Analysis of gesture-based scroll control for a living room photo browser

Hongchan Choi, Sukwon Chung, Tyler Crimm, Jieun Oh, Matt Sencenbaugh
{ hongchan, quantasw, tman7000, jieun5, msencenb } @stanford.edu

I. INTRODUCTION

We have two overarching motivations for the user study on our living room photo browser, specifically dealing with scrolling through a linear list of menu/photo items. First, we believe that **adding physics and animation into the GUI** (e.g. velocity-dependent scroll acceleration) will not only provide a natural "feel" to the gesture, but also allow users to utilize scrolling behaviors they are already familiar with through mobile devices usage. Second, we feel that having an **auditory feedback** -- in lieu of (or as an alternative to) having a real-time visual display of the skeleton -- can enhance system feedback; determining the most effective types of auditory feedback is valuable in designing gesture-based user interfaces.

More specifically, we hope to address the following questions in our study:

1. Is having animated physics in our graphical user interface helpful for users to navigate the menu to select desirable items? If so, how is it helpful? how do uses feel the system without it?
2. How do users feel when the system provides the reversed or overly-exaggerated directionality? what is the adequate level of physicality?
3. Is having auditory feedback in GUI helpful for users? what's the best level or scope of the auditory feedback? (*event-based* sonification that's triggered only when a gesture is recognized vs. *continuous sound* that maps to the position of relevant joints)
4. [Time/resource permitting: What is the best way to imply/suggest the available gestures at any given state? (small icons? a full-screen overlay? demonstration videos? voice instruction?)]

We hypothesize that (1) displaying animated transitions with physics can convey critical information such as where users are in the menu and what the available action are, and that (2) visual or auditory implication of "gesture-activating" elements can more efficiently direct users to appropriate actions and further provide feedback on how the system is registering the user's gestures.

II. METHODS

- **participants**

  We recruited 8+ participants, many of whom currently live in a family setting. The following is a more detailed demographics, grouped by family:

  - living on campus, a housewife, 30s, from Korea
  - living on campus, a preschooler, 4, from Korea

  - living on campus, a husband, 40s, from the US
  - living on campus, a wife, 40s, from the US
  - their 11 year old daughter

  - living on campus, a PhD student in Computer Music, from Canada

  - living on campus, female, a PhD student in MS&E, from Korea

- [if needed] living on campus, age 18-22, undergraduate students

Participants were recruited by cupcakes and coffee. :)

- **system setup**

Our setup is a mockup GUI based on C# XNA framework, with auditory feedback implemented using the ChucK programming language (http://chuck.cs.princeton.edu/). These two components are connected using the low-latency Open Sound Control protocol (http://opensoundcontrol.org/). Note: For the purpose of this user study only, we have isolated and encapsulated the "scroll menu" functionality from our full NodeJS + front-end architecture to simplify debugging and parameter changing.

We are testing different design variants of our system according to the following 2x2 matrix:

		variable 1: animated transitions and physics	
		no	yes
variable 2: auditory feedback	no	**(00)** **no animated transitions, no physics** **no auditory feedback**	**(01)** **animated transitions, physics** **no auditory Feedback**
	yes	(10) no animated transitions, no physics auditory feedback	(11) animated transitions, physics auditory feedback

- **environments**

The study is conducted in the family living room of the participant's home. For undergraduate participants, study is conducted in a dorm lounge.

- **tasks**

For each of the four conditions (00, 01, 10, 11 described above in the 2x2 matrix):
  - we show participants a linear menu of items labeled from 1 through 100. The default item in the center of screen is 50.
  - we ask participants to navigate to a "far away" number, say 90, and select that item.
  - from there, we ask participants to navigate to a "close" number, say 87, and select that item.
  - repeat the navigate-select task several times, for a range of distances.

- **data collected**

We will record the following data for each subject, condition, and task:

- **subject** ID
    - **condition**: 00, 01, 10, or 11
        - **distance** and **direction** instructed to scroll (e.g. “right” 40 items, if going from 50 to 90)
            - **time** it took from starting gesture to selecting the correct item
            - **number of “under” shooting** (i.e. didn’t quite reach the target item)
            - **number of “over”shooting** (i.e. went to far, and passed the target item)
            - “**velocity**” of the swipe gesture(s) used
        - follow-up **questions/ survey** to inquire experience

To ensure that we capture all the data needed, our study sessions will be video-taped.