Towards Establishing Consistent Proposal Binning Methods for Unimodal and Multimodal Interaction Elicitation Studies

Xiaoyan Zhou(✉), Adam S. Williams, and Francisco R. Ortega

Colorado State University, Fort Collins, CO 80525, USA
{Xiaoyan.Zhou,AdamWil,fortega}@colostate.edu

Abstract. More than two hundred papers on elicitation studies have been published in the last ten years. These works are mainly focused on generating user-defined gesture sets and discovering natural feeling multimodal interaction techniques with virtual objects. Few papers have discussed binning the elicited interaction proposals after data collection. Binning is a process of grouping the entire set of user-generated interaction proposals based on similarity criteria. The binned set of proposals is then analyzed to produce a consensus set, which results in the user-defined interaction set. This paper presents a formula to use when deciding how to bin interaction proposals, thus helping to establish a more consistent binning procedure. This work can provide human-computer interaction (HCI) researchers with the guidance they need for interaction elicitation data processing, which is largely missing from current elicitation study literature. Using this approach will improve the efficiency and effectiveness of the binning process, increase the reliability of us er-defined interaction sets, and most importantly, improve the replicability of elicitation studies.

Keywords: Elicitation study · Binning methods · Gestures proposals

1 Introduction

Elicitation is a form of participatory design where researchers ask novice users to produce input proposals that would execute various commands (i.e., referents) within a system (e.g., augmented reality displays, multi-touch surfaces) [24,29]. During this process the participant is shown referents one at a time and the participant produces input proposals that they find appropriate for executing those referents. This might look like showing a participant a virtual object using augmented reality technologies and asking them to produce a mid-air gesture that would cause the virtual object to move left [18,25].

During an elicitation study researchers often identify the consensus set of input proposals by quantifying the agreement between participants on input

M. Kurosu (Ed.): HCII 2022, LNCS 13302, pp. 356–368, 2022.
https://doi.org/10.1007/978-3-031-05311-5_25

proposals for a given referent [21,29]. In between labeling the raw input proposals and measuring participant agreement, researchers often bin the input proposals into equivalence groups based on their similarity of execution. Without binning proposals, the consensus set will be biased by individual differences and agreement calculations will return pessimistic estimates of agreement. As an example, without binning, a swipe right with the index and middle fingers would be counted as a distinct gesture from swiping right with the index alone. In actuality these two gestures are similar in execution, intent, and at times may have been produced by the same user for the same referent [18,27].

It is well known that the binning procedure is time-consuming due to the process of annotating the recorded videos [1,22], and error-prone because of the subjective human judgments involved [22]. Moreover, adopting different similarity criteria can lead to underestimated or overestimated agreement rates as a result of more or less permissive binning protocols [22]. In this paper, we present a formula that can assist experimenters with binning gesture proposals from participants through a unified binning procedure. This work uses mid-air gesture elicitation proposals, which can be extracted from both unimodal and multimodal interactions. This formula could be used equally well to bin gesture proposals from elicitation studies with different systems, such as virtual reality displays or multi-touch devices.

It is critical that experimental science is reproducible, replicable, and reapeatable. Researchers in HCI have been calling for more work focusing on the replication or reproduction of prior works [15,22]. In particular, for elicitation studies, adopting a standard data processing procedure and guidelines can accelerate the replication or reproduction process for future researchers. The necessity of a standard procedure for binning is increasing along with the rapid growth of elicitation studies. It is time to fill this gap and form a guideline for elicitation data processing.

2 Related Work

Based on recently published elicitation study literature reviews, over two hundred elicitation studies have been conducted in the last ten years [23]. Most of the previous works aimed to develop user-defined gesture sets [6,11,14,16,18,19, 25,27,29], and a few works have discussed the tools and criteria used for binning the interaction proposals in the studies [1,2,20].

2.1 Binning Tools

Several studies proposed new approaches to improve the efficiency of the binning process and reduce the effort required by researchers. Ali et al. proved that letting online crowd workers make the proposal's similarity comparison, then using automatic binning algorithms for sign classification, can produce similar results compared to using experts with less time spent on the task [1]. Nevertheless, the crowd-powered tool Crowdsensus proposed by Ali et al. only demonstrated the use of elicited voice commands in text string format. Vatavu introduced the dissimilarity-consensus method to replace the subjective binning criteria, and manual labeling with automated computation by a computer [20]. This method helps mitigate human error and subjective decisions on proposals similarity. However, using the dissimilarity consensus approach accurate skeletal and pose information is needed, which is not suitable for most elicitation studies' data captured in video format. Moreover, despite the computational method, subjectivity still exists during raw data preprocessing and extracting procedures. Anthony et al. developed a tool to identify similar gesture articulations and generate clusters automatically, and the agreement rates were reported after binning [2]. Nevertheless, the tool allows the researcher to edit and correct the cluster partition, making the binning process still subjective. While it is difficult to entirely avoid the subjective human judgments during the binning process, following the guidelines put forth here can help to mitigate it.

2.2 Binning Criteria

It is not uncommon that different elicitation studies adopt different criteria for evaluating the similarity of proposed interactions from participants [20,22]. Few works have reported the criteria they used for proposal binning in the study, which led to irreproducible study results [22]. Most of the elicitation studies have described the classification process in a broad manner [27–29]. Several works in augmented reality have mentioned the general binning approach or criteria was used in their studies [17,18,25]. Piumsomboon et al. defined "similar gestures" as the gestures that have identical static pose and path, or have different static poses but with consistent directionality [18]. Pham et al. adopted themes such as the analytic unit and binned gestures based on similar themes despite different executions (e.g. binned squishing with clapped hands, fingers, or fist) [17]. Williams et al. binned gestures according to the movement path and hand pose, and explained the definition of each hand pose presented in the study [25]. Many elicitation studies have used a taxonomy for gesture classification and binning [3,5,7,8,18,27].

Taxonomy based binning criteria are not always well suited for direct interaction elicitation study proposals. In most elicitation studies a large proportion of the referents revolve around direct manipulations (i.e., translations, rotations, and scaling). When eliciting direct manipulations, interaction proposals will often be impacted by the participants' notions of real world physics [9]. If the referent is *move left* the participants will most likely push the object to the left with their hand [17,18,27]. These proposals will often fall under a single category when using a taxonomy [10] (e.g., "direct manipulation"), making using taxonomy for binning difficult. This paper will not discuss the procedure of classifying mid-air gesture proposals based on taxonomy. Our study aims to provide researchers and designers with a guideline for executing the proposals binning procedure in an elicitation study, as shown in Fig. 1. With the structured criteria and procedure, more elicitation studies can be replicated, and the results can be reproduced.

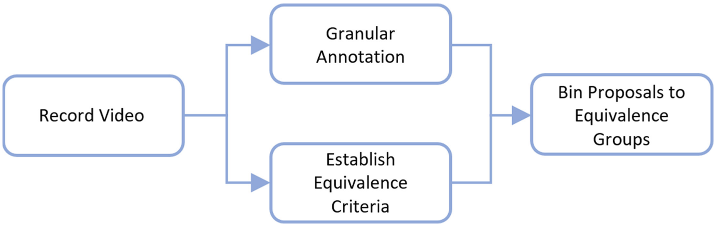

Fig. 1. Procedure to group interaction proposals into equivalence groups in an elicitation study

3 Data Collection

Often elicitation studies' data is captured via video recordings [14,22,25,26]. An exo-centric camera is commonly used to record participants' proposals. If the study used head-mounted displays (HMDs), an ego-centric camera can also be used. For data completion and analysis convenience, recommend experimenters adopt both ego-centric and exo-centric cameras to capture different views of interaction. Based on observations from our prior experiments' data with HMDs, participants have various preferences on hand position while performing manipulation tasks [25,26]. Some participants proposed physical interaction more than others, which means the gesture acted physically on the object, and their hands reached out to the objects. When participants proposed more metaphorical gestures for interaction, their hands were closer to their bodies during the manipulation. As a result, it is possible that the ego-centric camera is unable to capture the full view of the interaction due to the different positions of individual proposals. The exo-centric view could be very helpful, when the ego-centric camera does not capture the full hand movement. For example, while participants' hands are too close to the face and the movement is out of the view of the ego-centric camera, the exo-centric recording became the only source of participants' proposal. For exo-centric view, it is suggested to place the camera with a tripod on the front right or left corner that faces the participant. For the ego-centric view, the camera (e.g. GoPro) can be mounted on top of the head-mounted display (e.g. HoloLens 2). It should be noted that experimenters need to check the camera angle before starting the experiment. It is not unusual that the head-mounted display is unable to hold the camera on the top, or the camera mounting cause unbalance or discomfort to users. In that case, the experimenters could choose to use a head strap attaches to the camera and ask participants to wear it on their heads. The downside of this approach is the complexity of the camera angle set-up, since the head-mounted display could block the view of the ego-centric camera.

4 Video Annotation Procedure

Once the elicitation data has been collected a granular labeling pass should be done to convert the raw video to meaningful interaction information (Fig. 1). These labels should capture all relevant information about the gestures. For each interaction elicited, the hand used, fingers used, direction of movement, start time, stop time, changes in hand shape or pose, and the corresponding referent name should be recorded. Some additional interaction features that can be recorded depending on the studies' aims are outlined in the following sections.

4.1 Stroke Segmentation

A gesture generally starts with a preparation phase, moves into a stroke, then returns to the resting position [12]. For example, the hand was usually placed on the lap of the participant or on the desk surface with a relaxed hand pose. When the participant is ready to perform the gesture, the hand will hold up and leave the lap or desk first, then reach out to the target object or perform the interaction immediately. This means that experimenters need to find the boundaries of a gesture, and it also can be called the stroke of a gesture. According to McNeil, a stroke is considered the peak of effort for a specific gesture [13], which holds the meaningful content of the gesture. Based on previous experiment data, when the stroke happens, the direction of movement will change, or the speed of movement will show a sudden increase [26]. This can be observed while the experimenters go through each recorded video frame.

4.2 Gesture Information Recording

After the boundaries of the gesture are found, the video of the stroke can be segmented, or the screenshot of the stroke can be captured for later reference. This reference could be very useful while experimenters need to review or double-check certain proposed gestures. The start and end times of both gestures and strokes should be noted during video annotation for later analysis of stroke time distribution, or the relationship between strokes and other modalities of interaction. In addition, the experimenters should also record which hand and finger have been used for the interaction, how many hands were used, and what fingers were involved. These elements contain important information for verifying results from previous studies with mid-air gestures in augmented reality environment [17,18,25,26].

4.3 Stroke Coding

To code proposed gestures, experimenters need to define each hand shape before starting coding. For example, the "pointing" gesture could involve the index finger, index middle fingers, or all fingers from one hand. "Pinching" could include thumb and index finger, or thumb index and middle fingers. "Grasping" could be gathered five fingertips, and "grabbing" shows as a fist. During stroke coding, the hand pose and the path of hand movement are the main points that experimenters should pay attention to. In terms of hand pose, hand shape and palm and finger orientation should be noted. For example, moving an object up with the thumb and index finger holding the object should be coded as pinching with the palm facing forward and moving the hand along the Z-axis. In terms of movement path, experimenters should code proposed gestures based on which axis it moved along or around. For example, translations up and down should be both considered as movements along the Z-axis. The referents of translations left or right should be identified as movements along Y-axis. Moving an object towards or away should be recognized as movements along X-axis. The rotations clockwise and counterclockwise should be both coded as movements around X-axis. Rotating around Y-axis could describe the response to referents *pitch up* or *pitch down*, and rotating around Z-axis could represent proposals to referents *yaw left* or *yaw right*.

It is worth noting that the thumb position could be less meaningful compared to other fingers. We suggested experimenters consider the interaction theme while recording the thumb position for each proposal. The variants of thumb position in hand poses could be due to hand differences or individual habits. In general, the relaxed position for the thumb is to stick out instead of closing to the palm or gathering with other fingers. As a result, when a gesture does not require the thumb, it normally stays relaxed, which could look like the thumb is sticking out. For example, based on our previous data, some proposals with index finger pointing gestures also had the thumb pointed to the front [26]. In this case, we ignored the thumb pose and only recorded the pointing gesture executed by the index finger. However, with proposals involving pinching or grasping poses, which could be proposed for referents of translation or rotation, the experimenters should pay attention to every fingers' status including the thumb.

5 Interaction Proposal Binning

Once all of the collected video has been annotated the elicited interactions can be binned into equivalence groups based on pre-defined metrics of similarity. This reduced the space of elicited proposals to a smaller set of binned interaction proposals which can then be used during agreement rate and consensus calculations. Conceptually, the binning process smooths out interpersonal differences between interaction proposals. As an example, if binning processes were omitted, a three finger pinching gesture where the participants index, middle, and thumb come together to grab a virtual object would be considered different from a two finger pinching gesture where only the index and thumb are used. When considering how to implement the elicited interaction proposals into a interaction system, it makes more sense to say that those two gestures are the same. Their intent and motion trajectory's are similar. Outside of motion and intent, this step can be further justified when considering how the same user may use inconsistent numbers of fingers for the same gesture [18], making the difference between using the index and thumb or the index, middle, and thumb trivial.

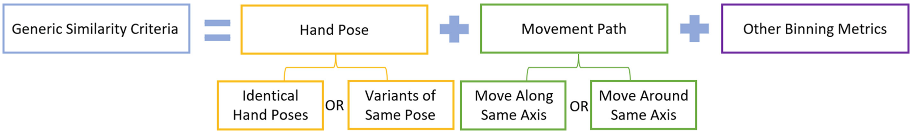

Fig. 2. Formula for binning gesture proposals

As shown in our formula (Fig. 2), while binning mid-air gesture proposals, we suggest the experimenter mainly focuses on the similarity of static hand pose and hand movement path. Bins are created based on the pairing of binning criteria. This might look like making bins for pinching hand poses moving along the x-axis, or grasping gestures that move vertically. Some bins may not be predictable a priori. To account for that researchers can expect to add new binning criteria while analyzing their results. For example someone might pantomime usage of a steering wheel for rotations. This gesture might necessitate a new equivalence group to be established. The experimenters can set up a worksheet to assign a name or number and provide a brief description to each bin for later reference. It is possible that the same gesture was proposed for different referents, and we recommend sharing bins through the binning procedure for all referents in the study. If the same gesture resulted in the same agreement rate with different referents, the experimenters could use further analysis or use other criteria to decide which referent the gesture is best suited for.

Fig. 3. Different proposals of the pinching gesture

Fig. 4. Interchangeable variants for pinching or grasping gestures

5.1 Hand Pose

The proposed gestures with identical or near identical hand poses and moving paths belong to the same bin (Fig. 5). For example, both gestures pinch an object with thumb and index finger then move the hand along the Y-axis. However, some near identical hand poses could look a bit different from others due to individual preferences, as shown in Fig. 3. For example, in prior work done by this lab, some participants proposed using blooming gestures with the palm face up to perform the create object task, and some participants opened their hands less wide than others [25,26]. For referent *move away*, among participants who adopted using an index finger to point at the object and moving along the X-axis, few participants held their pointing finger with their palm facing left instead of facing frontwards.

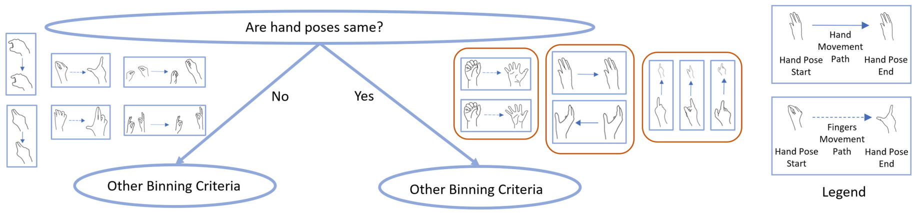

Fig. 5. Binning gesture proposals with identical hand poses criteria

As previous works pointed out [4,17,18,26], variations of similar hand poses were often elicited among participants. The variations could be due to the physical features or the size of the target object [17], and we are focusing on using the same objects for the entire elicitation study. According to prior studies, the variants of a single hand pose could come from the same participant or different participants [18], and the previous study has shown that participants preferred to perform the same gesture in a different manner [27]. In Piumsomboon et al.'s work on eliciting user gestures in augmented reality, variants of pinching and grasping gestures were similar and grouped into the same bin. The shapes or affordances of the objects used to elicit interactions can also impact the hand poses elicited. If the same object was used for all experiments, we suggested experimenters consider that variants of a single pose are often interchangeable, and that those variants should be grouped into the same bin if they have the same movement path (Fig. 4).

The most common variations of hand poses that were mentioned in prior studies involved a different number of fingers when gesturing [18,25,27]. As an example, the proposals for referent *rotate clockwise* in previous elicitation studies were performed by pinching gestures with two or three fingers, or grasping gestures with five fingers [25,26]. It was not due to different sizes or affordances of the object since the same object was presented to every participant. A similar case is that while participants performed *move away* with pointing fingers,

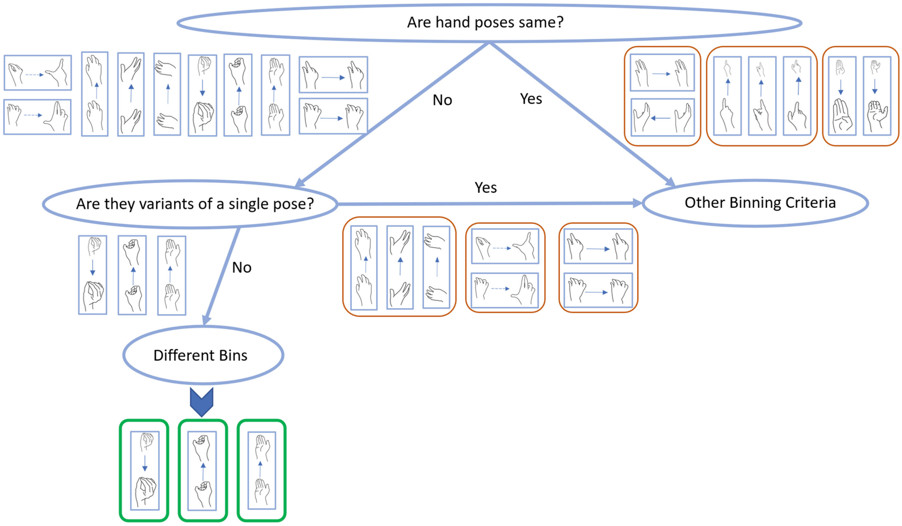

Fig. 6. Binning gesture proposals with identical hand poses and pose variants criteria

some chose to use both their index and middle fingers pointing at the object compared to others who used only their index finger. Another example would be the same pinching gestures with thumb and index finger, but different positions of the other three fingers. Some participants kept their other three fingers spread instead of together while performing pinch gestures. Furthermore, based on our elicitation data, we found another type of variation included different hand orientations. One example is that for the referent *move up* in our study, most participants have proposed grasping the object and moving their hand straight up, and we found three variants of the same hand pose with different hand orientations. The participants were grasping from the front of the object with palm facing front and fingers pointing up, left, or right (Fig. 6).

5.2 Movement Path

If the proposed gestures have identical or variant hand poses, experimenters can continue to check if the moving paths of proposals are the same before making the binning decision (Fig. 7). Based on the stroke coding results, if the moving path is along or around the same axis, we can put those gestures in the same bin. For example, pinching with thumb and index finger then moving hand along Y-axis can be grouped with grasping with five fingers then moving hand along Y-axis. However, pinching with a thumb and index finger then moving the hand along Y-axis should not be binned with grasping with five fingers and moving the hand along Z-axis then moving it along Y-axis. Not all gestures will move in a straight line or follow a directional axis. Some proposals for enlarging or shrinking objects, especially when the gestures involved both hands, were proposed with diagonal movement paths.

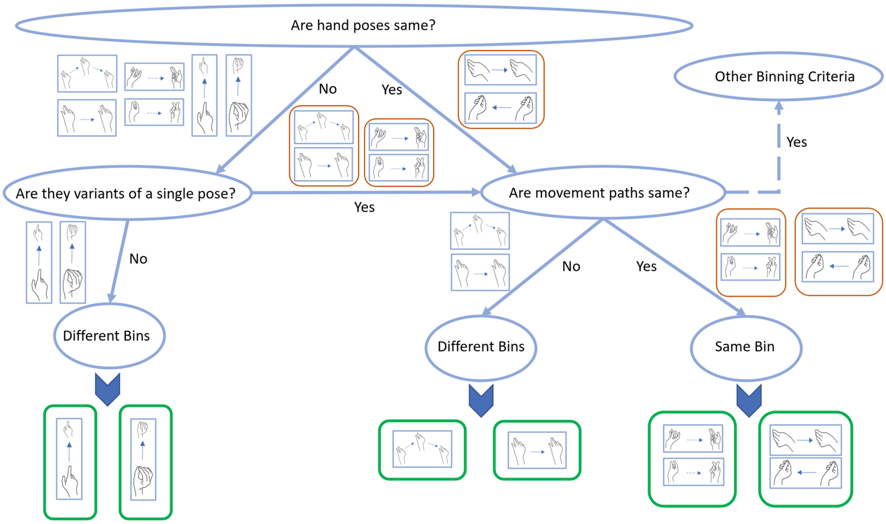

Fig. 7. Binning gesture proposals with identical hand poses, pose variants, and movement path criteria

6 More than Mid-Air Gestures

Our formula can be used equally well for elicitation studies outside augmented reality environments, because researchers can adopt our formula to bin gesture proposals from studies conducted with virtual reality headsets or multi-touch devices. For multi-touch devices, hand poses and movement path are still the crucial criteria for the binning procedure. The experimenters should reconsider the interchangeable variants of a single pose since the multi-touch devices could have a different definition for the number of fingers in a gesture counted as a single-point touch or a whole-hand touch. Based on Wobbrock et al.'s elicitation study on tabletop gestures [29], using 1–3 fingers should be considered as a single-point touch, and using four fingers or more could be considered as a whole-hand touch. It means that the experimenters can consider variants of touch gestures with 1–3 fingers as interchangeable during the binning process.

7 Limitation and Future Work

The formula presented in this paper focused on hand gesture proposals, and it is not suitable for other proposals generated by other body parts such as the head, foot, or arm. Since different body gestures have their features and limitations, it is difficult to use the same criteria for all elicitation studies' data processing. In future work, we are interested in perfecting the binning criteria for hand gesture proposals in elicitation studies. Another future direction regarding binning procedures is to explore further criteria for binning other body gestures and creating guidelines for researchers that can be used in different elicitation studies.

8 Conclusion

This paper proposed a formula that can help experimenters or researchers to bin gesture proposals in elicitation studies. In addition, we explained the necessity of a manual binning procedure and the formation of guidelines for elicitation data processing in HCI. To achieve this goal, this paper went through how experimenters might capture the elicitation data, annotate the recorded video, and what similarity criteria should be considered during the proposals binning process. We also presented real data from our previous elicitation study to demonstrate the possible variations of each gesture. We hope that this paper can be leveraged to create binning procedures for different categories of elicitation studies, such as studies involving gestures using different body parts or using different devices.

Acknowledgments. The authors would like to thank Brandon Kelly for the hand illustrations. This work was supported by the National Science Foundation (NSF) awards NSF 1948254 and 2037417.

References

1. Ali, A.X., Morris, M.R., Wobbrock, J.O.: Crowdsourcing similarity judgments for agreement analysis in end-user elicitation studies. In: Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST 2018). Association for Computing Machinery, New York (2018). https://doi.org/10.1145/3242587.3242621
2. Anthony, L., Vatavu, R.D., Wobbrock, J.O.: Understanding the consistency of users' pen and finger stroke gesture articulation. In: Proceedings of Graphics Interface 2013, pp. 87–94. Citeseer (2013)
3. Chen, Y.C., Liao, C.Y., Hsu, S.W., Huang, D.Y., Chen, B.Y.: Exploring user defined gestures for ear-based interactions. In: Proceedings of the ACM Human–Computer Interaction 4(ISS), November (2020). https://doi.org/10.1145/3427314
4. Danielescu, A., Piorkowski, D.: Iterative design of gestures during elicitation: understanding the role of increased production. arxiv preprint arxiv:2104.04685 (2022)
5. Dingler, T., Rzayev, R., Shirazi, A.S., Henze, N.: Designing Consistent Gestures Across Device Types: Eliciting RSVP Controls for Phone, Watch, and Glasses, pp. 1–12. Association for Computing Machinery, New York (2018). https://doi.org/10.1145/3173574.3173993
6. Dong, Z., Piumsomboon, T., Zhang, J., Clark, A., Bai, H., Lindeman, R.: A comparison of surface and motion user-defined gestures for mobile augmented reality. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA 2020), pp. 1–8. Association for Computing Machinery, New York (2020). https://doi.org/10.1145/3334480.3382883

7. Du, G., Degbelo, A., Kray, C.: User-generated gestures for voting and commenting on immersive displays in urban planning. Multimodal Technol. Interact. **3**(2) (2019). https://doi.org/10.3390/mti3020031.https://www.mdpi.com/2414-4088/3/2/31
8. Grijincu, D., Nacenta, M.A., Kristensson, P.O.: User-defined interface gestures: dataset and analysis. In: Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces (ITS 2014), pp. 25–34. Association for Computing Machinery, New York (2014). https://doi.org/10.1145/2669485.2669511
9. Jacob, R.J., et al.: Reality-based interaction: a framework for post-wimp interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2008), pp. 201–210. Association for Computing Machinery, New York (2008). https://doi.org/10.1145/1357054.1357089
10. Karam, M., Schraefel, M.C.: A Taxonomy of Gestures in Human Computer Interactions. s.n. (2005)
11. Leng, H.Y., Norowi, N.M., Jantan, A.H.: A user-defined gesture set for music interaction in immersive virtual environment. In: Proceedings of the 3rd International Conference on Human-Computer Interaction and User Experience in Indonesia (CHIuXiD 2017), pp. 44–51. Association for Computing Machinery, New York (2017). https://doi.org/10.1145/3077343.3077348
12. McNeill, D.: Hand and Mind: What Gestures Reveal about Thought. University of Chicago Press (1992)
13. Mcneill, D.: Gesture and Thought. The University of Chicago Press (2005). https://doi.org/10.7208/chicago/9780226514642.001.0001
14. Morris, M.R.: Web on the wall: insights from a multimodal interaction elicitation study. In: Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces (ITS 2012), pp. 95–104. ACM, New York (2012). https://doi.org/10.1145/2396636.2396651
15. Nebeling, M., Huber, A., Ott, D., Norrie, M.C.: Web on the wall reloaded: implementation, replication and refinement of user-defined interaction sets. In: Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces (ITS 2014), pp. 15–24. Association for Computing Machinery, New York (2014). https://doi.org/10.1145/2669485.2669497
16. Peng, Z., Xu, J.: User-defined gestures for taking self-portraits with smartphone based on consistency. In: Ahram, T., Falcão, C. (eds.) AHFE 2020. AISC, vol. 1217, pp. 316–326. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-51828-8_42
17. Pham, T., Vermeulen, J., Tang, A., MacDonald Vermeulen, L.: Scale impacts elicited gestures for manipulating holograms: implications for AR gesture design. In: Proceedings of the 2018 Designing Interactive Systems Conference, pp. 227–240. ACM (2018)
18. Piumsomboon, T., Clark, A., Billinghurst, M., Cockburn, A.: User-defined gestures for augmented reality. In: Extended Abstracts on Human Factors in Computing Systems (CHI 2013), pp. 955–960. Association for Computing Machinery, New York (2013). https://doi.org/10.1145/2468356.2468527
19. Vatavu, R.D.: User-defined gestures for free-hand tv control. In: Proceedings of the 10th European Conference on Interactive TV and Video (EuroITV 2012), pp. 45–48. Association for Computing Machinery, New York (2012). https://doi.org/10.1145/2325616.2325626

20. Vatavu, R.D.: The dissimilarity-consensus approach to agreement analysis in gesture elicitation studies. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI 2019), pp. 1–13. Association for Computing Machinery, New York (2019). https://doi.org/10.1145/3290605.3300454
21. Vatavu, R.D., Wobbrock, J.O.: 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 2015), pp. 1325–1334. Association for Computing Machinery, New York (2015). https://doi.org/10.1145/2702123.2702223
22. Vatavu, R.D., Wobbrock, J.O.: Clarifying agreement calculations and analysis for end-user elicitation studies. ACM Trans. Comput.-Hum. Interact. **29**(1) (2022). https://doi.org/10.1145/3476101
23. Villarreal-Narvaez, S., Vanderdonckt, J., Vatavu, R.D., Wobbrock, J.O.: A systematic review of gesture elicitation studies: what can we learn from 216 studies? In: Proceedings of the 2020 ACM Designing Interactive Systems Conference (DIS 2020), pp. 855–872. Association for Computing Machinery, New York (2020). https://doi.org/10.1145/3357236.3395511
24. Williams, A., Ortega, F.: A concise guide to elicitation methodology (2021)
25. Williams, A.S., Garcia, J., Ortega, F.: Understanding multimodal user gesture and speech behavior for object manipulation in augmented reality using elicitation. IEEE Trans. Visual. Comput. Graph. **26**(12), 3479–3489 (2020). https://doi.org/10.1109/TVCG.2020.3023566
26. Williams, A.S., Ortega, F.R.: Understanding gesture and speech multimodal interactions for manipulation tasks in augmented reality using unconstrained elicitation. In: Proceedings of the ACM Human–Computer Interaction, 4(ISS) (2020). https://doi.org/10.1145/3427330
27. Wittorf, M.L., Jakobsen, M.R.: Eliciting mid-air gestures for wall-display interaction. In: Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI 2016), pp. 3:1–3:4. ACM, New York (2016)
28. Wobbrock, J.O., Aung, H.H., Rothrock, B., Myers, B.A.: Maximizing the guessability of symbolic input. In: CHI 2005 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2005), pp. 1869–1872. Association for Computing Machinery, New York (2005). https://doi.org/10.1145/1056808.1057043
29. Wobbrock, J.O., Morris, M.R., Wilson, A.D.: User-defined gestures for surface computing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2009), pp. 1083–1092. ACM, New York (2009)