Project Gesture

Deliverable 2
Background Research

Michael Fretz
michael.fretz@zhdk.ch

Mentors
Max Rheiner
Stefano Vanotti

Spring 2011
University of the Arts Zurich
Department Design, Interaction Design

IAD Interaction Design hdk

INTRODUCTION

„A gesture[..] is any physical movement that a digital system can sense and respond to without the aid of a traditional pointing device.[..]“ [1] stated Dan Saffer in his book Designing Gstural Interfaces.

Have you ever found yourself, waving in front of a water tap in a public restroom until you recognize, that there is no automatic turn on mechanism? Gestures in public restrooms for using the water tab or the toilet flush are very common. People got used to it and don‘t think about it anymore while using it. Therefore using gestures to control interactive things will be the next obvious thing to come.
We as designers should always ask, is a gesture the right thing for the desired function? Is the use of a gesture the best and easiest method and will most people understand it easily? During my research I saw many gesture interfaces which are too complicated or not logically to use for an average user.

RELATED WORK

GIUC: A Gesture Interface for Ubiquitous Computing

GIUC [2] is a vision based gesture interface for ubiquitous computing environments. It uses a normal webcam to track the users hand. GIUC is based on a tracking and recognition algorithm combined with a particle filter algorithm. So far it is built to reconize six predefined gestures, each based on approximately 800 pictures. It has been tested for indoor environments.
One big advantage of this system is, that it requires nothing but a normal webcam. The disadvantage in this system is, that at the time of development, it was only able to run at 15 frames per second. For fast applications like games this might be too slow.

Marker-less Gesture Based Interaction for Design Review Scenarios

This prototype [3] uses computer vision methods to analyse camera images from a stereo camera setup in order to track 3 dimensional objects. The user can use gestures to control a visualisation software. The study showed, that gestural interfaces have a potential to increase the users efficiency by exploiting a far wider range of actions to manipulate a system, compared to a traditional interface. Using 3 dimensional movements to control a 3 dimensional interface makes it easier for users to understand what they are doing. The prototype showed an example of such a 3 dimensional interface without physically touching any device. A usability test was run with 17 participants. A short task together with a description about the handling of the gestural interaction was handed to the participants. All of them solved the task successfully. The overall result of this prototype shows that using gestures in 3 dimensional systems can be very rewarding.
One of the biggest advantage in this system is, that it does not need expensive devices, two cameras sufficient. I see great potential in the use of a 3 dimensional system combined with a 3 dimensional interface. A disadvantage of such a system clearly is, that the users have to learn the gestures before they can use the system. Learning a gestural set which is not intuitive can be frustrating.

Using a Depth Camera as a Touch Sensor

Microsoft Research explored depth-sensing cameras to detect touch on a tabletop [4]. Using this technology instead of a capacitive touch screen has the advantage that the surface doesn’t need to be instrumented or flat. An additional feature is the possibility to track the arms and hands of the users. The technique they used to track the data was Microsoft Kinect. In this setup the camera is above the surface. Its easily possible to track hover status and body parts of the user. The performance of such a system is not as good as it would be with a capacitive display but it‘s still good enough for a big variety of useful applications.
An advantage of this technology is, that it doesn‘t matter what kind of shape the display surface has. The tracking information can be useful and offer a wide range of opportunities for new applications. I see a disadvantage in projecting from above to a surface. Projection over the hands and arms can be quite strange while using the application. If a person leans over the surface all information behind the body cannot be tracked or displayed, this can lead to problems.

Information Book Beijing Planning Exhibition Hall

In the interactive information area at the Beijing Planning Exhibition Hall they have a virtual book. The shape looks like a huge real book. A projector displays information on it. An infrared camera tracks movements over the book. If a page turn gesture is performed a new site will be displayed. Using such an obvious pattern makes it easy for people to understand. One problem I saw during my observation was that most people don‘t read the displayed information anymore. Most users just performed page turns a couple of times. This project is very similar to Microsoft’s Using a Depth Camera as a Touch Sensor but with a lot less functions and gestures. Because it can only do one gestures and nothing else it‘s very easy for people to understand.

[p1] Virtual Book at Beijing Planning Exhibition Hall

Analysis of Natural Gestures for Controlling Robot Teams on Multi-touch Tabletop Surfaces

In this project [5] its mainly about the natural gestures of user and what gestures they would use for certain tasks. In an optimal environment, a normal user should be able to interact with the interface quickly and naturally without explicit instructions. The paper aims to find the most natural gestures for controlling robot teams, regardless of detectability or input technology. In user tests they tried to find the most common gesture for certain tasks.

Using a user centred gesture design is in my opinion the best way to get an easy and usable interaction with the software. One disadvantage of such an approach is that some cultures and people have different ideas how an gesture interaction should work. Therefore its hardly possible to find a solution which fits for all peoples.

User-Defined Gestures for Surface Computing

User defined gestures in the context of surface computing were analyzed by Jacob O. Wobbrock [6]. 20 participants generated over 1080 gestures for tabletop devices such as Microsoft Surface. Unlike most other gesture interaction studies, this study used non-technical users and let them design the gestures. Using a user generated gesture set instead of a system engineer set can lead to more problems in recognizing them on the technical side but will help the user to easily pick it up and use it.

One of my critique point of this study is, that the users could not change a behaviour after moving on to the next one. It could be quite likely that some of the gestures would have been more suitable for other functions.

Using Hands and Feet to Navigate and Manipulate Spatial Data

This project [7] is about an application to manipulate spatial data using hands and feet. In the example they built a geographical information system based on NASA‘s world. In addition they evaluated the difference between hand and hand & feet gestures to control the application. The users had to solve geospatial tasks and rate the overall experience afterwards.

Using different input methods like in this example is an interesting approach. Especially with a controller like the Nintendo Wii Fit Balance Board. But this interesting thing limits the usability too. Users need to be able to stand. For handicapped people this can be rather difficult or impossible.

g-stalt: a Chirocentric, Spatiotemporal, and Telekinetic Gestural Interface

The g-stalt project [8] is a 3 dimensional graphical space filled with over 60 cartoons. This movies can be viewed and rearranged using gestures. The system is a marker based system which tracks points on a glove. The software allows the user to navigate in a 3 dimensional graphical environment filled with video material. The videos and visual interface are projected to a large screen. While building the gesture set, they had a focus on real world gestures for certain behaviours. Whenever possible they used such gestures which led them to many gestures.

The main problem was the time to learn which gesture is responsible for which function. Using complex gestures for the control can be frustrating or disappointing.

Worlds of Information: Designing for Engagement at a Public Multi-touch Display

The project [9] is about an engaging multi user and multi touch display in public space. The focus of this project is engagement and group use of such a system. One of the problem they faced was the use of 3D with a multi touch display. Building gestures for a touch display to manipulate a 3 dimensional object was rather difficult and not all users did understand that.

Building a system with a complex abstract layer can make it difficult for inexperienced users. On the other hand having a 3 dimensional interface can engage new users. One thing, I really liked at this project is, that they build parts of the system and tested it on an exhibition. With the user feedback they got, they could improve the overall user experience.

Gestural Entertainment Center for Canesta

Kicker Studios built in 2008 an gesture controlled interface for an entertainment system [10]. They used a user centred approach and tried different gestures with participants. During the development process they looked for similar gesture patterns to reduce the size of the gesture language which users had to learn. In the design process they found out, that a Minority Report like interface is very tiring and to dramatic. The final interaction with the system focused on a easy to learn interaction using only a small set of gestures to control the system.

Gesture Space at ETH Library

In 2010, Kai Jauslin built for the ETH library a gesture based application [11] to display historic resources held at the library. It uses intuitive gestures to control the content. For his bachelor project at the university of the arts in Zurich he used to project all information on the floor. In his work for the ETH library he changed to a wall projection. The part I really like of his project is, that he did not focus on many gestures. Most gestures he used were mainly simple and could get adapted easily by a wide variety of people.

[p2] Gesturespace at ETH library

Microsoft Kinect for Xbox 360

Microsoft Kinect is a RGB camera combined with a infrared camera for depth sensing to interact with games and entertainment system without a classic controller. [12] It was developed by Microsoft for the Xbox 360 and is mainly used as game controller. To start the different games Kinect developed a gesture based interface and added some helpful tools. To start the tracking a wave gesture needs to be performed. This is indicated with a small animation. A small display always shows what the camera sees. If hands are detected they colour it on the small display and show a cursor on the screen. Nearly all information and actions are displayed using roll-over. There is no touch or press function to start a function. All "click" events will be started using a time based interaction. There is no push function for buttons integrated. All buttons are "magnetic" to make it easy for the user to activate it. For functions like pause a special gesture needs to be performed. This gesture is one of the few gestures which is not obvious but Microsoft explains it quite clear during start-up of the system. A swipe gesture is not as one would expect. To swipe to other content a buttons needs to be rolled over. If the hand is on rollover status, arrows indicate on which direction a movement of the hand will perform a change of the content. The usage of this way to perform a swipe gesture makes the system look quite slow. To see where the users position or hands are, a virtual avatar is displayed in the background of the interface. Every time a gesture is made, an icon displays it. During playing the games new interactions and gestures are possible, for example jump and movement of the hole body.
I think Microsoft Kinect Interface is quite a good example. There are many good points in it and they didn't use to many gestures to control the system. This makes it easy for a wide range of people to use. A drawback I see, is the missing click function. Not having this and only relay on time based activates makes this system look slow.

Easy Authoring for the Microsoft Kinect with Open Exhibits

The software [13] from Open Exhibits provides simple solution for gesture and flash based application. In their demo they showed different interactions using one or two hands. In one example they controlled a 360 degrees image using simple gestures. One hand is used to pan, two hands are used to zoom in and out. In another example they controlled Google Maps. The interaction and gestures for it are the same.
What I really like at this project is, that they provide an easy to use gesture set. Like the gesture Space project the gestures are not complicated to learn and people will pick it up quite easily.

Controlling PowerPoint Presentations With Kinect

Rafael Augusto Bassan has created a Microsoft Kinect controlled Application [15] to control PowerPoint. Simple gestures will change the slides. The only gesture he used was a swipe gesture. This makes it easy for everybody to control.

Gesture-based Fine Manipulation of a Surgical Tool using Kinect

The project [14] is a prototype for a gestural based surgical tool. It uses gestures to control a robot.
This could be a way how in the future surgeons can be performed over distance. The current state of the project is not yet as good that it could be used for a real surgeon. But I think it clearly shows the way how health care can be in the future. The gestures they used in this project are mostly logical. If the users one to grab something he has just to close the hand like we would do in real life. Using obvious gestures like that will make it easy for new user to adapt.

Moscow interactive shop window with gesture controls

VIVID Interactive produced an interactive shop window [16]. They used two basic gestures, swipe with one hand and pinch and spread with two hands. For pressing buttons they used the same system like Microsoft's Xbox. Rolling over a button and wait till a certain time is over to dispatch a click. In their example it seems very fast. One problem I see with such systems is that the content does not get as much attention similar to the Beijing City Planning Exhibition. People are more interested in the system than the content.

Kinemote Project - Kinect controlls Boxee

This project [17] uses a very basic gesture set to control a media centre. All gestures can be performed using only one hand. The system seems to be really fast. With a simple movement gesture in every direction the menu can be used. There is a click gesture included to activate a button. The sound level is controlled using up and down gestures. This project is a good example how an easy interface with a Kinect control can be developed. Because it uses only one hand it can be easily understand by a wide range of people. Compared to the official Microsoft Kinect controlled Xbox interface it seems to be very fast because they used click events instead of a time based button click.

[p3] Kinect controlled Boxee Media Centre

NATURAL USER INTERFACE

Natural User Interface are not natural

Currently it's all about Natural User Interface. Cheap technology make it possible to easily track gestures, body movements and speech. But behind all this buzzwords there is still a need of a good conceptual model, result and feedback. In GUIs there is a graphical guidance. All possible options are visible. Such a system can be easily learned through exploration. In Natural User Interface this is not always true. Different gestures apply for different cultures. Even though most common gestures like pinching and dragging are well known in different cultures there is still a problem with things like yes or no. Using a Natural User Interface with gestures must be learned and cannot easily be discovered. Physical gestures have other problems, using the hole body as an input can be difficult for handicapped people.
Gestural systems are not different from any other form of interaction design. They need to be well designed, have a good conceptual mode and an easy navigation through the application.

Natural Interface will be definitely play an important part in the future but it will need some time to develop for us to understand how to deploy them.
„Are natural user interface natural? No, But they will be useful." Dan Norman [18]

I agree with this paper. Some papers I read about gesture design are just too much focused on technology. Sometimes it seems that for some engineers it's not about the usefulness rather than the possibility. I think for designers it's very important to think about what makes sense and how people will use it.

Microsoft is Imagining a NUI future

Microsoft published on their blog a post [19] about their prediction for the future of natural user interface. Technology becomes more natural and intuitive. It's not only about multi-touch and speech sensors and technology. Future systems will combine different technologies and contextual awareness, 3D simulation and anticipatory learning. A future with an almost invisible technology and an easy interaction with such a system. Not only in the game industry will it become a major part, even in technology and health care will natural interfaces change a lot. One question about future development always comes up when predicting a futuristic scenario:"Is there a need, is the market ready for it, will it embrace?" In the field of natural user interface the answer is yes. Pools, ordered by Microsoft showed an huge interested in such technologies.

If we compare this blog post to the paper of Don Norman (Natural User Interface are not Natural) I think we can say that both papers predict a more natural user interface future. Where Microsoft is more focused on the technologies and possibilities, Don Norman predicts it a bit more abstract and critique.

GESTURES

In this section I would like to show the most common gestures currently used by systems. There are a few gestures which are very common and already established that we do not think about it anymore while using them. They became natural.

Tab or point to open, select or activate

Most touch based applications use the tab to send click events. It is a very common to open a function, select an item or activate it. Most touch screen mobile phone use this as one of the most used gesture. But if you look at controller free application like Microsoft Kinect this function is hardly used yet. „Pointing is the most natural gesture for selection." [20]

Drag and Drop

From GUIs on personal computers we know drag and drop functions. In a natural user interface such a function can be a very clear gestures as it can be directly transformed in how we move objects in real life.

Pinch to shrink and spread to enlarge

Since Apples iPhone and iPod Touch got really popular nearly everybody knows that the pinch or spread of fingers or hands gestures can change the size of an object. It got one of the most popular gestures without a direct transformation from real life.

Wave to activate

Waving is a simple gesture and has already a wide usage area. Not only is it needed to activate the user tracking on Microsoft's Kinect [21], it's also very common in public restrooms for the water tab, paper spender or the toilet flush.

CONCLUSION

During my research with gesture based applications I found that many applications are difficult to use. But all analyzed projects had a few interessting parts in it. For my bachelor project I want to build an easy to use and Kinect controlled application. I want to generate an easy to learn and user generated gesture set. What gesture makes sense and when.

Questions I would like to answer during this project:
- When does it make sense to use gesture?
- What would be an easy gesture set?
- What gesture will normal user use for common tasks?

LIST OF LITERATURE

1. Saffer, Dan (2008): Designing Gestural Interfaces. O'Reilly Media, Inc. ISBN 978-0-596-51839-4 Page: 2

2. GIUC: A Gesture Interface for Ubiquitous Computing 978-0-7695-3501-2/09 Dong Wang, School of Media Management, Communication University of China, Beijing 2009

3. Markerless Gesture Based Interaction for Design Review Scenarios 978-1-4244-4457-1 /09 Daniel Wickeroth, Paul Benölken, Ulrich Lang , Universität zu Köln

4. Using a Depth Camera as a Touch Sensor, Andrew D. Wilson, Microsoft Research, Redmond, WA 98052 USA, 978-1-4503-0399-6/10/11

5. Analysis of Natural Gestures for Controlling Robot Teams on Multi-touch Tabletop Surfaces, Mark Micire, Munjal Desai, Amanda Courtemanche, Katherine M. Tsui, and Holly A. Yanco University of Massachusetts Lowell, Department of Computer Science One University Avenue, Lowell MA 01854, USA {mmicire, mdesai, acourtem, ktsui, holly}@cs.uml.edu

6. Jacob O. Wobbrock
The Information School DUB Group University of Washington Seattle, WA 98195 USA wobbrock@u.washington.edu
Meredith Ringel Morris, Andrew D. Wilson
Microsoft Research
One Microsoft Way Redmond, WA 98052 USA {merrie, awilson}@microsoft.com

7. Spotlight on Works in Progress Session2 April 4-9, 2009 Boston, A, USA CHI2009
Johannes Schöning, Florian Daiber, Antonio Krüger, Michael Rohs

8. Jamie Zigelbaum, Alan Browning, Daniel Leithinger, Olivier Bau*, and Hiroshi Ishii
Tangible Media Group, MIT Media Lab Building E15, 20 Ames St. Cambridge, Mass. 02139 USA {zig, abrownin, daniell, ishii}@media.mit.edu
*InSitu, INRIA Saclay & LRI Building 490 Univ. Paris-Sud 91405 Orsay Cedex, France bau@lri.fr

9 Giulio Jacucci12, Ann Morrison1, Gabriela Richard3, Jari Kleimola1, Peter Peltonen1, 1 Lorenza Parisi4, Toni Laitinen1 Helsinki Institute for Information Technology HIIT, Aalto University, Finland, name.surname@hiit.fi 2School of Art and Design, Aalto University, Helsinki, Finland,
3Educational Communication and Technology, New York University, USA, name.surname@nyu.edu 4Facoltà di Scienze della Comunicazione, Sapienza Università di Roma, Italy

10. http://www.kickerstudio.com/canesta.html (call date: 19.02.2010)
http://www.youtube.com/watch?v=NwVBzx0LMNQ (call date: 19.02.2010)

11. http://www.library.ethz.ch/en/About-us/Projects/Gesturespace-ETH-Bibliothek
(call date: 16.02.2010)
http://www.nextension.com/eth-library/gesturespace2/
(call date: 16.02.2010)
http://www.nextension.com/zhdk/gesturespace/
(call date: 16.02.2010)
http://www.nextension.com/wp-content/uploads/2010/01/KJ_BA_Dokumentation_Web.pdf
(call date: 16.02.2010)

12. http://www.xbox.com/de-ch/kinect (call date: 16.02.2010)
http://en.wikipedia.org/wiki/Kinect (call date: 16.02.2010)

13. http://openexhibits.org/research/jims/123
(call date: 19.02.2010)
http://kinecthacks.net/easy-and-free-authoring-for-the-microsoft-kinect-with-open-exhibits/
(call date: 19.02.2010)

14. http://kinecthacks.net/gesture-based-fine-manipulation-of-a-surgical-tool-using-kinect/
(call date: 19.02.2010)

15. http://kinecthacks.net/controlling-powerpoint-presentations-with-kinect/
(call date: 19.02.2010)

16. http://www.kinect-hacks.com/kinect-news/2011/02/16/vivid-shopwindow-kinect-gesture-controls
(call date: 19.02.2010)

17. http://kinecthacks.net/kinect-controls-boxee/
(call date: 19.02.2010)

18. Don Norman http://www.jnd.org/dn.mss/natural_user_interfaces_are_not_natural.html
(call date: 19.02.2010)

19. http://blogs.technet.com/b/microsoft_blog/archive/2011/01/26/microsoft-is-imagining-a-nui-future-natural-user-interface.aspx

20. Saffer, Dan (2008): Designing Gestural Interfaces. O'Reilly Media, Inc. ISBN 978-0-596-51839-4 Page: 76

21. http://support.xbox.com/en-us/Pages/kinect/body-controller/default.aspx?step=walkthrough_content_2

LIST OF PICTURES

p1. persoanl photo, michael fretz

p2. http://www.nextension.com/eth-library/gesturespace2/

p3. http://www.geeky-gadgets.com/kinemote-uses-kinect-to-control-boxee-using-hand-gestures-2010-12-24/