Gesture Validation in Television Interfaces

Olabode O. Anise II

Auburn University
olabodeanise@auburn.edu

Abstract

Gesture recognition devices have become more prevalent in homes across the world as smart TV technology becomes more commonplace. The Clemson University Human-Centered Computing Lab has been working to develop a cable interface that is controlled exclusively by gesture.This project aims to validate the findings of a prior study which gathered gestures to control a cable interface.

I. Introduction

The typical way of interacting with a television is via remote control.

The current challenges facing television interaction involve the design of current standard remote controls. They feature numerous buttons that are rarely used by most consumers, thus preventing users from accessing all of the capabilities provided. Also, the configurations of remote controls do not line up with the interface of the televisions that they correspond with. The users have to constantly switch between viewing the information on the screen and searching the remote for different buttons to bring up the different menus.

In order to combat this many gesture recognition devices have been developed; however, they've fallen short because of the design of the gestures. Designers have developed gestures that aren't intuitive.

In the past, many studies have been done to see how users interact with gesture recognition devices. While those results have been duly noted, no study has gone to the lengths to confirm the gestures they identifed by another group of users. This study has done this.

II. Related Work

Hand gesturing has been one of the naturalistic and common ways of human-to-human communication. Any hand movement could be identified as a potential gesture:[8]. HCI researchers and practitioners have studied hand gestures in the persuit of designing input gestures that inherit everyday gestures qualities :[2]. Gesture scholars agree that their are different kind of gestures. Some differ in their occurance, such as those performed during speech, and some differ in their physical structure and how they are performed:[6]. Technology today has embraced this input modality and integrated it into their interaction design, taking into consideration the metrics that factor into the differences amongst user-defined gestures such as culture. Gestures frequency, viewpoint, rhythm, and description of motion are different amongst various cultures:[1, 4]. Research scholars had proposed a large number of gesture taxonomies in the literature:[6, 2],however, no standard guidelines for gesture taxonomies have been set for gesture interaction design.

Research shows that various gestures have often been defined by designers not users. Subsequently, users would have to learn these gestures:[9, 7, 3]. The latter makes the process of developing gestures unnatural. Despite the skillful design of such gestures, they are arbitrary gestures which have been designed specifically for reliable recognition:[5, 8]. This method of designing gestures might be suitable for prototyping. However, it is not useful for determining which gestures match those that would be chosen by users.

III. Methods

I. Participants

Fifteen paid participants volunteered for this study. Of the fifteen participants, five were male and ten were female. Five participants were African American, six were white, two were Hispanic or Latino, and two were Asian/Pacific Islander. The average age of the participants was approximately 20. Thirteen were right-hand dominant. Nine were at least slightly familiar with smart TV technology (1 very familiar, 3 moderately familiar, and 5 slightly familiar), and six were not familiar. The average familiarity with gesture recognition interfaces was 3, indicating a moderate familiarity. Most of the participants were students doing undergraduate research at Clemson in different areas.

II. Apparatus

The study was conducted in a conference room. Clips of the gestures that would be interpreted were projected on a 60 inch screen by a Toshiba projector. The video containing the coalesced clips was made with Windows Movie Maker. Responses from participants, demographic and subjective survey data were recorded on individual Asus B121 tablets.

III. Procedure

A second user study was conducted to confirm the generalized gestures that came from the first user study. The study was performed in a conference room with the subject responding on tablets. The study was presented as a means to examine the perceptions of gestures versus their desired intent. The subject signed a consent form and filled out pre and post surveys that gathered their demographics along with their familiarity with smart TV technology, preferred means of watching television, and thoughts on the experiment.

The actual experiment consisted of 21 different gesture clips. The participant was shown a gesture and then told to indicate all of the different actions that the gesture could be used to bring about. The participant could chose from any of the twenty-five actions. The participant repeated that sequence of steps for each of the 21 gesture clips. The whole experiment, including the pre and post-experiment surveys, lasted approximately 25 minutes.

IV. Results

Table 1: *Gesture Confirmation Score*

System Action	Confirmation Score
System Attention	12
Power Off	3
Power On	5
Volume Down	11
Volume Up	8
Mute	10
UnMute	0
Channel Up	1
Channel Down	2
Last Channel	4
Main Menu	2
Scroll Down	1
Scroll Up	7
Enter (generic)	11
Back (preserves change)	0
Page Down	0
Page Up	4
Cancel	3
Recorded TV Menu	2
Record (real time)	8
Play	10
Pause	7
Fast Foward	3
Rewind	0
Stop (Playback)	11

The values displayed in Table 1 show the confirmation score for each of the system actions that were displayed for the 15 participants. The confirmation score was calculated simply by recording each of the participants who correctly identified the gesture as that particular system action.

V. Conclusion

Overall the study was a success. Many of the gestures were validated by the user group. For the ones that weren't, the common theme seemed to be that those gestures were smaller and a lot more ambiguous.

References

[1] Brown, A. Gesture viewpoint in japanese and english: Cross-linguistic interactions between two languages in one speaker. *Gesture 8*, 2 (2008), 256–276.

[2] Grandhi, S. A., Joue, G., and Mittelberg, I. To move or to remove?: A human-centric approach to understanding gesture interpretation. In *Proceedings of the Designing Interactive Systems Conference* (New York, NY, USA, 2012), DIS '12, ACM, pp. 582–591.

[3] Malik, S., Ranjan, A., and Balakrishnan, R. Interacting with large displays from a distance with vision-tracked multi-finger gestural input. In *Proceedings of the 18th annual ACM symposium on User interface software and technology* (2005), ACM, pp. 43–52.

[4] Mauney, D., Howarth, J., Wirtanen, A., and Capra, M. Cultural similarities and differences in user-defined gestures for touchscreen user interfaces. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, 2010), CHI EA '10, ACM, pp. 4015–4020.

[5] Moscovich, T., and Hughes, J. F. Multi-finger cursor techniques. In *Proceedings of Graphics Interface 2006* (2006), Canadian Information Processing Society, pp. 1–7.

[6] Poggi, I. From a typology of gestures to a procedure for gesture production. In *Gesture and sign language in human-computer interaction*. Springer, 2002, pp. 158–168.

[7] Rekimoto, J. Smartskin: an infrastructure for freehand manipulation on interactive surfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2002), ACM, pp. 113–120.

[8] Wobbrock, J. O., Morris, M. R., and Wilson, A. D. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2009), ACM, pp. 1083–1092.

[9] Wu, M., and Balakrishnan, R. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proceedings of the 16th annual ACM symposium on User interface software and technology* (2003), ACM, pp. 193–202.