

Identifying the Usability Factors of Mid-Air Hand Gestures for 3D Virtual Model Manipulation

Li-Chieh Chen¹, Yun-Maw Cheng², Po-Ying Chu¹, and Frode Eika Sandnes³

¹Department of Industrial Design, Tatung University, Taipei, Taiwan
{lcchen, juby}@ttu.edu.tw

²Graduate Institute of Design Science, Department of Computer Science and Engineering,
Tatung University, Taipei, Taiwan
kevin@ttu.edu.tw

³Oslo and Akershus University College of Applied Sciences, Oslo, Norway
Frode-Eika.Sandnes@hioa.no

Abstract. Although manipulating 3D virtual models with mid-air hand gestures had the benefits of natural interactions and free from the sanitation problems of touch surfaces, many factors could influence the usability of such an interaction paradigm. In this research, the authors conducted experiments to study the vision-based mid-air hand gestures for scaling, translating, and rotating a 3D virtual car displayed on a large screen. An Intel RealSense 3D Camera was employed for hand gesture recognition. The two-hand gesture with grabbing then moving apart/close to each other was applied to enlarging/shrinking the 3D virtual car. The one-hand gesture with grabbing then moving was applied to translating a car component. The two-hand gesture with grabbing and moving relatively along the circumference of a horizontal circle was applied to rotating the car. Seventeen graduate students were invited to participate in the experiments and offer their evaluations and comments for gesture usability. The results indicated that the width and depth of detection ranges were the key usability factors for two-hand gestures with linear motions. For dynamic gestures with quick transitions and motions from open to close hand poses, ensuring gesture recognition robustness was extremely important. Furthermore, given a gesture with ergonomic postures, inappropriate control-response ratio could result in fatigue due to repetitive exertions of hand gestures for achieving the precise controls of 3D model manipulation tasks.

Keywords: Mid-air hand gesture, 3D virtual model manipulation, Scaling, Translation, Rotation

1 Introduction

Manipulating 3D digital contents through mid-air hand gestures is a new paradigm of Human Computer Interaction. In many applications, such as interactive and virtual product exhibition in public spaces and medical image displays in surgery rooms, the

tasks may include scaling, translation, and rotation of 3D components. In order to ensure the natural mapping between controls and displays, the selections of mid-air hand gestures for different tasks were based on the metaphors of physical object operations [1] or the gestures from user-elicitation studies [2]. Given a consensus hand gesture type for a specific task, the performance of gesture recognition and control could still be influenced by many factors, such as the moving speed and trajectory of hands, the occlusion of fingers due to hand pose changes and transitions, as well as the individual differences of performing a specific hand gesture. Since the characteristics of diverse gestures could result in different challenges, it is necessary to identify usability factors for specific gestures through experiments. Therefore, the objective of this research is to study the usability factors of hand gestures for 3D digital content manipulations.

2 Literature Review

With the benefits of natural and intuitive interactions, and free from sanitation problems in public spaces, the applications of mid-air hand gestures included interactive navigation systems in museum or virtual museum [3, 4], medical or surgical imaging system [5, 6, 7, 8], large display interactions [9], interactive and public display [10], and 3D modelling [11, 12, 13, 14]. Based on previous research, mid-air hand gestures could be analyzed in five gesture types: pointing, semaphoric, pantomimic, iconic, and manipulation [15]. Based on the number and trajectory of hands, mid-air gestures could be classified as one or two hands, linear or circular movements, and different degrees of freedom in path (1D, 2D, or 3D) [16]. Gesture vocabulary was dependent on the context [17, 18]. For example, the gestures for short-range human computer interaction [19] and TV controls [20] were reported to be different. For mid-air 3D object manipulation, natural gestures were necessary for accurate control tasks of scaling, translation, and rotation [21]. While choosing an appropriate mid-air hand gesture, it is necessary to consider the mental models of target users [22], reduce the workload [23], and increase gesture recognition robustness [24]. Although design principles could be derived from the literature, the factors that influence the perceived usability of gestures for specific tasks should be identified through experiments.

3 Experiment

In order to investigate the usability factors of mid-air hand gestures for manipulating 3D virtual models, an experimental system was constructed by modifying the sample programs of Intel RealSense 3D Camera with Unity 3D Toolkit (Fig. 1). In a laboratory with illumination control, the participant stood on the spot in front of a 100-inch projection screen, at a distance of 240 cm. During the experiments, the 3D virtual car model was projected on the screen. Each participant completed the tasks for scaling the car, translating the car seat, and rotating the car with respect to the vertical axis using designated hand gestures as follows.

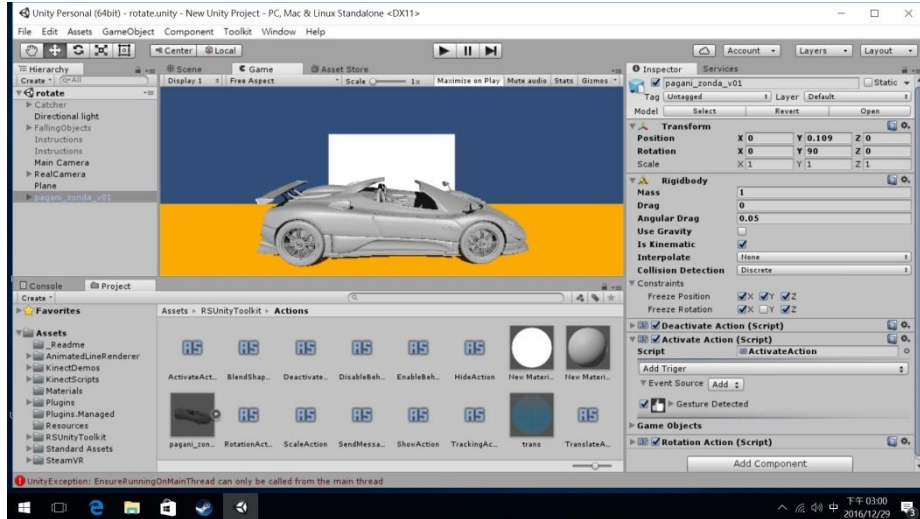


Fig. 1. A 3D Car Model in the Unity 3D System

Among the diverse gesture types, “grab and move” and bimanual “handle bar metaphor” were reported as the intuitive gestures for object manipulation tasks [1, 2, 25]. In addition, users often preferred using gestures resembling physical manipulation for wall-sized displays [26]. Therefore, the two-hand gesture with grabbing while moving apart/close to each other was applied to enlarging/shrinking the 3D virtual car (Fig. 2). The one-hand gesture with grabbing while moving up/down, left/right, or forward/backward was applied to translating a car seat (Fig. 3). The two-hand gesture with grabbing while moving relatively along the circumference of a horizontal circle, a handlebar metaphor, was applied to rotating the car (Fig. 4). The characteristics of these gestures and referenced literature were summarized in Table 1.

Table 1. Designated Hand Gestures for Experiment Tasks

Tasks	No. of Hands	Hand Pose	Hand Orientation	Hand Motion and Trajectory	Referenced literature
Scaling	Two hands	Starting from open palm then grab (while moving)	Facing to each other	Moving apart/close to each other	[1][25]
Translation	One hand		Facing to the target	Moving up/down, left/right, or forward/backward	[2][25]
Rotation	Two hands		Facing down	Moving relatively along the circumference of a horizontal circle	handle bar metaphor [1][25]

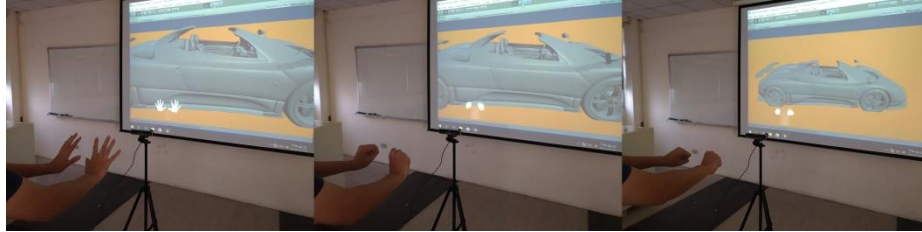


Fig. 2. Enlarging/Shrinking the 3D Virtual Car

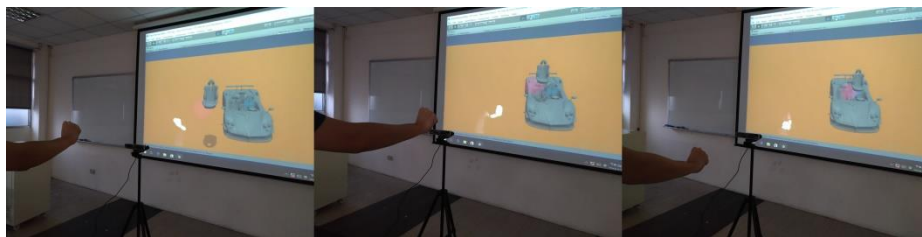


Fig. 3. Translating the Car Seat to the Target Position



Fig. 4. Rotating the Car with respect to the Vertical Axis

In the experiment, an Intel RealSense 3D Camera (F200) was used to extract the positions and movements of 22 joints on each hand skeleton. With the Intel RealSense SDK, basic gestures, such as spread fingers and fist could be recognized (Fig 5). Spread fingers and fist were the static gestures of opening palm and grabbing, respectively. Therefore, it was expected to discriminate the transitions from opening palm to grabbing, and vice versa.

The camera was placed between the participant and the screen. The distance to the participant was adjusted with respect to the arm length. The height was adjusted to the elbow height of each participant.

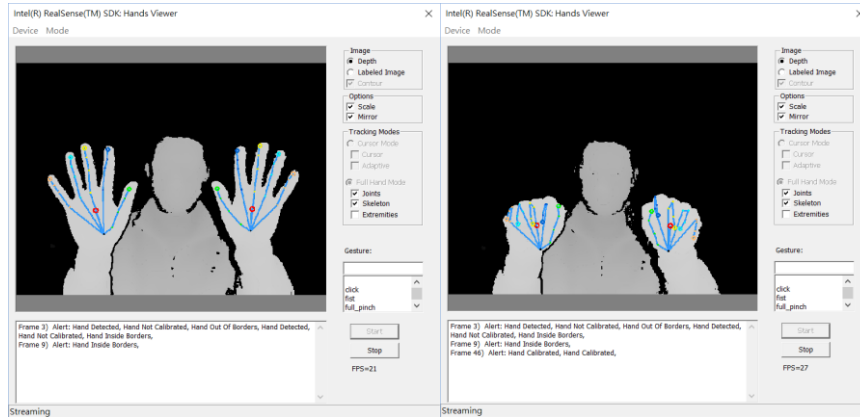


Fig. 5. The Spread-Fingers and Fist Gestures Identified by the Intel RealSense SDK

4 Results and Discussions

Seventeen students, 7 female and 10 male, were invited to participate in the experiment. They studied in either the Ph.D. program of Design Science or the Master program of Industrial Design, with the age range from 22 to 37 (mean: 26.12, standard deviation: 4.65). All participants had the experiences of using 3D modelling software and smartphones with touch gestures. In the experiment, they were asked to apply gestures to carry out scaling, translation, and rotation tasks.

After completing each task, they were asked to evaluate the gesture using a 7-point Likert scale, indicating the degree of agreement, from the perspectives of acceptance to performing in public, comfort, smoothness of operation, easy to understand, easy to remember, informative feedback, correctness of system response, appropriate control-response ratio, and overall satisfaction (Table 2). The result of ANOVA indicated that there were significant differences in user evaluation among these usability criteria. The gesture for scaling was considered as the one needed to be improved in smoothness of operation, correctness of system response, control-response ratio, and overall satisfaction. The gesture for translation yielded the lowest score in the smoothness of operation. The gesture for rotation yielded the lowest score in appropriate control-response ratio.

Tables 3, 5, and 7 listed the reported usability problems with respect to scaling, translation, and rotation tasks, respectively. Two-hand gestures for scaling and rotation tasks caused more usability problems in failure of gesture recognition, lag in system response, gesture detection range, control-response ratio, and fatigue. Based on the original comments from participants, quick movements or rapid pose changes of two hands were the major causes of system failures. Evidently, the original detection range was not wide or deep enough for the natural and linear motions of both hands. The default control-response ratio needed to be adaptive for precise controls.

In addition, the participants were encouraged to offer user-defined gestures for each task. The alternative gestures were listed in Tables 4, 6, and 8. The alternative gestures for scaling (Table 4) included using one hand with posture change (from open palm to grab or pinch) or using two hands to form two corners of a rectangle boundary and slightly adjust the boundary size (Fig. 6). These gestures featured the benefit of requiring less movement range. The alternative gesture for translation was to employ pinching, instead of grabbing (Table 6). The alternative gestures for rotation ranged from steering wheel metaphor, one-hand circular movement, or a more complicated gesture with the first hand staying still as a rotation axis and the second hand moving circularly with respect to the first hand (Table 8).

Table 2. The Evaluation of Gestures and System Performance for Different Tasks

Evaluation Criteria	Scaling		Translation		Rotation	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Acceptance to performing in public	5.41	(1.12)	6.12	(1.11)	4.94	(2.05)
Comfort	4.53	(1.59)	4.94	(1.60)	4.18	(1.55)
Smoothness of operation	3.12	(1.80)	4.18	(1.78)	4.18	(1.67)
Easy to understand	6.18	(1.01)	6.12	(1.54)	5.88	(1.32)
Easy to memorize	6.24	(0.83)	6.35	(1.00)	6.00	(1.46)
Informative feedback	5.41	(1.58)	5.41	(1.54)	6.00	(1.06)
Correctness of system response	3.82	(1.78)	4.24	(1.68)	4.18	(1.74)
Appropriate control-response ratio	3.65	(1.58)	4.24	(1.35)	4.12	(1.27)
Overall Satisfaction	3.82	(1.47)	4.59	(1.42)	4.41	(1.42)

Table 3. The Usability Problems of Two-Hand Grabbing while Moving Apart/Close for Scaling

Problem Types	Original Comments from Participants
Failure of gesture recognition	After enlarging or shrinking, the interactive mode could not be released through opening palm.
	Sometimes, open palm or fist were not recognized by the system.
	Sometimes, the grabbing gesture for initializing scaling was not detected.
	If two hands were too close to each other, the switch between enlarge and shrink become unstable.
Lag in system response	The scaling was lagged or discontinued while two hands were moving quickly.
	The detection of hand movements was discontinued at some time.
Gesture detection range	The detection area was too narrow. Sometimes, it was easy to move out of the range.
Control-Response ratio	It was difficult to reach desired size.
	Sometimes, it was not easy to control the size while scaling up.

Problem Types	Original Comments from Participants
	The mapping between the distance of two hands and the scaling factor of enlarging or shrinking was not clear.

Table 4. Alternative Gestures for Scaling

No. of Hands	Hand Pose	Hand Orientation	Hand Motion and Trajectory
One hand	Open palm for scaling up; Grab for scaling down;	Facing toward the object	Staying still
One hand	Open palm for scaling up; Pinch for scaling down;	Facing toward the object	Staying still
Two hands	Both in L handshape (ASL); forming two corners of a rectangle boundary	One facing forward; the other facing backward	Moving apart for scaling up; Moving close for scaling down;

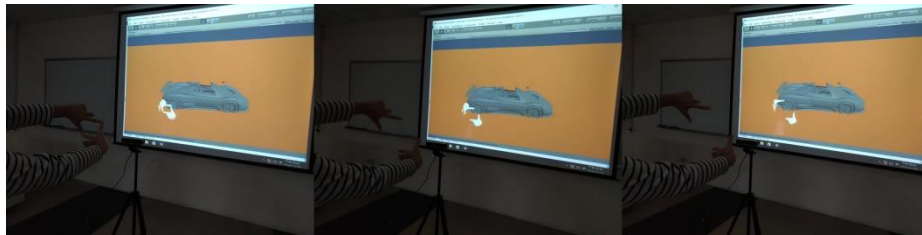


Fig. 6. Alternative Gestures for Two-Hand Scaling (Enlarging/Shrinking)

Table 5. The Usability Problems of One-Hand Grabbing while Moving for Translation

Problem Types	Original Comments from Participants
Gesture Discrimination	The gesture was too similar to the gestures of other operations in previous experiences, such as grabbing and moving forward/backward for zoom out/in of a 2D image.
Control-Response ratio	High gain of xyz movements made positioning difficult.
	It was difficult to move to precise position.
	The depth movement was not easy to control.

Table 6. Alternative Gestures for Translation

No. of Hands	Hand Pose	Hand Orientation	Hand Motion and Trajectory
One hand	Pinch while moving; Open palm for releasing and stop moving	Facing toward the object	Moving from the original position to the target position

Table 7. The Usability Problems of Two-Hand Grabbing while Moving Relatively on a Circumference of a Horizontal Circle for Rotation

Problem Types	Original Comments from Participants
Failure of gesture recognition	The direction of rotation was inconsistent while two hands were crossing to each other in front of the body.
	Sometimes, the gesture was not detected.
Lag in system response	The rotation was not smooth.
Gesture detection range	The system was too sensitive to the height of both hands.
Control-Response ratio	It was difficult to rotate to the desired angle. Sometimes overshoot.
Fatigue	Fatigue while operating repetitively to reach a desired angle.

Table 8. Alternative Geatures for Rotation

No. of Hands	Hand Pose	Hand Orientation	Hand Motion and Trajectory
One hand	Open palm	Facing down	Moving along the circumference of a horizontal circle
Two hands	Holding a virtual steering wheel,	Facing to each other	Moving relatively with a constant distance along the circumference of a horizontal circle
Two hands	First hand fist; Second hand fist for rotating, open palm for stop rotation.	Facing to each other	First hand staying still in front of the body and serving as a rotation axis; Second hand moving along the circumference of a horizontal circle with respect to the first hand;

5 Conclusion

In this research, the usability factors of mid-air hand gestures for 3D virtual model manipulations were identified. The results indicated that the width and depth of detection ranges were the key factors for two-hand gestures with linear motions. For dynamic gestures with quick transitions and motions from open to close hand poses, ensuring gesture recognition robustness was extremely important. Furthermore, given a gesture with ergonomic postures, inappropriate control-response ratio could result in fatigue due to repetitive exertions of hand gestures for achieving the precise controls of 3D model manipulation tasks. These results could be used to inform the development team of vision-based mid-air hand gestures and serve as the checking lists for gesture evaluation.

Acknowledgement. The authors would like to express our gratitude to the Ministry of Science and Technology of the Republic of China for financially supporting this research under Grant No. MOST 105-2221-E-036-009.

References

1. Song, P., Goh, W. B., Hutama, W., Fu, C., & Liu, X. (2012). A handle bar metaphor for virtual object manipulation with mid-air interaction. Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12. doi:10.1145/2207676.2208585
2. Groenewald, C., Anslow, C., Islam, J., Rooney, C., Passmore, P., & Wong, W. (2016). Understanding 3D Mid-Air Hand Gestures with Interactive Surfaces and Displays: A Systematic Literature Review. Proceedings of the 30th International BCS Human Computer Interaction Conference, BCS HCI 2016, Bournemouth University, Poole, UK, 11-15 July 2016. doi:10.14236/ewic/hci2016.43
3. Hsu, F., & Lin, W. (2012). A multimedia presentation system using a 3D gesture interface in museums. *Multimedia Tools and Applications*, 69(1), 53-77. doi:10.1007/s11042-012-1205-y
4. Caputo, F. M., Ciortan, I. M., Corsi, D., De Stefani, M., & Giachetti, A. (2016). Gestural Interaction and Navigation Techniques for Virtual Museum Experiences. Zenodo. <http://doi.org/10.5281/zenodo.59882>
5. O'Hara, K., Gonzalez, G., Sellen, A., Penney, G., Varnavas, A., Mentis, H., Criminisi, A., Corish, R., Rouncefield, M., Dastur, N., & Carrell, T. (2014). Touchless Interaction in Surgery. *Communications of the ACM*, 57(1), 70-77
6. Rosa, G. M., & Elizondo, M. L. (2014). Use of a gesture user interface as a touchless image navigation system in dental surgery: Case series report. *Imaging Science in Dentistry*, 44(2), 155. doi:10.5624/isd.2014.44.2.155
7. Rossol, N., Cheng, I., Shen, R., & Basu, A. (2014). Touchfree medical interfaces. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. doi:10.1109/embc.2014.6945140
8. Hettig, J., Mewes, A., Riabikin, O., Skalej, M., Preim, B., & Hansen, C. (2015). Exploration of 3D Medical Image Data for Interventional Radiology using Myoelectric Gesture Control, Eurographics Workshop on Visual Computing for Biology and Medicine (2015)
9. Chattopadhyay, D., & Bolchini, D. (2014). Understanding Visual Feedback in Large-Display Touchless Interactions: An Exploratory Study. [Research Report] IUPUI Scholar Works, Indiana University.
10. Ackad, C., Clayphan, A., Tomitsch, M., Kay, J. (2015). An In-the-Wild Study of Learning Mid-air Gestures to Browse Hierarchical Information at a Large Interactive Public Display, UBIComp '15, SEPTEMBER 7–11, 2015, OSAKA, JAPAN
11. Vinayak. & Ramani, K. (2015). A gesture-free geometric approach for mid-air expression of design intent in 3D virtual pottery. *Computer-Aided Design*, 69, 11-24. doi:10.1016/j.cad.2015.06.006
12. Vinayak. & Ramani, K. (2016). Extracting hand grasp and motion for intent expression in mid-air shape deformation: A concrete and iterative exploration through a virtual pottery application. *Computers & Graphics*, 55, 143-156. doi:10.1016/j.cag.2015.10.012
13. Cui, J., Fellner, D. W., Kuijper, A., & Sourin, A. (2016). Mid-Air Gestures for Virtual Modeling with Leap Motion. *Distributed, Ambient and Pervasive Interactions Lecture Notes in Computer Science*, 221-230. doi:10.1007/978-3-319-39862-4_21
14. Nakazato, K., Nishino, H., & Kodama, T. (2016). A Desktop 3D Modeling System Controllable by Mid-air Interactions. 2016 10th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS). doi:10.1109/cisis.2016.80
15. Aigner, R., Wigdor, D., Benko, H., Haller, M., Lindlbauer, D., Ion, A., Zhao, S., Koh, J. (2012). Understanding Mid-Air Hand Gestures: A Study of Human Preferences in Usage of

Gesture Types for HCI, Microsoft Research Technical Report MSR-TR-2012-111. <http://research.microsoft.com/apps/pubs/default.aspx?id=175454>

16. Nancel, M., Wagner, J., Pietriga, E., Chapuis, O., & Mackay, W. (2011). Mid-air pan-and-zoom on wall-sized displays. Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11. doi:10.1145/1978942.1978969
17. LaViola Jr., Joseph J. (2013). 3D Gestural Interaction: The State of the Field, ISRN Artificial Intelligence, Volume 2013, Article ID 514641.
18. Pisharady, P.K. and Saerbeck, M. (2015). Recent methods and databases in vision-based hand gesture recognition: A review, Computer Vision and Image Understanding, Volume 141, December 2015, Pages 152–165, Pose & Gesture
19. Pereira, A., Wachs, J. P., Park, K., & Rempel, D. (2015). A User-Developed 3-D Hand Gesture Set for Human-Computer Interaction. Human Factors, 57(4), 607-621. doi:10.1177/0018720814559307
20. Choi, E., Kim, H., & Chung, M. K. (2014). A taxonomy and notation method for three-dimensional hand gestures. International Journal of Industrial Ergonomics, 44(1), 171-188. doi:10.1016/j.ergon.2013.10.011
21. Mendes, D., Relvas, F., Ferreira, A., & Jorge, J. (2016). The benefits of DOF separation in mid-air 3D object manipulation. Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology - VRST '16. doi:10.1145/2993369.2993396
22. Cui, J., Kuijper, A., Fellner, D. W., & Sourin, A. (2016). Understanding People's Mental Models of Mid-Air Interaction for Virtual Assembly and Shape Modeling. Proceedings of the 29th International Conference on Computer Animation and Social Agents - CASA '16. doi:10.1145/2915926.2919330
23. Nunnari, F., Bachynskiy, M., & Heloir, A. (2016). Introducing postural variability improves the distribution of muscular loads during mid-air gestural interaction. Proceedings of the 9th International Conference on Motion in Games - MIG '16. doi:10.1145/2994258.2994278
24. Smedt, Q. D., Wannous, H., & Vandeborre, J. (2016). Skeleton-Based Dynamic Hand Gesture Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). doi:10.1109/cvprw.2016.153
25. Fonseca, F., Ferreira, A., Mendes, D., Jorge, J. & Araújo, B. (2013). 3D mid-air manipulation techniques above stereoscopic tabletops. ISIS3D Workshop in conjunction with ITS'13, October 6, 2013, Scotland, UK.
26. Wittorf, M. L., & Jakobsen, M. R. (2016). Eliciting Mid-Air Gestures for Wall-Display Interaction. Proceedings of the 9th Nordic Conference on Human-Computer Interaction - NordiCHI '16. doi:10.1145/2971485.2971503