Shop Window Control

Bibiana Bayer, Kerstin Blumenstein, Thomas Ederer, Stefanie Größbacher, David Mayerhuber, Sabrina Rockenschaub, Grischa Schmiedl, Carina Skladal

Fachhochschule St. Pölten

dm131526@fhstp.ac.at

Abstract

The research of this paper deals with digital shop windows and their clients interacting with it. The glass pane between user and system prevents direct contact with the screen and thus forces the user to actively interact via gesture control with the system. With help of a self-developed prototype of a gesture driven digital shop window control system, a usability-test was realized to test different gestures and their effect on clients.

1 Introduction

A typical shop window for real estates is cobbled with printed ads of houses or flats. People pass by without even recognising that there is something to look at. What if the use of new technologies could enhance this kind of shop windows? What technologies could be used? And most importantly, how can people interact with a system through a pane of glass?

ShopWindowControl was built to catch the attention of pedestrians. But that's just the beginning. The ShopWindowControl also gives pedestrians the chance of individually filter ads of a real estate agent to meet their own specific needs. For example, if someone wants to buy a flat for about € 100,000 in Vienna, they can choose not to get any ads for houses in another city. Thus, ShopWindowControl provides users a list of real estate ads that can be filtered. The users can see different details and pictures of every object in the list and are able to save interesting offers onto their own smartphone or tablet. And that should work with the usage of 3D gestures. Figure 1 shows one of the first concept pictures we made for ShopWindowControl. It should illustrate how we think modern shop windows for real estates could look like in the future.

Figure 1 Concept for ShopWindowControl

For tablet or smartphone interaction the usage of 2D gestures turned out to be the most effective. In case of touch-less interactions it seems that 3D interfaces are evolving. But until now there are only a few products like some of the games for Microsoft Xbox that were able to successfully hit the mass market (cf. Bigdelou et al. 2012).

In this work we explore different technical solutions for a 3D gesture based shop window controlling and how people interact with them.

2 Related Work

Touch-less interaction with devices has been a field of interest for decades, though, as mentioned in the introduction, only a hand full of products based on those researches have hit the mass market.

In the past few years a lot of attempts to make touch-less interaction possible have been conducted. The researches are based on different technologies eg. webcam-based attempts like in the works of Nguyen et al. (2011), Dardas & Petriu (2011) and Mazumdar et al. (2013). Then there are the

Leap-Motion-based attempts like the works of Potter et al. (2013) and Sutton (2013), and Kinect-based attempts like the work of Bigdelou et al. (2012).

It is important to know how people react to an interactive shop window and how they notice and learn the interaction with it. There are some works regarding this knowledge and covering common problems about it (cf. Ning et al. 2011; Müller et al. 2012; Ten Koppel et al. 2012).

Another problem was to decide the kinds of gesture, which should be used to control the interactive shop window. Therefore the work of Pagán (2012) and Perry et al. (2010) were explored and the most efficient gestures for our prototype chosen, which were simple gestures like swiping left and right or lifting one arm up.

3 Empirical Research

3.1 Implementation of the functional prototype

The prototype should resemble the prospective product as close as possible. It should already use the desired hard- and software as well as the programming architecture. One of the main testing points is to turn the attention towards all possible technical constraints of the hardware like solar radiation, penetrability of the glass, other pedestrians who could possibly disturb the recognition of gestures, and of course accuracy and coverage of the sensor.

The testing scenario was an interactive shop window for real estate agents. The screen should show various estates for passers-by to skim through, give the possibility to switch to a details page, where users could look at photos and a map of the surrounding area, and finally offer a way of transferring the desired estate to users' smartphone or e-mail.

The whole set-up took place in an open area with a lot of activity at the University of Applied Sciences in St. Pölten to simulate the crowded streets in front of shop windows and to show if pedestrians cause perturbation with the gesture recognition system.

3.2 Preliminary Decisions

To build a thorough foundation for implementing and building the first prototype, it was necessary to clarify questions such as which hardware to use for gesture recognition and which system or programming language should be used for displaying the front end.

The two most frequently used gesture recognition systems are the Microsoft Kinect and the Leap Motion. The research showed that the Microsoft Kinect would have the most potential to serve the desired purposes for various reasons. In Figure 2 you see pictures of our pre-test where we tested the Leap Motion and the Microsoft Kinect v1, which we used for our prototype.

Figure 2 Technical pre-tests

The most crucial decision aspect was the action range. The Microsoft Kinect has a much wider range from 0.4 to 4 meters whereas the Leap Motion only has a small sight angle directly above the hardware from only a few centimeters. The main feature that was used in our prototype was the skeletal tracking of the Microsoft Kinect. It enables a user to be represented as a number of joints without any required calibration. These joints are body parts such as head, neck, shoulders, arms and legs and represent their actual and real-time position by its 3D coordinates (cf. Zhang 2012). Knowing where these body parts are and how they move in space is crucial for implementing a gesture detection method. Figure 3 shows how the skeletal tracking of the Microsoft Kinect looks like.

To decrease development time, we implemented the frontend of our prototype with web technologies. The gesture detection now runs a service, which sends its data to the webpage via sockets.

Figure 3 Technical prototype

The prototype consists of one big screen, the Microsoft Kinect above of it and a PC, which runs the gesture recognition and the frontend – everything placed behind a display window. The screen shows adverts of estates in the neighborhood. The user is able to flick through the adverts, to retrieve additional details and to transfer information about one estate onto the user's smartphone.

Gestures

The prototype should be operated through 3D touch-free gestures that do not have restrictions to the *x*- and *y*-axis (cf. Pagán 2012). This means that there exists a wide range of possibilities of input and interaction for the user. But as the work of Perry et al. (2010) shows, the gestures carried out in public should not be too overt and over-dimensioned because passers-by may be too embarrassed to perform them. Hence, it was decided to use only simple, swipe-related gestures, which count to the most basic 3D gestures.

All in all, there are four different types of gestures implemented in the final prototype:

- swipe left with left or right hand to display next object
- swipe right with left or right hand to display previous object
- swipe up with both hands to show details
- swipe down with either both hands or the left or right hand to hide details.

Direction-giving images in form of silhouettes mimicking the gesture-to-use are placed throughout the user interface, in order to give the passers-by an idea of how to use the application and which gesture to perform as seen in figure 4, at the bottom of the interface.

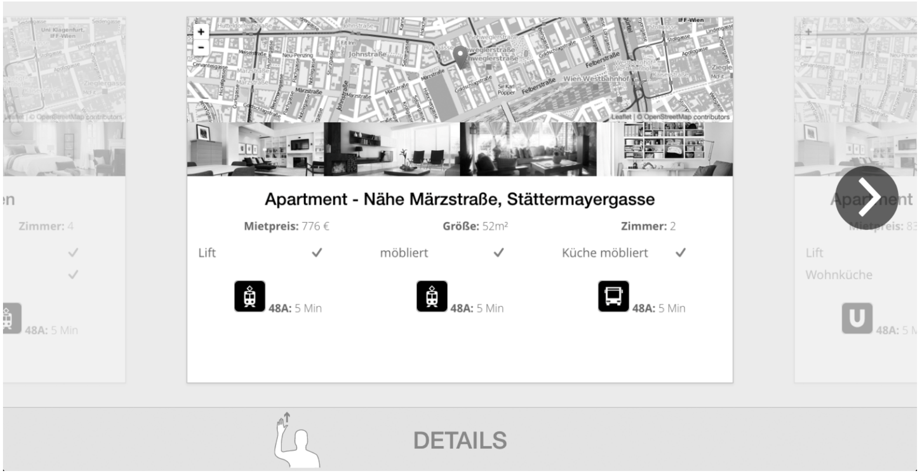

Figure 4 ShopWindowControl Interface

Technical implementation

The front-end is a simple designed webpage made with HTML5 and SASS (Syntactically Awesome Style Sheet). Its functionality has been built with AngularJS. The gesture detection runs separately inside a C# program, which streams the events to the webpage via sockets. Gestures are found by iterating over segments of different gestures and the system fires an event if one of them is completed. The webpage then handles the incoming event.

Qualitative empirical social research

To test the prototype, a qualitative empirical social research in the context of a laboratory analysis was conducted. The aim of the research was to analyze the efficiency and simplicity of gesture control, even with people that are not mandatorily technophile. Participants were between 21 and 30 years of age, from both genders and different social classes.

The probands were divided into two different groups consisting of six people each:

- The first group did not get any help or instructions on how to use or interact with the system, though they did know that the project was called “interactive shop window”.
- The second group got the information that the system is operated through gestures and also got some examples about what a gesture could look like.

Figure 5 shows the whole setup of the empirical social research. A probands stands in front of a huge monitor, which is equipped with a Microsoft Kinect performing a swipe gesture.

Figure 5 Proband during testing

Body of research

Every participant had to answer some demographic questions, such as age, hobbies and occupation. It was also necessary to ascertain if some of the probands had any technical know-how or, concerning the first test group, were familiar with gestures. Afterwards they had to solve a set of tasks corresponding to the predefined test manual. While solving the tasks, a testing director accompanied the user for supervising and observing purposes at all times. Additionally, the whole research was filmed and audio recorded to offer the possibility of better analyzing the research process afterwards. The test persons were asked questions about difficulty and problems they had during the completion of the tasks.

Tasks

The proband was given three tasks to fulfil. Everything of importance was documented for the analysis, like facts about needed help from one of our group members, the gestures the person tried, and if those gestures were appropriate to the task. For further development of our application, we also wanted to know about the user experience and therefore, we documented if people were feeling good with the given gestures and our prototype.

- Task 1: You are standing in front of an interactive shop window and you want to look through the displayed estates. Correct gesture: swipe left/right
- Task 2: You found an estate that you are interested in. Select it and display its details. Correct gesture: swipe up/down
- Task 3: Have a look at the pictures of this estate. Correct gesture: swipe up/down

Statistic Analysis

As mentioned before, twelve people were tested, half of them female. The age of participants varied between 21 and 30 years.

Six people were given information about the application, while the second group had to fulfil the tasks without information about the gestures. The average time for one test in general was 20 minutes, there were no significant differences between the two groups.

Execution of task 1 was no problem for the group of people with information (second group) and the feedback about the gesture was good. People said the gesture is intuitive because it is similar to gestures on touch-devices. Half of the group without information (first group) immediately found out which gesture they had to use, two people tried touch-and-swipe first and only one person had to be given information about the gesture, because otherwise she wouldn't have been able to fulfil the task. They told us the gesture is easy to learn, but it might be easier if you do not have to strike out that much. Furthermore the displayed arrows confused them and were the reason, that they thought the display is a touch-table.

Task 2 was not a problem, either. The testers with information handed to them (second group) were able to make the right gesture at first try. The response was positive as they considered the gesture is easy to use, though some of the participants said that a long click could be useful for this kind of information-display.

Some of them said that the shown mouse cursor (which was displayed as a symbol of a hand) was affecting their way of thinking, because it reminded them of touch devices. In the information-lacking group there were two people who needed help as they didn't realise there is a symbol showing them how to execute the right gesture. Their feedback included the wish for more information about the gesture.

The third and last task seemed to be the most difficult of all of them. In both groups one person was totally confused about the task and about the gesture. Two people from the group with information (second group) wanted to execute a long click, in the other group (first group) two people needed help with the gesture. In general the people were confused because they did not recognize that there is more than one picture shown on that page. The mouse cursor was affecting them again and they told us that more feedback and information would be considered desirable.

4 Results

During our research, the development of the prototype and the user experience test we figured out some facts that are interesting for our future development of the application and are further described in the following paragraph.

4.1 User-Experience

The group of people that was given information about the application and the gestures (second group) did not really have problems with executing the tasks and finding out how to use the application, but in general we found they did not finish faster than the other group without information (first group). That group found easily that the application is executed by gesture recognition and not by touching on it.

The symbols showing the gesture to use was helping all of the users to finish tasks easily. The more symbols and help we showed them, the easier it was. Some of the users would not have been able to finish most of the tasks without those helping symbols, only the first gesture (swiping left and right) was intuitive for them in the first instance.

The mouse-cursor was confusing as it gave respondents the illusion that they are able to drag and drop elements.

4.2 Technical

Using the Microsoft Kinect is the best decision for this use case and for our prototype. Still the Microsoft Kinect has some problems as well. Especially the gesture recognition through a window was not working when we were testing at our university, most likely because of the fire protection glass, which blocks some of the infrared rays.

5 Discussion

We found that it doesn't matter to people if they know they have to use 3D gestures or not. We expected a time difference between group one and group two but there was none. It was enough for the probands to see illustrations that show which gestures they can perform. For example lifting an arm to view details of a real estate.

Another important thing is performance. If people make a gesture and nothing happens they get confused and think they did something wrong and start trying gestures randomly.

In the future, we will develop a user interface that shows the user at any time which gestures are possible through the help of illustrations. Additionally it should react more fluently to the users' input.

A potential obstacle for using the Microsoft Kinect was that it doesn't operate well in broad sunlight because of the use of structured infrared light to measure depth. With the release of the Xbox One, Microsoft also released the Kinect v2, which now uses a time-of-flight camera and is supposed to work better – even in outside environments. It features better recognition of users with more joints than its predecessor and can track a higher number of users simultaneously.

After adding these features, we will take this second prototype to a real shop window and test it under a natural environment with real pedestrians. This should help us to test it under nearly natural preconditions and get insights about the technical limitations, but also to get a broader excerpt of the future target group.

References

Bigdelou, A.; Benz, T.; Schwarz, L. & Navab, N. (2012): Simultaneous categorical and spatio-temporal 3D gestures using Kinect. In: *2012 IEEE Symposium on 3D User Interfaces (3DUI)* (pp. 53–60). doi:10.1109/3DUI.2012.6184184

Dardas, N. H. & Petriu, E. M. (2011): Hand gesture detection and recognition using principal component analysis. In: *2011 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA)* (pp. 1–6). doi:10.1109/CIMSA.2011.6059935

Mazumdar, D.; Talukdar, A. K. & Sarma, K. K. (2013): Gloved and free hand tracking based hand gesture recognition. In: *1st International Conference on Emer-*

ging Trends and Applications in Computer Science (ICETACS) (pp. 197–202). doi:10.1109/ICETACS.2013.6691422

Müller, J.; Walter, R.; Bailly, G.; Nischt, M. & Alt, F. (2012): Looking Glass: A Field Study on Noticing Interactivity of a Shop Window. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 297–306). New York, NY, USA: ACM. doi:10.1145/2207676.2207718

Nguyen, T. T. M.; Pham, N. H.; Dong, V. T.; Nguyen, V. S. & Tran, T. T. H. (2011): A fully automatic hand gesture recognition system for human-robot interaction (pp. 112–119). ACM Press. doi:10.1145/2069216.2069241

Ning, T.; Müller, J.; Wacharamanotham, C.; Alt, F.; Walter, R. & Borchers, J. (2011): No Need to Stop: Menu Techniques for Passing by Public Displays: http://www.gillesbailly.fr/publis/BAILLY_CHI11.pdf <23.10.2014>.

Pagán, B. (2012): New Design Practices for Touch-free Interactions: http://uxmag.com/articles/new-design-practices-for-touch-free-interactions <10.11.2013>.

Perry, M.; Beckett, S.; O'Hara, K. & Subramanian, S. (2010): WaveWindow: Public, Performative Gestural Interaction. In: *ACM International Conference on Interactive Tabletops and Surfaces* (pp. 109–112). New York, NY, USA: ACM. doi:10.1145/1936652.1936672

Potter, L. E.; Araullo, J. & Carter, L. (2013): The Leap Motion Controller: A View on Sign Language. In: *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration* (pp. 175–178). New York, NY, USA: ACM. doi:10.1145/2541016.2541072

Sutton, J. (2013): Air Painting with Corel Painter Freestyle and the Leap Motion Controller: A Revolutionary New Way to Paint! In: *ACM SIGGRAPH 2013 Studio Talks* (Article No. 21). New York, NY, USA: ACM. doi:10.1145/2503673.2503694

Ten Koppel, M.; Bailly, G.; Müller, J. & Walter, R. (2012): Chained Displays: Configurations of Public Displays Can Be Used to Influence Actor-, Audience-, and Passer-by Behavior. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 317–326). New York, NY, USA: ACM. doi:10.1145/2207676.2207720