# Gestures for Large Display Control

Wim Fikkert*, Paul van der Vet, Gerrit van der Veer, and Anton Nijholt

Human Media Interaction, University of Twente,
P.O. Box 217, 7500 AE Enschede, The Netherlands
{f.w.fikkert, p.e.vandervet, g.c.vanderveer, a.nijholt}@utwente.nl
http://hmi.ewi.utwente.nl

**Abstract.** The hands are highly suited to interact with large public displays. It is, however, not apparent which gestures come naturally for easy and robust use of the interface. We first explored how uninstructed users gesture when asked to perform basic tasks. Our subjects gestured with great similarity and readily produced gestures they had seen before; not necessarily in a human-computer interface. In a second investigation these and other gestures were rated by a hundred subjects. A gesture set for explicit command-giving to large displays emerged from these ratings. It is notable that for a selection task, tapping the index finger in mid-air, like with a traditional mouse, scored highest by far. It seems that the mouse has become a metaphor in everyday life.

**Keywords:** Human-centered computing, user interfaces, input devices and strategies, intuitive hand gestures, large display interaction.

## 1 Introduction

Physically large, digital surfaces can supply richly detailed yet diverse content in our everyday environments, whether it is at home, at the office or in a public space [1]. Such surfaces can be embedded in walls, floors, furniture and other physical objects. Ubiquity characterizes the availability of these large digital surfaces in future environments. A wide range of sensors enable us users to interact with such surfaces through very diverse modalities, resulting in an interactive exchange of information between human and computer. Depending on the settings, not all modalities will be suited for interacting with the digital surfaces, also known as 'displays'. For example, (parts of) the displays are not always within reach of the user; preventing touch input. Likewise, speech input is not always desirable, for example, in shopping malls or during conversations.

In this paper, we focus on large display interaction through hands gesturing from beyond arm's length [2]. The goal of interaction is to extract information from the display by navigating through its digital contents, not through a strictly menu-based approach per se [3]. The large display interaction may be obtaining the latest fashion trends in a shopping mall to supporting discussions with detailed results in project-based teamwork to entertainment and games [4]. The

* Corresponding author.

S. Kopp and I. Wachsmuth (Eds.): GW 2009, LNAI 5934, pp. 245–256, 2010.

common denominator for such large display interfaces as we address them is that they offer diverse, detailed and structured information for comprehensive access and that they require explicit command-giving. Communicative modalities in proactive displays such as body posture or eye-contact are beyond our scope [2].

The contribution of this work is a gesture set that comes naturally and with which commands can be issued to a large interactive display with ease [5]. A typical way to interact with large displays through gesturing is to introduce a gesture set that is designed to accommodate the sensors used [6]. Such a gesture set can be difficult for users to learn and use [7]. We consider gestures that come naturally to be intuitive but we recognize that this may have very diverse causes ranging from strong metaphors in everyday life and work, which, we feel, by now includes the indoctrination by decades of mouse-based interfaces.

This work is structured as follows. Section 2 starts with a categorization of the commands that are present in an interactive system. In Section 3 we describe an initial study aimed to discover which gestures are made by uninstructed users. Section 4 then describes how we consulted an international population on gesture representations for issuing specific commands. We discuss our findings in Section 5 where we also describe future work for validating this gesture set.

## 2 Commands to Issue with Gesturing

Explicit command-giving is the basis for the large display interfaces that we focus on. Our aim is not to design new interaction paradigms but to discover how existing large display interfaces might be controlled with the hands. However, it is not clear which elementary commands are at the basis of this interface.

Navigation, selection and manipulation have been mentioned as the elementary tasks in an interface [9]. However, such tasks are often formed by chunking together more fundamental tasks [10]. Buxton argues that human-computer interaction should be regarded from a more human perspective rather than from the device or system [8]. This approach can describe interface tasks at a more generic level. Buxton proposed a three-state model to represent the interactions such as point, select and drag for indirect devices such as the mouse, see Figure 1. For instance, a one-button mouse can be represented to be out-of-range (state #0) when the user is not touching it, tracking (state #1) when the user is moving it and dragging (state #2) when pressing the button. Selection is done with a quick 1-2-1 state transition. The precise meaning of these three states varies (slightly) with the device or interaction technique that is being represented. For example, a stylus is out-of-range when it is lifted from its

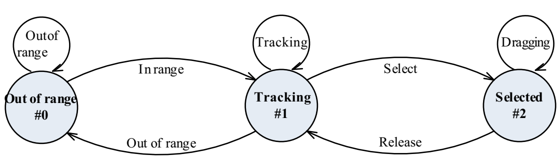



**Fig. 1.** Buxton's [8] three-state model for graphical input. Picture adapted from [8].

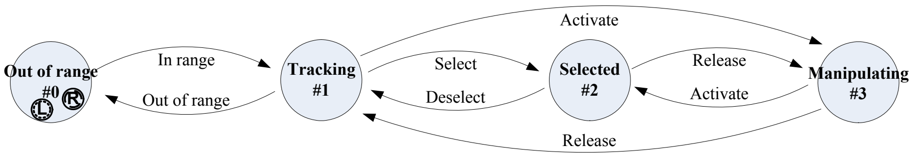



**Fig. 2.** Our four-state model for direct free-hand graphical input, extending the three-state model proposed by [8]. The dynamic manipulation state (#3) represents diverse tasks such as resize and activate. Tokens represent the left (L) and the right (R) hand.

tablet, a joystick has no out-of-range state (#0) because it keeps tracking when untouched and a buttonless joystick does not have a dragging state (#2).

Buxton's model can also describe direct input devices that work directly on the display surface. For such devices, for example, light pens and touch surfaces, a special case of the model applies with a direct transition between the #0 (passive tracking) and #2 (dragging) states because the system does not know what is being pointed at until contact [8]. Looking at the hands gesturing, Buxton's three-state model describes such interfaces adequately but not fully. The user is out-of-range when not addressing the screen (#0). By addressing the screen, the system switches to the track-state (#1) from which selection is possible (#2). Manipulation of any (selected) contents is, however, not always a chunking of these three steps. For example, resizing a selected object [2], activating a selected object [11] or performing a 'right-mouse'-like action on a selected object [10], cannot be described. A five-state model was proposed in which new states are added for each new mouse-button [10]. We propose a similar solution: adding a dynamic state (#3) to Buxton's three-state model to generically encompass manipulation tasks, see Figure 2. Guiard describes the human motor system with two motors that represent the hands [12]. These motors cooperate to form a kinematic chain with the non-preferred hand serving as a reference frame for the preferred hand. Based on Petri nets, we use tokens to represent the hands ($R$ and $L$). These tokens move between states separately or together, capturing unimanual and bimanual interactions. Repeat transitions as in Buxton's model [8] are omitted: tokens remain in a state until an explicit transition.

## 3 Gestures That Come Naturally

To discover which gestures are considered natural by potential users we asked uninstructed users to issue commands through gesturing.

### 3.1 Methodology

Our participants were asked to manipulate a topographic map of our university's local surroundings. Participants were not expected to have knowledge of the local topography. Figure 3 shows the map and the Wizard of Oz set-up that was used[1].

[1] A user is fooled to believe that she is in actual control while, in fact, the wizard is.


**Fig. 3.** The whole and a part of the map that the participants had to manipulate and the set-up for this investigation. The operator sat in the back, out of view.

Three increasingly difficult assignments were executed in which the map view needed to be moved to a specific location on a specific zoom level. Assignments were to locate and display one or more specific town(s) and to position the view port so that these target(s) would fill the screen. Participants could issue two commands to complete each assignment successfully: pan and zoom in/out. We omit the tracking state in our four-state model here because the user is either not interacting when out-of-range(#0) or she is panning (#2) or zooming (#3).

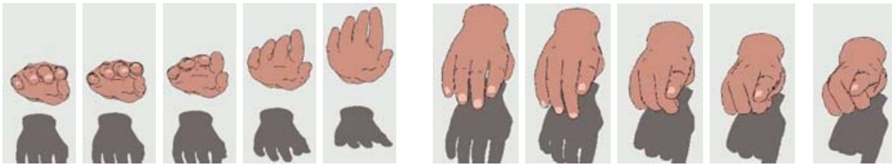
The wizard was introduced as a technician who would perform minor adjustments to a working gesture interface during a brief speak-out-loud session at the beginning of each trial. Note that during this phase, the operator chose the coupling of gesture and state-change, including the out-of-range state. The operator was instructed to respond only to hand movements and to ignore any verbal commands. Visual feedback consisted of the map panning or resizing.

Each trial was video taped with a camera facing the subject. The recorded trials were annotated in ASCII Stokoe[2] and analyzed based on occurrences in Anvil [13]. ASCII Stokoe is designed for sign language annotations. We expanded the annotation with additional symbols `&` and `Z` that represent hand movements with the same hand shape and orientation and the circumfix `S(..)` represented synchronized bimanual movements. We abstracted our annotations based on the assumption that similar gestures have a similar meaning [2]. For example, differences between hand orientation (slightly upwards compared to fully upwards) and hand shapes (cupped hand versus slightly stretched hand) are generalized. The operator also mentioned to have interpreted the gestures in this manner.

### 3.2 Results

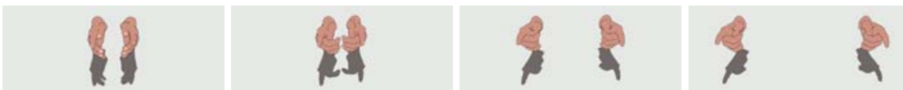
Nine students and colleagues from our group took part in this within-subjects study. They were on average 27 years old ($\sigma = 6$), ranging from 19-36 years. One participant was female, eight were male. All participants were right handed. On a 1-3 Likert-style scale, our participants were proficient with computers ($\mu = 2.8$, $\sigma = .4$), the internet ($\mu = 3.0$, $\sigma = 0$) and map applications ($\mu = 2.8$, $\sigma = .4$). They were somewhat familiar with the local topography ($\mu = 2.1$, $\sigma = .6$).

[2] `http://www.speakeasy.org/~mamandel/ASCII-Stokoe.txt`, last checked Sep. 2009.

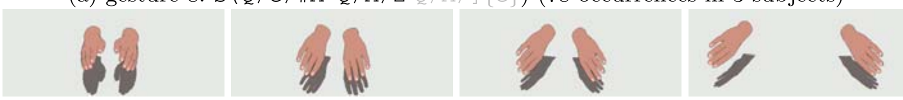


(a) gesture 2: Q/Mf/d (b) gesture 3: Q/C/],f{B} Q/B/& Q/B/#,t{C}

**Fig. 4.** The two most occurring gesture for panning (notation in gray is not depicted): (a) pointing hand away from the user towards the display that moves around, relaxing the hand would release (136 occurrences in 5 subjects) and (b) relaxed/cupped hand to stretched hand that moves around for panning, relaxing or retracting the hand would release (222 occurrences in 8 subjects)



(a) gesture 8: S(Q/C/#A Q/A/Z Q/A/]{C}) (78 occurrences in 3 subjects)



(b) gesture 9: S(Q/C/],fB Q/B/Z Q/B/#,t{C}) (74 occurrences in 4 subjects)

**Fig. 5.** The two most occurring gestures for zooming (notation in gray is not depicted): (a) relaxed/cupped hands to pointing hands that move apart for zooming out and move together for zooming in, relaxing and retracting the hands would release and (b) relaxed/cupped hands to stretched hands that move apart for zooming out and move together for zooming in while relaxing and retracting the hands would release

We identified 14 pan and 13 zoom gestures which we reduced to 3 distinct pan and 6 distinct zoom gestures by generalizing. For panning, two gestures (IDs 2 and 3, see Figure 4) were observed to occur significantly ($p = .01$) more often and in most users. In gesture 2 the hand changes from relaxed to stretched at the beginning of the movement and back to relaxed at its end as if to press a button. In gesture 3 the hand closes with an extended index finger and back to a relaxed hand. Gestures 1 and 3 are similar but then the whole hand is closed.

For the zoom task, we observed a similar although less pronounced ($p = .07$) distinction; gestures 8 and 9 (see Figure 5) occur more for zooming than the other 4 observed gestures. These two gestures are, in fact, very similar still; differing in the hand shape only like in the pan task. Gesture 8 grabs and releases the canvas while gesture 9 will explicitly stretch the hand during the zoom movements. We identified 3 gestures made with one hand (by 5 subjects) and 3 others that were made with both hands (by 8 subjects). Subjects were consistent in their choice for using either one or two hands; some participants used both gestures.

### 3.3 Wrap-Up

The gestures that we observed differ mostly in the preparation (start) and retraction (end) phases of the gesture phrase [14]. We found that in the stroke phase the movements can more or less be directly used as parameter changes for panning and zooming; the hand shape during these movements does not matter much. Our subjects explicitly marked the start and end of their gesture by changing their hand shape from rest to a flat hand or pointing hand for panning and two flat or pointing hands for zooming. Subjects remarked in a post-test interview that their gesture choices were often based on their knowledge of 'mainstream' gesture interfaces such as the Apple iPhone or that they had mimicked movements that they remembered from science fiction movies such as "Minority Report"[3]. We are unsure how this has influenced our results here but it is clear that users readily accept such 'predefined' gesturing as a natural form of interacting. In addition, our subjects consistently apply the same idiosyncratic combinations of gesture and command with a great deal of similarity between users. This leads us to believe that it is possible to construct a set of gesture-commands for large display control that comes natural to the users.

## 4 Gestures to Issue Commands

Gesture interfaces will make use of a broader command set than just the two that we explored in our Wizard of Oz investigation. Section 2 already introduced three elementary gestures (point, select, deselect). Here we add three more gestures (activate, deactivate, resize) that address our fourth state in this model. We performed an investigation to discover which gestures are suited for these tasks.

### 4.1 Methodology

Three experts cannot think up a gesture set that twenty laymen can [15]. In order to fully appreciate or dislike a gesture for issuing a command we argue that the user needs to have experienced it in a working system. Getting subjective results then becomes cost-ineffective as opposed to gathering such information from a large user group online. We must rely on the extent to which these users can imagine operating the large display in the manner that we show them; placing an unknown bias on the results. This bias is expected to be minimal but that will need to be proven by comparing the results to those from a working interface.

In an online investigation[4] we asked our participants to rate gestures for each of our 6 commands. Gestures were selected from literature but also from science fiction movies and commercial gesture interfaces such as the Apple iPhone. In total, we selected 24 gestures for the 6 commands mentioned above. The commands were ordered in a predefined sequence because users would need to make up their mind first, for example, about how they would point before they could select. The gestures, then, were completely randomized per command.

[3] `http://www.imdb.com/title/tt0181689`, last checked Sep. 2009.
[4] `http://fikkert.net/experiment.php`, last checked Sep. 2009.

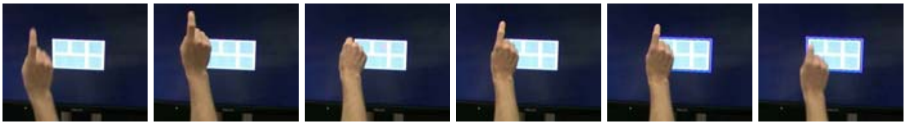


**Fig. 6.** *AirTap* [6]

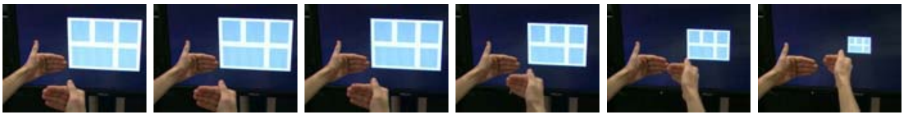


**Fig. 7.** *ReferencedPullPush* from the movie Minority Report

Some examples of the gestures that we selected follow. For pointing (3 gestures) we included *RayCasting* [16] in addition to two more indirect approaches. For select (5 gestures) *AirTap* [6] and *ThumbTrigger* [17] were selected. Deselect had 4 gestures including *SelectOther* where another object is selected to deselect the current one (from MS Windows). Activate and deactivate were taken together (8 gestures) so that gestures that activate mirror those that deactivate, following [6]. We included *AirTap* and *ThumbTrigger* again but also added drawing a triangular 'play' shape that is well-known from audiovisual equipment in the household. For resize (4 gestures) we introduced moving two fingers apart as with the Apple iPhone. Also, *ReferencedPullPush*, from the movie Minority Report, showed the dominant hand serving as a reference for resizing with the distance to the other hand defining the amount, see Figure 7.

Gestures were presented in a video clip showing both hand movements and the response of the interface. The interface was an abstracted system that responded solely through visual feedback. Before the participant filled out the questionnaire we explained this interface and how it would respond. The video clips, see Figures 6 and 7, were shot in a mocked-up setting. Participants were asked how they score gestures (on 7-point Likert-style scales) intuitiveness ('1: very difficult' – '7: very intuitive'), physical effort ('1: little effort' – '7: much effort') required and if they would gesture in this way ('1: no way' – '7: for certain').

### 4.2 Results

99 subjects from four Western-European countries (The Netherlands, Belgium, Germany and The United Kingdom) participated in this investigation. Participants were on average 28 years old (ranging 20-60 years, $\sigma = 8$ years). Our subjects needed roughly 25 minutes to complete the entire trial. In our subject population, 22 subjects were female and 77 were male with 25 subjects who held a Bachelor degree, 53 a Masters, 12 a PhD and 9 had no degree. Our subjects were very familiar with (online) video clips in which gesture interfaces play a role ($\mu = 5.0$, $\sigma = 1.7$, with '1: unfamiliar (never seen one)' – '7: very familiar (regularly)'). They were proficient with the Apple iPhone ($\mu = 4.5$, $\sigma = 1.9$) but

**Table 1.** Description of our trials data ($N = 2376$)

| | intuitiveness | physical effort | would use |
|---|---|---|---|
| mean | 4.44 | 3.60 | 3.89 |
| std. deviation | 1.701 | 1.626 | 1.811 |
| variance | 2.892 | 2.643 | 3.280 |
| kurtosis | -0.979 | -0.769 | -1.135 |
| skewness | -0.238 | 0.340 | 0.039 |

only moderately so with PDAs, smart phones and similar hand-helds ($\mu = 3.7$, $\sigma = 1.6$).

The results of our analyses for normality on the collected trials indicate that we cannot assume a normal distribution of our data. A D'Agostino-Pearson $K^2$ analysis gave $K^2 = 80.504$ ($p < .01$) on intuitiveness, $K^2 = 85.119$ ($p < .01$) on physical effort and $K^2 = 68.474$ ($p < .01$) on whether the participant would use that gesture. This is, as is often the case in count-based data, caused by a ceiling or floor effect. We applied the non-parametric alternative to ANOVA, Kruskal-Wallis H, to assess if there is a significant difference between gestures while pairs were compared with a Mann-Whitney U analysis. Our trials data are further described in Table 1 where skewness and kurtosis are reported; the trials data are mostly deformed due to kurtosis.

**Point.** We found significant differences between gestures for intuitiveness ($\chi^2 = 106.098$, $p < .01$), physical effort ($\chi^2 = 61.827$, $p < .01$) and whether the participant would use this gesture ($\chi^2 = 138.275$, $p < .01$). Our Mann-Whitney U analysis results show that there are significant differences between the three gestures. *RayCasting* scored significantly higher on intuitiveness and on 'would use' than both other gestures. The difference was less pronounced regarding physical effort. Fourteen subjects commented that this is highly intuitive but that it would be fatiguing in the long run. Five subjects wondered how to stop pointing and proposed to stop when pointing off-screen.

**Select.** There was a significant difference between gestures for intuitiveness ($\chi^2 = 98.816$, $p < .01$), physical effort ($\chi^2 = 58.266$, $p < .01$) and whether the participant would use this gesture ($\chi^2 = 80.725$, $p < .01$). Mann-Whitney U analyes show that *AirTap* significantly scores best on intuitiveness, physical effort and 'would use'. In addition, *FistGrab* (closing the hand to a fist) and *ThumbTrigger* score significantly lower than *Encircling* (drawing a circle around the target with the on-screen pointer) for physical effort. Six subjects commented that they did not like the mouse-metaphor of *AirTap* while three subjects disliked the gun-metaphor of *ThumbTrigger*.

**Deselect.** We found a significant difference between gestures for intuitiveness ($\chi^2 = 47.743$, $p < .01$), physical effort ($\chi^2 = 22.817$, $p < .01$) and whether the participant would use this gesture ($\chi^2 = 51.914$, $p < .01$). Mann-Whitney U analyes show that *DropIt* (opening the hand palm-down as if dropping something on the floor) and *SelectOther* scored significantly better on intuitiveness (higher), physical effort (lower) and 'would use' (higher) than both *RetractToRest* and *JerkyRetract* where the hand retracted to rest or in a jerky way, respectively. There was no significant score difference between *DropIt* and *SelectOther* on these three topics. Twelve subjects indicated that they found selecting another target very familiar from Windows$^{TM}$. One of these subjects commented that although it is familiar, he would prefer 'some' other gesture.

**Activate and Deactivate.** There was a significant difference between gestures for intuitiveness ($\chi^2 = 140.976$, $p < .01$), physical effort ($\chi^2 = 121.518$, $p < .01$) and whether the participant would use this gesture ($\chi^2 = 154.250$, $p < .01$). Mann-Whitney U analyes show that *AirTap* combined with an exit cross scored significantly better on intuitiveness, physical effort and 'would use' than all other proposed gestures. Overall, we can distinguish three groups of gestures: the best gestures are *AirTap* and *AirTap* combined with the exit cross, the worst gestures are drawing 'play' and 'stop' shapes and using (de)activation zones [6]. The other gestures score in between with no significant differences. Eight subjects commented that it was confusing to use the *AirTap* to both activate and select.

**Resize.** We found a significant difference between gestures for intuitiveness ($\chi^2 = 74.200$, $p < .01$), physical effort ($\chi^2 = 64.381$, $p < .01$) and whether the participant would use this gesture ($\chi^2 = 64.117$, $p < .01$). Mann-Whitney U analyis results show no significant score difference between *FingersApart* and *HandsApart* on intuitiveness and 'would use'. In these two gestures either the fingers of one hand or the two hands are moved apart. However, with respect to physical effort, *HandsApart* scored significantly poorer (35%). *ReferencedPull-Push* scored significantly poorer on intuitiveness, physical effort and 'would use' with respect to the other three gestures. Subjects often commented (19 subjects) that *FingersApart* would be too hard to do on large displays were the larger scale is an issue. Likewise, 9 subjects commented that both *FingersApart* and *Hands-Apart* needed some way to start resizing: "needs clicking". Only three subjects mentioned the Apple iPhone as the source of these resizing gestures.

### 4.3 Wrap-Up

We gathered subjective ratings on 24 gestures for 6 interface commands from a large population with a similar background in our online investigation. We found significant preferences for a specific gesture to issue a command. Based on these findings we can construct a gesture set to issue commands with gestures that come naturally. Users expect a gesture-based interface to allow them to point directly at a target using pixel-precise ray-casting. For selecting, *AirTap* mimics clicking a mouse button very precisely even though no actual button can be pressed [6]. The gesture preferred for deselecting also leans heavily on existing interfaces where another target is selected to deselect the current target. The gesture used for selecting was preferred to be used for activating and deactivating targets as well. Although some subjects indicated that they found it confusing to have the same gesture for two tasks (*AirTap* for select and activate), we believe that this simplifies the interface which is indicated by the strong preference for this gesture. This also follows from the comments made by our subjects that they missed some way to 'click' for the resize command. For resizing, our subjects found that moving their fingers or hands apart was the most intuitive while some subjects wondered if moving two fingers apart would scale sufficiently to large displays. Common remarks throughout our study were how to start and stop gesturing, for example, for pointing. Like we found in our Wizard of Oz investigation, it seems that a way should be found to explicitly mark where a gesture begins and ends: when do we move to and from the null-state (out-of-range)? This is perhaps more delicate than simply pointing off-screen.

## 5 Discussion and Future Work

Interacting with physically large, digital surfaces through explicit command-giving can be done through the hands gesturing. A gesture is, in some cases, more suited or desired than other modalities such as touch or speech. In a Wizard of Oz investigation we found that uninstructed users independently generate the same gestures for a limited command set. The stroke phase in a gesture directly changes a parameter (zoom or pan) and the gesture was explicitly started and ended by alternating between a tensed and a relaxed hand respectively. Our follow-up study showed that such agreement is also found in a large user group for a broader and more complete command set. Moreover, similar commands, for example, select and activate, should be issued by the same gesture to simplify the interface as much as possible. Our results so far indicate that the participants in both our investigations were influenced by existing interfaces such as MS Windows and the Apple iPhone. Although biased to some extent, we feel that this made these users more open-minded to new forms of interaction that we aim to build. However, it will also have hindered them to look past existing solutions and to imagine what it really would be like to control an interface such as the one we have suggested. That is why we did not expect the *AirTap* [6] gesture to be the best candidate for the select and activate commands; users will become fatigued easily, feeling like having 'gorilla arms' when they have to extend their arms while interacting. Moreover, *AirTap* builds on the familiar mouse with tangible feedback that is not there any more, something none of our subjects mentioned. The gesture set we discovered so far will adequately control large displays but it needs to be consolidated in a working environment where users can truly experience the interaction.

We propose to continue this line of research with a third investigation in which our previous results are validated. A working prototype allows users to experience the gestures for our command set and thus provide us with a more in-depth opinion. The interactions should last long enough for the user to appreciate the gesture and to comment on it. We aim to select a focus group of participants from our online investigation so that these users are already familiar with both the gestures and how to issue commands. Using a pair of data gloves and a position tracking system we can detect both hand locations and hand shapes. The whole gesture set from our online investigation will be included to adequately validate our earlier findings. The interface needs be more elaborate than the map or abstracted applications that we used previously so that users will repeatedly issue commands through gesturing. We propose a list of search-tasks through a collection of photos where information from multiple photos needs to be cross-referenced. For example, we envision a virtual tourist application where a visitor of our university will have to look for and select landmarks on the university and to couple them to the university map.

It is surprising in some way to discover that the familiarity of the mouse has such a strong impact on our findings. We readily accept that these results would be very different when consulting a user group that does not have this type of experience, from a different culture or even from another social group. However, the standard Windows-Icon-Mouse-Pointer paradigm has, over the past decades, indoctrinated most users of the systems that we aim to control. In that respect, we might even argue that the author of this work was born after the invention of the mouse and that he, like many other users of large display systems, grew up with this input device. We feel that the mouse has become such a strong metaphor that, even though a natural interface might be defined otherwise [1], it has by now become an everyday metaphor that drives the formation of a gesture set for explicit command-giving through hand gestures.

## Acknowledgements

This work is part of the BioRange program carried out by the Netherlands Bioinformatics Centre (NBIC), which is supported by a BSIK grant through the Netherlands Genomics Initiative (NGI).

## References

1. Oviatt, S., Cohen, P.: Perceptual user interfaces: multimodal interfaces that process what comes naturally. Communications of the ACM 43(3), 45–53 (2000)
2. Fikkert, W., D'Ambros, M., Bierz, T., Jankun-Kelly, T.: Interacting with visualizations. In: Kerren, A., Ebert, A., Meyer, J. (eds.) GI-Dagstuhl Research Seminar 2007. LNCS, vol. 4417, pp. 77–162. Springer, Heidelberg (2007)
3. Wahlster, W.: SmartKom: Foundations of Multimodal Dialogue Systems. Cognitive Technologies, vol. XVIII. Springer, Heidelberg (2006)
4. Nijholt, A., Reidsma, D., Poppe, R.: Games and entertainment in ambient intelligence environments. In: Aghajan, H., Delgado, R., Augusto, J.C. (eds.) Human-Centric Interfaces for Ambient Intelligence. Elsevier, Amsterdam (2009)
5. Hummels, C., Smets, G., Overbeeke, K.: An intuitive two-handed gestural interface for computer supported product design. In: Wachsmuth, I., Fröhlich, M. (eds.) GW 1997. LNCS (LNAI), vol. 1371, p. 197. Springer, Heidelberg (1998)

6. Vogel, D., Balakrishnan, R.: Distant freehand pointing and clicking on very large, high resolution displays. In: Proceedings of the 18th annual ACM symposium on User interface software and technology (UIST 2005), pp. 33–42. ACM Press, New York (2005)
7. Wexelblat, A.: Research challenges in gesture: Open issues and unsolved problems. In: Wachsmuth, I., Fröhlich, M. (eds.) GW 1997. LNCS (LNAI), vol. 1371, pp. 1–11. Springer, Heidelberg (1998)
8. Buxton, W.: A three-state model of graphical input. In: Proceedings of the IFIP TC13 Third Interational Conference on Human-Computer Interaction (INTERACT 1990), pp. 449–456. North-Holland Publishing Co., Amsterdam (1990)
9. Bowman, D., Kruijff, E., LaViola, J., Poupyrev, I.: 3D User Interfaces: Theory and Practice. Addison Wesley Longman Publishing Co., Inc., Redwood City (2004)
10. Hinckley, K.: Input technologies and techniques. In: Sears, A., Jacko, J. (eds.) Handbook of Human-Computer Interaction: fundamentals, evolving technologies and emerging applications, pp. 151–168. Lawrence Erlbaum Associates Inc., Hillsdale (2006); Revision of 2002 chapter with lots of new material
11. Ahlstroem, D., Alexandrowicz, R., Hitz, M.: Improving menu interaction: a comparison of standard, force enhanced and jumping menus. In: Proceedings of the SIGCHI conference on Human Factors in computing systems (CHI 2006), pp. 1067–1076. ACM, New York (2006)
12. Guiard, Y.: Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. Journal of Motor Behavior 19(4), 486–517 (1987); Slightly edited version of an article originally published
13. Kipp, M.: Gesture Generation by Imitation - From Human Behavior to Computer Character Animation. PhD thesis, Saarland University, Saarbruecken, Germany, Boca Raton, Florida (December 2004)
14. McNeill, D.: Hand and mind: What gestures reveal about thought. University of Chicago Press, Chicago (1992)
15. Wobbrock, J., Morris, M., Wilson, A.: User-defined gestures for surface computing. In: Proceedings of the 27th international conference on Human factors in computing systems (CHI 2009), pp. 1083–1092. ACM, New York (2009)
16. Bolt, R.: "put-that-there": Voice and gesture at the graphics interface. SIGGRAPH Computer Graphics 14(3), 262–270 (1980)
17. Grossman, T., Wigdor, D., Balakrishnan, R.: Multi-finger gestural interaction with 3d volumetric displays. In: Proceedings of the 17th annual ACM symposium on User interface software and technology (UIST 2004), pp. 61–70. ACM Press, New York (2004)