 7. (Interaction instructions appear)	All test participants read the instructions.	All test participants read the instructions.
Part 8. (Look at a property)	Three out of five test participants performed the right gesture (grab and pull out) at this stage, two test participants tried to grab and pull down the property object to the phone zone, which was incorrect.	Four out of six test participants performed the right gesture (grab and pull out) at this stage, two test participants tried to grab and pull down the property object to the phone zone, which was incorrect.
Part 9. (Find out more detailed information)	Two out of the five test participants did some kind of flip gesture, the remaining three performed some kind of swipe gesture.	Four out of the six test participants did some variant of the flip gesture, the remaining three performed some sort of swipe gesture.

continued on the next page

	<b>Group A (Inexperienced)</b>	<b>Group B (Experienced)</b>
Part 10. (Flip back the property object)	The test participants that had performed correctly previously also performed correctly in this part of the test as well. When the test participants that had performed the wrong gesture in the previous part had been told what they had done wrong, they all performed the right gesture.	The test participants that had performed correctly previously also performed correctly in this part of the test as well. When the test participants that had performed the wrong gesture in the previous part had been told what they had done wrong, they all performed the right gesture.
Part 11. (Save property object to phone)	Four out of the five inexperienced test participants performed the correct gesture, one person tried to swipe down the property object to the phone zone.	Four out of the six experienced test participants performed the correct gesture, two people tried to swipe down the property object to the phone zone.

## Clothing use case

Table 7.3: The table presents the qualitative results for the Clothing use case.

	<b>Group A (Inexperienced)</b>	<b>Group B (Experienced)</b>
Part 1. (Start and facial recognition)	Not a problem for any of the test participants.	Not a problem for any of the test participants.
Part 2. (Wave to start)	All test participants performed the wave gesture without any trouble.	All test participants performed the wave gesture without any trouble.
Part 3. (Department choice)	Out of the five test participants, four used an open palm as a “push” gesture, one used the index finger, like on a touch screen.	Out of the six test participants, four used an open palm as a “push” gesture, two used the index finger, like on a touch screen. The same conclusion can be drawn for the people in this group who used the index finger as a push gesture.
Part 4. (Interaction instructions appear)	All the test participants read the instructions without any problem, a few asked some questions, which were answered without compromising the test.	All the test participants read the instructions without any problem, one or two test participants were talking during the time when the instructions were displayed on the screen but they did not miss any information.
Part 5. (Change two pictures)	All test participants performed the correct gesture.	Not a problem either, intuitive for everyone to use the swipe gesture to switch inspirational pictures.
Part 6. (Look at the green shirt)	Two test participants had trouble remembering that they needed to grab and pull out an object to be able to find out more about it, one person used the push gesture and the other grabbed and dragged the object to the phone zone instead.	In the experienced group there were two test participants that used the push gesture instead, the rest performed the correct gesture.
Part 7. (Find a new picture of the shirt)	A majority of the test participants used the swipe gesture to change pictures, One person tried to grab and pull the object.	All test participants used the correct swipe gesture to switch pictures of the shirt.

continued on the next page

	<b>Group A (Inexperienced)</b>	<b>Group B (Experienced)</b>
Part 8. (View the red shirt)	One person did the gesture that was intended for this part of the interaction, which was grab and pull to the main view. Everyone else tried to either swipe up the red shirt to the main view or push the shirt to select it.	The experienced group had a 50/50 distribution on how many performed the intended gesture whilst the other half did similar actions as the inexperienced group.
Part 9. (Save shirt to phone)	Nearly everyone performed immaculately at this stage and did the grab and pull gesture correctly, only one person tried to swipe the object to the phone zone.	All but one person grabbed and pulled the object correctly, the one who did not, grabbed and pulled the object straight down rather than to the designated confirmation zone.
Part 10. (Remove shirt from main view)	One person did the intended push gesture, the other four performed different variants of the swipe gesture.	Three out of the six experienced test subject used the push gesture, whilst the other three did as the inexperienced group did, which was to try to swipe away the object from the view.

## 7.4.2 Quantitative Results

The quantitative results from the usability tests were based on two different surveys, one for each use case the test participants tested. Each survey had four questions each and the questions that were answered were scaled from 1-6 meaning. Zero meaning that the test subject disagreed with the question/claim, the number six meaning that the test subject completely agreed with the question/claim. The reason why the scale was between 1-6 was to avoid middle answers, the subjects were made to take a position if they agreed more or disagreed more rather than picking a middle choice if they felt uncertain what number that should represent their feelings. The graphical representations depict the average value of the answers from the inexperienced (group A) and inexperienced group (group B).

The survey questions can seem a bit biased but that is the point of the questions. It is crucial to investigate questions such as “I think gesture based interaction could be a fun way to interact with interfaces”, otherwise there is a risk of building a system or a product without analyzing if potential users/customers even want to use it. This aspect was not only important to investigate for the thesis’s sake but also interesting for Crunchfish to see if a product in this field might be commercially viable.

## Realtor Use Case

Before the answers of the surveys were analyzed, it was quite easy to predict that the experienced test participants might enjoy a gesture based interface since they work with gestural technology on a daily basis. However, when analyzing the results from the surveys, both groups answered very similarly to all questions, indicating that new and first time users like and enjoy gesture based HMI as well.

All test participants got the feeling that a graphical representation of the gestures would be a useful tool in the use case-prototypes they tested. Since the tests were designed in such a way that it was desired to observe how the test participants spontaneously performed certain gestures, per definition they did not know how to perform an instructed gesture, which in its turn might have lead to them feeling a bit lost due to the intended lack of tips and information the interface provided.

It was predicted that the grab and flip gesture was going to be tricky to understand and perform by the test participants and that prediction seemed to have been correct.

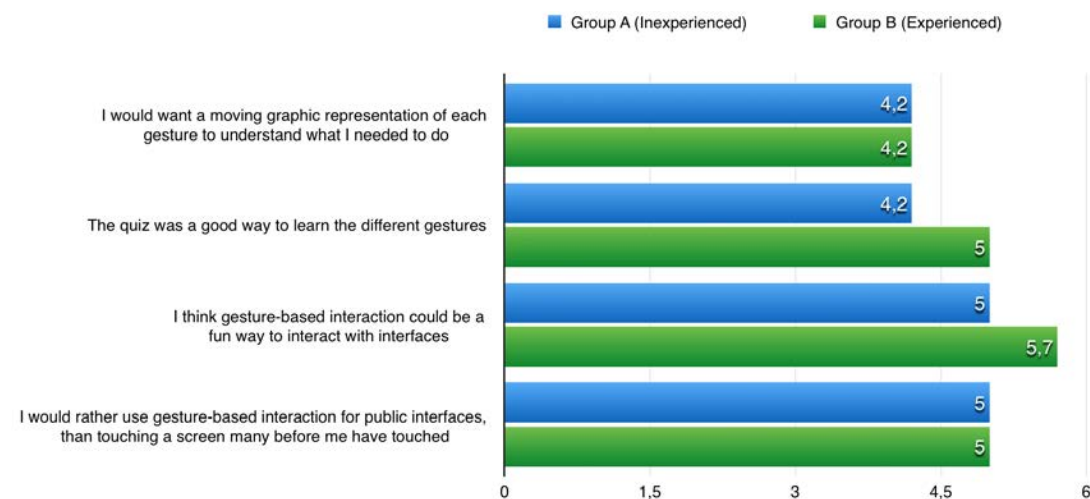


Figure 7.1: The picture illustrates the results of the quantitative data correlating to the Realtor use case's survey answers.

## Clothing Use Case

The clothing use case had almost the same answers regarding the need for graphical representations of the gestures as the Realtor use case. As mentioned before, the reason behind this result might be the tests designs regarding the desire to observe all the test participants individual spontaneous gestural behavior.

The rest of the survey questions answers were pretty straight forward, the test participants seemed to like gestural technology and the use cases.

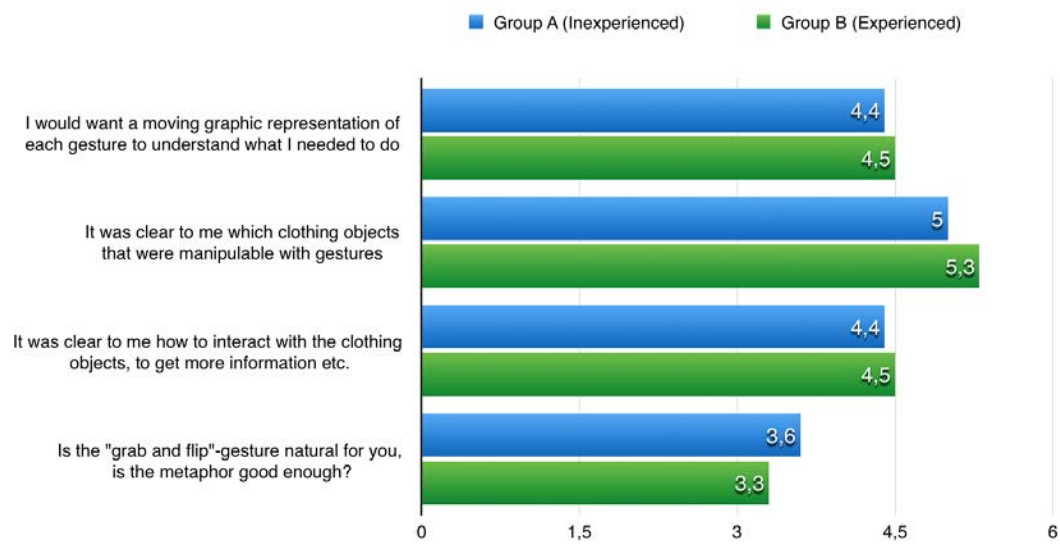


Figure 7.2: The picture illustrates the results of the quantitative data correlating to the Clothing use case's survey answers.

# Chapter 8

## Discussion

This chapter revolves around discussing the results and prototypes, related to the thesis goals. As a result of the fact that the thesis has been written in a process oriented way, some things that are discussed below has been mentioned in previous parts of the thesis in results-and evaluation sections.

### 8.1 General discussion of the results

#### 8.1.1 Test participants

Due to the limited time frame, the test participants were chosen from our group of friends that are tech savvy, as well as co-workers at Crunchfish who work with gestures every day. In retrospect, would the time have sufficed it would have been interesting to perform the usability tests on a group of people without engineering backgrounds, which would have then represented the inexperienced group. However, it turned out very well with the co-workers from Crunchfish as the experienced group. The reason for this is that the amount of people who have knowledge and work experience with gesture-based technology is small and it was interesting to pit a highly experienced group against an inexperienced group to observe and analyze their performances when doing the usability tests. Additionally, the fact that both the groups quantitative and qualitative results pointed at the fact that they liked this type of gesture-based interface indicated that new users as well as old and experienced ones both appreciate the use cases and the prototypes, which is a great result for this Master's thesis.

#### 8.1.2 Usability testing

The group that performed the best was clearly the experienced group, as a whole they made fewer mistakes than the inexperienced group. However, the performance times of the tests have been ignored even though one of the purposes of the usability tests was to time the test participants' performances. The reason why the test times have been ignored is the fact that a few test participants talked and was making jokes during the tests, which distorted the time results.

As predicted, the experienced group performed many of their gestures in a non-spontaneous way, they were prone to using a few of Crunchfish's gestures as well as generally a static way of interacting. However, when looking at the usability tests they still performed many of the gestures in a correct manner but it was clear that they tried to avoid a few pitfalls in gesture-based



technology that they know could exist in such a system.

The inexperienced group was not inhibited to the same extent as the experienced group, which meant that they could perform the tasks and gestures in a more spontaneous manner, which they also did to a larger extent in comparison to the experienced group. However, since they were inexperienced in the gestural technology field, they as well made a few mistakes during the usability tests.

Even though the transferal of learning effect was supposed to be balanced out to some extent by using the within subjects design, it was abundantly clear that all test participants, regardless of which test any subject began with, the performance in the second test was always better. The test participants explained in the semi-structured interviews that they were able to see patterns and similarities between the both use cases, which meant that having experience from a previous test immediately made the test participants perform better. This was not only true for the inexperienced group but also for the experienced group.

The example that follows depicts a scenario from the usability tests which is believed to be an issue or a hindrance for new gesture users, at least when using the Realtor use case in its current form. Some of the inexperienced test participants ignored the instructions presented on the screen in connection to the second Home matcher quiz question, the conclusion involves the theory that the test participants were becoming more confident with gesture interaction at this point in the process and unknowingly ignored the screen's gestural instructions as they felt that they already had mastered the previous gesture and intuitively used that gesture instead of paying attention to the new instructions presented on the screen. However, at the end of the tests the test participants had learned the different gestures as a result of the quiz.

An interesting example of how the test participants performed the grab and flip gesture was the fact that the test participants that had done the Clothing use case before the Realtor use case were prone to grab and twist their closed hand when wanting to perform the grab and flip gesture. The test participants that belonged to this group of people were asked why they did the grab and flip-gesture in the way they did, and they each said independent of each other that they had learned that most previous gestures had begun with a "grab" and naturally thought that the flip operation also should begin with a grab. This result was the desired one and it also proves the notion that more experience with gesture interaction corresponds to a better understanding regarding gesture-based technology. However, the other test participants that did the flip gesture with an open hand or in any other way, were not believed to be wrong, they simply mapped the instruction to a different motion.

In the Clothing use case when the test participants were given the task to put back the shirt from where they had grabbed and pulled it out from, the wanted result was for them to grab and push back the object to its original position. However, most of the test participants, inexperienced as well as experienced, tried to swipe away the clothing object to a corner of the screen. This result is quite interesting since the test participants had forgotten that the swipe gesture in this setting meant to switch pictures of the clothing item in question. However, since the Clothing use case was a mock-up all the functionalities had not been implemented, if they had been implemented the test participants would have received feedback that the swipe-gesture switched pictures of the clothing object hence making the test participants understand the fact that the swipe gesture was not meant for removing objects from the main view.

### 8.1.3 Prototype evaluations from a usability perspective

Some kind of graphical tip or representation of a gesture could be needed at a few points in both the prototypes. The graphical tip should appear when the user is performing a gesture or a movement that the camera system does not recognize. As a result of the prototypes being Keynote mock-ups, real feedback and other graphical tips that could help a user understand the product better was not in place. Even though the help tips that should be implemented in a finalized gesture-based interface were absent in the use case prototypes, the test participants performed well and the original expectations of the prototypes were met, both when looking at the qualitative data as well as the quantitative data. Some mistakes were made along the way by the test participants but the good news was that the users quickly learned and corrected their mistakes when performing a new gesture or when testing the secondary use case prototype.

### 8.1.4 Mid-Fi prototypes

The two Mid-Fi prototypes that represented the final product, which visualized the ideas and technology were not ultimate, as there was not enough time to alter them after the last user tests. Time and the lack of advanced enough gesture-recognition technology limited the outcome. Since the prototypes only implemented certain functionalities, which were necessary for the user tests, the prototypes could not be played with. This was a problem as the prototypes had to be developed by only trying a few gestures and functionalities, leaving the rest of the design elements passive, thus not testable to see if they were good from an interaction perspective or not. Naturally, we would have wanted to get feedback on the entire prototype, covering all the theoretical possibilities of the prototypes.

Owing to the fact that the prototypes were remote controlled, the outcome and preciseness of the gesture animations were impacted by human error. Sometimes, the Wi-Fi connection would lag, causing the animation to play after the gesture had been carried out. This detracted from the overall feeling and made it obvious that the prototypes were not actually registering the users' gestures. The actual design of the interfaces for the Mid-Fi prototypes was simple, but conducted the elements necessary to test the gesture interaction. Would there had been another iteration, the images in e.g. the Clothing use case would have been changed to images that were more suited to the test objectives of the user tests.

## 8.2 Industry relevance

Better prototypes could potentially have been created, would the usability testing have included a lot more people. The Realtor prototype was shown to different realtor firms, making them aware of touchless technology and the use case. The reactions received were very positive, the realtor firms never having heard of using gestures to attract attention to the displays and by displaying big, manipulable, objects potentially driving up property sales. As mentioned in the interview results, different kinds of stores were excited about the potential, which makes the product applicable to many different sectors. The thesis could be used as a basis on which to produce a Hi-Fi prototype, and ultimately a final product, given that the Mid-Fi prototypes were so well received.

Similar work has been produced by Shiratuddin and Wong (2011), where a gestural framework was created for architectural design in 3D. Creating designs requires a lot more precision than the actions performed in the use cases developed in this Master's thesis, however, it is interesting to note that the conclusions are similar regarding what type of gestures work. This Master's

thesis has stressed the importance of the gestures being natural, rather than created as in the case of Microsoft Kinect (Microsoft 2015a).

Malizia and Bellucci (2012) as well as Norman (2010) and their views on NUI's have been great influences when developing the gestures for this Master's thesis. Their thoughts on NUI's and the importance of making the gestures feel spontaneous and natural to use for anybody were crucial in the development phase. It seems like this goal was met as the previously mentioned inexperienced test group caught on to the gestures after just a couple of minutes of testing the rather simple prototypes and could utilize the transferal of learning effect as an advantage.

Palacios and Romano (2008) cited in Shiratuddin and Wong (2011) have developed 14 gestures, some of which are used in the architectural design framework. The gestures focus on direct manipulating objects, similar to the gestures described in this thesis project, the difference being the metaphors and approach to gestures. Since this thesis revolves around gestures in a shopping context, the metaphors were gathered from a shopping point of view. As an example, take the grab and pull gesture (see 5.2) where the metaphor of grabbing an item of a shelf and placing it in the shopping cart is used. The gestures used by Shiratuddin and Wong (2011) are naturally focused on 3D design, making shopping metaphors unsuitable.

Possible improvements in the field of gestural interaction could be made by user testing different sets of gestures, e.g. the gestures developed for this Master's thesis, in different scenarios and use cases. This way, the gestures and the metaphors they are based on could be analyzed outside the thought area of application, thus determining a gesture's suitability in a general application. However, taking a gesture out of its context might make it unnatural, thus contradicting our results that gestures must be natural in order to work. Perhaps the only way to make gesture interaction truly natural is to have different gestures for different contexts, rather than trying to create a "one fits all" type of framework. Shiratuddin and Wong (2011) mention that extensive usability testing was not carried out on the gestural framework, leaving out a final verdict on the suitability of the developed gestures. Furthermore it is mentioned that more studies would have to be made, subsequently creating a standardized library for gesture interaction. Drawing from the experience gained throughout this Master's thesis, it is the belief of the authors that such a standardized library might not be viable, at least not as a natural user interface.

### **8.3 Continued work**

In order for the use cases to be implemented into fully functional products, technology advances are necessary. During the thesis, it was assumed that the gesture recognition never failed, that the camera could detect all the users' faces and that the information transfer from screen to phone would work seamlessly, without delays or hick-ups. Moreover, to properly study the effects of online and physical shopping, more detailed studies would have to be performed and analyzed.

# Chapter 9

## Conclusions

The purpose of this chapter is to conclude the work carried out for the Master's thesis.

After having conducted the extensive theoretical background studies, it was made clear how the development of the touchless prototypes should be done. All research that had been done pointed towards the fact that a touchless gesture-based interface needed to be designed in such a way that a user's gestures had to be as spontaneous as possible.

### Conclusions regarding the problem statements

This section is meant to summarize the results and analysis with regards to the problem statements.

- *What are present day's Natural User Interface technologies strengths and limitations?*  
A strength in today's Natural User Interface technologies is that they provide a benchmark and a proof of concept of the possibilities in the field of gesture interaction.

However, when looking at today's technology and technical prerequisites, it is not really possible to build a fully spontaneous gesture-based NUI. Until the technology is ripe, research, mock-up designing, prototyping and usability testing like this thesis has covered, is a good way to get a head start in the building of gesture-based interfaces. When the technology has caught up, implementing such a system is made easier when the ground work regarding the usability aspects has been done.

- *What technical, as well as cognitive factors need to be taken into account when designing a gesture-based interface?*  
As mentioned in the previous section, today's technology that is used for handling gestures is quite underdeveloped and the gesturized applications are in many cases a bit too advanced for the interaction to work seamlessly. The key for today's gesture-based interaction is to keep things simple and not try to build features that do not really work. Users will only become irritated and lose faith in the product.

Another problem with most of today's gesture-based systems and products is that users are required to learn the gestures for the specific product they are using at the moment. For this reason it was crucial to develop both the use case's interfaces and gestures in such a way that the interaction felt natural and spontaneous for any user, experienced or inexperienced. However, for any new system a user is going to use, there is a required learning period, nobody is fully learned from birth. The objective was therefore to build the system so that the learning and understanding curve would grow exponentially with the time a user used the system. Even though there were some issues that were made clear during the usability testing, the results of the designed prototypes and their gestures at the end was positive.

- *Could gesture-based interactive shopping windows be a motivation for online shoppers to buy more items in physical stores?*

This question has to remain a bit unanswered since it is very hard to draw any certain conclusions at this point in time. Interactive shopping windows need to become a standard in stores for a while so that a comparison between online shopping and physical shopping can be made and clear results can point towards either direction. However, the surveys and interviews this Master thesis has conducted, point towards the fact that physical shopping might get an upswing if these interactive shopping windows are realized. Although it is believed that most people might return to physical stores in a larger extent if the prices become more competitive.

- *Are the built prototypes good representations of the use cases they implement and the final product they aspire to be the basis for?*

The results from the user tests pointed to the users being satisfied and having a generally positive view towards both prototypes, even though they were not perfectly made, the users enjoyed the idea of gesture-based shopping.

- *What kind of gestures are natural for users and how should they be applied in products that belong to the field of human machine interaction?*

When developing this particular problem statement, the theoretical background studies had yet to be done, meaning that it was believed that there were some gestures that were natural and some that were not. After finishing this thesis it can be concluded that this problem statement is a bit misleading. As mentioned many times already in the thesis, the only gestures that are actually natural, in the word's true meaning, are the gestures that each person and individual performs spontaneously, meaning that it is very difficult to design and build a perfect natural gesture-based interface. Nevertheless, if the interface and gestures are designed in such a way that a user feels that he or she is able to manipulate objects and make interaction decisions without feeling restricted, even though he or she follows some instructions, the interaction can be seen as relatively natural. It is always hard to design systems for every potential user's needs and specifications

# Bibliography

- Alderman, Naomi (2015). "HoloLens: Get Ready To Mix The Real And The Virtual In A Mind-Blowing New World". In: URL: <http://www.theguardian.com/technology/2015/feb/09/hololens-microsoft-virtual-world-augmented-reality-> (visited on 03/03/2015).
- Amato, Andrew (2013). "The Physical Limitations of the Oculus Rift with its 21-Year-Old Inventor". In: URL: <http://bostinno.streetwise.co/2013/11/04/the-physical-limitations-of-the-oculus-rift-with-its-21-year-old-inventor/> (visited on 05/22/2015).
- Apple (2015). *Siri*. URL: <https://www.apple.com/ios/siri/> (visited on 02/26/2015).
- Ballmer, Steve (2010). *CES 2010: A Transforming Trend – The Natural User Interface*. URL: [http://www.huffingtonpost.com/steve-ballmer/ces-2010-a-transforming-t\b\\\_416598.html](http://www.huffingtonpost.com/steve-ballmer/ces-2010-a-transforming-t\b\_416598.html) (visited on 02/26/2015).
- Bilton, Nick (2015). "Why Google Glass broke". In: URL: [http://www.nytimes.com/2015/02/05/style/why-google-glass-broke.html?\\_r=0](http://www.nytimes.com/2015/02/05/style/why-google-glass-broke.html?_r=0) (visited on 03/05/2015).
- Blénessy, Órs-Barna (2015). *Erghis Sphere - Simplifying Touchless Interaction*. URL: <https://www.indiegogo.com/projects/erghis-sphere-simplifying-touchless-interaction> (visited on 03/04/2015).
- Bonsor, Kevin (2001). "How Augmented Reality Works". In: URL: <http://computer.howstuffworks.com/augmented-reality.htm> (visited on 02/16/2015).
- Bryson, Steve (2013). *Virtual Reality: A Definition History - A Personal Essay*. URL: <http://arxiv.org/abs/1312.4322> (visited on 02/16/2015).
- Carmack, John (2013). "Kinect is "Fundamentally a Poor Interaction"; PS Move has "Fundamental Advantages"". In: URL: <http://www.dualshockers.com/2013/08/03/john-carmack-kinect-is-fundamentally-a-poor-interaction-ps-move-has-fundamental-advantages/> (visited on 03/05/2015).
- Cass, Stephen (2015). "Google Glass, HoloLens and the real future of Augmented Reality". In: URL: <http://spectrum.ieee.org/consumer-electronics/audiovideo/google-glass-hololens-and-the-real-future-of-augmented-reality> (visited on 03/04/2015).
- Christensson, Per (2012). *NUI*. URL: <http://techterms.com/definition/nui> (visited on 02/26/2015).
- Colgan, Alex (2015). *Leap Motion*. URL: <http://blog.leapmotion.com> (visited on 03/04/2015).
- Cook, Liz (2015). "Brainstorming - Generating Many Radical, Creative Ideas". In: URL: <http://www.mindtools.com/brainstm.html> (visited on 06/14/2015).
- Crawford, Stephanie (2010). *How Microsoft Kinect Works*. URL: <http://electronics.howstuffworks.com/microsoft-kinect.htm> (visited on 02/25/2015).
- Crunchbase (2014). *Leap Motion*. URL: <https://www.crunchbase.com/organization/leap-motion> (visited on 03/03/2015).

- Dredge, Stuart (2014). "Facebook closes its \$2bn Oculus Rift acquisition. What next?" In: URL: <http://www.theguardian.com/technology/2014/jul/22/facebook-oculus-rift-acquisition-virtual-reality> (visited on 02/24/2015).
- Erghis Technologies (2015). *Introducing The Sphere*. URL: [http://erghis.com/?page\\_id=641](http://erghis.com/?page_id=641) (visited on 03/04/2015).
- Evans, Clare (2014). "10 Examples of Augmented Reality in Retail". In: URL: <http://www.creativeguerrillamarketing.com/augmented-reality/10-examples-augmented-reality-retail/> (visited on 02/17/2015).
- Garber, Lee (2013). "Gestural Technology: Moving Interfaces in a New Direction". In: *Computer* 46.10, pp. 22–25.
- Gates, Bill (2011). *The Power of the Natural User Interface*. URL: <http://www.gatesnotes.com/About-Bill-Gates/The-Power-of-the-Natural-User-Interface> (visited on 02/25/2015).
- Gilbert, Ben (2015). "I experienced 'mixed reality' with Microsoft's holographic computer headset, 'HoloLens'". In: URL: <http://www.engadget.com/2015/01/21/microsoft-hololens-hands-on/> (visited on 03/04/2015).
- Gillies, Marco and Andrea Kleinsmith (2014). "Non-representational Interaction Design". In: *Contemporary Sensorimotor Theory*. Ed. by John Mark Bishop and Andrew Owen Martin. Vol. 15. Studies in Applied Philosophy, Epistemology and Rational Ethics. Springer International Publishing, pp. 201–208. URL: [http://dx.doi.org/10.1007/978-3-319-05107-9\\_14](http://dx.doi.org/10.1007/978-3-319-05107-9_14) (visited on 03/05/2015).
- Goldman, David (2012). "Google unveils 'Project Glass' virtual reality glasses". In: URL: [http://money.cnn.com/2012/04/04/technology/google-project-glass/?source=cnn\\_bin](http://money.cnn.com/2012/04/04/technology/google-project-glass/?source=cnn_bin) (visited on 03/05/2015).
- Google (2015). *Voice Actions*. URL: [https://support.google.com/glass/answer/3079305?hl=en&ref\\_topic=3063233&rd=1](https://support.google.com/glass/answer/3079305?hl=en&ref_topic=3063233&rd=1) (visited on 02/26/2015).
- Goth, Gregory (2011). "Brave NUI world". In: *ACM SIGSOFT Software Engineering Notes* 36.
- Graziano, Dan (2012). "Microsoft sells over 67 million Xbox 360s, 19 million Kinects". In: URL: <http://bgr.com/2012/05/30/xbox-360-kinect-sales/> (visited on 02/26/2015).
- Hearst, Marti A. (2011). "'Natural' search user interfaces". In: *Communications of the ACM* 54.11, p. 60.
- Helander, Martin G., Thomas K Landauer, and Prasad V. Prabh (1997). *Handbook of Human-Computer Interaction*. 2nd ed. North-Holland.
- Hewlett-Packard (2015). *Control apps with the wave of a hand*. URL: <http://www8.hp.com/us/en/ads/envy-leap-motion/overview.html> (visited on 03/04/2015).
- Hutchinson, Lee (2013). "Hands-on with the Leap Motion Controller: Cool, but frustrating as hell". In: URL: <http://arstechnica.com/gadgets/2013/07/hands-on-with-the-leap-motion-controller-cool-but-frustrating-as-hell/> (visited on 03/04/2015).
- Jack, Eric P. and Thomas L. Powers (2013). "Shopping Behaviour and Satisfaction Outcomes". In: *Journal of Marketing Management* 29.13-14.
- Juang, B H and Lawrence R Rabiner (2005). "Automatic Speech Recognition – A Brief History of the Technology Development". In: *Elsevier Encyclopedia of Language and Linguistics*.
- Kickstarter (2012). *Oculus Rift: Step into the Game*. URL: <https://www.kickstarter.com/projects/1523379957/oculus-rift-step-into-the-game/description> (visited on 02/23/2015).
- Kuchera, Ben (2015). "Sceptical of HoloLens? It's time to rewatch how Microsoft sold us the Kinect". In: URL: <http://www.polygon.com/2015/1/21/7868351/hololens-kinect-microsoft-windows-10> (visited on 02/25/2015).

- Leap Motion (2014). *Product*. URL: <https://www.leapmotion.com/product> (visited on 03/03/2015).
- LeClair, Dave (2014). "SMI announces eye-tracking upgrade for Oculus Rift". In: URL: <http://www.gizmag.com/eye-tracking-oculus-rift/34878/> (visited on 03/04/2015).
- Limer, Eric (2013). "Leap Motion Controller review: Waiting for the future to catch up". In: URL: <http://gizmodo.com/leap-motion-controller-review-waiting-for-the-future-t-817895912> (visited on 03/04/2015).
- Malizia, Alessio and Andrea Bellucci (2012). "The artificiality of natural user interfaces". In: *Communications of the ACM* 55, p. 36.
- Metz, Rachel (2013). "Look before you Leap Motion". In: URL: <http://www.technologyreview.com/news/517331/look-before-you-leap-motion/> (visited on 03/04/2015).
- Microsoft (2015a). *Common Kinect Navigation Gestures On The Xbox One*. URL: <http://support.xbox.com/en-US/xbox-one/kinect/common-gestures> (visited on 02/25/2015).
- (2015b). *Kinect for Windows*. URL: <http://www.microsoft.com/en-us/kinectforwindows/develop/Win8Compatibility.aspx> (visited on 02/25/2015).
- (2015c). *Kinect picture*. URL: <https://www.medialon.com/product/microsoft-kinect-for-windows-6-0-1/> (visited on 05/27/2015).
- (2015d). *Microsoft HoloLens*. URL: <http://www.microsoft.com/microsoft-hololens/en-us> (visited on 03/03/2015).
- (2015e). *The era of holographic computing is here*. URL: <http://www.microsoft.com/microsoft-hololens/en-us?video-url=vdeHHP> (visited on 03/03/2015).
- (2015f). *Use Kinect voice commands with Xbox One*. URL: <http://support.xbox.com/en-US/xbox-one/kinect/voice-commands> (visited on 02/26/2015).
- Moorhead, Patrick (2014). "Wearables Have A Long Way To Go To Be Mass Consumer Markets". In: URL: <http://www.forbes.com/sites/patrickmoorhead/2014/05/27/wearables-have-a-long-way-to-go-to-be-mass-consumer-devices/> (visited on 03/06/2015).
- (2015). "Microsoft HoloLens Gets Face Wearables Right". In: URL: <http://www.forbes.com/sites/patrickmoorhead/2015/01/28/microsoft-hololens-gets-faces-wearables-right/> (visited on 03/03/2015).
- Norman, Don (2010). *Natural User Interfaces Are Not Natural*. URL: [http://www.jnd.org/dn.mss/natural\\\_user\\\_interfa.html](http://www.jnd.org/dn.mss/natural\_user\_interfa.html) (visited on 03/05/2015).
- O'Brien, Terrence (2013). "Microsoft's new Kinect is official: larger field of view, HD camera, wake with voice". In: URL: <http://www.engadget.com/2013/05/21/microsofts-new-kinect-is-official/> (visited on 02/25/2015).
- Oculus VR (2015). *About Oculus*. URL: <https://www.oculus.com/company/> (visited on 02/23/2015).
- Pescarin, S. et al. (2013). "NICH: A preliminary theoretical study on natural interaction applied to cultural heritage contexts". In: *Digital Heritage International Congress (DigitalHeritage), 2013*. Vol. 1, pp. 355–362.
- Pheasant, Stephen and Christin M. Haslegrave (2006). *Anthropometry, Ergonomics And The Design of Work*. Boca Raton, FL: Taylor & Francis.
- Rempel, David, Matt J. Camilleri, and David L. Lee (2014). "The design of hand gestures for human-computer interaction: Lessons from sign language interpreters". In: *International Journal of Human Computer Studies* 72.10-11, pp. 728–735.
- Robertson, Adi and Michael Zelenko (2014). *VOICES FROM A VIRTUAL PAST*. URL: [http://www.theverge.com/a/virtual-reality/oral\\_history](http://www.theverge.com/a/virtual-reality/oral_history) (visited on 02/23/2015).
- Rubin, Peter (2014). "The Inside Story of Oculus Rift and How Virtual Reality Became Reality". In: URL: <http://www.wired.com/2014/05/oculus-rift-4/> (visited on 02/23/2015).



- Schnipper, Matthew (2014). "SEEING IS BELIEVING: THE STATE OF VIRTUAL REALITY". In: URL: <http://www.theverge.com/a/virtual-reality/intro> (visited on 02/23/2015).
- Shiratuddin, Mohd Fairuz and Kok Wai Wong (2011). "Non-Contact Multi-Hand Gestures Interaction Techniques for Architectural Design in a Virtual Environment". In: *Proceedings of the 5th International Conference on IT & Multimedia at UNITEN (ICIMU 2011) Malaysia*.
- Singh, Mona and Munindar P. Singh (2013). "Augmented reality interfaces". In: *IEEE Internet Computing* 17.6, pp. 66–70.
- Statista (2015a). *Annual B2C e-commerce sales in the United States from 2002 to 2013 (in billion U.S. dollars)*. URL: <http://www.statista.com/statistics/271449/annual-b2c-e-commerce-sales-in-the-united-states/> (visited on 03/31/2015).
- (2015b). *Online retail expenditure in the United Kingdom (UK) from 2000 to 2013 (in billion GBP)*. URL: <http://www.statista.com/statistics/283165/online-retail-expenditure-in-the-united-kingdom-uk/> (visited on 03/31/2015).
- Stevens, Tim (2010). "Microsoft Kinect Gets Official, Video Chat Announced". In: URL: <http://www.engadget.com/2010/06/13/microsoft-kinect-gets-official/> (visited on 02/25/2015).
- Swider, Matt (2015). "Google Glass Review". In: URL: <http://www.techradar.com/reviews/gadgets/google-glass-1152283/review> (visited on 03/03/2015).
- Szabó, Katalin (1995). *Metaphors and the user interface*. URL: <http://www.katalinszabo.com/metaphor.htm> (visited on 03/05/2015).
- Terdiman, Daniel (2013). "HP embeds Leap Motion gesture control tech in 11 computers". In: URL: <http://www.cnet.com/news/hp-embeds-leap-motion-gesture-control-tech-in-11-computers/> (visited on 03/04/2015).
- Total Immersion (2015). "The Future Of Augmented Reality". In: URL: <http://www.t-immersion.com/augmented-reality/future-vision> (visited on 02/22/2015).
- VJs Mag (2015). *Leap Motion*. URL: <http://www.vjsmag.com/shop/leap-motion/> (visited on 05/27/2015).
- Wikipedia (2015a). *Augmented Reality*. URL: [http://en.wikipedia.org/wiki/Augmented\\_reality](http://en.wikipedia.org/wiki/Augmented_reality) (visited on 02/16/2015).
- (2015b). *Google Glass*. URL: [http://en.wikipedia.org/wiki/Google\\_Glass](http://en.wikipedia.org/wiki/Google_Glass) (visited on 03/05/2015).
- (2015c). *Leap Motion*. URL: [http://en.wikipedia.org/wiki/Leap\\_Motion](http://en.wikipedia.org/wiki/Leap_Motion) (visited on 03/02/2015).
- (2015d). *Oculus VR*. URL: [http://en.wikipedia.org/wiki/Oculus\\_VR](http://en.wikipedia.org/wiki/Oculus_VR) (visited on 02/23/2015).
- (2015e). *Virtual Reality*. URL: [http://en.wikipedia.org/wiki/Virtual\\_reality](http://en.wikipedia.org/wiki/Virtual_reality) (visited on 02/22/2015).
- Wirth, Werner et al. (2007). "A Process Model of the Formation of Spatial Presence Experiences". In: *Media Psychology*. Vol. 9. 3, pp. 493–525.
- Yao, Yuan and Yun Fu (2014). "Contour Model-Based Hand-Gesture Recognition Using the Kinect Sensor". In: *IEEE Transaction on Circuits and Systems for Video Technology* 24.11, pp. 1935–1944.
- Yarm, Mark (2014). "7 Ways the Oculus Rift Could Change Entertainment as We Know It". In: URL: <http://www.rollingstone.com/culture/news/7-ways-the-oculus-rift-could-change-entertainment-as-we-know-it-20140415> (visited on 05/22/2015).
- Yee, Alaina (2015). "MICROSOFT'S HOLOLENS HEADSET FEELS LIKE SCIENCE-FICTION COME TO LIFE". In: URL: <http://www.ign.com/articles/2015/01/24/microsofts->

[hololens-headset-feels-like-science-fiction-come-to-life?page=1](#) (visited on 03/03/2015).

# Chapter 10

## Glossary

aBubbl: Crunchfish AB have an application called aBubbl that connects users' devices, which lets information be shared between the devices.

AR (Augmented Reality): A live view or feed of the physical-and real world is presented but with elements that are augmented or supplemented by a computer-generated sensory input, e.g. sound, video, graphics or GPS data.

Generation Y: Refers to people born roughly between the late 1980s and the mid 1990s.

Gesture recognition: In computer science, physical gestures recognized by a computer algorithm.

HCI: Human-Computer-Interaction. HCI refers to how a person/user interacts with a computer through a interface.

Instructions: The general instructions given by the interface.

Interface: The visual representation of the Mid-Fi prototypes.

Hi-Fi: High Fidelity, a highly functional and nearly finished final product.

HMD: Head mounted display, a display used in e.g. virtual reality.

HMI:Human-Machine-Interface. The goal of this interaction type is for the human to efficiently control and operate the machine he or she is interacting with, whilst the machine, simultaneously, gives the user informative and high quality feedback aiding the users decision making process.

Holography and holograms : Holography is a technique used to project three-dimensional images, so called holograms.

Home matcher: The quiz game in the Realtor use case.

Humanoid: Human-like.

Lo-Fi: Low fidelity, a prototype with no real functionality, e.g. paper and drawings of a system.

Mid-Fi: Medium fidelity, a prototype with some functionality, not fully implemented but still has the appearance that some sort of implementation has been done.

Mockup: The way the prototypes were implemented/presented (i.e. paper mockups, digital mockups, Keynote mockups etc.).

Multi-touch interaction: Multi touch is a term used to describe a surface's ability to know and recognize the presence of more than one or two points of contact.

NUI: Natural User Interface. A user interface that appears natural to the users.

Objective: The thing that the user is supposed to do at a certain stage of a usability test.

Prototype: The different "products" created in the during the thesis project, Lo-Fi and Mid-Fi.

Touchless: Technology that uses non-physical input (e.g. gestures), rather than physical input such as touch screen as the interaction medium.

Transfer of learning effect: If a test subject performs usability tests and is able to perform better with time thanks to experience.

UI (User interface): A graphical representation of an interface with which the user can interact.

VR (Virtual Reality): Virtual reality is an environment simulated through a computer, which is capable to stimulate physical presence in places modelled from the real world or invented and imagined worlds.

# Appendix A

## Personas and Scenarios

### A.1 Clothing use case

#### Persona

Name: Eric

Age: 23 years old

Status: Single, at the start of his career as a software developer at a startup company in Stockholm.

#### Scenario

The summer is getting closer and Eric wants to update his wardrobe. In his mind he has a few items he knows he wants to buy, when he passes the Gant store in the town where he lives, he sees their new interactive shopping window and becomes intrigued and curious.

He walks towards the shopping window and stops in front of it. The shopping window's camera, which is facing outwards, registers his face and presents a welcome screen. Since Eric has his aBubble activated on his phone, the shopping window creates a connection to Eric's phone. The welcome screen then tells Eric to wave his hand to start the interaction. Eric waves his hand and the screen proceeds to give him two choices, one alternative is to browse the men's clothing department and the other alternative is to browse the women's clothing department. The screen tells him to use his palm and a push gesture to select what department he wants to browse. Eric uses the push gesture to select the men's department.

When the men's department has been selected the screen switches to a slideshow with only male models wearing clothes that a customer can buy in the physical store. On this slideshow Eric sees the shoes he wants and the screen comes with a tip saying "grab the product you are interested in and pull it straight out, or swipe to change slideshow". Eric follows the instructions and grabs the shoes and pulls them out. The shoes turn into an information page that covers nearly the entire screen. The information screen consists of a few pictures of the product, which sizes remain and the price. On the bottom of the screen there are pictures of products similar to the chosen one. Eric sees another tip given by the screen, a small animated hand that grabs the air and turns itself with a text saying "grab and turn your hand to receive more information". Eric is intrigued by the new gestural opportunity and follows the instructions. The information

screen flips and reveals its backside with more detailed information about the product.

After Eric has read the detailed information he pulls the whole shoe object to his phone zone and proceeds to enter the physical store to buy the shoes.

## A.2 Realtor use case

### Persona

Name: Michael

Age: 26 years old

Status: Single and working at an investment bank in Stockholm. Looking for a new bigger two-bedroom apartment to move to from his one-bedroom apartment.

### Scenario

Michael just got a promotion and since he's been thinking of moving to a bigger apartment for a while, he thinks this would be a good time to start looking.

Michael passes by the realtor on his way home from work. As he's passing by the windows he notices pictures of properties on display, one apartment in particular catches his eye. He stops, and looks at the image on the screen. As he's standing there, a message appears on the screen: "Hi there! Wave your hand to begin the search for your dream home". Michael, waves his hand and is greeted by a new screen with two alternatives: The Home Matcher or continue to browse. The Home Matcher alternative consists of two images, one is of a big villa outside Juan-les-Pins, France, the other is of a house in a small Swedish village. Michael is intrigued by the images displayed by the Home Matcher alternative, and thus he pushes his hand towards the screen, where the home matcher alternative is located.

A new screen appears with two images, one of wine bottles, the other a bottle and glass of beer. There is a text message asking Michael to select the item that best fits him by grabbing and dragging it to a drag zone on the bottom of the screen. Michael, being a wine lover, selects the wine. He grabs the image by closing his hand, while it is hovering above the image. The image bounces a bit, denoting that it has been selected. Michael then drags the image down to the drag zone and releases. Subsequently, a new screen appears.

The second screen is similar to the first screen, but instead shows images of a beach with palm trees and a ski resort respectively. Instead of grabbing and dragging, the text message states that Michael should grab and flip the image that best fits him. Michael grabs the beach image, like in the first screen, and flips it over. A third screen appears with two images, the left one of a hamburger and the right one of a plate of sushi. The third screen entails swiping in the direction of the image that Michael feels suits him best. Michael, being a fan of a good burger, swipes to the left to select the burger.

After the three screens, a screen with four images of four different properties appear, which are based on the images Michael selected in the previous quiz. A message pops up stating that the objects can be examined further by grabbing and pulling them out, as well as a message about dragging objects to a phone icon in order to save them to his mobile phone. The first object in the top left corner instantly sparks Michael's interest, a two-bedroom apartment in Vasastan, Stockholm. Michael notices that the object seem to hover slightly above the background. He

figures that he should use the gestures that he was taught earlier in the Home Matcher, and grabs the first object. The object bounces and becomes a bit larger, Michael pulls his hand towards him and the object enlarges even more. He pulls his hand all the way towards him and the object now covers a large part of the screen. He can see that the size, price, and address all fit what he's looking for. Michael uses the "grab and flip"-gesture to flip the image and as the object is flipped over, more images and information is displayed. Michael notices that the apartment has an elevator and was built in 1899. Content with what he has seen, Michael wonders how he can get in contact with the realtor. The realtor's contact information is on the back side of the object, but it is after closing hours and Michael is really eager to get in touch with the realtor ASAP. Michael remembered the information before about saving object to his phone, so Michael flips back the object, looks at the price again, and then drags the object to the phone icon. He feels his phone vibrating in his pocket, picks it up and sees a message telling him that he has saved the object and that the realtor will contact him for more details. There is also an option to view the object from home, on his phone or computer. Perfect, Michael thinks because then he can show it to his mom and dad later as they are coming over for dinner.

# Appendix B

## Test objectives

### Test objectives for the Clothing use case

1. Follow the instructions given on the screen when facial recognition has been completed.
2. Select the men's department.
3. Read the instructions presented on the screen.
4. View the next inspirational image to the left, twice.
5. Find out more about the item on sale.
6. See if you can find another image of the shirt.
7. View the red shirt located in the suggestion bar.
8. Add the red shirt to your shopping cart.
9. Put the shirt back and count the number of items in the inspirational image that are viewable.



**Test objectives for the Realtor use case**

1. Follow the instructions given on the screen when facial recognition has been completed.
2. Use the instructed gesture to select the Home matcher-quiz.
3. Choose the wine option by using the gesture instructed on the screen.
4. Choose the beach option using the gesture instructed on the screen.
5. Choose the burger option using the gesture instructed on the screen.
6. Read the instructions presented on the screen.
7. Find out more information about the property located in Nice, France.
8. Find out detailed information about the property.
9. Save the property to your phone.
10. Put back the property object to its original position.