Selection and Manipulation Whole-Body Gesture Elicitation Study in Virtual Reality

Francisco R. Ortega* Colorado State University

Katherine Tarre† Florida International University

Mathew Kress‡ Florida International University

Adam S. Williams§ Colorado State University

Armando B. Barreto¶ Florida International University

Naphtali D. Rishe‖ Florida International University

Abstract

We present a whole-body gesture elicitation study using Head Mounted Displays, including a legacy bias reduction technique. The motivation for this study was to understand the type of gesture agreement rates for selection and manipulation interactions and to improve the user experience for whole-body interactions. We found that regardless of the production technique used to remove legacy bias, legacy bias was still found in some of the produced gestures.

Index Terms: Gesture Elicitation—Gestures—Virtual Reality;

1 Introduction

The abundance of Virtual Reality (VR) and head-mounted displays (HMDs) has allowed developers to ask question to improve **direct manipulation** for Natural User Interfaces (NUIs). We delve into the following questions: (1) What gestures are appropriate for selection and manipulation in a VR environment? (2) Does the size of an object effect how participants interact with it? (3) Would legacy bias still be present after using a reduction legacy bias method? This whole-body gesture elicitation study provides the findings and observations to these questions and additional recommendations.

This user study provides the following contributions. First, using the production legacy bias reduction method participants provided different gestures for the environment under consideration that would have not otherwise occur; however, gesture legacy is still apparent in the study. Second, this study allows for user input (based on their favorite gestures). We believe that understanding selection and manipulation, specifically with different objects is critical for the applications that will be developed in the near future for HMDs and similar devices. Third, after the study, we have found some indication that production may not be an effective legacy bias reduction technique. Finally, we have validated that the size of the object may influence the type of gesture, as also found in [5].

Elicitation studies are critical to understand gestures and preferences from users. However, we understand that this practice has generated debate and confusion. For example, legacy bias reduction techniques [3] have been suggested but very little evidence has been shown that these techniques may provide an improved gesture set. We used the gesture elicitation methodology introduced by Wobbrocks et al. [7] and refined by Vatavu and Wobbrock [6]. We have chosen this methodology because it is currently the most empirically tested gesture elicitation methodology in use.

*e-mail: fortega@colostate.edu
†e-mail: ktarr007@fiu.edu
‡e-mail: mkres006@fiu.edu
§e-mail: AdamWil@colostate.edu
¶e-mail: Barretoa@fiu.edu
‖e-mail: ndr@acm.org

The main two objectives of a gesture elicitation study are to collect a gesture set from the users and to understand user behavior [7]. The popularity of gesture elicitation is reflected in a variety of studies, ranging from 3D travel using multi-touch to mid-air gestures [4]. This is not the first study conducted for whole-body interaction. Connell et al. conducted a whole-body study with children [1]. Besides the common problem found with legacy device bias, they found low-agreement rates, which provides some similarities to our study.

2 User Study and Evaluation

The experimenter asked the subject to complete an entry questionnaire. Participants were given 5 minutes to get used the environment while wearing the HMD. Then, subjects were asked to performed an elicitation for two referents as part of their training. These were different from the actual referents used during the experiment. Next referents were asked in randomized order. For each referent subjects were asked to provide multiple gestures (production). Each time they performed the gesture, the environment would execute the referent (e.g., an object was created). For each referent the experiment would ask them to rate their preference among the gestures produced. At the end of the experiment an exit-questionnaire was conducted. The 23 participants considered were composed of 10 women and 13 men with an average age of 21. All of our participants were right-handed and just over 50% of them had no experience with either Microsoft Kinect or VR HMDs. For the exit questionnaire, participants were asked a series of questions on a scale of 1 to 7, with 1 being the minimum and 7 being the maximum. The questions included, extent of participants' mental immersion experience (95% of participants rated 3 to 5), general enjoyableness of the experiment (over 95% rated 5 or above), level of the environmental responsiveness to the initiated actions (70% of participants rated 6 or above).

We used Windows 10, Unity Game Engine 2017.1, HTC Vive HMD (without the controllers), and Microsoft Kinect version 2. The Kinect was used to record data but not to recognize gestures. We also use GOPro 4 camera to record the user gestures for analysis. The experimenter made note of the gestures in addition to the recorded data already mentioned.

The **dataset** consists of over 1,000 gestures obtained from a total of 30 participants. Participants were encouraged to provide multiple gestures and then select their favorite for each referent. Referents, such as rotate, move, destroy, and create were seen twice using visualization of a cube and a wall, in order to see how participants would respond to different sizes of items in the same environment. In some instances, users were unable or unwilling to develop a gesture for the referent. For this reason, a reduced dataset was used for analysis, including only 23 participants (out of 30) who provided at least one gesture for all referents. This reduced dataset considered gestures identified as favorites by the participants for each referent. We derived **Overall** agreement, which considered the most repeated gestures (as if it was a gesture elicitation study without production).

The **gesture agreement** was considered using the new formula by Vatavu and Wobbrocks [6] because in the former approach [7],

2019 IEEE Conference on Virtual Reality and 3D User Interfaces
23-27 March, Osaka, Japan
978-1-7281-1377-7/19/$31.00 ©2019 IEEE

Table 1: Summary for Production, Favorites, and Overall Gesture Set

Referent	Production	Favorite	Overall
Z_Scale	Accordion	Accordion	Accordion
Z_RotateWall	2H-S. Wheel	2H-S. Wheel	2H-S. Wheel
Z_RotateCube	Doorknob	Doorknob	Doorknob
Z_Rotate	2H-S. Wheel	S. Wheel	2H-S. Wheel
Z_MoveWall	2H-Push	2H-Push	2H-Push
Z_MoveCube	2H-Push	Push	2H-Push
Z_Move	2H-Push	2H-Push	2H-Push
Y_Scale	Vert. Accordion	Vert. Accordion	Vert. Accordion
Y_RotateWall	Swipe Right	Swipe Right	Swipe Right
Y_RotateCube	Swipe Right	Swipe Right	Swipe Right
Y_Rotate	Swipe Right	Swipe Right	Swipe Right
Y_MoveWall	2H-Swipe Down	Swipe Down	2H-Swipe Down
Y_MoveCube	Swipe Down	2H-Swipe Down	Swipe Down
Y_Move	Swipe Down	Swipe Down	Swipe Down
X_Scale	Push	Push	Push
X_RotateCube	Swipe Down	Swipe Down	Swipe Down
X_RotateWall	Push	2H Swipe Down	Push
X_Rotate	Swipe Down	Swipe Down	Swipe Down
X_MoveCube	Swipe Right	Swipe Right	Swipe Right
X_MoveWall	Swipe Right	Swipe Right	Swipe Right
X_Move	Swipe Right	Swipe Right	Swipe Right
Select	Push	Push	Push
Destroy_Cube	Kick	Push	Kick
Destroy_Wall	Punch	Punch	Punch
Destroy	Kick	Punch	Kick
Create_Wall	Push	Draw Square	Push
Create_Cube	Push	Push	Push
Create	Push	Push	Push

Legend: S. Wheel: Steering Wheel; Light gray rows are grouping of referents

the formula held that even with zero agreement, gestures agreed with themselves. In other words, referents with zero agreement did not have an agreement rate of zero. The new approach will have lower agreement rates. In addition, the former approach does not account for the effects of sample size.

This study revealed a surprising amount of variation in the agreement rates for the different referents. When considering the full gesture sets for each referent. The agreement rates range from 4.43%, for Create Cube, to 29.15% , for Z Scale. Interestingly, in most cases, the agreement rate for the representation of the cube seemed to be higher than that of the wall. When we group the referents without consideration for visual effects used, it is clear that agreement rates for X, Y, and Z Move referents are higher than those for rotations. This has manifested in other gesture elicitation studies, such as [4]. In addition, translation (e.g., using swipe gestures) gestures are common in existing devices (e.g., iPhone) and are an indication that production may not always remove legacy bias (see Discussion section).

In the **favorite** gestures set, all referents have 23 gestures, one per participant, with the exception of the combined referents, which have 46 observations. It can be observed that the agreement rates have increased significantly with the highest agreement rate for Z Scale rising to 67.98%. In the **overall** agreement (most repeated gestures) set, all referents have 23 gestures, one per participant, with the exception of the combined referents, which have 46 observations. The agreement rates increased significantly with the highest agreement rate for Z Scale, rising to 75.5%.

3 Discussion

A summary of the gestures with the highest agreement rate per gesture set can be seen in Table 1. One of the most noticeable findings is the presence of gestures that would be considered legacy, such as Swipe. However, a few gestures not always found in typical devices (e.g., iPhone or iPad) were present. One of them is the accordion gesture. While this gesture may be considered an evolution from the legacy pinch gesture, the movement was very appropriate for the environment presented in this study. Another study found that production did not reduce legacy bias [2]. In the case of both this study and [2], there is not enough evidence to say definitively, but there is a trend towards finding that production is not an ideal legacy bias reduction technique.

Another observation is the variation between agreement rates with a plurality of them having low agreement rates. While Z Scale had the highest agreement rates for the Favorite and Overall gesture set, this wasn't the case for the X and Y axes. A possible explanation is that it is harder for users to think in 3D, in particular in a true 3D stereoscopic system, such as those rendered in HMDs. For Production agreement rates, the rates involved with cube referents were higher than those dealing with the walls (except for X Move). This could be due to that the cube was significantly smaller than the wall and it was cognitively easier for participants to interact with.

Some common gestures emerged. For example, for Scaling the Accordion gesture was a favorite with a high agreement rate. The steering wheel and doorknob gestures were common for Z Rotate. The question still remains if we should even care about legacy bias? When we started the study, we were firmly convinced that some reduction technique was needed. Today, we are no longer convinced.

4 Future Work and Conclusion

The study showed that legacy bias reduction method may not help gesture elicitation studies. In addition, We found some indications that people will generate different gestures for the same referent depending on the size of the object. This effect has also been found by Pham et. al [5]. For future work, the question about legacy bias needs to be explored further. What other methods are there to reduce legacy bias? When do we need to reduce legacy bias?

Acknowledgments

This material is based in part upon work supported by the National Science Foundation under Grant Nos. I/UCRC IIP-1338922, III-Large IIS-1213026, MRI CNS-1429345, MRI CNS-1532061.

References

[1] S. Connell, P.-Y. Kuo, L. Liu, and A. M. Piper. A wizard-of-oz elicitation study examining child-defined gestures with a whole-body interface. In *Proceedings of the 12th International Conference on Interaction Design and Children*, IDC '13, pp. 277–280. ACM, New York, NY, USA, 2013.

[2] P. Koutsabasis and C. K. Domouzis. Mid-Air Browsing and Selection in Image Collections. In *the International Working Conference*, pp. 21–27. ACM Press, New York, New York, USA, 2016.

[3] M. R. Morris, A. Danielescu, S. Drucker, D. Fisher, B. Lee, m. c. schraefel, and J. O. Wobbrock. Reducing legacy bias in gesture elicitation studies. *interactions*, 21(3), 2014.

[4] F. R. Ortega, A. Galvan, K. Tarre, A. Barreto, N. Rishe, J. Bernal, R. Balcazar, and J. L. Thomas. Gesture elicitation for 3d travel via multi-touch and mid-air systems for procedurally generated pseudo-universe. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 144–153, March 2017.

[5] T. Pham, J. Vermeulen, A. Tang, and L. MacDonald Vermeulen. Scale impacts elicited gestures for manipulating holograms: Implications for ar gesture design. In *Proceedings of the 2018 Designing Interactive Systems Conference*, DIS '18, pp. 227–240. ACM, New York, NY, USA, 2018.

[6] R.-D. Vatavu and J. O. Wobbrock. Formalizing Agreement Analysis for Elicitation Studies. In *the 33rd Annual ACM Conference*, pp. 1325–1334. ACM Press, New York, New York, USA, 2015.

[7] J. O. Wobbrock, M. R. Morris, and A. D. Wilson. User-defined gestures for surface computing. In *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009.