Assisting Gesture Interaction on Multi-Touch Screens

Writser Cleveringa[1,2] **Maarten van Veen**[1,2] **Arnout de Vries**[1] **Arnoud de Jong**[1] **Tobias Isenberg**[2]
[1]TNO Groningen [2]University of Groningen
{writser.cleveringa | maarten.vanveen | arnout.devries | arnoud.dejong}@tno.nl isenberg@cs.rug.nl

ABSTRACT
In this paper we present our ongoing work on multi-user touch interfaces with specific attention to assisting gesture interaction by using gesture previews.

Author Keywords
Multi-touch, multi-user interface, information visualization, gesture recognition, usability, use cues.

INTRODUCTION—MULTI-TOUCH RESEARCH AT TNO

TNO is the national research organization in the Netherlands that applies scientific knowledge with the aim of strengthening the innovative power of industry and government. In this paper we briefly introduce two projects from TNO that address collaboration on tabletop devices, after which we present our research on gesture previews. This project aims to increase the usability of gestural interfaces and to support collaboration on multi-touch screens and is conducted together with the University of Groningen.

Multi-Touch Interaction in Control Rooms

One part of TNO's research concerns control room situations. Control rooms are pre-eminently places where large quantities of data streams from diverse sensor networks and sensor sources come together. Important decisions have to be made based on these data streams and critical systems, sometimes even lives depend on these decisions. Our work explores the use of multi-touch technology and data visualizations for future control room designs where multiple users work together.

We are developing a flood control room application for multi-touch tabletop devices. This application is based on a general multi-touch framework which uses Touchlib[1] for finger recognition. Our framework allows developers to create multi-touch software using Windows Presentation Foundation, a graphical subsystem of the .Net framework. Currently the focus lies on developing a collaborative environment in which all participants can present their knowledge and expertise. The main entrance of the application is a shared map that shows relevant information needed for the tasks at hand. Participants can control data filters and simulations and navigate using intuitive gestures and finger movements.

The regional crisis center for the province of Groningen is participating in our research to determine how information overload (caused by a combination of communication channels and sensor sources) can be handled by using information visualization and multi-touch interaction. A scenario was developed for a flood crisis and this will be simulated in a crisis exercise in October 2008. The water board district, emergency services, municipalities, and other related organizations will work together in this flood crisis situation, using a multi-touch table and our interactive visualization tools. This is a real-life practice of a multi-touch screen and one that can make a big difference in people's lives.

In developing this application some aspects of collaboration have surfaced that are worth mentioning. Multi-touch tables do not support the identification of individuals which introduces challenges in working with objects and tools that are presented on the display. A traditional toolbar in which users can select tools and work with them is not applicable in this setting. As a result the application now uses personalized floating tools that can be dragged and used anywhere on the map. A gestural interface could make this process easier. For example, gestures could be used to summon new toolboxes.

Multi-Touch and Serious Gaming for Education

Research at TNO also focuses on innovation in education. To investigate whether the combination of multi-touch interaction and serious gaming can be of added value to the educational sector we are developing a game for children of age 8 to 12 with a Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS). The game will teach these children the basic social skills needed for working together.

PDD-NOS is a disorder from the autism spectrum resulting in difficulties with social interaction [1]. Because of these problems it is very challenging for people suffering from PDD-NOS to work in groups. Starting in 2003, the Dutch educational system has changed from classical teaching to more problem driven education, where students tackle problems in groups. This cooperative learning is hard for children with PDD-NOS and has resulted in a reduced number of children who are able to make the transition from a special elementary school to a normal high school [2]. To make

[1] See http://www.nuigroup.com/touchlib/ .

Figure 1: Children playing our educational game.

this transition more likely again it is very important to teach young children basic collaboration skills.

Current training programs that focus on social interaction are expensive and mostly given in group form. Because these programs are not taught at school the transfer of knowledge learned to the classroom is non-existent. The use of computers could allow inexpensive training at school as part of the curriculum. In addition, according to C. Tenback from RENN4,[2] computers have a tremendous attraction to children with PDD-NOS. The children are more concentrated and better motivated. Also the way tasks have to be executed, according to strict rules, fits these children well. But working at the computer is usually done alone and stimulates individual behaviour. Only one child can control the keyboard and mouse at the same time. A multi-touch screen can solve this problem because it supports multi-user interaction. With the use of a multi-touch screen we are trying to teach children with PDD-NOS to physically work together in a way that is best suited to them.

In collaboration with teachers from the Bladergroenschool[3] and with RENN4 we are developing a game for children. In this game pairs of children have to build a rocket together. The children have to solve math problems to collect parts and inventory for the rocket, count the number of passengers, mix fluids, and defend their rocket against the competition (Figure 1). The game consists of six levels with different types of math problems. While the task is to solve these problems, each of the levels is designed to teach the children a specific skill necessary for working together such as turn taking. The players start at opposite ends of the multi-touch table with their own part of the screen and their own task. But playing through the levels, the tasks merge and the players have to start cooperating to be able to complete them.

We are currently implementing the game which will be tested at a special elementary school in January 2009. Using the data collected from the game and a questionnaire we are hoping to answer the question whether the game can help teach children with PDD-NOS how to collaborate and whether the combination of multi-touch and serious gaming can be valuable to the education sector.

[2]RENN4 is an institution for special education support.

[3]The Bladergroenschool is a special education elementary school.

ASSISTING GESTURE INTERACTION

In both our control room project and our serious gaming project it proves to be challenging to design an intuitive and effective user interface. Traditional WIMP interfaces are designed for single users; yet in our projects we desire simultaneous input from multiple users. A solution that does meet this requirement is a gestural interface. In the past, people have shown that gestural interaction is possible on multi-touch screens [10]. The question arises whether such an interface is easy to use. For instance, users may have to draw an S-shaped gesture to invoke a menu in a given environment. A problem here is that users have to know this gesture in advance, otherwise they will have a hard time figuring out how to invoke the menu because many gestures have no intrinsic meaning. This also leads to difficulties in learning gestures, which depends on the number of available gestures and their complexity. Especially on multi-touch screens gestures can be complex since gestures on these devices are not limited to a single touch. In fact, Morris et al. are experimenting with cooperative gestures, i. e. gestures that can be drawn by multiple users working together [5].

In such environments it is hard for novice users to deduce what gestures are available to them. The only fact that all users, including novices without prior knowledge, know about touch screens is that "they can be touched." This forms the starting point of our research on gesture interaction for multi-touch screens. We focus on improving gesture interaction on touch-enabled surfaces by providing previews of possible gestures or possibilities to complete gestures, both for single-user gestures and collaborative gestures. In this context it is important to first define a gesture on a touch screen. A few people have given limited definitions of what a gesture is [9, 11] but these definitions are often informal and not focused on touch screens. In the next section we present a formal definition of a touch screen gesture. After that we introduce our concept of gesture previews and discuss the benefits of gesture previews in collaborative settings.

Gesture Definition

It is our opinion that a formal definition of a gesture is essential if one wants to store and exchange gestures, develop tools and frameworks that can simplify development of gestural interfaces, or support similar gestural interfaces on a wide range of devices. We define a gesture based on the following two properties: First, every gesture can be represented by a set of measured positions in space. This explains why one cannot perform gestures on a keyboard: a keyboard does not detect motion. Input devices that do detect motion and thus allow users to perform gestures include mice, digital gloves and camera tracking systems and touch screens. Second, these positions are measured over time. The way a gesture is drawn over time can change the meaning of a gesture. This leads us to the following definition of a gesture:

A set of measured points P in space and a corresponding set of time intervals T between measurements.

Our definition might seem obvious but a precise definition is necessary to store and work with composed gestures (as introduced below), cooperative gestures [5], and previewing

Figure 2: Work in progress. At this point the table should indicate that a triangle can be completed.

of incomplete gestures. This definition encapsulates what a gesture is, but not its meaning. The meaning of a gesture is derived from its context. The majority of input devices can measure more than movements. A mouse has buttons that can be pressed. Some touch screens can make a distinction between touches of different shapes, e. g. distinguish between different hand postures [4]. These measurements, together with the state of the digital device, provide the context that defines the meaning of a gesture.

Gesture Previews

We propose a solution for helping users with drawing gestures. Our solution is based on the fact that not only motion but also timing forms an essential part of gestures, as is discussed in the previous section. We are working on an interface that detects if a user hesitates while drawing a gesture. When this occurs, an overlay is shown that helps them with selecting and drawing a gesture that matches their goal. Vanacken et al. were among the first to explore a similar idea for touch screens [3]. They introduced the idea of using virtual instructors on touch screens: on-screen avatars that can demonstrate possible interactions. However, it is our opinion that such a help system is not suitable in all settings and often a less intrusive technique seems better. Our overlay shows what gestures are available and what actions are coupled to those gestures, only when this is desired by users.

The goal of our project is to demonstrate that an overlay can be designed which helps users of multi-touch screens to make complex gestures (Figure 2). The overlay should:

- inform the user of possible gesture finishes,
- make clear to users when and how cooperative gestures can be initiated,
- give users the possibility to cancel gestures,
- indicate what actions are coupled to gestures that can be performed,
- increase the detection rate of gesture recognizers by showing users what motions are available, and
- not limit expert users who already know how to work with the gestural interface.

To achieve these goals a gesture recognition algorithm is needed that can match gestures while they are being drawn. This requires an algorithm that is not computationally expensive. Wobbrock et al. [9] introduced a simple yet fast solution for recognizing gestures. However, their algorithm is invariant in regard to orientation and scale. This can be troublesome in collaborative environments. Take the example where a user draws an arrow on a map in a control room. A rotation-invariant matching algorithm can only detect that the drawn gesture looks like an arrow, but cannot infer the direction in which this arrow is pointing. This is problematic, because the control room application might need this direction to deduce which object the arrow is pointing to. A solution for this problem could be to store all rotations performed in the matching algorithm and use these to infer the angle of the gesture as it was originally drawn. A challenge here is that the non-uniform scaling of gestures can “skew” gestures, making the inferred angle not very accurate.

The second problem with the solution proposed by Wobbrock et al. is that their algorithm can only recognize completed gestures [9]. We propose a solution where complex gestures are “composed” from simpler gestures. An example would be a triangle gesture that is composed of three connected line gestures. By using composed gestures one can construct a finite state machine that can guess what gesture is being drawn. With such a solution it is possible to provide previews on-the-fly to users and to undo parts of a gesture. This will be computationally expensive. Therefore, we propose a system that only shows previews when a user hesitates while drawing a gesture. This can be as soon as a user touches the screen or after the user has drawn a partial gesture. Further research is required to determine a suitable delay for showing previews. The previews should help hesitating users but not distract expert users.

If a set of gestures is available that is too large to display at once, different previews can be shown over time. By starting with showing the most common gestures and ending with rare ones we keep the average time a user has to wait low. The meaning of a gesture is made clear to users by an icon or animation, displayed at the end of the path of a gesture. An action can be performed both by drawing the path leading to this icon or by touching the icon with another finger. This way novice users can use the gesture previews as a menu structure. At the same time, they see what gestures they can draw to achieve the same goal even faster. When a user removes his or her finger from the interface without finishing a complete gesture and without touching a gesture endpoint, nothing will happen and the preview overlay will disappear. This way users can intuitively cancel initiated gestures.

Collaborative Gestures and Application

Scott et al. have shown [7] that, for collaborative interaction on tabletop displays, it is important to support interpersonal interaction and to provide shared access to physical and digital objects. It is our opinion that gesture previews can help to adhere to these guidelines. A gesture preview does not have to be limited to a single user. The state machine introduced in the previous section can be extended to support collaborative gestures, i. e., gestures performed by multiple users to achieve a single goal. The overlay introduced in the previous section can show cooperative gestures when a second user touches the screen. For example, in the control room scenario, decision-making can be supported by allowing all users to draw votes on the screen. A firetruck will only be sent out if all users draw a "yes" gesture on the screen. Such a voting process is not very intuitive for novice users. Our overlay can overcome this challenge by showing users how to cast their votes as soon as someone is hesitating in the manner indicated in the previous section. Morris et al. [5] state that a problem with cooperative gestures is that they are sometimes drawn accidently. A possible solution to this problem is to only allow cooperative gestures to be drawn by users when the touch screen is showing an overlay. Then any user can cancel a cooperative gesture by lifting his or her finger from the screen at any time as described previously.

Our research regarding the combination of cooperative gestures and gesture previews is still ongoing. We plan to demonstrate gesture previews in a game setting. Gesture previews will show users where and how they can interact with the game environment. A possible idea is a game similar to Tower Defense.[4] Novice users can use the gesture preview overlay to become familiar with the game interface and can use it to build simple towers. Expert users do not need the overlay for this and have learned the gestures shown in the overlay to build structures much faster. In addition, expert users can use the overlay to draw complex cooperative gestures together with other players, which are required to finish more difficult stages in the game. Scores are calculated such that collaboration is favored over individual work. Besides being fun, the game can serve as a starting point for usability studies of gesture previews and cooperative gestures.

DISCUSSION AND FUTURE WORK

Initial feedback suggests that gesture previews can be used to make working with gestures easier for novice users and that they support and encourage cooperation on tabletop devices. We are currently experimenting with the mentioned game setting. It would be interesting to see whether our solution is also applicable in other domains, e. g., in the control room and educational settings that our group is also focusing on. Another area that requires further research is the recognition of incomplete gestures. Our proposed solution is based on the $1 Recognizer [9]. An interesting question is whether other types of gesture recognizers, for example ones based on the Rubine classifier [6] or on Hidden Markov Models [8], are better suited for this task. Finally, we would like to perform a formal usability study. Our current feedback has been received mostly from expert users. Yet we think that one of the main benefits of gesture previews is that they make touch screens easier to use for novices. We want to verify this in a controlled setting. In this setting we can also test our hypothesis that gesture previews increase the accuracy with which users draw gestures, which in turn increases the detection rate of gesture recognizers.

CONCLUSION

Our work on control rooms and educational games revealed that designing a touch interface is challenging, especially when collaboration between multiple users is desired. In this paper we have introduced our ongoing work on gesture previews. Gesture previews make working with touch screens easier for novice users and support and encourage collaboration. At the same time, they do not limit or distract expert users. To achieve our goal we are experimenting with composed gestures and on-the-fly gesture detection. It is our hope that these concepts can be used to develop a next generation of software that makes working on large touch screens easier and more intuitive.

[4] See http://en.wikipedia.org/wiki/Tower_defense .

REFERENCES

1. A. P. Association. *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR Fourth Edition (Text Revision)*. American Psychiatric Publishing, July 1994.
2. J. Besseling, B. Hagen, S. Andriessen, A. te Peele, M. Crone, L. Kok, and E. de Vos. Toename gebruik ondersteuning voor jongeren met een gezondheidsbeperking. Technical report, TNO, 2007.
3. K. L. K. C. Davy Vanacken, Alexandre Demeure. Ghosts in the Interface: Meta-user Visualizations as Guides for Multi-touch Interaction. In *Proc. IEEE Tabletop*, pp. 87–90, Los Alamitos, 2008. IEEE Computer Society.
4. T. Isenberg, M. Everts, J. Grubert, and S. Carpendale. Interactive Exploratory Visualization of 2D Vector Fields. *Computer Graphics Forum*, 27(3):983–990, May 2008.
5. M. R. Morris, A. Huang, A. Paepcke, and T. Winograd. Cooperative Gestures: Multi-User Gestural Interactions for Co-located Groupware. In *Proc. CHI*, pp. 1201–1210, New York, 2006. ACM Press.
6. D. Rubine. Specifying Gestures by Example. *SIGGRAPH Computer Graphics*, 25(4):329–337, 1991.
7. S. D. Scott, K. D. Grant, and R. L. Mandryk. System Guidelines for Co-located, Collaborative Work on a Tabletop Display. In *Proc. ECSCW*, pp. 159–178, Norwell, MA, 2003. Kluwer Academic Publishers.
8. H. Winkler. HMM-based Handwritten Symbol Recognition Using On-line and Off-line Features. In *Proc. ICASSP*, pp. 3438–3441, 1996.
9. J. O. Wobbrock, A. D. Wilson, and Y. Li. Gestures Without Libraries, Toolkits or Training: A $1 Recognizer for User Interface Prototypes. In *Proc. UIST*, pp. 159–168, New York, 2007. ACM Press.
10. M. Wu and R. Balakrishnan. Multi-Finger and Whole Hand Gestural Interaction Techniques for Multi-User Tabletop Displays. In *Proc. UIST*, pp. 193–202, New York, 2003. ACM Press.
11. M. Wu, C. Shen, K. Ryall, C. Forlines, and R. Balakrishnan. Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces. In *Proc. IEEE Tabletop*, pp. 183–190, Los Alamitos, 2006. IEEE Computer Society.