Design of a Portable Gesture-Controlled Information Display

Sebastian Loehmann, Doris Hausen, Benjamin Bisinger, Leonhard Mertl
University of Munich (LMU), HCI Group, Amalienstrasse 17, 80333 Muenchen, Germany
sebastian.loehmann@ifi.lmu.de, doris.hausen@ifi.lmu.de,
bisinger@cip.ifi.lmu.de, mertll@cip.ifi.lmu.de

ABSTRACT

Stopping by at somebody's office can be frustrating if the required person is absent. To offer visitors additional information in this case, we built a gesture controlled public display. We applied a user-centered design approach and as first step evaluated basic parameters, such as desired tracking area and preferred gestures. We incorporated these results in a standalone working prototype and achieved natural, intuitive gestures with a recognition rate above 80%.

Author Keywords

Freehand gestures, public display, user-centered design

INTRODUCTION

Students frequently come to our research lab with question about open topics for their bachelor thesis and visitors come by to chat about our research projects. However, oftentimes we are in a meeting or not in our office at all. Therefore, we propose to install a small public screen at the office door, displaying relevant information about staff members. For interactions with the display, we opted for freehand gestures over touch because this (1) keeps the display clean and without smudge, avoiding the need to constantly clean it; (2) considers hygiene issues and (3) allows us to explore this design space. In this paper, we give first insights into our design process, the definition of a gesture set and a preliminary prototype (see Figure 1). We follow a user-centered approach, involving users early in the development.

RELATED WORK

Gestural interaction with large screens has been introduced by Bolt in the late 1970s to support voice control of user interfaces, concluding "a gain in naturalness and economy of expression" [2].

Figure 1. Prototype of a gesture-controlled information display which can be placed next to office doors.

Two decades later, CHARADE [1] allowed for the manipulation of presentation slides by performing only gestures, underlining the support of direct manipulation of user interface elements. In recent years, several studies on freehand interaction with large displays, such as distant pointing and clicking [8], have been conducted. Wachs et al. [9] provide an overview of pros and cons, tracking technologies and application areas for freehand gestures.

Moreover, several projects focused on smaller information displays on office doors. They offer visitors location and calendar details of the owner, the possibility to leave messages and retrieve private information after authenticating themselves [6]. Cheverest et al. [3] installed small information screens at their office doors and were able to send short messages, which were then visible on their display.

PRE-STUDY: GESTURE SET AND DESIGN SPACE

Following a user-centered approach, we involved users very early in the design process. In a pre-study, before starting the implementation of the system, we observed and interviewed 16 students and staff members while interacting with a screen using freehand gestures.

Setup

In order to avoid possible distractions, we used a picture frame as dummy (see Figure 2), representing the tablet computer we used for later prototyping. First, we asked participants to take the frame and place it on the wall at a height that seemed comfortable for them. Then, we invited them to navigate in a virtual picture gallery by asking to "go to the next picture" or "scroll through the selection of pictures". We requested participants to only use touchless hand and arm movements that they consider appropriate for the current task.

During the study, we recorded all interactions, including the placement of the display dummy, on video from two perspectives. Through post-study video analysis, we were able to measure (1) the distances of the person and the hand to the display, (2) the interaction space used by each participant for each gesture, (3) the actual gestures they used as well as (4) the time they needed to perform each gesture. In order to achieve 'natural' results, we did not inform participants about our intention to take these measurements.

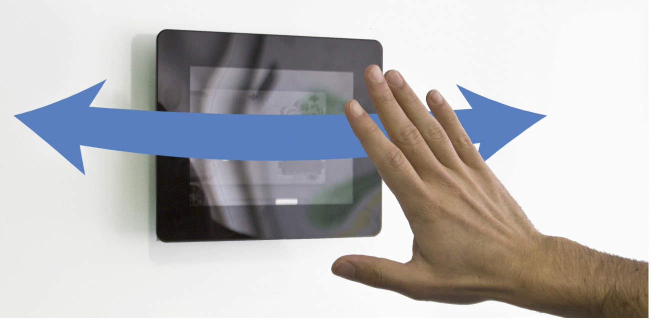


Figure 2. Dummy frame used for the pre-study to evaluate how participants want to interact with such a display.

Results

16 students and researchers from our lab, seven of them female, participated in our study. They were between 22 and 30 years old, with an average of 25 years. Six participants had at least some experience with freehand gestures before the study, mostly due to using the Microsoft Kinect for gaming. Four of them were PhD students from our lab, who dealt with gestures in their research on a theoretical and practical level at some point of their research. All participants owned a smartphone or tablet and used touch gestures daily.

As Figure 3 shows, the average distance between display and upper body was 51 cm. Within this range, 27 cm were actually used to perform gestures, with a minimal distance of 5 and a maximum distance of 32 cm to the display. 87% of all gestures were performed with a minimal distance of 15 cm. While executing gestures, the hand exceeded all four sides of the display by 10 cm on average. The duration of performing a gesture was 247 ms on average (minimum: 150 ms; maximum: 400 ms). Participants placed the display in an average height of 88% relative to their own size.

For both, vertical and horizontal, 94% of the performed gestures were swiping gestures in the corresponding direction. Qualitative results from the interviews confirmed this trend and revealed the relation to swiping gestures used for touch devices.

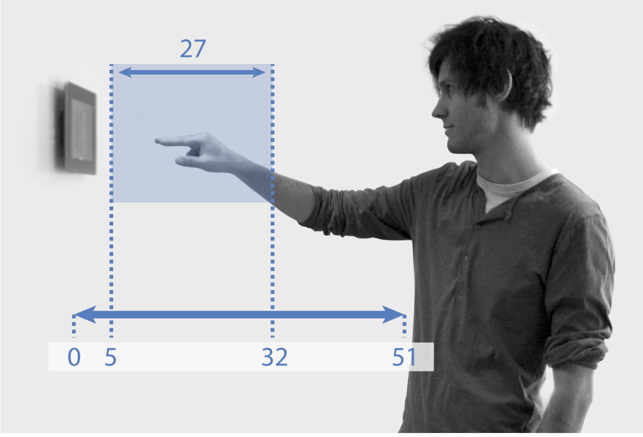


Figure 3. Preferred interaction space as found in the pre-study (in centimeters).

PROTOTYPE

Taking results from the pre-study into account, we built a first prototype. Our goal was a portable stand-alone system without a connection to a computer.

Hardware

The main component of the prototype is a 7" tablet PC running Android. To track hand movements in front of the display, we use Sharp infrared distance sensors, three on each side [5]. We chose to use six sensors to cover the vertical range of performed gestures during the pre-study: some participants performed the gestures right in front of the display, others just on the lower edge (probably trying to not occlude the display contents). To cover the horizontal area in which gestures were performed during the pre-study, we installed the sensors at an angle of about 45° towards the display's center (see Figure 4).

To compensate the jitter of the sensor values, we use capacitors and resistors in terms of hardware as well as a filter in our software.

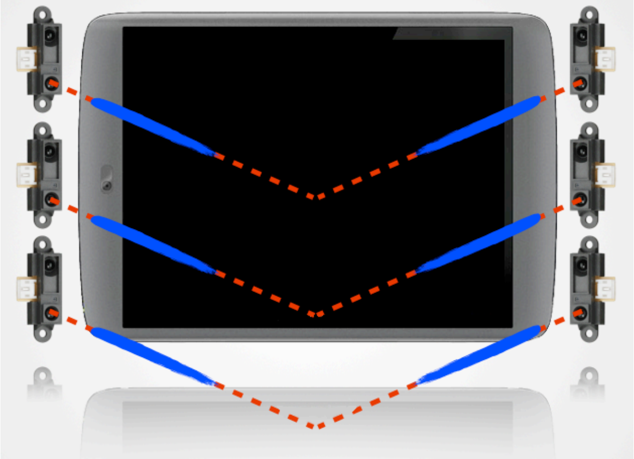


Figure 4. Arrangement of six infrared distance sensors around the display. Sensors are installed at an angle of about 45°.

To read and process the data produced by the sensors, we added an Arduino Mega ADK microcontroller board, which is connected to the tablet via USB. Any Android application can now access the data communicated by the Arduino via the Android Open Accessory protocol.

Housing

Using acrylic glass and a laser cutter, we built a chassis for the hardware. The front covers the display, which can therefore not be touched. A tinted foil, which does not block infrared light, hides the sensors. The Arduino and all wires are attached to the back wall of the housing and are thus not visible either. The only visible chord powers the Arduino and the tablet. Consequently, the prototype is portable and can be placed in arbitrary locations. Prototype can easily be fixed to any wall using power strips.

Software

The sensors can recognize objects that are between 4 and 30 cm away, receiving values between 0 and 650. By using a formula (a non-linear function), the values can be converted into the actual distance.

The Arduino reads values from all six sensors at a frame rate of 25 Hz. For each frame, a vector of six values (one depth sample per sensor) is sent to the tablet. An Android application implements the Dynamic Time Warping [7] algorithm to classify gestures executed in different speeds, as seen in the pre-study. To determine if one of four possible swiping gestures was performed, the DTW compares the data to recorded values of 64 sample gestures, 16 for each gesture, calculating the distance based nearest neighbor. Finally, the application updates the user interface according to the identified gesture.

STUDY: EVALUATING THE IMPLEMENTATION

In order to evaluate the prototype and to collect data for further development, we conducted a user study with 20 students and staff members. Please note that five of the staff members have already participated in the preliminary study. While one could argue that this bias can potentially influence the results, we think that there is no bias because (1) we did not inform participants about the results of the study or the following implementation of the system and (2) in the first study we used a nonfunctional display dummy in contrast to the fully functional interface in this study. Additionally, due to the user-centered approach, it was our purpose to include staff members into both studies as they will be future users of the system.

Setup

We implemented two applications: One showing a basic interface representing three staff members (see Figure 5) and information on their research topics.

Horizontal gestures navigate the staff and vertical gestures navigate the staff's information. We used this app for training purposes and to collect qualitative feedback on the look-and-feel of the prototype. We told participants that the interaction with the display works contactless, but not which gestures they would be able to use. They had now five minutes time to explore the app content.

The second application was implemented to measure the tracking accuracy of the prototype. We displayed arrows and asked participants to perform the swipe gesture in the corresponding direction. Each participant performed each swipe 20 times, for a total of 80 gestures per user, in a randomized order. After each gesture, we gave visual feedback by moving the content out of sight in the direction of the recognized swipe. After the study, we analyzed recorded videos to calculate tracking accuracies and a confusion matrix.

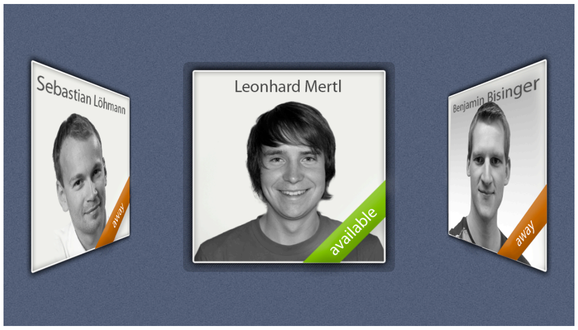


Figure 5. User interface for the first task of the user study. Horizontal swiping: Navigation through staff members. Vertical Swiping: Information about their research projects.

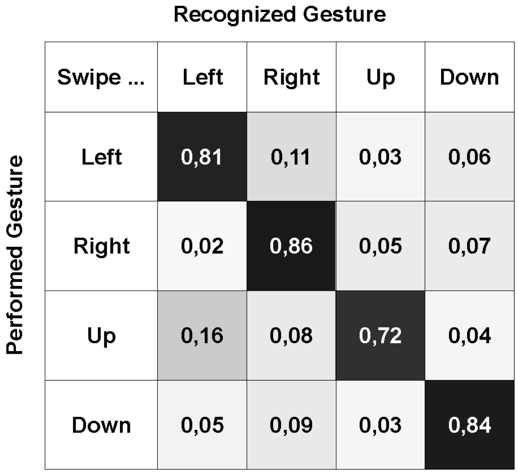
Results

20 students and staff members, seven of them female participated in our study. They were between 20 and 32 years old, with an average age of 25. Two of them were left-handed. Nine participants were familiar with the Microsoft Kinect for gaming.

The analysis of 1600 recorded gestures (20 participants x 4 gestures x 20 trials per gesture) lead to an average recognition rate slightly above 80%. Two participants, using gestural interfaces on a regular basis achieved higher rates of 88% and 90%. Recognition rates for each gesture were 81% for a swipe to the left, 85% to the right, 72% up and 84% down (see Figure 6).

Discussion & Implications

Our design process proved to be valuable, as participants – without instructions – immediately started to interact with the prototype using swiping gestures. We further believe that recognition rates were positively influenced by first exploring a suitable interaction space in front of the display. Looking closer at the average recognition rate for each individual trial from first to last performed gesture, the results improve from 75% to 85%. Thus, shortcomings of the gesture tracking can partly be compensated by the training effect.



Recognized Gesture

Performed Gesture

Swipe ...	Left	Right	Up	Down
Left	0,81	0,11	0,03	0,06
Right	0,02	0,86	0,05	0,07
Up	0,16	0,08	0,72	0,04
Down	0,05	0,09	0,03	0,84

Figure 6. Confusion Matrix showing recognition rates for all four swiping gestures - left, right, up and down.

The relatively low recognition rate of 72% for the up gesture is due to the forearm that is still in the tracked area when the movement is in fact already completed, causing segmentation problems. Gestures that were performed rather fast resulted in false recognition due to the maximum sampling rate of the sensors. Generally, observations made during the exploration of the design space can be in conflict with constraints caused by the tracking technology, indicating a need to find trade-offs between both when prototyping gestural interfaces.

Another interesting observation was the influence of the distance between participants' hands and the display. Similar to the trade-off mentioned above, some users positioned their body unexpectedly far away from the prototype, causing the tracking to fail. Interestingly, they seemed to notice this issue without further clues and started to approach the display until they noticed a reaction of the system. We like to conduct further studies in order to find out how users explore the functionality of such a system. We also conclude that affordances need to be implemented, showing how and which gestures can be performed.

FUTURE WORK

Next steps include the improvement of the hardware. Due to the way we implemented the tracking algorithm, we can reorder the sensors by placing one above and one below the display to receive higher recognition rates, without the need to change the software. Taking new tracking technology into account, we are trying to replace the distance sensors with the Leap Motion Controller. This sensor is suitable for gesture tracking, but the current incompatibility between Leap and Android contradicts our design goal of a standalone device. As another alternative, we investigated the potential of the camera integrated into the tablet. When testing different libraries for gesture tracking using this approach, we noticed that (1) the tracking is unreliable as soon as other body parts (like the user's head) is visible and moving and (2) this kind of tracking becomes rather complex when extending the gesture set beyond swiping and thus causes delays due to the limited computing power of the tablet.

We will add further gestures that came up during the pre-study and are interested in how this influences recognition accuracy with different tracking technologies. Additionally, we will work on a more mature UI and add a wireless connection to the application to be able to add content on the fly. Considering the UI, a major issue will be the 'findability' of the possible gestures. In a new version of the interface, we try small icons (e.g. arrows) on the edges of the screen to indicate, where and how additional content can be obtained. Another challenge is to show first time users that the display can be controlled via gestures instead of touch. One approach is to utilize the distance of the hand to the display: when the hand gets to close, we progressively dim the display until the contents become invisible as soon as the display is touched.

We finally plan a long-term study by installing the prototype at our office door for several months to explore how users interact with the interface and its contents.

REFERENCES

1. Baudel, T., and Beaudouin-Lafon, M. Charade: remote control of objects using free-hand gestures. Comm. of the ACM 36, 7 (1993), 28-35.
2. Bolt, R. A. "Put-that-there": Voice and gesture at the graphics interface. ACM SIGGRAPH Comput. Graph. 14, 3 (1980), 262-270.
3. Cheverst, K., Dix, A., Fitton, D., Friday, A., and Rouncefield, M. Exploring the utility of remote messaging and situated office door displays. MobileHCI, (2003), 336-341.
4. Fitton, D., Cheverst, K., Kray, C., Dix, A., Rouncefield, M., and Saslis-Lagoudakis, G. Rapid prototyping and user-centered design of interactive display-based systems. Pervasive Computing 4, 4 (2005), 58-66.
5. Kratz, S., and Rohs, M. HoverFlow: expanding the design space of around-device interaction. MobileHCI, (2009), 8 pages.
6. Nguyen, D. H., Tullio, J., Drewes, T. and Mynatt, E. D. Dynamic Door Displays. Georgia Institute of Technology GVU Technical Report GIT-GVU-00-30, (2000).
7. Salvador, S., and Chan, P. Toward accurate dynamic time warping in linear time and space. Intelligent Data Analysis 11, 5 (2007), 561-580.
8. Vogel, D., and Balakrishnan, R. Distant freehand pointing and clicking on very large, high resolution displays. UIST, (2005), 33-42.
9. Wachs, J. P., Kölsch, M., Stern, H., & Edan, Y. (2011). Vision-based hand-gesture applications. Comm. of the ACM 54, 2 (2005), 60-71.