A Comparative Study on Distant Free-Hand Pointing

Ondřej Poláček, Martin Klíma, Adam J. Sporka
Faculty of Electrical Engineering, Czech Technical University in Prague,
Karlovo nám. 13, 12135 Praha 2, Czech Republic
{polacond, xklima, sporkaa}@fel.cvut.cz

Pavel Žák, Michal Hradiš, Pavel Zemčík, Václav Procházka
Graph@FIT, Brno University of Technology
Božetechova 1/2, 612 66 Brno, Czech Republic
{izakpa, ihradis, zemcik}@fit.vutbr.cz
xproch12@stud.fit.vutbr.cz

ABSTRACT
In this paper we present a comparative study of free-hand pointing, an absolute remote pointing device. Unimanual and bimanual interaction were tested as well as the static reference system (spatial coordinates are fixed in the space in front of the TV) and novel body-aligned reference system (coordinates are bound to the current position of the user). We conducted a point-and-click experiment with 12 participants. We have identified the preferred interaction areas for left- and right-handed users in terms of hand preference and preferred spatial areas of the interaction. In bimanual interaction, the users relied more on dominant hand, switching hands only when necessary. Even though the remote pointing device was faster than the free-hand pointing, it was less accepted probably due to its low precision.

Categories and Subject Descriptors
H.5.2 [**Information interfaces and presentation**]: User Interfaces – *Input devices and strategies*

General Terms
Experimentation, Measurement, Performance.

Keywords
Free-Hand Pointing, Remote Pointing Device, Mouse, Bimanual Interaction, Fitt's Law, Kinect, Wii

1. INTRODUCTION
Since Bolt's famous Put-That-There system [1], the distant pointing techniques and devices have made a large progress. Today we witness a great commercial success of computer games that are controlled by body motion and limb gestures, which is a very intuitive and responsive way of input. Playing such games requires a system that is capable of tracking the current position and motion of the user. In our work, we investigated performance of two existing commercial systems, *Kinect* and *Wii*, in terms of speed and pointing precision.

In our study we target two scenarios where the user can navigate the user interface distantly and no direct touch is required. The first scenario is an Interactive TV. Both *Kinect* and *Wii* are intended to be used with a TV as a special remote control device. Besides playing games, they both can be used to control a user interface of an Interactive TV. The other scenario is a large public display presenting information such as indoor navigation, planned events, booking of rooms and seats, etc. in a public building like a library. Let us further imagine that the display is large enough to let several users work at the same time. In this case, we can introduce a dedicated user interface for each detected user. As the user moves along the large display, the UI controls are following her or him.

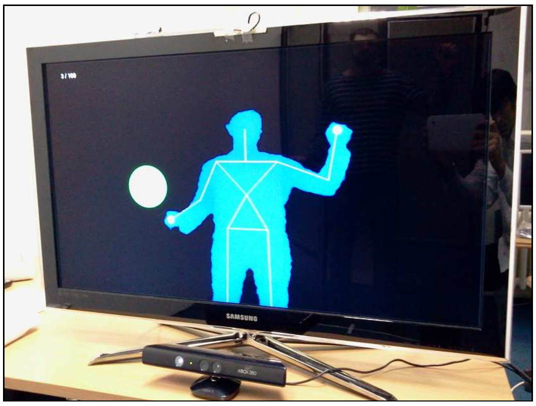


Figure 1. A silhouette of the user interacting bimanually

2. RELATED WORK
Several works exist that have focused on the remote large display navigation, gesture recognition, and remote pointing. The Fitt's law [2] is a well described model and it has been proven, that it works for touch as well as for remote pointing with Wii [3]. Test of two remote input devices for Interactive TV [4] resulted in conclusion that the mouse is strongly preferred technique. According to Riek et al. [5], the Fitt's law is surprisingly not violated in bimanual aiming task.

The distant display pointing techniques have been elaborated for menu selection problem [6] or button pressing [7]. Clark et al. [8] explored multimodal interfaces for distant and close interaction.

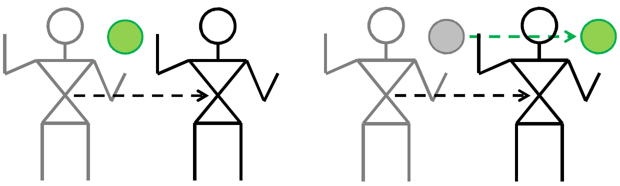
Several different techniques have been introduced for pointing. Some of them (e.g. [9,10]) are based on computer vision using regular cameras and are able to handle both pointing and touching while not requiring any additional devices held by the user. *Samsung ES8000* is an example of a commercial product operated by gestures. Others (e.g. [11, 12]) require additional hardware which can be an obstacle for use in our proposed scenarios. Vogel et al. [13] used a commercial motion tracking system that allowed emulating click by precise finger gestures.

3. OUR SOLUTION

Our free-hand pointing solution utilizes motion sensing device *Microsoft Kinect* originally developed for *Xbox 360* game console. The device is placed below a screen and it monitors area in front of the screen. *OpenNI* library (http://www.openni.org/) was used to track the user. Compared to professional motion tracking systems, this solution is much cheaper but less precise.

A silhouette of the user is displayed on the screen and two cursors are shown on the user's palms (see Figure 1). As the palm moves the cursor moves accordingly on the screen. Click operation is performed by moving palm forward or backward the screen by 3 inches (approx. 7.5 cm). These movements resemble pushing and pulling a button.

Besides a user interface, in which elements do not change the position (fixed reference system, further referred to as *static*), we also implemented a user interface, in which the reference system follows the center of the user's body (body-aligned reference system, further referred to as *dynamic*). In other words, whole user interface moves on the screen along with the user's motion. The difference between these two user interfaces is depicted in the Figure 2.



a. b.

Figure 2. a. Static interface – the UI element, depicted as green circle, do not change the position as the user moves. **b.** Dynamic interface – the UI element moves along with the user.

4. EXPERIMENT

The aim of the experiment was to compare four versions of the free-hand pointing interface with a remote pointing device and a PC mouse. The PC mouse was the only indirect pointing device and it was used only as a baseline method. We expected the mouse to outperform the other mentioned methods. The following interaction methods were prepared for the comparison test:

- *Free-hand dynamic bimanual*. The user interface was following the position of the participant who could use both hands for pointing.
- *Free-hand dynamic unimanual*. The user interface was following the position of the participant, who used only dominant hand for pointing.
- *Free-hand static bimanual*. The participant could use both hands for pointing.
- *Free-hand static unimanual*. The participant used only dominant hand for pointing.
- *Remote pointing device*. *Nintendo Wii* remote control was used. The remote control has the ability to determine its position using a built-in infrared camera that monitors infrared LED diodes mounted on top of the TV, called sensor bar. The mouse cursor was directly manipulated by the device. Mouse click was actuated by pressing the trigger on the bottom of the device.
- *Mouse*. A common computer optical mouse was used. *Enhanced pointer precision* option in *MS Windows* was turned off.

The performance of each method was measured using a sequence of 100 point-and-click target acquisition tasks. Each participant was requested to aim and click on circles that appeared one-by-one on the screen. There were three levels of size (32, 64, and 128 px in diameter) and three levels of distance between the successive circles (128, 256, and 512 px). Both levels were being randomly selected during the task. The objective data collected in the task were processed into three indicators as follows:

- *Point time*, which is the duration needed to move the pointer to a circle.
- *Click time*, which is the duration needed to perform a click operation on a circle.
- *Selections per turn*, which is a number of how many times the pointer crossed the border of the target in one turn (the point-and-click task). This measure is a quantification of precision of the interaction method.

The point-and-click experiment ran on a personal computer connected to a 40" Samsung LCD TV with resolution 1280x720 pixels. All participants were told to stand 6 feet (approx. 1.8m) far from the TV.

12 participants took part in the experiment. They were recruited from the university staff (4 women, 8 men, mean age=34, SD=8.6) and were all experienced computer users, however they had no previous experience with *Kinect* or *Wii* systems. Ten participants were right-handed. Within-subject experimental design was used, the factor was the choice of the interaction method.

Participants were asked to complete the sequence of 100 tasks for each input method. The order of presentation was counterbalanced using a Latin square to compensate for a learning effect. The entire session (all six series) lasted for approximately 30 minutes. After completing all tasks, the participants were asked to complete a questionnaire.

5. RESULTS

5.1 Speeds

The measured values in the experiment are summarized in the Table 1. The ANOVA test and pairwise t-test with Bonferroni correction were used to find statistically significant differences in mean quantities among the methods. The task times suggest that pointing device is slightly faster than free-hand pointing. However, the only significant difference in the task times ($F(5,66) = 11.9$, $p<<.001$) is between the mouse and all other methods as expected. Using dynamic interface is a bit slower than static as a distant circle gets even farther while the user trying to reach it. The static unimanual interface is significantly faster than the dynamic unimanual interface for both pointing ($F(5, 7032) = 207.0$, $p<<.001$) and clicking (($F(5, 7032) = 61.7$, $p<<.001$). Bimanual interaction is faster than unimanual and it is significantly faster in dynamic interfaces in the point and click times. Remote pointing device was faster than all the free-hand pointing methods in point time but it is less precise as selection per turn suggests. Compared to the static interfaces, the pointing device is even significantly less precise ($F(5, 7032) = 48.9$, $p<<.001$).

Table 1. Mean times and standard deviations (SD) for each measured value and modality.

	dynamic bimanual		dynamic unimanual		static bimanual		static unimanual		remote pointing device		mouse	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
Task time [s]	215.3	26.0	263.4	85.6	201.8	40.6	213.2	50.6	186.4	52.5	110.7	13.3
Point time [ms]	1073	428	1173	552	1054	432	1102	438	889	419	685	226
Click time [ms]	1022	1184	1188	1269	961	1054	962	1033	953	914	432	225
Selections per turn	1.64	1.3	1.7	1.3	1.5	1.1	1.5	1.1	1.7	1.2	1.1	0.3

5.2 Hands usage

In case of bimanual interaction, participants relied more on their dominant hand. The Figure 3 shows ratio between left and right hand for the right-handed participants on the screen divided into 9x5 grid. The number in each cell indicates the percentage of use of the right hand in that region. Note that in front of the body, the participants used mostly right hand (70-77%). For the left-handed participants the situation is mirrored. Figure 4 shows mean point times for each region when using right hand in static unimanual interaction. Mean times on the right of the screen are generally shorter than those on the left.

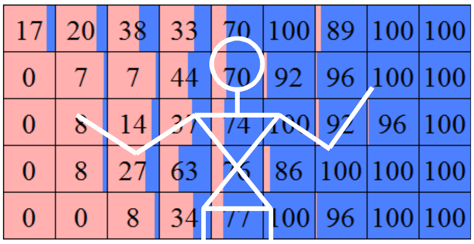


Figure 3. Ratio of right hand usage in different positions on the screen. Red and blue colors (light and dark grey on B/W print) correspond to left and right hands respectively

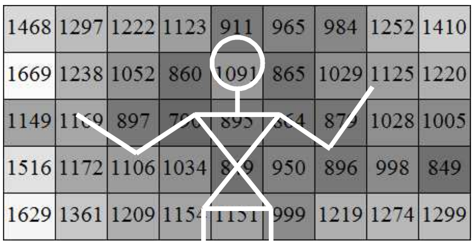


Figure 4. Mean point times in different positions on the screen when using right hand in static unimanual interaction.

5.3 Fitt's law

Fitt's law [2] provides model of human movement in simple pointing task. It predicts the time required to point (movement time, MT) at a target. The model can be mathematically formulated as follows:

$$MT = a + bID \quad (1)$$

$$ID = \log_2(2A/W) \quad (2)$$

where a and b are device-dependent parameters, ID is index of difficulty of a combination of target distance (or amplitude, A), and width W. In our experiment, there were three different amplitudes (distances between successive circles) as well as target widths (circle diameter), which yielded 5 different indices of difficulty.

In order to find models for each method, a linear regression was performed. The results are shown in the Figure 5. The Table 2 shows a and b values as well as correlation of a model for each method. Note, that correlations are high even for bimanual interaction methods. We believe that this is caused by participant's strategy of changing hands – they tend to use one hand until a circle was too difficult to reach. Then they switched hands and continued in the same manner. The bimanual interaction actually consisted of series of unimanual interaction.

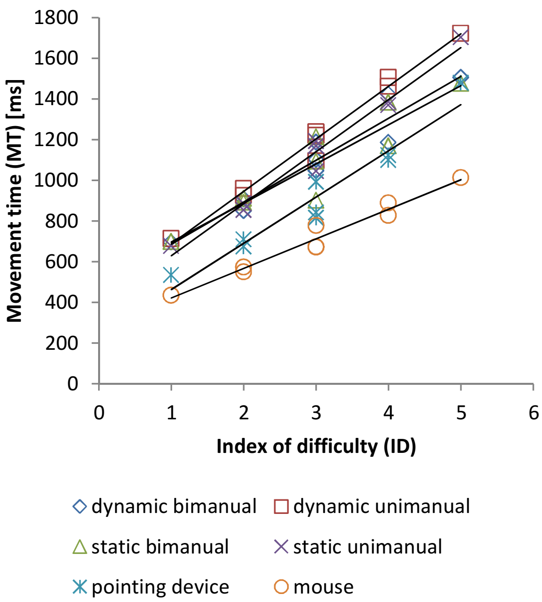


Figure 5. Linear regression between the pointing time (MT) and index of difficulty (ID)

Table 2. Parameters of Fitt's law model for each method.

	A	B	correlation (r)
dynamic bimanual	481.7	205.9	.962
dynamic unimanual	430.5	258.2	.990
static bimanual	504.2	192.3	.926
static unimanual	373.9	255.8	.988
pointing device	235.4	227.5	.968
mouse	275.5	145.6	.981

5.4 Subjective results

The participants were asked to rate all the methods in terms of perceived speed, comfort and accuracy on a five-item Likert scale (1 – certainly yes, 5 – certainly no). The median values of collected data are depicted in the Figure 6. Static interfaces were rated better than dynamic ones and they were also rated better than the pointing device. However, the mouse received the best scores among all the methods.

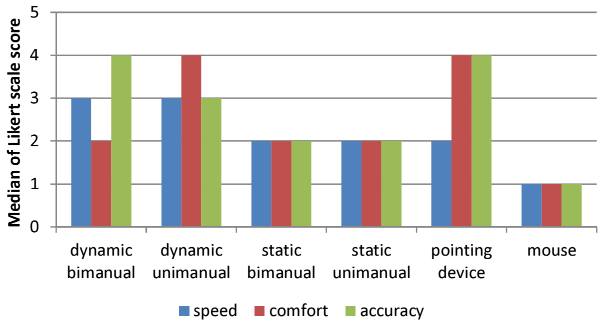
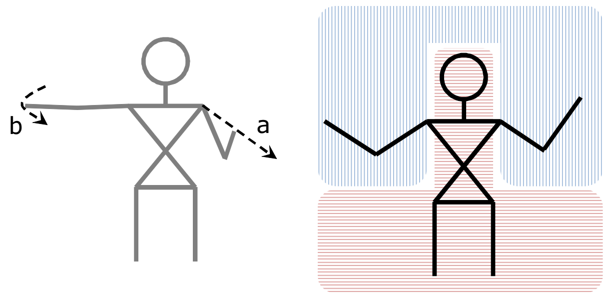


Figure 6. Subjective scores

Free-hand pointing was quite intuitive for all participants but the click operation was found cumbersome. As already mentioned, click was triggered by moving palm forward or backward the screen by 3 inches. This was found difficult to perform. We observed that the instruction "*move your palm forward*" was interpreted by the participants as palm movement in a straight line defined by shoulder and current position of palm (see Figure 7a). When the arm is stretched, the meaning of the command is to rotate the arm forward (or backward) in the shoulder.



a. **b.**

Figure 7. a. Meaning of command "move the hand forward" for bent (a) and stretched (b) arm positions. **b.** Regions suitable (blue vertical lines) and unsuitable (red horizontal lines) for interaction.

We also asked the participants to comment comfortable and uncomfortable hand positions when using free-hand pointing. The regions suitable for interaction are depicted in Figure 7b by blue horizontal lines. The less comfortable positions are shown in red vertical lines. Almost all participants complained about interaction in front of their head as they did not see the screen properly. Six participants did not like to interact in front of their body, four of them complained about lower positions of the hand. On the other hand, they mostly appreciated positions besides their body and higher positions. In unimanual interaction they complained about interaction on the other side of their body.

6. CONCLUSION

The comparative study has shown that free-hand interaction with dynamic interfaces is slower and less accepted by the users than static interfaces. Even though remote pointing device was faster than free-hand pointing, it was less accepted probably due to its low precision. The study has also shown that in bimanual interaction, the users rely more on dominant hand and they switch hands only when necessary. We have also identified suitable regions for free-hand interaction.

We believe that these results can be applied for future designs of user interfaces for both Interactive TV and public large displays. In the future work we would like to aim at finding actual usability of a dynamic interface, not only its performance.

7. ACKNOWLEDGMENTS

This work has been supported by the EU FP7 project TA2: Together Anywhere, Together Anytime ICT-2007-214793, grant no 214793, and the grant OHK3-048/10 Automatically generated user interfaces in nomadic applications.

8. REFERENCES

[1] Bolt, R. 1980. "Put-that-there": Voice and gesture at the graphics interface. Proc. of SIGGRAPH '80, ACM Press, 262-270.

[2] MacKenzie, I.S. 1995. Movement Time Prediction in Human-Computer Interfaces. Readings in human-computer interaction (2nd ed.), Morgan Kaufmann, 483-493.

[3] McArthur, V., Castellucci, S. J., MacKenzie, I. S. 2009. An empirical comparison of "wiimote" gun attachments for pointing tasks. Proc. of EICS'09, ACM Press, 203-208.

[4] MacKenzie I.S., Juosh, S. 2001 An Evaluation of Two Input Devices for Remote Pointing. In Engineering for Human-Computer Interaction, Springer-Verlag, 235-250.

[5] Riek, S., Tresilian, J.R., Mon-Williams, M., Coppard, V.L., Carson, R.G. 2003. Bimanual aiming and overt attention: one law for two hands. Experimental Brain Research, vol. 153, Springer-Verlag, 59-75.

[6] Bailly, G., Walter, R., Müller, J., Ning, T., Lecolinet, E. 2011. Comparing Free Hand Menu Techniques for Distant Displays Using Linear, Marking and Finger-Count Menus. Proc. of INTERACT'11, Springer-Verlag, 248-262.

[7] Cesar, P., Chorianopoulos, K. 2009. The Evolution of TV Systems, Content, and Users Toward Interactivity. Foundations and Trends in Human–Computer Interaction, vol. 2, Now Publishers, 373-395.

[8] Clark, A., Dunser, A., Billinghurst, M., Piumsomboon, T., Altimira, D. 2011. Seamless Interaction in Space. Proc. of OzCHI '11, ACM Press, 88-97.

[9] Freeman, W. T., Weissman, C. 1995. Television control by hand gestures. International Workshop on Automatic Face- and Gesture- Recognition, IEEE Computer Society,179-183.

[10] Schick, A., Camp, F., Ijsselmuiden, J., Stiefelhagen, R. 2009. Extending touch: towards interaction with large-scale surfaces. Proc. of ITS '09, ACM Press, 117-124.

[11] Parker, K., Mandryk, R. L., Inkpen, K. M. 2006. Integrating Point and Touch for Interaction with Digital Tabletop Displays. IEEE Computer Graphics and Applications, vol.26, IEEE Computer Society, 28–35.

[12] Bi, X., Shi, Y., Chen, X. 2006. uPen: A Smart Pen-like Device for Facilitating Interaction on Large Displays, Proc. of Conference on Horiz. Interactive Human-Computer Systems, IEEE Computer Society, 160-168.

[13] Vogel, D., Balakrishnan, R. 2005. Distant freehand pointing and clicking on very large, high resolution displays. Proc. of UIST'05, ACM Press, 33-42.