A STUDY ON THE USE OF GESTURES FOR LARGE DISPLAYS

António Neto and Carlos Duarte
LaSIGE, Faculty of Sciences of the University of Lisbon, Campo Grande, Lisbon, Portugal

Keywords: Gestures, Large Displays, Interaction Techniques, Touch Screens, User-centered Studies.

Abstract: Large displays are becoming available to larger and larger audiences. In this paper we discuss the interaction challenges faced when attempting to transfer the classic WIMP design paradigm from the desktop to large wall-sized displays. We explore the field of gestural interaction on large screen displays, conducting a study where users are asked to create gestures for common actions in various applications suited for large displays. Results show how direct manipulation through gestural interaction appeals to users for some types of actions, while demonstrating that for other types gestures should not be the preferred interaction modality.

1 INTRODUCION

Nowadays, with pixels getting cheaper, computer displays tend toward larger sizes. Wall sized screens and other large interaction surfaces are now an option for many users and this trend raises a number of issues to be researched in the user interface area. The simplistic approach of transferring the main interaction concepts of the classic WIMP (Window, Icon, Menu, Pointer) design paradigm, based on the traditional mouse and keyboard input devices, quickly led to unexpected problems (Baudisch, 2006).

If we put together the possibilities opened up by the recent "touch revolution" and the transition we have been witnessing for 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 classical WIMP paradigm in large screen displays.

Gestures are a central element of communication between people and its importance has been neglected in traditional interfaces. Interfaces based on gesture recognition offer an alternative to the use of traditional menus, keyboard and mouse, paving the way to different approaches to "object manipulation". Being able to specify objects, operations or other parameters with a simple gesture is intuitive and has been the focus of much research. Some results have already been seen in the design of interfaces for physically challenged people and for commercial products such as pen-based tablet computers, PDAs and smartphones.

In this paper, we describe the problems of using the classic WIMP paradigm on large wall sized displays, and the advantages and disadvantages of using gestural interaction to overcome these problems. Our main objective is to explore the field of gestural interaction on large screen displays. To this end we conducted a study where users were asked to create gestures for common actions in everyday applications and scenarios, while interacting with a SMARTBoard, a large touch-controlled 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.

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

2 RELATED WORK

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.

Neto A. and Duarte C. (2009).
A STUDY ON THE USE OF GESTURES FOR LARGE DISPLAYS.
In *Proceedings of the 11th International Conference on Enterprise Information Systems - Human-Computer Interaction*, pages 55-60
DOI: 10.5220/0001986600550060

2.1 Adaptations of WIMP

On large screens, higher mouse accelerations are used in order to traverse the screen reasonably quickly. High density cursor (Baudisch et al, 2003) helps users keep track of the mouse cursor by filling in additional cursor images between actual cursor positions. Drag-and-pop (Collomb et al, 2005) is an extension of the traditional drag-and-drop. This technique provides users with access to screen content that would otherwise be hard or impossible to reach. As the user starts dragging an icon towards its target, drag-and-pop responds by temporarily moving all potential target icons towards the current cursor location. Snap-and-go (Baudisch et al, 2005) is a technique that helps users align objects and acquire very small targets. With traditional snapping, placing an object in the immediate proximity of a snap location requires users to temporarily disable snapping to prevent the dragged object from snapping to the snap location.

2.2 Innovative Interaction Techniques

Unlike what happens with the adaptations of the traditional techniques of the WIMP paradigm, some research projects explore innovative interaction techniques. Barehands (Ringel et al, 2001) describes a free-handed interaction technique, in which the user can control the invocation of system commands and tools on a touch screen by touching it with distinct hand postures. Shadow Reaching (Shoemaker et al, 2007) is an interaction technique that makes use of perspective projection applied to a shadow representation of a user. This technique was designed to facilitate manipulation over large distances and enhance understanding in collaborative settings. The "Frisbee" (Khan et al, 2004) is a widget and an interaction technique for interacting with areas of a large display that are difficult or impossible to access directly.

The advantages of these techniques are that they are specially built for large displays, not inheriting the classic paradigm problems. The main disadvantage is the time spent by users getting used to these techniques.

2.3 Gesture Interaction Techniques

Gesture interaction has been explored by many researchers. Hover Widgets (Grossman et al, 2006) creates a new gesture command layer which is clearly distinct from the input layer of a pen interface. Cao and Balakrishnan (2003) describe a gesture interface using a wand as a new input mechanism to interact with large displays, and demonstrate a variety of interaction techniques and visual widgets that exploit the affordances of the wand, resulting in an effective interface for large scale display interaction. Epps et al (2006) present the results of a study that offers a user-centric perspective on a manual gesture-based input for tabletop human-computer systems, and suggestions for the use of different hand shapes in various commonly used atomic interface tasks. Rekimoto (1998) explores the variety of vocabulary that gestures provide. While the mouse has a limited manipulation vocabulary (e.g. click, double click, click and drag, right click), hands, fingers and gestures provide a much more diverse one (e.g. press, draw a circle, point, rotate, grasp, wipe, etc).

As new computing platforms and new user interface concepts are explored, the opportunity for using gestures made using pens, fingers, wands, or other path-making instruments is likely to grow, and with it, interest from user interface designers in using gestures in their projects.

3 MOTIVATION AND GOALS

Through the participation on a project aimed at managing group therapy through multiple devices (Carriço et al, 2007) we managed to conduct simulations of group therapy sessions involving, amongst other devices, one SMARTBoard. During those sessions we noticed that, with a persistent regularity, people emphasized the difficulties of interacting with the SMARTBoard. These complaints were mainly due to the lack of a traditional physical input channel, such as a mouse or a keyboard. This motivated us to undertake this study. We believe the lack of traditional inputs may be minimized with proper gestural interaction, which would increase productivity and usability, and could even correct some of the classic WIMP paradigm gaps in large screens. Accordingly, we aimed at exploring the possibilities offered by touch interfaces, particularly on large screens.

Initially we focused on gestural interaction with the surface of the device. We conducted a study to define a set of gestures for certain actions that can be used in single surface, non multi-touch, interaction scenarios. The main goal was to study the advantages of using gestural interaction on large screens, in opposition to the standard WIMP paradigm interaction techniques. Additionally, we also aimed at determining which scenarios and actions are more adequate to gestures. One of the objectives was to define a set of gestures, in order to solve some of the existing problems with the standard WIMP interaction technique in large interactive displays, envisioning the creation of a prototype that mitigates the mentioned problems and limitations.

4 STUDY

This study tried to find out, in an interaction context with a non multi-touch large screen, which applications could benefit from the use of gestural interaction, which actions within those applications are able to be carried out through gestures and what gestures are most appropriate, intuitive, and comfortable for each action.

We began by defining a set of scenarios that could benefit from gesture interaction on large screens. Afterwards, for each scenario, we specified the main actions that could be achieved by gestures, and finally created a procedure with several tasks for each scenario, so the users could create a gesture that they believed to make sense for each one of these tasks. Based on the information gathered from the several gestures registered along the process, we tried to find gestural patterns in order to justify the adequateness of the gestures created for each action within the defined scenarios.

Twelve participants took take part in the experiment. The average participant age was 24 years. All of them were regular computer users, with previous experience of the applications used in the study scenarios. None of the participants had 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.

4.1 Scenarios and Action Set

We chose five scenarios (figure 1): object manipulation on the desktop, windows manipulation on the desktop, image visualization, multimedia player, and 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).

We left out some applications, such as word processors or spreadsheets, because, besides being highly dependent on keyboard operation, they also are not suitable applications for large screen scenarios. We did not choose an internet browser as a scenario because gestural interaction with this kind of application is already used to navigate and manage windows in the main browsers such as Mozilla Firefox, Opera or Internet Explorer.

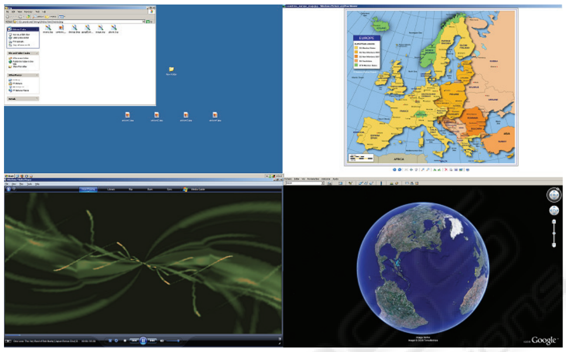

Figure 1: The 5 Scenarios, (1) Objects and (2) windows manipulation on the Desktop (top left), (3) Image Visualization (top right), (4) Multimedia Player (bottom left), (5) Google Earth (bottom right).

We have specified a set of actions that seemed the most common for each scenario, taking into account the large screen context. In the case of object manipulation: create a folder, delete, copy, paste, cut, move, compress and print; in the windows manipulation scenario: minimize, maximize, restore, and close; in the image visualization scenario: zoom in, zoom out, next, previous, rotate clockwise, rotate counter-clockwise, and print; in the multimedia player scenario: play, pause, stop, next, previous, volume up, volume down and mute; in the Google Earth scenario: zoom in, zoom out, rotate clockwise (cw), rotate counter-clockwise (ccw), tilt up, tilt down, find, and placemark.

4.2 Procedure

During the whole study, the participants were given total interaction freedom to execute the gestures. In other words, there were no explicit restrictions on the creation of the gestures for every action considered within each scenario, besides not being able to use multi-touch interaction. This limitation is imposed by the SMARTBoard capabilities. However, since SMARTBoards are extensively used in a multitude of real life scenarios, we do not feel this makes the results less relevant.

Every user was involved in an individual session on which he/she participated in the five possible scenarios. At the beginning of each session, the purpose of the study was explained to the participants. The participants were given a paper sheet with all the tasks to be performed during the procedure, ordered by the five scenarios, and asked to create gestures to complete each task. The paper sheet's role was two-folded. On one hand it was the list of actions to perform, and on the other it worked as an implicit restriction to hand usage together with the pen provided to the participants, attempting to avoid multi-touch gesture interaction. No time limits were imposed to the participants.

The SMARTBoard captured and registered the gestures made by each participant. Participants were also asked to describe the gesture and to reason about it. The sessions were filmed and snapshots of each gesture were registered. At the end of each session participants filled a questionnaire where they were asked to rate the adequateness of each action when performed with the mouse, the keyboard, and the gesture they chose (in a 5 point scale, with 1 being the least adequate, and 5 the most adequate and intuitive interaction).

5 RESULTS

A total of 420 gestures were performed in 12 sessions, averaging 35 gestures per session. Participants took an average of approximately 29 minutes per session, with 45 seconds per gesture. The remaining time was used for explanations and questions. In the following paragraphs we first discuss the set of actions for which at least 50% of the participants made similar gestures and, afterwards, we discuss the actions for which the participants failed to agree on appropriate gestures.

5.1 Actions with Similar Gestures

Table 1 presents, for each scenario, the action to be performed, a representation of the gesture made by the majority of the participants, the percentage of participants who made it and the time it took participants to think up and draw the gesture. The last 3 columns of the table are the questionnaire ratings, representing how adequate users feel each modality is to perform each action.

The actions performed by the participants can be grouped into two categories, according to their execution complexity in the WIMP paradigm. The first category groups the less complex actions, which are the ones achieved through a simple mouse action or key press (e.g. "move" or "delete" in the object manipulation scenario). The second category groups the more complex actions, where the user has to navigate through several menus or perform complex combinations of key presses (e.g. "print" in the image visualization scenario or "placemark" in the Google Earth scenario). It can be seen from the results that actions in the first category are ones where users have a bigger agreement on the gesture to employ, and also a shorter creation time. The second category actions have, correspondingly, higher response times, and users fail to reach consensus on what gestures can best represent them.

Table 1: Gestures agreed by the study participants.

Action	Gesture	%	Time	Mouse	Key	Gest
Object manipulation						
New Folder		75%	0:45	4	2	5
Delete		92%	0:10	4	5	4
Copy		50%	0:28	4	4	4
Move		100%	0:14	4	2	4
Compress		67%	0:35	3	2	4
Window manipulation						
Minimize		67%	0:35	5	2	4
Maximize		67%	0:14	5	2	4
Close		92%	0:15	5	4	4
Image manipulation						
Zoom In		92%	0:30	4	3	5
Zoom Out		92%	0:10	4	3	5
Next		75%	0:15	4	5	4
Previous		75%	0:07	4	5	4
Rotate cw		83%	0:08	4	2	5
Rotate ccw		83%	0:06	4	2	5
Media player						
Play		92%	0:10	4	3	4
Pause		92%	0:15	4	3	4
Stop		92%	0:13	4	3	4
Next Item		100%	0:08	4	3	4
Previous		100%	0:05	4	3	4
Vol. Up		75%	0:18	4	3	3
Vol. Down		75%	0:06	4	3	3
Mute		50%	1:30	4	3	4
Google Earth						
Zoom In		75%	0:10	4	2	4
Zoom Out		75%	0:08	4	2	4
Rotate cw		92%	0:10	3	2	5
Rotate ccw		92%	0:07	3	2	5

Actions like next/previous, zoom in/out, rotate cw/ccw, all have the same gesture mapping on the five scenarios, which lead us to believe they are standard gestures valid across different applications. Many of the gestures chosen by the participants resulted from their experience with traditional interactions. Take for instance the example of the gestures chosen for the multimedia player scenario, in this case the participants based their gestures in a classic remote control of a multimedia device, like a DVD player.

5.2 Actions with Dissimilar Gestures

For some actions, we could not find a gesture pattern. Table 2 includes these actions, together with the corresponding questionnaire results. Figure 2 present some of the gestures the study participants drawn for some of these actions.

Table 2: Actions without a defined gesture pattern.

Scenario	Action	Mouse	Key	Gest
Object Manipulation	Paste	4	4	4
	Cut	4	4	3
	Print	3	4	3
Window Manipulation	Restore	5	2	3
Image Manipulation	Print	3	3	3
Google Earth	Find	4	3	3
	Tilt Up	4	2	4
	Tilt Down	4	2	4
	Placemark	4	3	4

Actions with lower agreement on the adequate gesture are also the ones where the participants spent more time. This leads us to believe these actions are less intuitive and possibly not adequate to accomplish by gestural interaction.

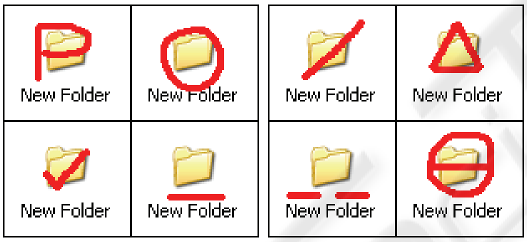

Figure 2: Examples of gestures used for the action Paste (left) and Cut (right), on the Object Manipulation Scenario. The gesture was made over the folder the document should be pasted into.

6 DISCUSSION

Despite the novelty of using gestures as an interaction method, according to the questionnaires' results, some of the actions were more intuitive to perform through gestures than through the more familiar keyboard and mouse interaction. The results of the questionnaires show that, for some actions, users prefer gestures, especially in the object manipulation, image visualization and Google Earth scenarios. In the following paragraphs we put forward several considerations which might contribute to explain this phenomenon, and discuss other aspects relevant to the adoption of gesture based interaction on large interactive surfaces.

Direct manipulation moved interfaces closer to real world interaction by allowing users to directly manipulate the objects presented on screen rather than instructing the computer to do so by typing commands or selecting menu entries. Gestural interaction techniques can push interfaces further in this direction. They increase the realism of interface objects and allow users to interact more directly with them, by using actions that correspond to daily practices within the non digital realm, and thus felling more realistic and intuitive. In the three scenarios where participants preferred gestural interaction over the mouse and keyboard interaction, gestures are a more direct fit and more closely related to real world interaction. On the object manipulation scenario the interaction is performed directly on the object's graphical representation. On the image manipulation scenario, actions are performed on the image itself as if they were done on printed photos. In the Google Earth scenario, the interactions are direct, as if the user was handling non-digital objects, like an earth globe or a map.

There are actions for which gestures seem to be the most appropriate interaction technique. By mapping gestures to every day control commands which are used in several applications, users save time that is traditionally required to select these commands in menus or to remember key combinations. However, according to the information on the questionnaires and to the user comments registered along the different sessions, if the action is triggered by a simple button, such as the mute volume of a media player, unless the button is too small, or poorly positioned on the interaction surface, the gestural interaction, even if the gesture is very simple, does not represent an advantage over the traditional approach.

For analogue controls, such as volume increase and decrease, gestural interaction is a good alternative to discrete buttons. A gesture, in addition to representing an action, can represent the intensity of this action. Take for instance the case of rotating an image in the image manipulation scenario, where the angle of rotation will equal the gesture's angle.

Through the questionnaires' results it is possible to infer what modalities are favoured by the participants for the different actions and scenarios. Keyboard is only favoured for the actions where a shortcut is available, like cut, copy, paste, delete and print, independently of the application. This can be interpreted as an indication of users favouring speed of execution in the interface. The mouse seems to be preferred for actions which involve direct manipulation of an interface control. The actions which users prefer to accomplish with a mouse include all actions of the multimedia player and the window manipulation scenarios and the action find in the Google Earth scenario. The common factor of these actions is that they are triggered by a simple mouse click over a button or an icon. Finally, gestures were preferred when users needed to interact directly with an object. This was best saw in the image manipulation and the Google Earth scenarios. When users needed to manipulate the images or the 3D globe, through rotation, zooming or tilting, they felt gestures were the more adequate interaction mean. This is also the case for object moving in the object manipulation scenario.

7 CONCLUSIONS

In this paper we presented a study conducted in order to understand what actions could benefit from the addition of gesture based interaction on large displays. Additionally, the study aimed at defining a set of gestures for certain actions that can be used in non multi-touch large surface interaction scenarios, by allowing the users creative freedom to design such gestures.

The results of this study show that gestural interaction can solve some of standard WIMP paradigm problems on large screens, especially for some actions within scenarios where large screens are typically used. The study's results also show that it is a mistake to assume that gestural interaction is a good solution to trigger all actions. This is corroborated by the low gestural agreement results found for several actions, which leads us to believe that these actions are less intuitive and inadequate to accomplish by gestures.

The classic WIMP paradigm was not originally designed for large screens or for systems without common interfaces 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.

This is the first step towards a better use of gestural interaction on large screens. We plan to hold further studies with support for multiple surfaces and multi-touch technology, creating an additional set of gestures that allows 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

Baudisch, P., 2006. Interacting with Large Displays. In *IEEE Computer*, 3(39), p. 96-97, IEEE Press.

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.

Baudisch, P., Cutrell, E., Hinckley, K., Eversole, A., 2005. Snap-and-go: Helping Users Align Objects without the Modality of Traditional Snapping. In *Proc. CHI*, p. 301-310, ACM Press.

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.

Carriço, L., Sá, M., Duarte, L., and Carvalho, J., 2007. Managing Group Therapy through Multiple Devices. Human-Computer Interaction. In Proc. HCII, p. 427-436, LNCS 4553/2007, Springer.

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.

Epps, J., Lichman, S., Wu, M., 2006. A study of hand shape use in tabletop gesture interaction. In *CHI Extended Abstracts*, p. 748-753, ACM Press.

Grossman, T., Hinckley, K., Baudisch, P., Agrawala, M., Balakrishnan, R., 2006. Hover widgets: using the tracking state to extend the capabilities of pen-operated devices. In *Proc. CHI*, p. 861-870, ACM Press.

Khan, A., Fitzmaurice, G., Almeida, D., Burtnyk, N., Kurtenbach, G., 2004. A remote control interface for large displays, In *Proc. UIST*, p.127-136, ACM Press.

Rekimoto, J., 1998. A multiple device approach for supporting whiteboard-based interactions. In *Proc. CHI*, p. 344-351, ACM Press.

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.

Segen, J., Kumar, S., 2000. Look ma, no mouse! In *Commun. ACM*, 43(7). p. 102-109, ACM Press.

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.