Natural Gesture Interaction with Accelerometer-Based Devices in Ambient Assisted Environments

Kawa Nazemi, Dirk Burkhardt, Christian Stab, Matthias Breyer, Reiner Wichert, and Dieter W. Fellner

Fraunhofer Institute for Computer Graphics Research, Fraunhoferstraße 5, 64283 Darmstadt, Germany
{kawa.nazemi,dirk.burkhardt,christian.stab,matthias.breyer, reiner.wichert,d.fellner}@igd.fraunhofer.de

Abstract. Using modern interaction methods and devices enables a more natural and intuitive interaction. Currently, only mobile phones and game consoles which are supporting such gesture-based interactions have good payment-rates. This comes along, that such devices will be bought not only by the traditional technical experienced consumers. The interaction with such devices becomes so easy, that also older people playing or working with them. Especially older people have more handicaps, so for them it is difficult to read small text, like they are used as description to buttons on remote controls for televisions. They also become fast overstrained, so that bigger technical systems are no help for them. If it is possible to interact with gestures, all these problems can be avoided. But to allow an intuitive and easy gesture interaction, gestures have to be supported, which are easy to understand. Because of that fact, in this paper we tried to identify intuitive gestures for common interaction scenarios on computer-based systems for uses in ambient assisted environment. In this evaluation, the users should commit their opinion of intuitive gestures for different presented scenarios/tasks. Basing on these results, intuitively useable systems can be developed, so that users are able to communicate with technical systems on more intuitive level using accelerometer-based devices.

Keywords: gesture-based interaction, computer-human-interaction, human-centered interfaces.

1 Introduction

Modern interaction systems, particularly in ambient assisted environments, are using modern interaction styles, orienting on the human factors. This allows systems to provide easy to use interfaces, which the user can use without reading manuals or documentations. A good example for this aspect is the iPhone or the iPad from Apple. These devices provide a well-designed and easy-to-use graphical user interface and an intuitive useable interaction with finger gestures. But also entertainment-based platforms and consoles are focusing increasingly on these modern styles of interaction. On the one hand it is more intuitive for the users and on the other hand it is closer and therefore more natural to the performances in the real world. The first entertainment console was the Nintendo Wii, which had an amazing success because of its innovative intuitive accelerometer-based device, with which games could played by dynamic gestures. Next to this controller also a balance board exists, using which the user can surf or ride a virtual snowboard, so that the gaming is similar to the real world activity.

In general the interaction using modern interaction methods becomes easier. In difference to the established devices like mouse or keyboard, with the new controllers every user comes clear from the first time of use. Especially computer novices have now an easier access to these systems. Next to technical aspect in using such devices, there is also a healthy once. Older people are not able to move a computer mouse as softly as needed to select icons for starting a single program. For them it is also hard to use a normal remote control e.g. for the television, because of the small buttons, which have to be pressed for changing the program channel. For those older people it will be helpful, if they can use a more natural and easier interaction style. This can be achieved with a gesture-based interaction. With gestures, seniors can interact through menus and systems in general without facing problems of small buttons or to small written channel numbers on a remote control for the television. They also do not get overcharged, because of having too much buttons for controlling a system. So the support of gestures can be a helpful development, for easy interaction with systems.

In this paper we will present the results of an evaluation, to determine gestures, which are intuitive for common interaction scenarios and tasks on computer-based systems. We will concentrate on gestures, which should be performed with an accelerometer-based interaction device, like the WiiMote. Next to computer systems, the results are also interesting for similar other systems, like televisions, because the most of the scenarios, like scrolling or zooming are also useful for those systems.

2 Overview and Related Works

In the communication between people, gestures are playing an important role. Because with gestures, the spoken word can better be pictured so that a committed aspect can better be described additionally. But the importance of gestures is also indicate able by observing the environment, where a speech communication is impossible or is restricted because of surrounding annoying sources e.g. noise on a construction area. On the role field of an airport or on a construction area it is hard to give all needed commands for a communication by voice so that for the most commands different gestures are existing and will be used. The gesture commands are understood by both communication partners and are less error susceptible than a communication only with speech.

Similar to such working areas, also the communication between computers and users can be made easier by using more intuitive interaction methods like a gesture-based interaction. Especially in ambient assisted environments a more natural form of interaction is used, so that also older people are able to communicate with the technologies and systems.

2.1 Gestures Depends on Cultural Back-Ground

Gestures are a common additive in the non-verbal communications. Next to the verbal communications, gestures will help to commit an opinion better convey to the audience, because with gestures different aspects can better be sketched additionally. A human being learns gestures during his entire life, beginning at the time when he is baby. During these learning phases a person learns a dataset of different gestures. Some gestures are used without noticing it for instance in discussions and some of them are used directly like pointing gestures to an item.

Another person is only able to understand all the gestures, if both communication partners have the same common ground [1] for the used gestures. But this common ground depends also on the cultural background [2]. Perhaps in different countries different gestures for the same meaning can be exists and also same gestures with different meaning are possible. This is a relevant fact, by designing and creating a system for different countries, which should be controlled with gestures.

2.2 Throw and Tilt Interaction

Dachselt et. al. [3] developed a system, which allows an intuitive interaction with media data. This system allows using tilt gestures for an interaction with mobile applications or distant user interfaces. The mobile phone, which contains accelerometers, acts as remote control. Furthermore, throwing gestures can be used to transfer media documents and even running interfaces to a large display like a beamer.

However, this system uses simple gestures for basic functionalities, like transmitting data to another technical device. And also the list of functions is limited to general understandable gestures, but these functions does not lean on the use of established platforms with common functions, like selecting, moving or zooming in a windowed application.

2.3 Hand-Drawn Gestures

Next to interactions by selection menu options and similar things, it is necessary to provide functionalities, using which a user can easily enter texts or edit paragraphs etc. But for a user it can be more efficiently, if such tasks can be accomplished with gestures and not by entering a command or finding an option in a menu. To enter texts with gestures, basing on traditional way of writing a letter with a pen, a lot of positive implementations and evaluations were done. But it is also possible to provide an intuitive to use of editing options with gestures for working on the text [4].

By using gestures in writing texts and also editing texts or paragraphs a user can fast create documents, without knowing much about computers or a specific writing program.

The results of such techniques are necessary, because for shorter texts it can be easier to write a text with gestures, then with a keyboard – especially for non-computer experts. The aspect of writing and editing texts is a specific field and provides no generalizable interaction for common interaction tasks with a technical system like scrolling through a document.

2.4 Postures and Gestures in Human-Computer-Interaction

Most gesture interaction technologies are basing on the idea of identifying a performed posture or gesture, which is more complex than only a simple movement. Especially in the non-verbal communication between people, gestures can be used to convey information without using words in an intuitive way [5]. Most of these interactions will be done with hands and arms, but it is also possible to interact with gestures, performed with the head or the whole body.

Because of the use of gestures in the all-day life, gestures are easy to use for users and so they are also intuitive for them. It is also intuitive, if users have to perform the gesture with a device in his hand that identifies performed gestures or postures, like the WiiMote. The challenge in using such more complex gesture lays in the high effort in developing recognition systems, which are able to detect gestures with a good recognition rate.

3 Experiment for Determining Intuitive Gestures Using Accelerometer-Based Interaction Devices

The experiment we made for this paper has as main goal to identify typical and intuitive gestures for typical interaction scenarios for instance scrolling through a presented document or text in general. Next to the gestures, we also expect to determine prevalent gestures, which are similar or equally performed. This is a relevant point, because technical systems must support those gestures, which are commonly used by users, but it is often impossible to invest the effort for regarding all possible gestures that people suggest as intuitive once.

In the following chapter we describe the evaluation environment and for which scenarios we ask the participants for intuitive gestures.

3.1 Overview

To determine intuitive gestures for common scenarios, we present such a scenario in an example screenshot to the participants. Then a participant gets time for trying different gestures. At the end of this phase, he has to commit that gesture, where he thinks it is for him the most intuitive once. As answer only gestures are allowed. The evaluation encompasses only solvable and commonly occurring scenarios.

All activities of the participant will be recorded in two different ways. At first, we store all activities with two cameras, from two different positions as an informal observation. On the other hand, we store significant peaks of the data-stream from the accelerometers of the WiiMote as formal observations. During the interpretation, with these two kinds of stored results it is possible to compare between the performances of the users with each other, whether they are the same meaning, are just looking equal (but are different in detail, e.g. one participant performed gesture with a light rotated controller) or are in fact different.

3.2 Evaluation Environment

The experiment is performed in a room, which provides enough space to perform also bigger gestures. The participant stands in front of a monitor, where an example context/scenario in a screenshot will be shown. An interviewer stands next to the monitor and describes the situation and the tasks that have to be performed.

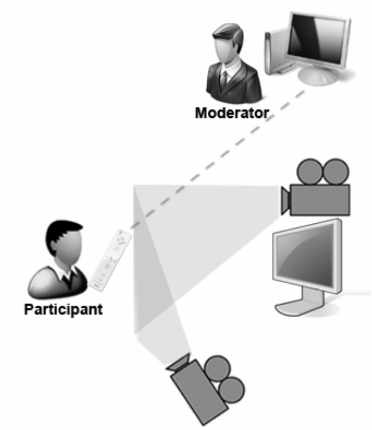


Fig. 1. The evaluation environment, in which the participant is getting observed. Furthermore all performed gestures with the WiiMote are stored on the computer.

During the evaluation, a camera next to the monitor (in front of the user) and a camera right to the user is recording all activities. The WiiMote, with which the participant is performing the gestures sends the data-stream over a Bluetooth-connection to the computer, on which the software (Figure 2) [6] records the peaks of the data-stream and store them on the computer.

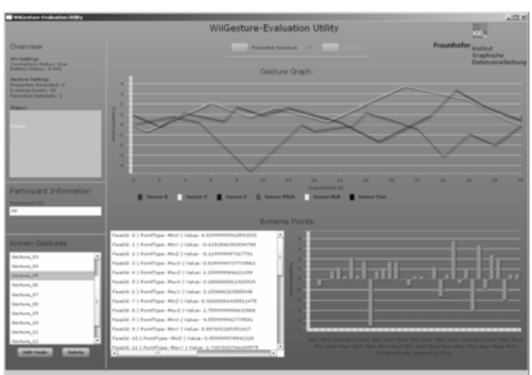


Fig. 2. Screenshot of the Evaluation-Tool, which stores the data from the accelerometers of the WiiMote to performed gestures by a participant.

3.3 Participants

The experiment was performed with 26 participants with European cultural background. The participants were middle-aged people. Because of the fact that we were going to determine intuitive gestures for common interaction scenarios on a technical system, it was not required to evaluate older people e.g. seniors. The described tasks were also abstract and so also older people would be able understand them. Eight of participants had major experience in playing and using the Nintendo Wii and its controller. 10 of the test persons had less experience and 8 did not have any experience with the Nintendo Wii. By making the evaluation with participants who have no or only fewer experiences with the WiiMote, we regard that the participants are open-minded and are not influenced by the default used gestures in Wii-games etc.

All the test persons were regularly working with a PC, thus they knew the presented situations, for which they should find intuitive gestures.

3.4 Evaluation Tasks

During the evaluation, several interaction situations were presented, where the participants have to perform a gesture by interacting with the WiiMote-controller. The gestures which can be performed were not restricted so that the test persons could use the same gesture in more than one task and with an unlimited complexity.

Graph-Navigation

In a presented relation graph (e.g. for presenting semantically information), the participants had to move between different nodes along presented edges. The nodes were placed in a square-layout. The tasks were:

1. Select the node right to the actual selected bottom-left node.
2. Select the node above to the actual selected bottom-right node.
3. Select the node left to the actual selected top-right node.
4. Select the node below to the actual selected top-left node.

Simple Dialog

In the next situation the participant had to choose an option in a simple dialog. In the dialog the participant had only two options, like it is normally used in small requests e.g., “Do you really want to delete the selected items?”:

5. accept the dialog
6. decline the dialog

Zooming Function

In the next two situations, a very small and an oversized visualization were present. The participants had to perform the following tasks:

7. enlarge the visualization (Zoom-in)
8. minimize the visualization (Zoom-out)

History/Undo Function

Some applications provide the possibility to undo actions or e.g., in web-browsers the user can go back to a previous visited website. So the participants get the tasks:

9. go backwards / undo
10. go forwards / redo

Scrolling in Windows

To present more information on small physical screens like on mobile devices, most application use scrollbars. This is a widely used feature with a well understood metaphor. So the participants had to perform the following tasks:

11. scrolling to the right
12. scrolling to the left
13. scrolling to the bottom
14. scrolling to the top

3.5 Accomplishment

Before performing gestures, the participants had to fill out a form for socio- and demographical information and questions about their experiences, especially with the Wii and the WiiMote, because this could influence the performances of the gestures and therefore it is important for the later analysis.

After that, the participants got some screenshots to the tasks, to allegorize the situation and what is mentioned by the tasks. The participant then had to perform the gesture most appropriate to solve the task. All of the tasks and performed gestures captured by the two cameras and also the data-stream of the accelerometers of the WiiMote were stored on a PC. The above-mentioned tasks are in a special order, so that similar tasks like navigation within visualization and scrolling in applications will noticed as completely different tasks by the participants. This avoids unintended dependencies in the gesture performances.

At the end of the practical part of this evaluation, the participants appraised the practical part and estimated the appropriateness of a gesture-based interaction for the different situations.

Fig. 3. Screenshot of the Analysis-Tool for the data-stream of the WiiMote.

At last, all of the recorded information was analyzed. For a better understanding of some gestures, the data-stream from the WiiMote helps to image, how a gesture was exactly performed by the participant (Screenshot of the Analysis-Tool is presented in (Figure 3). The data-stream also helps to see, if gestures were performed equally or if they are only similar.

4 Results

After the evaluation was finished, we analyzed the results, if performed gestures to the same task were equal or similar. If so, they will be summed. In the following we want to show the results of the performed gestures. To present the performed gestures, we sketch the gestures in 2D-graphs. So it is possible, to understand the made performances. As orientation how a gesture was performed in the room in real and later sketched in 2D-graphs, the mapping is shown in Figure 4.

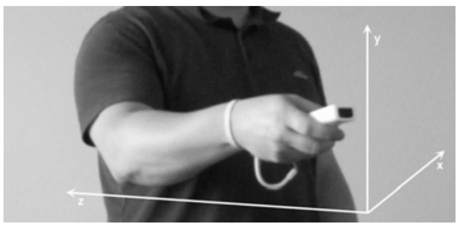


Fig. 4. Interpretation of performed gestures, which are sketched in the follow graphs.

Next to sketch of the performed gestures, the distribution is calculated absolutely and relatively. Especially for the complex tasks, more different gestures could be identified. Also during the evaluation the participants needs more time in cogitation for an intuitive gesture.

The detailed results are presented on the last pages of this paper, because of the needed space.

5 Discussion

The results of this evaluation show, that the numbers of different gestures increases if the tasks become more complex and allow a wider space for more creativity. So for example, if the user wants to scroll, the gestures are limited to gestures orientating to the side, in which they should be scrolled. In comparison by the gesture for accepting a dialog, there is wide range of possible gestures, like the side-position of the button or an abstract metaphor like check. It will be interesting to analyze, if it is possible, to determine on which degree of complexity the users will commit different gestures for the same command.

In focus of creating systems for a user-centered interaction, it is necessary to support those gestures, which users suspect they are intuitive for them. But especially for the more complex gestures many gestures have to be regarded for only one command. The higher the possibilities of gestures that have to be recognized, the effort increases too. So it is important to find a weighted solution between supported gestures and an acceptable essential effort. For that, further researches are required to identify the satisfactory level. Next to this question, it is also useful to solve, if it is useful to support more complex gestures – because by giving the user the freedom for performing personal intuitive gesture it can also happened that users spending more time in thinking for gestures, than using them for an easier interaction. This circumstance can also end in overstraining the users.

Another aspect is to support only one gesture for each command, like it is applied by mobile devices or multi-touch interaction under Microsoft Windows 7. So an evaluation will be useful, if such a limited predefining comes along with a user-centered interaction or maybe it will be a break. And furthermore, if it is really useful for providing a free and individualize interaction by let defining gestures depending on the users' perception.

6 Conclusion

In this paper we presented the results of an evaluation, to determine intuitive gestures for different common interaction scenarios on computer-based systems. The users were asked to define intuitive gestures for different interaction situations. After that, all the performed gestures for a single situation were analyzed, if they were similar. The results are presented in an overview, categorized by the different tasks. With these results a system can be created, which will support intuitive gestures for typical scenarios. It is also possible to identify that more effort is need in supporting and implementing gestures, where the scenario has a potentially wide range of possible gestures. Especially if the gestures becomes more general and allow more freedom in the performance, the number of different gestures increases.

We also noticed that in this area many further questions comes up, especially in defining general rules for good implementations of gesture-recognition systems. But to spend this effort in answering these questions is useful because most of the participants have answered that they imagine that a gesture-based interaction will make the interaction with technical systems easier. In perspective to the older people this is an approach that provides simple interaction interfaces to seniors and provide them so the possibility of using modern technical systems intuitively thus making their all-day life easier.

References

1. Hofemann, N.: Videobasierte Handlungserkennung für die natürliche Mensch-Maschine-Interaktion. Diss., Technische Fakultät of the University Bielefeld (2007)
2. Rehm, M., Bee, N., André, E.: Wave like an Egyp-tian: accelerometer based gesture recognition for cul-ture specific interactions. In: Proceedings of the 22nd British HCI Group Annual Conference on HCI 2008: People and Computers Xxii: Culture, Creativity, interaction, vol. 1, pp. 13–22. British Computer Society, Swinton (2008)

3. Dachselt, R., Buchholz, R.: Natural throw and tilt in-teraction between mobile phones and distant dis-plays. In: Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems, pp. 3253–3258. ACM, New York (2009)
4. Wolf, C.G., Samuels, P.M.: The use of hand-drawn gestures for text editing. International Journal of Man-Machine Studies 27(1), 91–102 (1987)
5. LaViola, J.J.: A Survey of Hand Posture and Gesture Recognition Techniques and Technology. Technical Report. UMI Order Number: CS-99-11., Brown University (1999)
6. Burkhardt, D., Nazemi, K., Bhatti, N., Hornung, C.: Technology Support for Analyzing User Interactions to Create User-Centered Interactions. In: Proceedings of HCI International 2009, pp. 3–12. Springer, Heidelberg (2009)

A Result Details

Like it was described in chapter 4, the details for the results are presented here in this section. The results are categorized into the different kinds of tasks, which were described in chapter 3.4. For all performed gestures to a single task, we analyzed the absolute and relative distribution.

A.1 Graph-Navigation

Move from bottom-left node to the bottom-right one:

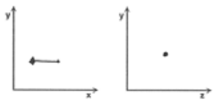
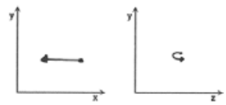
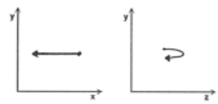
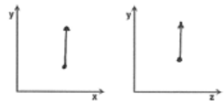
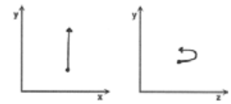
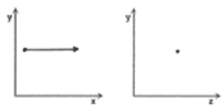
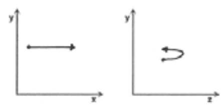
Gesture #1a	**Gesture #1b**	**Gesture #1c**
		
18x	6x	2x
69%	23%	8%

Fig. 5. Two gestures performed differently for the same sub-task: movement to the right-side.

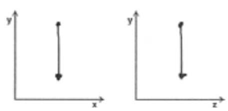
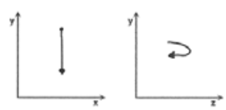
Move from bottom-right node to the top-right one:

Gesture #2a	**Gesture #2b**
	
23x	3x
88%	12%

Move from top-right node to the top-left one:

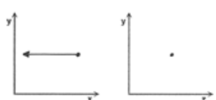
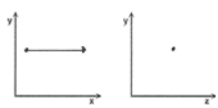
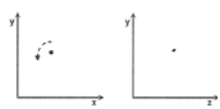
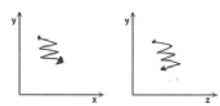
Gesture #3a	**Gesture #3b**
	
22x	4x
85%	15%

Move from top-left node to the bottom-left one:

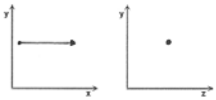
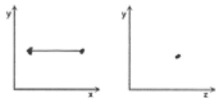
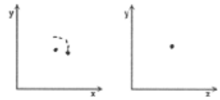
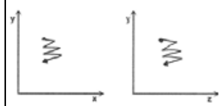
Gesture #4a	**Gesture #4b**
	
23x	3x
88%	12%

A.2 Scrolling in Windows

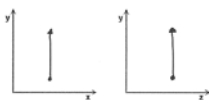
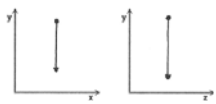
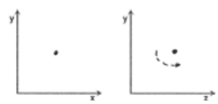
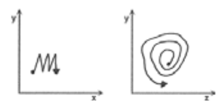
Scrolling to the right side (Scrollbar move rightwards):

Gesture #5a	**Gesture #5b**	**Gesture #5c**	**Gesture #5d**
			
16x	5x	3x	2x
62%	19%	11%	8%

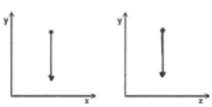
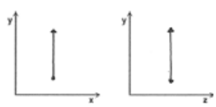
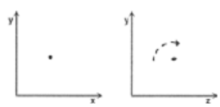
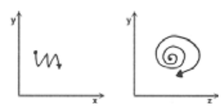
Scrolling to the left side (Scrollbar move leftwards):

Gesture #6a	Gesture #6b	Gesture #6c	Gesture #6d
			
16x	5x	3x	2x
62%	19%	11%	8%

Scrolling to bottom (Scrollbar move downwards):

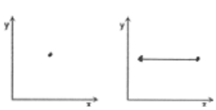
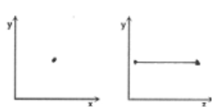
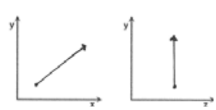
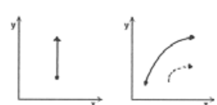
Gesture #7a	Gesture #7b	Gesture #7c	Gesture #7d
			
12x	10x	3x	1x
46%	38%	12%	4%

Scrolling to top (Scrollbar move upwards):

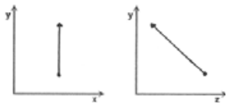
Gesture #8a	Gesture #8b	Gesture #8c	Gesture #8d
			
12x	10x	3x	1x
46%	38%	12%	4%

A.3 Zooming Function

Zoom-in a visualization (enlarge it):

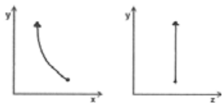
Gesture #9a	Gesture #9b	Gesture #9c	Gesture #9d
			
12x	5x	3x	2x
46%	19%	11%	8%

Gesture #9e



2x
8%

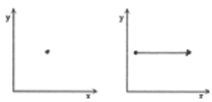
Gesture #9f



2x
8%

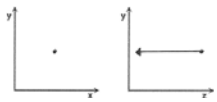
Zoom-out a visualization (minimize it):

Gesture #10a



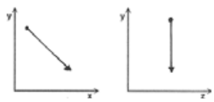
12x
46%

Gesture #10b



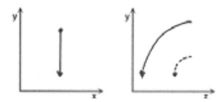
5x
19%

Gesture #10c



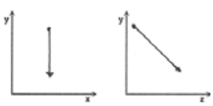
3x
11%

Gesture #10d



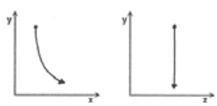
2x
8%

Gesture #10e



2x
8%

Gesture #10f

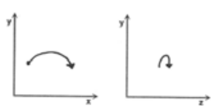


2x
8%

A.4 History/Undo Function

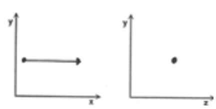
Go back in history/undo an applied action:

Gesture #11a



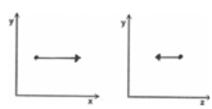
8x
30%

Gesture #11b



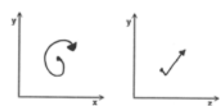
6x
23%

Gesture #11c



3x
11%

Gesture #11d



2x
8%

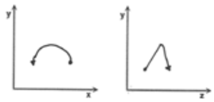
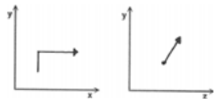
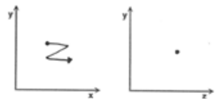
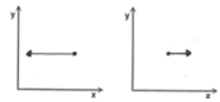
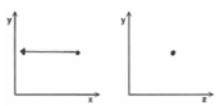
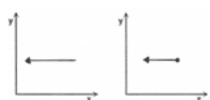
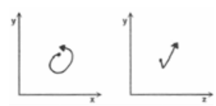
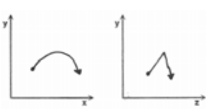
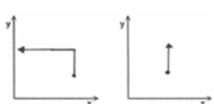
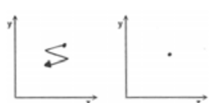
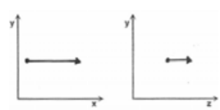
Gesture #11e	Gesture #11f	Gesture #11g	Gesture #11h
			
2x	2x	2x	1x
8%	8%	8%	4%

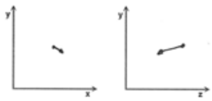
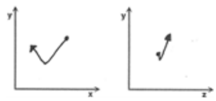
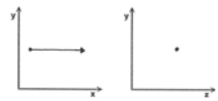
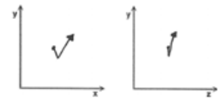
Fig. 6. Three of eight performed gestures for the sub-task “back-in-history”

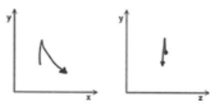
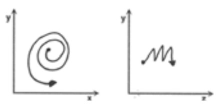
Go forward in history/redo an action:

Gesture #12a	Gesture #12b	Gesture #12c	Gesture #12d
			
8x	6x	3x	2x
30%	23%	11%	8%
Gesture #12e	**Gesture #12f**	**Gesture #12g**	**Gesture #12h**
			
2x	2x	2x	1x
8%	8%	8%	4%

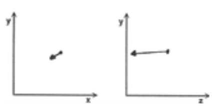
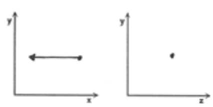
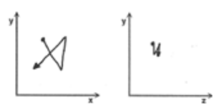
A.5 Simple Dialog (Accept/Decline a Request)

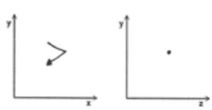
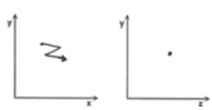
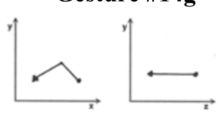
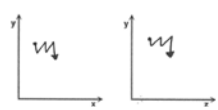
Accept the dialog:

Gesture #13a	Gesture #13b	Gesture #13c	Gesture #13d
			
pointing gesture			
11x	5x	4x	3x
42%	19%	15%	12%

Gesture #13e	Gesture #13f
	
2x	1x
8%	4%

Decline the dialog:

Gesture #14a	Gesture #14b	Gesture #14c	Gesture #14d
			
pointing gesture			
9x	4x	3x	3x
35%	15%	11%	11%

Gesture #14e	Gesture #14f	Gesture #14g	Gesture #14h
			
2x	2x	2x	1x
8%	8%	8%	4%