Gestures in Human-Computer-Interaction

Sarah Muser

Abstract— Gesture control is very important in todays life when it comes to interacting with several devices, which can be seen as a positive step forward as this way of interaction is more natural than e.g. using a mouse.

This paper is meant as an introduction to gestures with regard to Human-Computer-Interaction(HCI). It provides a definition and different classifications of gestures (according to function and alternatively division into continuous and discrete) followed by a brief overview of some technologies that were important for the development of gestures in HCI. They are separated in those using multi-touch and those using mid-air gestures. Further it will enlarge shortly upon possible future fields of application for gestures in HCI.

Index Terms—multi-touch gestures, gesture-based interaction

✦

1 INTRODUCTION

The most commonly used input devices for computers are still mouse and keyboard, even though they restrict the possibilities of interaction with a computer to a minimum. But with further development and research attention is shifting to another way of Human-Computer-Interaction(HCI): Gesture control. On one side there are multi-touch gestures which are getting more and more established in the world market due to the development of smartphones and tablet-PCs. On the other hand there are mid-air gestures where there are already some application possibilities, however they haven't reached broad mainstream yet.

2 GESTURES IN EVERYDAY LIFE

To understand how gestures can be used in interfaces we will first have a look at what gestures are in general and how we use them in our everyday life. A definiton of gesture is provided by Kurtenbach and Hulteen(1990) [3, P. 1]:

"A gesture is a motion of the body that contains information. Waving goodbye is a gesture. Pressing a key on a keyboard is not a gesture because the motion of a finger on its way to hitting a key is neither observed nor significant. All that matters is which key was pressed".

There are plenty of gestures and they can differ a lot. They can be performed with objects, such as pointing at something or manipulating an object, or without, like waving or using sign language. Some may contain more information than others, some we are aware of when using and some we aren't.

3 CLASSIFICATIONS OF GESTURES

Gestures can be classified in different ways. The following are just a few classifications that are useful in consideration of multi-touch and mid-air gestures.

3.1 According to function

Cadoz proposes three hand gesture functions (1994) [17]:

- Ergotic: Gestures to modify and transform the environment. There's energy transfer in this action.
- Epistemic: Gestures to get knowledge from the environment. E.g. the information received by the tactile sense.
- Semiotic: Gesture to convey information to the environment.

- *Sarah Muser is currently studying Media Informatics at the Ludwig-Maximilians-University of Munich, Germany, E-mail: Sarah.Muser@campus.lmu.de*
- *This research paper was written for the Media Informatics Proseminar in 2015, supervised by Dr. Julie Wagner, Prof. Florian Alt and Prof. Andreas Butz.*

As semiotic gestures are especially important here because HCI is all about conveying and receiving information to and from the computer, we will further categorize them according to Rime and Schiaratura (1991) [3, P. 2]:

- Symbolic: Gestures that have a single meaning, such as the "OK" gesture and furthermore gestures of sign languages.
- Deictic: Gestures of pointing.
- Iconic: Gestures that are used to convey information about the size, shape or orientation of an object.
- Pantomimic: Gestures the speaker makes while mimicking a particular movement of some invisible tool or object in his hand.

McNeill (1992) adds beat gestures and cohesives to this taxonomy [3, P. 3].

- Beat gestures: The gestures we are using collaterally in the rhythm of the spoken. They convey just minimal or no extra information.
- Cohesives: They are used to physically tie different parts of a narrative together.

Naturally gestures are intimately related to speech. Depending on the information a gesture contains and if a gesture can be understood by itself or if it has to be in context with speech these gestures can be arranged. This arrangement is described in Kendons' Continuum [3, P. 3]:

$Gesticulation(Beat, Cohesive) \rightarrow Language - Like(Iconic) \rightarrow Pantomimes(Pantomimic) \rightarrow Emblems(Deictic) \rightarrow SignLanguage(Symbolic)$

From left to right the gestures contain more information and are therefore less dependent on speech.

3.2 Continuous and discrete

A possible classification especially for multi-touch gestures is the differentiation in continuous and discrete gestures [2, P. 3].

- Continuous: Special gestures that must be known by the user in advance. E.g. pinching and flicking gesture.
- Discrete: These gestures are self-revealing as usually additional graphical information is provided. E.g. tapping on a displayed button or on a virtual keyboard.

4 EVOLUTION OF GESTURE-DRIVEN DEVICES

As gestures are natural for us and computers are firmly integrated in our life, it appears obvious to have the ambition to interact with them via gestures. But the usage of gestures in HCI is limited by the available technologies in the first place. In the following two subsections are some picked technologies capable of gesture-sensing showing their development through time. For a more complete list of the evolution of multi-touch devices see [2] and of devices using mid-air gestures see [3].

4.1 Interactive Surfaces

The development of devices with possible touch input already started in the second half of the 1960s. As this subsection concentrates on devices using multi-touch gestures we start in 1982 when most likely the first multi-touch system for Human-Computer-Interaction was developed [2]:

- 1982: A Flexible Machine Interface (Nimish Meta, Universitiy of Toronto) [9] [2]
  - First multi-touch system for Human-Computer-Interaction working with a camera
  - Able to caption shadows of fingers and objects
- 1983: Videoplace - Table configuration (Myron Krueger) [2] [3]
  - Capable of tracking hands and fingers of multiple people
  - Number of different configurations, including table and wall
  - Touch-based from users' perspective in table configuration (not really because captured by camera)
  - Rich set of gestures
  - Very innovative
  - Influential in development of multi-touch gestures (e.g. pinch-to-zoom)
- 1984: Multi-Touch Screen (Bob Boie, Bell Labs, Murray Hill NJ) [2]
  - First multi-touch screen (not tablet)
  - Transparent capacitive array of touch sensors overlaid on a CRT
  - Manipulation of graphical objects
  - Excellent response time
- 1985: Multi-Touch Tablet (Input Research Group, University of Toronto) [2]
  - Capable of sensing an arbitrary number of simultaneous touch inputs
  - Captures location and degree of touch
  - Uses capacitance
  - Thinner and simpler than camera-based systems
- 1991: Digital Desk (Pierre Wellner, Rank Xerox EuroPARC, Cambridge) [2]
  - Front projection tablet top system
  - Uses optical and acoustic techniques
  - Senses hands and fingers, and certain objects (paper-based controls and data)
  - Possible gestures: two finger scaling and translation of graphical objects, using a pinching gesture or a finger from each hand, amongst others.
- 1998: Fingerworks (Newark, Delaware) [2]
  - Produced a number of multi-touch-able touch tablets, including the iGesture Pad
  - Many multi-point / multi-finger gestures supported
  - The company was acquired by Apple in 2005
- 1999: Portfolio Wall (Alias Wavefront,Toronto On, Canada) [2]
  - Digital cork-board for presenting images
  - Just single-touch but entirely based on finger touch gestures
  - Introduction of many gestures that we know today from touch-based mobile devices (e.g. flick-gesture)
- 2001: Diamond Touch (Mitsubishi Research Labs, Cambridge MA) [2]
  - Capable of distinguishig between different hands and fingers
  - Various and rich gestures possible
- 2004: Neonode N1 Mobile Phone (Stockholm, Sweden) [2]
  - First smartphone with finger swipe-type gestures, not just tapping
  - "Haptic" tactile-vibratory feedback
- 2007: Apple iPhone [2]
  - Smartphone with a soft touch-based interface
  - Provides gestures like pinch-to-zoom or flicking through images (from Alias Portfolio Wall)
  - Establishes smartphones on the "mainstream-market"
- 2007: Microsoft Surface Computing (later versions called Microsoft PixelSense) [2]
  - Interactive table surface
  - Senses multiple fingers, hands, and objects and their position on the surface
  - Brings the results of research and development to mainstream
- 2010: Apple iPad [12]
  - Accomplishes the success of tablet-computers

4.2 Devices using mid-air gestures

- 1983: Videoplace - Wall configuration(Myron Krueger) (also see section 4.1) [2] [3]
  - Real time image processing of live video of the user
  - Dynamic natural gestures (e.g. pointing), no training required
- 1984: The Media Room (Bolt and Negroponte's Architecture Machine Group) [3] [7]
  - Large room with a back projection panel on wall
  - Capable of sensing speech, gesture and eye movement (e.g. "Put-that-there"-gesture)
  - Pointing gestures measured through magnetic position sensing device on user's wrist
- 1987: The Dataglove (Zimmerman, Lanier, Blanchard, Bryson and Harvill)
  - First commercially available hand tracker
  - Fiber optic cables with cracks in it
  - When fingers bent light leaks through cracks
- 1989: The CyberGlove (Kramer)
  - Also measurement of sideway movement of fingers
  - More accurate bend sensing
- 1999: REA (Cassell, MIT Media Laboratory) [3] [5]
  - Computer generated humanoid
  - Senses the user through cameras and audio input
  - Capable of speech with intonation, facial display and gestural output
  - Rea asks and answers questions, allows to be interrupted, takes turn when possible
- 2010: Microsoft Kinect (Microsoft) [14] [10]
  - Senses mid-air gestures and speech
  - Meant to be a replacement of a game controller
- 2013: Leap Motion Controller (Leap Motion, Inc.) [8] [6] [15] [1]
  - Little sensor device using mid-air gestures, meant to be a replacement of the mouse
  - A Leap Motion-equipped tablet by MotionSavvy can interpret American Sign Language
- 2015: Microsoft HoloLens (Microsoft) [11]
  - Smart glasses containing a Windows 10 computer
  - Capable of gaze detection, voice recognition and sensing mid-air gestures
  - Creates an augmented reality environment

5 OUTLOOK

Also in the future many application fields are conceivable for gesture control. Multi-touchable surfaces will carry on spreading in almost all parts of our lifes. Especially the use of displays using multi-touch or mid-air gestures or even both in public places could be an important topic for the future as research in this field already takes place [13] [16]. Other interesting fields for mid-air gestures are virtual and augmented reality [4] which is already partly shown with the development of the Microsoft Hololens [11]. In the gaming industry virtual reality has a promising future and combined with the use of mid-air gestures gaming experience would become even more "real". The company Leap Motion is actually taking big steps towards this goal which currently can be seen on their website [8].

6 CONCLUSION

Research and development of gesture controls for over 30 years made it possible that we can now interact with devices in a more natural and intuitive way than in the past. Gesture controls already have many application possibilities and will play an even bigger role in our technnological future.

REFERENCES

[1] S. Buckley. Motionsavvy uses gesture recognition to give sign language a voice. `http://www.engadget.com/2014/05/17/motionsavvy-uses-gesture-recognition-to-give-sign-language-a-vo/`, 2014. visited 13.08.2015.

[2] B. Buxton. Multi-touch systems that i have known and loved. `http://www.billbuxton.com/multitouchOverview.html`, 2014. visited 13.08.2015.

[3] B. Buxton and M. Billinghurst. Gesture based interaction (haptic input draft, chapter 14). `http://www.billbuxton.com/input14.Gesture.pdf`, 2011. visited 09.08.2015.

[4] M. Elgan. Why in-air gestures failed, and why they'll soon win. `http://www.computerworld.com/article/2599359/emerging-technology/why-in-air-gestures-failed-and-why-theyll-soon-win.html`, 2014. visited 12.08.2015.

[5] J. C. et al. Embodiment in conversational interfaces: Rea. `http://www.ru.is/~hannes/publications/CHI1999.pdf`, 1999. visited 12.08.2015.

[6] T. Foster. Will these guys kill the computer interface as we know it?

[7] S. Huber. Entwicklung und untersuchung eines räumlichen interaktionskonzepts für die navigation in zoombaren benutzerschnittstellen, m.sc. thesis, universität konstanz, fachbereich informatik und informationswissenschaft. `http://hci.uni-konstanz.de/downloads/thesis_huber.pdf`, 2013. visited 11.08.2015.

[8] I. Leap Motion. `https://developer.leapmotion.com/gallery/category/vr`. visited 13.08.2015.

[9] Mehta and Nimish. A flexible machine interface, m.a.sc. thesis, department of electrical engineering, university of toronto supervised by professor k.c. smith., 1982. visited 09.08.2015.

[10] Microsoft. Kinect for windows. `https://www.microsoft.com/en-us/kinectforwindows/`. visited 13.08.2015.

[11] Microsoft. Microsoft hololens.

[12] S. Sai. History of touch interface. `http://www.cirolab.org/blog/wp-content/uploads/2010/06/History-of-Touch-Interface.pdf`, 2010. visited 10.08.2015.

[13] J. Schöning. Do we need further "multi-touch" touch affordances? `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.187.3213&rep=rep1&type=pdf`. visited 13.08.2015.

[14] T. Stevens. Kinect for windows sdk beta launches, wants pc users to get a move on. `http://www.engadget.com/2011/06/16/microsoft-launches-kinect-for-windows-sdk-beta-wants-pc-users-t/`, 2011. visited 13.08.2015.

[15] D. Terdiman. Leap motion: 3d hands-free motion control, unbound.

[16] R. Walter, G. Bailly, and J. Müller. Strikeapose: Revealing mid-air gestures on public displays. `http://joergmueller.info/pdf/CHI13WalterStrikeapose.pdf`, 2013. visited 10.08.2015.

[17] M. Wanderley. Gestural channel and the function of gesture. `http://recherche.ircam.fr/equipes/analyse-synthese/wanderle/Gestes/Externe/Sensor_Feedb/node6.html`, 1999. visited 09.08.2015.