GestAnalytics: Experiment and Analysis Tool for Gesture-Elicitation Studies

Oğuz Turan Buruk
Koç University – Arçelik
Research Center for Creative
Industries (KUAR)
Istanbul, Turkey
oburuk@ku.edu.tr

Oğuzhan Özcan
Koç University – Arçelik
Research Center for Creative
Industries (KUAR)
Istanbul, Turkey
oozcan@ku.edu.tr

DIS'17 Companion, June 10-14, 2017, Edinburgh, United Kingdom

ACM ISBN 978-1-4503-4991-8/17/06.
http://dx.doi.org/10.1145/3064857.3079114

Abstract
Gesture-elicitation studies are common and important studies for understanding user preferences. In these studies, researchers aim at extracting gestures which are desirable by users for different kinds of interfaces. During this process, researchers have to manually analyze many videos which is a tiring and a time-consuming process. Although current tools for video analysis provide annotation opportunity and features like automatic gesture analysis, researchers still need to (1) *divide videos into meaningful pieces*, (2) *manually examine each piece*, (3) *match collected user data with these,* (4) *code each video* and (5) *verify their coding*. These processes are burdensome and current tools do not aim to make this process easier and faster. To fill this gap, we developed "GestAnalytics" with features of simultaneous video monitoring, video tagging and filtering. Our internal pilot tests show that GestAnalytics can be a beneficial tool for researchers who practice video analysis for gestural interfaces.

Author Keywords
User-Elicitation, Gestures, Video Annotation, Video Analysis, Gesture Elicitation, User Centered Design

ACM Classification Keywords
H.5.2 Evaluation/methodology – Screen Design - GUI

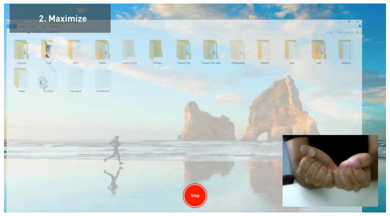

Figure 1: Recording Screen of GestAnalytics

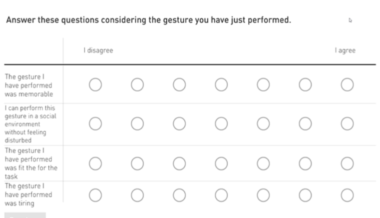

Figure 2: Questionnaire Screen

Introduction

Gestural interfaces encapsulate many different modalities such as mid-air gestures [11], on-skin gestures [1] or surface gestures [13]. Since these modalities are novel interaction methods, how these control types will be taken by users are unknown. Therefore, to understand the user preferences about new types of gestural controls, researcher often use user-elicitation studies [1,8,9,10,12]. In these studies, researchers want users to generate some gestures which they think are appropriate for specific tasks. For example, they show the name of animation of a task such as "accept" and want users to produce gestures which they think fit with this "accept" action.

Our internal elicitation studies [1,4] indicate that, to document and analyze this process, researchers must (Issue1) *video/sound record* the whole process and *divide these recordings into meaningful pieces,* (I2) *get user feedback and questionnaire data* (if applicable), (I3) *match this data with video pieces*, (I4) *manually watch and code each video* to extract similarities and differences and (I5) *verify the coded data*. This process is quite time-consuming [7] and we believe that the current assistive tools do not accelerate and ease it. To feel this gap, we developed GestAnalytics which regulates this process covering both the experiment and analysis procedures.

Current tools which are used in gesture elicitation studies provide advanced annotation features to researchers. Two annotation tools, ELAN [2] and ANVIL [3], are the most common tools for user-elicitation studies. These tools are designed for adding notes and annotations to the specific parts of the videos and extracting them as text files. They provide many options and alternative settings and are considered as successful annotation tools. Another tool, GestureAnalyzer [7], was also specifically developed for gesture analysis and capable of tracking mid-air gestures and automatically analyze and illustrate them. However, this tool is specialized only for mid-air gestures and relies on existing technology. Thus, many studies using Wizard-of-Oz [5] or focusing on gesture elicitation for non-existing technologies [1,12] cannot benefit from this application. Moreover, none of these tools solves the mentioned issues.

To overcome these issues, we developed an analysis and an experiment tool for gesture analysis called GestAnalytics. During the experiment, it automatizes the process of matching the questionnaire data with specific videos. Moreover, it saves video pieces after each task is completed and eliminates the need for dividing videos after the experiment. During the analysis, it allows researchers to monitor up to 50 videos simultaneously and eases the process of comparing different videos for extracting their similarities and differences. It also provides a tagging tool for taxonomical coding and lets researchers filter all videos according to these tags. In this paper, we explain the features of GestAnalytics, how these features may help researchers and discussed the shortcomings and potentials based on our pilot tests.

GestAnalytics

Experiment Tool

Experiment Tool of GestAnalytics mainly automatizes the video/sound recording and matches the related questionnaire data to each video. In the conventional workflow, researchers video record the whole process and make participants fill a questionnaire after each

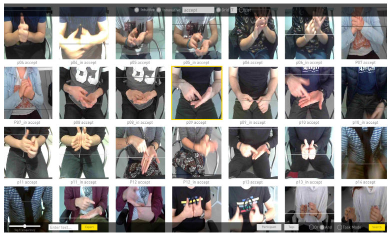

Figure 3: Simultaneous Video Monitoring

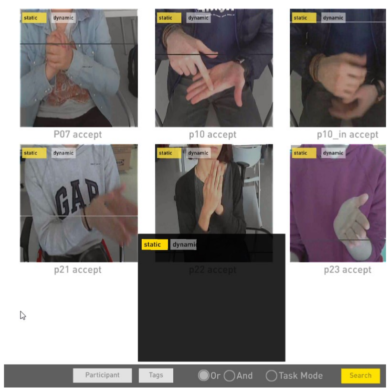

Figure 4: Some Videos which are filtered with "static" tag

task in a separate medium. In a randomized test, they have to keep track of each task name and match these with videos later on. Moreover, they have to watch each video to mark the parts that refers to gestures and divide these videos accordingly. By eliminating this time-consuming process, GestAnalytics records a single video for each task and save this video along with the questionnaire items and the answers by labeling them with the "name of the task and the participant." This data can be exported as a ".csv" file which can easily be edited with analysis software such as SPSS. With this new workflow model, researchers can gain time by skipping video dividing and data matching processes.

Analysis Tool

SIMULTANEOUS VIDEO MONITORING (FIGURE 3)

Gesture elicitation studies usually focus on extracting gesture sets which are preferable by users. For placing the most preferable gestures, the agreement scores of gestures for one task are calculated. This can only be done by identifying the similar gestures by examining each video. In a 20-participant user test with 20 tasks, a researcher has to open 400 video files and annotate or take notes about each to a separate medium. In GestAnalytics, 20 videos which belong to a task can be monitored at the same time, ordered in a customized matrix grid and zoomed in and out to examine it closely. Moreover, researchers can unmute to sounds of any videos to listen to participants' ideas if think-aloud protocol was applied during the experiment. Monitoring these videos simultaneously saves researchers time since they do not need to open each and every video file and more importantly create the opportunity to compare many videos at the same time.

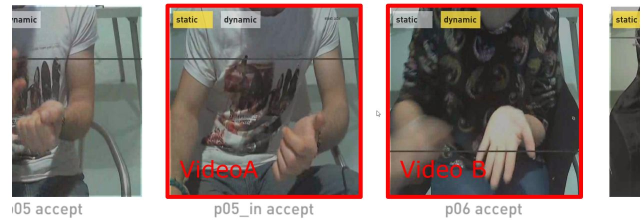

Figure 5: Tagging Feature

TAXONOMICAL TAGGING

Gesture elicitation studies, besides extracting gesture sets, can yield results such as gesture taxonomies. For taxonomizing a gesture set, each video needs to be tagged with the related taxonomy item. In GestAnalytics, tags can be added to the videos on-the-fly and when a taxonomy tag is added, it creates a slot in each video to be marked. In Figure 5, you can see two different videos which are tagged with different taxonomy items. In this example, gesture in the Video A has the "static" item while the one in the Video B has the "dynamic" item. Still, all videos have inactive items, that can be activated easily with a single click. In this tagging system, researchers can tag videos quickly and double-check their coding easily by comparing it to other videos. For further operations, tagging information can be exported as a ".csv" file.

FILTERING (FIGURE 4)

GestAnalytics lets researchers examine the videos per task in the default mode. However, tagging for taxonomy creation is a process with a lot of back-and-forth to make sure that each video is tagged with the correct taxonomy item. It is a burdensome process to check each video file to verify the tags. However, GestAnalytics allows researchers to filter and view all videos belonging to a tag or multiple tags at the same time. In this way, it is easier to notice faulty information about a taxonomy item. Additionally, videos also can be filtered per participant for making examinations about a specific participant.

Discussion and Conclusion

Gesture elicitation studies are common practices in the field and we believe that a specialized tool for this practice will be beneficial for many researchers in this area. We used GestAnalytics in one of our elicitation studies [1] and our observations suggest that it eases this burdensome process remarkably. Moreover, it also prevents many mistakes that can be raised due to the human involvement. For example, we experienced that matching the questionnaire data or naming them per tasks can be problematic and open for mistakes in our other similar studies [6]. Therefore, besides allaying the time-consuming process, GestAnalytics can help researchers to come up with more valid results with less faulty data. Moreover, filtering feature also will help researchers to double-check taxonomy items and decrease the erroneous information.

We believe that current tools for assisting the experiment and the analysis process for gesture elicitation studies fail to address main problems that researchers face in the process. Current tools offer many advanced annotation tools and even automatic analysis of gestures from videos. Still, dividing long videos into meaningful small parts, matching the questionnaire data, examination of videos and the verification process are very time-consuming processes and these have to be manually done by researchers to ensure the quality of the research project. Therefore, we developed a new tool, GestAnalytics, which assists researchers in these problematic areas which were not solved by existing studies and software. Although, in many research projects, customized software is created according to nature of the research data, our project proposes an interface framework that can also be adopted by customized software for overcoming the mentioned issues. A demo of GestAnalytics which includes our research data can be downloaded from the link (bit.ly/gestanalytics). We also plan to present GestAnalytics as an open source software as we improve the user interface for other researchers' use.

Limitations and Future Work

GestAnalytics is developed in Unity3D, which is a game engine. Unity3D is not a tool which is optimized for video viewing, thus the performance of the software can increase if it is moved to another platform which performs better for video rendering.

Continuous video recording, which should be divided into pieces afterwards, is a must for studies which requires the videos of uninterruptible events such as a conversation environment. Therefore, we also plan to improve GestAnalytics with tools speeding up the video dividing process. Moreover, although it prevented mistakes in coding and data/video matching, several videos were accidentally stopped by participants before recording the gestures in our study. Therefore, we need to implement a timer which prevents to stop recording before a certain period.

Percentage and agreement score calculation are quite common practices in gesture-elicitation studies. Hence, we plan to equip GestAnalytics with tools that can do these kinds of basic calculations without the need of a external software. After presenting the software to the field, we will continue improving the tool per feedbacks of the users.

References

1. İdil Bostan, Oğuz Turan Buruk, Mert Canat, et al. 2017. Hands as a Controller: User Preferences for Hand Specific On-Skin Gestures. *Proc. DIS '17*, ACM.
2. H Brugman and A Russel. 2004. Annotating Multimedia/ Multi-modal resources with ELAN. *Proc. LREC '04*: 2065–2068.
3. Harry Bunt, Michael Kipp, and Volha Petukhova. 2012. Using DiAML and ANVIL for multimodal dialogue annotation. *Proc. of LREC 2012*: 1301–1308.
4. Oğuz Turan Buruk and Oğuzhan Özcan. 2014. DubTouch: exploring human to human touch interaction for gaming in double sided displays. *Proc. NordiCHI '14*, 333–342.
5. Sabrina Connell, Pei-Yi Kuo, Liu Liu, and Anne Marie Piper. 2013. A Wizard-of-Oz elicitation study examining child-defined gestures with a whole-body interface. *Proceedings of the International Conference on Interaction Design and Children*: 277–280.
6. Hayati Havlucu, Mehmet Yarkın Ergin, İdil Bostan, Oğuz Turan Buruk, Tilbe Göksun, and Oğuzhan Özcan. 2017. It Made More Sense: Comparison of User-elicited On-Skin Touch and Freehand Gesture Sets. *Proc. HCII '17*.
7. Sujin Jang, Niklas Elmqvist, and Karthik Ramani. 2014. GestureAnalyzer: Visual Analytics for Pattern Analysis of Mid-Air Hand Gestures. *Proc. SUI '14*: 30–39.
8. Meredith Ringel Morris, Jacob O Wobbrock, and Andrew D Wilson. 2010. Understanding users' preferences for surface gestures. *Proc. GI '10*: 261–268.
9. Chaklam Silpasuwanchai and Xiangshi Ren. 2015. Designing concurrent full-body gestures for intense gameplay. *International Journal of Human-Computer Studies* 80: 1–13.
10. Ying-Chao Tung, Chun-Yen Hsu, Han-Yu Wang, et al. 2015. User-Defined Game Input for Smart Glasses in Public Space. *Proc. CHI '15*: 3327–3336.
11. Radu-Daniel Vatavu and Ionut-Alexandru Zaiti. 2014. Leap gestures for TV: insights from an elicitation study. *Proc. TVX '14*: 131–138.
12. Martin Weigel, Vikram Mehta, and Jürgen Steimle. 2014. More Than Touch : Understanding How People Use Skin as an Input Surface for Mobile Computing. *Proc. CHI '14*, 179–188.
13. Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. *Proc. CHI '09*: 1083.