Comparing Gestures and Traditional Interaction Modalities on Large Displays

António Neto and Carlos Duarte

LaSIGE, Faculty of Sciences of the University of Lisbon,
Campo Grande, Lisboa, Portugal
aneto@lasige.di.fc.ul.pt, cad@di.fc.ul.pt

Abstract. Interfaces based on gesture recognition offer a simple and intuitive alternative to the use of traditional menus, keyboard and mouse. In this paper we explore the field of gestural interaction on large screen displays, presenting the results of a study where users employ gestures to perform common actions in various applications suited for large displays. Results show the actions for which gestural interaction is a best asset compared to traditional interaction.

Keywords: Gestures, Large Displays, Touch Screens, User-centered Study.

1 Introduction

The importance of gestures, despite being a central element of communication between people, has been neglected in traditional interfaces. Nowadays, with the possibilities opened up by the recent "touch revolution" and the transition we have been witnessing in the past few years to large screen displays, we are now able to explore how the use of gestural interaction can contribute to overcome the problems with the WIMP (Window, Icon, Menu, Pointer) paradigm in large screen displays.

There are two different types of research on interaction techniques for large displays. Those who seek to adapt the interaction techniques of the WIMP design paradigm to large displays and those who innovate and break from this classic paradigm. Some examples of adapted techniques are High density cursor [1], which helps users keep track of the mouse cursor, and Drag-and-pop [3], an extension of the traditional drag-and-drop. Some of the innovative techniques are Barehands [5], a free-handed interaction technique, where the user can control the invocation of system commands and tools on a touch screen by touching it with distinct hand postures, and Shadow Reaching [7], an interaction technique that makes use of perspective projection applied to a shadow representation of a user.

The field of gesture interaction has also been explored by many researchers. For example, Segen and Kumar presented Gesture VR [6], a system using vision tracking of hand gestures for spatial interaction; Cao and Balakrishnan [2] describes a gesture interface using a wand as a new input mechanism to interact with large displays.

Our current work aims at exploring the field of gestural interaction on large screen displays. For this purpose we have conducted a study where users used gestures for common actions in everyday applications, while interacting with a SMARTBoard, a large touch screen that works with a projector and a computer, and is used in face-to-face or virtual settings in education, business and government scenarios.

This paper contributions include results of gestures usage for typical actions on today's computer applications, a characterization of the applications and actions that makes them more or less suited for gestural interaction, and recommendations for gestural interfaces' designers.

In this study we used a set of gestures based in a previous study [4] that was aimed at defining a set of gestures for certain actions that can be used in non multi-touch large surface interaction scenarios.

2 Study

One of the goals of this study was to find out, in an interaction context with a single-touch large screen, which applications could benefit more from the use of gestural interaction, and which actions within those applications are able to be carried out through gestures.

Ten persons, with an average age of 24 years, participated in the study. All of the participants were regular computer users, with previous experience of the applications used in the study, and no prior experience with large displays or gestural interaction.

The experiment was conducted on a front projection SMARTBoard, with 77" (195.6 cm) active screen area, connected to a laptop PC running Windows XP SP3 with screen resolution of 1400 by 1050 pixels. The SMARTBoard supports tracking a single point in its interaction surface.

We chose four scenarios: (1) object and (2) windows manipulation, (3) image visualization, and (4) Google Earth. These everyday life scenarios are representative of settings that might benefit from the use of gestural interaction on large displays (e.g. they all can be envisioned in the context of meetings and brainstorm sessions).

Afterwards, for each scenario, we specified the actions that could be achieved by gestures, and finally, using the gestures previously defined for each action [4], implemented a gesture based interface.

Every user participated in an individual session. At the beginning of each session the purpose of the study was explained to the participants. The participants were asked to perform all actions with gestures (without being told what the "correct" gesture was) and without gestures (using the virtual keyboard and virtual mouse). All interaction was performed directly on the SMARTBoard's surface.

During each session the number of attempts and the time that the participants took to find out the right gesture for each action was measured. The timeout given to the participants to discover the gesture was one minute. After that time, if they hadn't discovered it, the gesture was exemplified.

At the end of the sessions, participants filled a questionnaire where they were asked to rate the adequateness of each action when performed with gestures and without gestures (in a 5 point scale, with 1 being the least adequate and 5 the most).

The tables bellow present, for each scenario, the action to be performed, a representation of the corresponding gesture (arrows next to the gesture indicate the direction of the movement), the average number of attempts and the average time in seconds it took participants to find out the right gesture for each action. The next two columns of the table are the questionnaire ratings, representing how adequate users feel each modality (M – mouse and keyboard, G – gesture) is to perform each action. The final column, presents the results of paired-samples t-tests comparing the adequacy of both modalities for performing the required action. Results are statistically significant for t above 2.262 (shown in bold).

Table 1. Object Manipulation Scenario

Action	Gest	#	T	M	G	t
NewFolder		**2.4**	**23.3**	3.2	4.2	**3.35**
Delete		**1.3**	**8.2**	3.4	4.0	1.76
Copy		**1.2**	**8.0**	3.3	4.3	**2.74**
Paste		**1.1**	**6.0**	3.3	4.0	2.09
Cut		**5.3**	**55.0**	3.4	2.2	-2.57
Compress		**1.3**	**8.2**	3.0	4.1	2.09

Table 2. Google Earth Scenario

Action	Gest	#	T	M	G	t
Find		**1.5**	**10.9**	3.0	3.2	0.56
Placemark		**3.5**	**21.9**	3.4	3.8	1.31
Rotate		**1.1**	**6.0**	3.4	4.7	**4.99**
Zoom		**1.2**	**5.8**	3.8	4.4	1.33
Mode 3D		**1.9**	**9.8**	2.9	4.5	**2.95**

Table 3: Image Visualization Scenario

Action	Gest	#	T	M	G	t
Next/Prev		**1.3**	**6.9**	3.5	4.8	**3.54**
Zoom		**1.5**	**7.4**	3.5	4.5	**2.53**
Rotate		**1.1**	**4.4**	3.6	4.4	**3.21**
Print		**1.7**	**12.5**	3.9	4.1	0.56

Table 4: Windows Manipulation Scenario

Action	Gest	#	T	M	G	t
Minimize		**2.0**	**17.2**	3.6	4.5	2.21
Maximize		**2.2**	**19.2**	3.6	4.4	2.23
Close		**1.3**	**6.9**	3.8	3.8	0.00

3 Discussion

Based on the average time and number of attempts that participants took to find the gesture for each action, we can conclude, as expected, that these two factors are connected. The actions for which participants took less time were also the ones that needed fewer attempts to discover the right gesture ('rotate', 'zoom', 'next/prev', 'copy' and 'paste'). For the actions participants took more time they also needed more attempts to find the right gesture ('new folder', 'cut' and 'placemark').

Analysing the questionnaires' results, we can conclude that the image visualization scenario is the one with best results for gestural interaction. It can be observed that participants preferred gestural interaction when the gestures are a more direct fit and more closely related to real world interaction. On the image manipulation scenario, actions preferred are performed on the image itself as if they were done on printed photos. In the Google Earth scenario, the interactions preferred are direct, as if the user was handling non-digital objects, like an earth globe or a map.

We can verify that there is a direct relation between the factors measured in each session, time and number of attempts, and the questionnaires' results. In general, the actions for which participants discover the gesture quicker and in fewer attempts are also the ones for which gestural interaction had best results in the questionnaires, and vice-versa.

Most of the results of this paper are in line with the results of previous studies [4]. Some of these results are concerned with the most intuitive actions to be performed using gestures and with the scenarios where gestural interaction is preferred by participants. One exception is the questionnaire results concerning the windows manipulation scenario. In previous studies participants preferred the traditional interaction to manipulate windows, which was not verified in this study. We think this difference may be due to the fact that in previous studies, participants didn't try to make these actions with traditional interaction on large screens, where despite being easily triggered by buttons, these buttons are usually small and poorly positioned in the interaction surface, which leads to gestural interaction being a better option, once it's an 'anywhere' interaction.

The classic WIMP paradigm was not originally designed for large screens or for systems without common interaction modalities such as mouse and keyboard. The use of gestural interaction does not replace these interfaces but could, if well implemented, minimize the problems and limitations introduced by their absence and improve the user interaction. One example where the direct manipulation afforded by WIMP interfaces may not be the easier option is when a user tries to delete a file. With direct manipulation the user drags the file to the trash bin. This may be the easier option when the interaction surface is small, the file is close to the trash bin, and the trash bin is not covered by a window. Performing a gesture over the iconic representation of the file is easier when these conditions are not met.

This is part of a set of studies towards the goal of a better use of gestural interaction on large screens. We plan to hold further studies with support for multiple surfaces and multi-touch technology, allowing the cooperation between multiple users on different surfaces. In the future we will develop a prototype, based on the results of this and further studies, to explore the possibilities of open cooperation between multiple surfaces through gestural interaction.

References

1. Baudisch, P., Cutrell, E., Robertson, G., 2003. High-Density Cursor: A Visualization Technique that Helps Users Keep Track of Fast-Moving Mouse Cursors. In *Proc. INTERACT*, p. 236-243, ACM Press.
2. Cao, X., Balakrishnan, R., 2003. VisionWand: Interaction Techniques for Large Displays using a Passive Wand Tracked in 3D. In *Proc. UIST*, p. 193-202, ACM Press.
3. Collomb, M., Hascoet, M., Baudisch, P., Lee, B., 2005. Improving drag-and-drop on wall-size displays. In *Proc. GI*, p. 25-32, ACM Press.
4. Neto, A., Duarte, C., 2009. A Study on the Use of Gestures for Large Displays. In *Proc. ICEIS*, Springer.
5. Ringel, M., Berg, H., Jin, Y., Winograd, T., 2001. Barehands: implement-free interaction with a wall-mounted display. In *Proc. CHI Extended Abstracts*, p. 367-368, ACM Press.
6. Segen, J., Kumar, S., 2000. Look ma, no mouse! In *Commun. ACM*, 43(7). p. 102-109, ACM Press.
7. Shoemaker, G., Tang, A., Booth, K., 2007. Shadow reaching: a new perspective on interaction for large displays. In *Proc. UIST,* p. 53 -56, ACM Press.