

Thesis proposal: gesture previews

Introduction

Modern touch-screen devices can detect and respond to gestures performed by users. An example is the slider displayed on an iPhone (see figure 1). Thanks to the layout, arrow and text of this interface element it is intuitively clear to most users what gestures they can perform on this element.

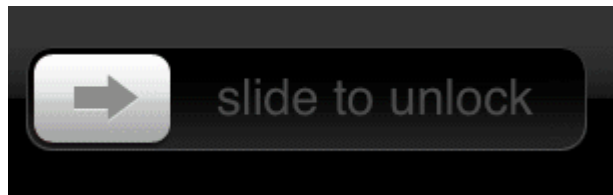


Figure 1: iPhone slider

Another application of gesture detection is text recognition. In this case it is also clear to the user what gestures can be performed, assuming the user is familiar with the expected alphabet. The big challenge for developers here is to correctly “guess” what letter a user is trying to enter. This is not trivial at all: for example the letter Z looks a lot like the number 2.

In other situations another problem arises: the user might not know what gestures are available. A good example is the NUI Horizon touch-screen available at TNO. In the demonstration software a central menu can be activated at any time with the gesture shown in figure 2. If you don't know this gesture you can spend a lot of time trying to find it. Even if you see someone else performing this gesture it is hard to replicate it yourself.



Figure 2: NUI Horizon menu gesture

Now this example gesture is very simple. The situation gets extra challenging if more gestures are available and / or if the gestures are more complex. Imagine a multi-touch interface where you collaborate with several people to manipulate a three-dimensional object. Here gestures are not limited to a single finger. In fact, a complex gesture could be performed by multiple users at the same time. It becomes increasingly difficult to solve these problems:

- How should users know what gestures are available?
- How should the device know what gesture is meant by the users?

In my thesis I propose a possible solution to these problems.

My solution

I propose a solution where users are given visual feedback to demonstrate what gestures are available. This solution should help multiple users to perform complex gestures together. If you start performing a gesture the device will show an overlay with feedback to its users. This overlay could:

- show pulsing circles / dots where other users (or your other fingers) can start collaborating with you on a complex gesture.
- show gestures that your fingers can follow.
- show what actions are coupled to the possible gestures. An example could be an icon at the end of a gesture.

The advantages of this overlay are that:

- users can always see what gestures are available to them.
- users can “follow” a gesture preview with their fingers. This makes their gestures more precise which in turn should improve the detection rate of their gestures.

Now this feedback can be presented in a wide variety of ways. Gestures can be shown in different colours, line widths, line types and line patterns. Gesture previews could even use animation and motion to enhance their usability. In my thesis I want to show that an overlay can be designed that helps users of multi-touch screens to make complex gestures and I want to explore different visualization methods. My main research issue would be:

Assisting users of multi-touch screens with performing complex gestures

Note that I am specifically interested in gestures to be performed by *multiple users*. Other issues that I would like to visit in my thesis are: What kind of processing power is required? How to provide feedback to users that stand around a multi-touch device at arbitrary positions? Is it possible / desirable to use hardware acceleration for generating an overlay? What other research has been done in this field?

I plan to deliver:

- A thesis containing related work, my concept and realization, an evaluation of my idea as well as my implementation work and ideas for future projects.
- A framework for showing gesture previews (preferably extending the framework that is available at TNO).
- A game that uses gesture previews. This would be a game where users have to cooperate: they have to enter complex gestures together. I believe this is a new gameplay format that is unique to multi-touch devices. I have brainstormed with several of the readers of this document about a game in which multiple users have to control a complex device together, for example a catapult. I think the usability of such a game could be improved a lot by showing gesture previews.

Future work

I'm also interested in measuring the effectiveness of an overlay. Are gestures made by users that are assisted with an overlay better recognizable by a computer? This would be a case study. Can graphics hardware help with detecting gesture patterns (for example by using a programmable shader)? But these two issues are probably too much work for me to work on during a single thesis project.