Understanding Face Gestures with a User-Centered Approach Using Personal Computer Applications as an Example

Yenchin Lai[1], Benjamin Tag[2], Kai Kunze[3], Rainer Malaka[4]
[1,4]The University of Bremen,[2]The University of Melbourne, [3]Keio University
[1]ylai@uni-bremen.de, [2]benjamin.tag@unimelb.edu.au, [3]kai@kmd.keio.ac.jp, [4]malaka@tzi.de

ABSTRACT

While face gesture input has been proposed by researchers, the issue of practical gestures remains unsolved. We present the first comprehensive investigation of user-defined face gestures as an augmented input modality. Based on a focus group discussion, we developed three sets of tasks, where we asked participants to spontaneously produce face gestures to complete these tasks. We report our findings of a user study and discuss the user preference of face gestures. The results inform the development of future interaction systems utilizing face gestures.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; *Empirical studies in interaction design*; *Gestural input*.

KEYWORDS

Face gesture, user-defined gesture, user-centered, elicitation study

ACM Reference Format:
Yenchin Lai[1], Benjamin Tag[2], Kai Kunze[3], Rainer Malaka[4]. 2020. Understanding Face Gestures with a User-Centered Approach Using Personal Computer Applications as an Example. In *AHs '20: Augmented Humans International Conference (AHs '20), March 16–17, 2020, Kaiserslautern,Germany.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3384657.3385333

1 BACKGROUND AND MOTIVATION

Various functional face-based input systems (wearable, desktop-based, *etc.*) have been proposed and diverse recognition algorithms with different sensing methods (electrooculography, camera-based, *etc.*) have been created and examined [1, 3–6]. Until today, one of these input modalities have yet been (fully) accepted in our everyday lives. When Lyons proposed a mobile face gesture interface [3], he pointed out an unsolved face gesture issue, namely "what kind of action-effect mappings result in comfortable, easy to learn interfaces?" As a broad discussion about face gesture design is still scarce, we are presenting findings of a study that aims at gaining a more comprehensive understanding of face gestures, and to contribute to answering this question.

AHs '20, March 16–17, 2020, Kaiserslautern, Germany

ACM ISBN 978-1-4503-7603-7/20/03.
https://doi.org/10.1145/3384657.3385333

2 METHODOLOGY

We conducted a study containing of a focus group that informed a user-defined gesture study [8].

2.1 Focus Group

We recruited 11 participants (4 women, 7 men), aged between 22 and 38 years (Mean = 27, SD = 4.95), who were knowledgeable in Human-Computer Interaction (HCI). The participants were invited to the study in two groups of five and six respectively. After collecting written consent from all participants, a moderator instructed the group and posed questions concerning potential applications and user scenarios for using human face to interact with computer systems. For each question, participants were given about five minutes to write down ideas on post-it notes. The ideas were shared with and discussed by the other group members. Each session lasted about 90 minutes.

53 different application ideas were created during the study. We grouped them into six categories: Socializing (3), LifeLog (13), Auto Suggestions (5), Functional Controls (26), Privacy and Security (4), Game (2). Also, *face gesture* and *facial expression* were differentiated together with the participants. *Facial expression* is considered as an implicit input and *face gesture* is an explicit input with an unequivocal intention to manipulate computer systems. Head movements were not considered face gestures in this work. 26 ideas from Functional Controls and three ideas from Privacy and Security were using *face gestures.* The ideas in the other categories did not include explicit input; they were, therefore, excluded. We created three target applications based on the most mentioned ideas.

2.2 User-defined Gesture Study

Three computer applications with four commands each were used as target interactions in this study: 1) **E-mail Inbox** with commands Next, Mark, Open, and Close; 2) **Media Player** with commands Play, Pause, Fast Forward, and Fast Backward; 3) **Document Reader** with commands Zoom In, Zoom Out, Scroll Down, and Scroll Up. We built simplified GUIs for each application and recorded all effects of the commands in separated short videos.

We recruited 20 participants (9 females, 11 males), aged between 22 and 30 years (Mean = 24.95, SD = 2.11) for this study. After the participants gave written consent, *inter alia* to receive permission to record participants' faces on video for later analysis, the leading researcher explained the purpose and procedure of the study. The study was composed of three application sessions with short post-task questionnaires and a post-study interview session. The observer instructed each participant to use one visible facial movement as a gesture. After an effect was shown in a short video, participants performed the best-fit gesture in their opinions to the

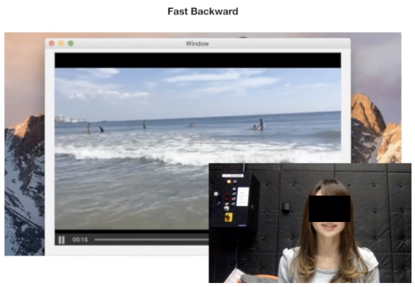

Figure 1: A command, Fast Forward, is presented to a participant. The participant's face movement is video-recorded.

command (Figure 1). The participants then answered a questionnaire rating their own gesture mappings.

The questionnaire contained four questions for each command. The first question was to understand the emotional factors because emotion-related face gestures (*e.g.* smile as explicit input) were often discussed in related works [4, 5]. The participants had to answer whether any emotional valence and arousal levels were related to the gesture mappings. An emotion assessment scale with self-assessment manikin (SAM) [2] was used to rate the valence and arousal when the answer was yes. The other three questions, presented as 7-point Likert Scales (1= "Entirely Disagree" to 7 = "Entirely Agree"), asked the participants if the gestures were *well-matched*, *easy to perform*, and *socially acceptable*. In the post-study interview session, participants were asked to state their opinions about using this input modality.

3 RESULTS

We collected 240 gestures in total (12 commands * 20 participants) and presented the most frequently used gesture mappings shown in Figure 2 along with agreement rate [7]. A taxonomy of all collected gestures by facial parts and movements is presented in Figure 3. **Upper Facial Part** (used 75%), includes gestures performed with the eyebrows, eyelids and eye gaze; **Middle Facial Part** (used 2 %) includes nose and cheek movements; lips, mouth and chin movements are included in the category of **Lower Facial Part** (used 23 %). Interestingly, no gestures were made using two or more facial parts and all the most frequently used gesture mappings were included in **Upper Facial Part**. Many participants chose eye movements containing directional information as gestures, *e.g.* Look Up and Down for Scroll Up and Down (N = 9), Look Right for Fast Forward (N = 5).

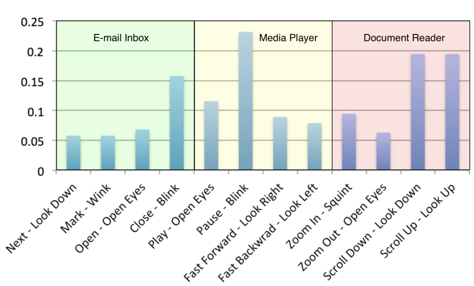

Figure 2: User-defined gestures with agreement rate.

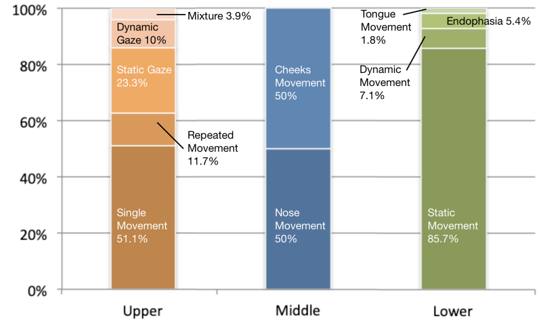

Figure 3: Face Gesture Taxonomy

When the participants were asked whether emotional factors were included in the gesture mappings, 64.17% of the gesture mappings were answered no. It showed that the majority of the defined gestures were not considered emotional. And we did not find any valence-arousal trends in the remaining 35.83%, that answered yes. For the other self-assessment ratings, the participants considered the gestures were well-matched (Mean = 5.2, SD = 1.27), easy to perform (Mean = 5.9, SD = 1.21) and socially acceptable (Mean = 5.5, SD = 1.46).

In the post-study interview, we observed other preferences and opinions of the participants. Some mentioned that they used Upper Facial Part to perform gestures because of its subtlety. P17 mentioned that he chose eye movements because he felt he must look weird if he used other facial parts such as moving the mouth. Similarly, P15 selected Upper Facial Part gestures and said "... I tried to avoid looking weird because I don't want to use them (weird gestures) in public." Furthermore, we asked for participants' opinions on using face gestures. Overall, the participants considered face gestures to be an optional input modality to support the users when hands are occupied, rather than a main interaction (N = 9). Also, some (N = 3) considered face gestures convenient because of its subtlety so they could check contents on computers covertly.

4 FUTURE WORK AND CONCLUSION

We reported a comprehensive investigation of user-centered face gesture study. We conclude that users prefer to use upper facial movements to perform gestures because the movements are subtle, and can include directional information. Also, fewer gestures were emotion-related and the emotion-related ones were not identical. Simple and functional gestures are preferred for computer controls.

Based on the results of our focus group, we used personal computer applications as examples to discuss user-defined face gestures in this study. We will further investigate the other applications with explicit face gesture input in order to have a thorough understanding of face gestures from the users' perspective. Finally, we hope the exploration will further drive researchers and designers to develop practical face gestures in future interaction systems.

5 ACKNOWLEDGMENTS

We would like to thank all participants. This work was partially funded by the Klaus Tschira Stiftung (KTS).

REFERENCES

[1] Toshiyuki Ando, Yuki Kubo, Buntarou Shizuki, and Shin Takahashi. 2017. CanalSense: Face-Related Movement Recognition System Based on Sensing Air Pressure in Ear Canals. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, 679–689.

[2] Margaret M. Bradley and Peter J. Lang. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25, 1 (1994), 49 – 59.

[3] M. J. Lyons. 2004. Facial gesture interfaces for expression and communication. In *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583)*, Vol. 1. 598–603 vol.1.

[4] Denys J. C. Matthies, Bernhard A. Strecker, and Bodo Urban. 2017. EarFieldSensing: A Novel In-Ear Electric Field Sensing to Enrich Wearable Gesture Input Through Facial Expressions. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, 1911–1922.

[5] David Rozado, Jason Niu, and Martin Lochner. 2017. Fast Human-Computer Interaction by Combining Gaze Pointing and Face Gestures. *ACM Trans. Access. Comput.* 10, 3 (Aug. 2017), 10:1–10:18.

[6] Junichi Shimizu, Juyoung Lee, Murtaza Dhuliawala, Andreas Bulling, Thad Starner, Woontack Woo, and Kai Kunze. 2016. Solar System: Smooth Pursuit Interactions Using EOG Glasses. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*. ACM, 369–372.

[7] Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. 1325–1334.

[8] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. [n. d.]. User-defined Gestures for Surface Computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, 1083–1092.