Check for updates

A Comparative Study of Hand Gesture Recognition Devices in the Context of Game Design

Ahmed S. Khalaf
Play & Interactive Experiences for Learning Lab
Computer Science Department
New Mexico State University
khalaf@nmsu.edu

Sultan A. Alharthi
Play & Interactive Experiences for Learning Lab
Computer Science Department
New Mexico State University
salharth@nmsu.edu

Igor Dolgov
Psychology Department
New Mexico State University
id@nmsu.edu

Phoebe O. Toups Dugas
Play & Interactive Experiences for Learning Lab
Computer Science Department
New Mexico State University
phoebe.toups.dugas@acm.org

ISS'19, November 10–13, 2019, Daejeon, Republic of Korea

ACM ISBN 978-1-4503-6891-9/19/11.
http://dx.doi.org/10.1145/3343055.3360758

Abstract
Gesture recognition devices provide a new means for natural human-computer interaction. However, when selecting these devices for games, designers might find it challenging to decide which gesture recognition device will work best. In the present research, we compare three vision-based, hand gesture devices: Leap Motion, Microsoft's Kinect, and Intel's RealSense. We developed a simple hand-gesture based game to evaluate performance, cognitive demand, comfort, and player experience of using these gesture devices. We found that participants' preferred and performed much better using Leap Motion and Kinect compared to using RealSense. Leap Motion also outperformed or was equivalent to Kinect. These findings suggest that not all gesture recognition devices can be suitable for games and that designers need to make better decisions when selecting gesture recognition devices and designing gesture based games to insure the usability, accuracy, and comfort of such games.

Author Keywords
Gesture recognition devices; games; hand gesture; Leap Motion; Kinect; RealSense.

Introduction & Background
The proliferation of gesture-recognition technology enables players to interact with games in new ways [8]. Such devices enable players to use different parts of their body as a

controller, which increases physical immersion and enables designers to explore novel game mechanics and interaction methods [4].

Our hands are one of the most expressive parts of our body that can be used to interact with the world around us [9]. To detect the different hand movements and gestures, several technologies have been developed to be used for different domains and applications, including in education [19], entertainment [3], and training games [14]. However, the accuracy and performance of these technologies vary greatly [1], and in some cases these technologies even can be affected by how they are being used and designed for.

Prior studies have addressed how gesture-based interaction compares to more traditional approaches [11], as well as investigated the performance of specific devices [2]. For example, Sambrooks and Wilkinson [11], conducted a comparative study between Microsoft Kinect, mouse, and touchscreen, in which they concluded that the performance of Microsoft Kinect was much worse than mouse and touch screen for simple computer tasks. Furthermore, Seixas et al. [12] compared between Leap Motion, mouse, and touchpad, and found that Leap Motion performed poorly in pointing tasks compared to both mouse and touchpad. These studies together provide insights into the performance and usability of mid-air gesture device compared to traditional input modalities for 2D pointing tasks.

These studies together point to the advantages and disadvantages of these different input modalities, and provide insights into the performance of these different gesture recognition devices. However, they do not focus on vision-based gesture devices, do not provide enough insight into how such devices can be used in games, and are thus unable to assist game designers into making sound decisions on designing and working with hand-gesture devices in games.

Within the context of games, prior work investigated the performance and usability of different hand-gesture devices [6]. Correctly performing gestures in games can be part of the main mechanics, challenges, and experience of playing these games [17]. However, hand-gesture recognition devices need to be responsive, intuitive, and comfortable for them to be used successfully.

While all of these prior research provide an understanding of how gesture recognition devices can be used in games, and shed light on the challenges of using these devices, they do not provide game designers with an understanding of the differences between these gesture recognition devices and what they need to consider when designing games using this type of interaction.

Thus, when selecting these devices, game designers might find it challenging to decide which hand-gesture recognition device will work best for their game. The present research enables game designers to understand the differences between these devices, how to select them, and how to incorporate them within their games. Based on our motivation and prior research, we evaluate three commonly used gesture recognition devices: Leap Motion (LM), Intel RealSense (RS), and Microsoft Kinect (MK). This comparison develops an understanding of the main factors that influence their performance and use in games.

Based on our study of prior work [7, 10, 12, 16, 5, 13, 18], we hypothesized that LM would outperform MK and that MK would outperform RS in games across the performance and accuracy.

Methods
To compare between LM, RS, and MK, we run a pilot study. A within-subjects design was used consisting of one within-

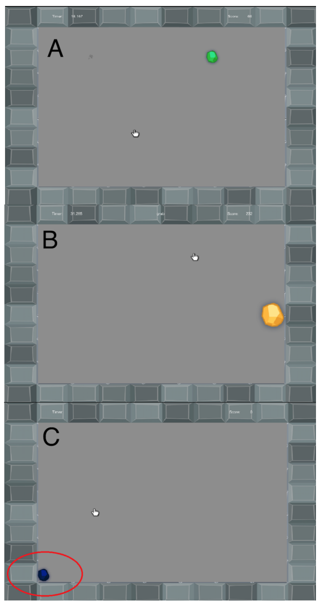

Figure 1: The hand-gesture game *Handy*. A: Players need to collected the green small object to increase their score in the game. B: players also need to collect large objects. C: in some cases, these objects are placed in difficult areas that might require players to be more precise when moving their hand and making the hand gesture.

subjects factor, device type, with three levels: (a) LM, (b) MK, (c) RS.

Participants
Three students (all male; mean age = 34.66; SD = 3.51) from New Mexico State University participated in the experiment.

Measures
Game completion time accounts for how long it took players to move from each gem to the next, performing the gesture accurately. We captured this data automatically with a logging system built into the game software, which separates how long it took the player to collect each goal.

Gesture error rate was determined by observing the player during gameplay. One researcher performed the observation and recorded false positives and negatives from the gesture system.

To gauge subjective experience, after each session, we asked participants about the perceived accuracy of each device, perceived comfort, preference, and overall experience.

Research Artifact: ***Handy***
To compare between the three hand-gesture devices, we designed *Handy*, a hand-gesture single-player game, in which the player uses their hand movements and gestures to play. The goal is to collect 36 objects that appear in a sequence and come in two different sizes (i.e., small, large) (Figure 1). These objects need to be collected by placing the player's embodiment over the gem and using a "grab" gesture as fast as possible. The player's hand position is presented as a hand-shaped cursor embodiment on a 2D space. To successfully collect one object in the game, players need to perform the following:

- move their hand in any direction using hand movements with an open palm gesture and try to position their hand over the visible game object; then
- perform the "grab" gesture to successfully collect the game object.

On each play through, the player first has 6 randomly positioned targets, which do not count for score, followed by 30 that are pre-positioned. The pre-positioning was accomplished with an algorithm to randomly position the objects a consistent distance apart; the results were saved and used for each game round. This design enables the player to have a tutorial with the new device, followed by a consistent set of targets. Players are scored based on the time it takes them to collect each object. The gesture devices can be interchanged in the game for the purpose of this study.

In each condition, the participants played the game with a different hand-gesture device. These vision-based hand-gesture recognition devices were selected based on their popularity and wide use within games:

- **Leap Motion**[1]: a small device for detecting hand position and gestures released in 2012. It is placed on a surface in front of a monitor, then can detect hands above it using IR sensors (Figure 2).
- **Intel RealSense**[2]: a device for tracking human bodies, hands, and faces released in 2015. We use the F200 model, which uses an IR depth camera and RGB camera (1080p resolution) to capture imagery 20–120cm from the device (Figure 2.
- **Microsoft Kinect v2**[3]: a motion sensing input devices by Microsoft for Xbox 360 and Xbox One video game consoles

[1] https://www.LeapMotion.com
[2] https://downloadcenter.intel.com/product/92255
[3] https://developer.microsoft.com/en-us/windows/kinect

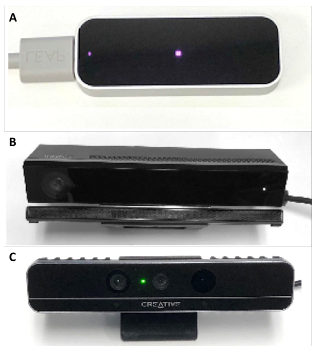

Figure 2: All the vision-based gesture recognition devices compared in this study. A: Leap Motion; B: Microsoft Kinect v2 (Windows Version); C: Intel RealSense (F200).

and Windows PCs, originally introduced in 2009, then upgraded in 2013 [15] (Figure 2).

Study Protocol
Before the beginning of the experiment, participants signed a consent form. Participants were then asked to complete the demographics questionnaire and asked about their prior experience with hand-gesture devices. Participants then played each of the three rounds of the game, with each of the different hand-gesture recognition devices. Between each round, they completed two Likert-scale-based questions. When participants finished the last session of the experiment, they were asked open ended question to assess their overall experience.

Preliminary Results

In this section, we present both the quantitative and qualitative preliminary results from our mixed methods within-subjects user study and discuss the main initial findings.

Performance
Repeated-measures analyses of variance (ANOVAs; IV: device) were used to evaluate the impact of using LM, RS, and MK on game performance, including completion time (overall, large objects, small objects) and error rate. The main effect of device was significant across all of the behavioral measures; effect sizes were all very large. Pairwise comparisons showed RS to be worse than LM and MK across all measures ($p < 0.05$ in call cases). Pairwise comparisons also showed LM outperformed MK in measure error rate ($p < 0.05$). In addition, pairwise comparisons showed that there is no significant difference between LM and MK in measures of completion time (overall, large objects, small objects).

Perceived Comfort and Accuracy
We considered the responses to the questionnaire and coded the response values 1–5, with positive responses being higher. Repeated-measures ANOVAs (IV: device) were used to evaluate the impact of using LM, RS, and MK on measures of perceived comfort and accuracy. The main effect of device was significant across both measures and the effect sizes were both large. Pairwise comparisons showed RS to be significantly worse than LM and MK across both measures ($p < 0.05$), whereas LM and MK did not differ from one another.

Player Experience
We examined the participants reflections on their experiences using these different hand-gesture devices in the game. Players stated that the accuracy of the three gestures devices varied:

> *I think leap motion is more accurate the other devices* [P3]

Players reported that they were more comfortable during the game when they used certain devices.

> *I am more comfortable with Microsoft Kinect.* [P1]

These preliminary results suggest that mid-air hand gesture devices are quite different, hard to design for, and do not support the design of cross-platform games. Game designers need to find novel solutions to overcome some of the challenges they may face when designing hand-gesture based games to provide players with an overall positive experience. The objective of this current study and research is to compare between hand gesture recognition devices in the context of games and to provide insights and point to solutions for designing hand gesture-based games. Our preliminary results show that LM has better performance and accuracy compared to MK and RS, and players felt more comfortable when they played the game using LM and MK.

Conclusion and Future work

In this study, we investigated three different vision-based gesture devices, LM, RS, and MK, in a computer game context to understand the differences in accuracy and performance between these devices. We found that participants preferred and performed better using the LM and MK than using the RS device.

For future work, we plan to run a full user study to compare between LM, RS, and MK. Also, we plan to expand the number of included gesture devices to better understand the main factors that may influence how they can be used in games (e.g., Myo armband, TAP). In addition, to understand if these devices pose different cognitive demands on player, we will use the NASA Task Load Index (NASA-TLX) to assess cognitive workload.

REFERENCES

1. Mohamed-Ikbel Boulabiar, Gilles Coppin, and Franck Poirier. 2014. The Issues of 3D Hand Gesture and Posture Recognition Using the Kinect. In *Human-Computer Interaction. Advanced Interaction Modalities and Techniques*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 205–214.
2. Diana Carvalho, Maximino Bessa, Luis Magalhães, and Eurico Carrapatoso. 2015. Performance Evaluation of Gesture-based Interaction Between Different Age Groups Using Fitts' Law. In *Proceedings of the XVI International Conference on Human Computer Interaction (Interaccion '15)*. ACM, New York, NY, USA, Article 5, 7 pages. DOI: http://dx.doi.org/10.1145/2829875.2829920
3. Harrison Cook, Quang Vinh Nguyen, and Simeon Simoff. 2015. Enabling Finger-Gesture Interaction with Kinect. In *Proceedings of the 8th International Symposium on Visual Information Communication and Interaction (VINCI '15)*. ACM, New York, NY, USA, 152–153. DOI: http://dx.doi.org/10.1145/2801040.2801060
4. Ali Erol, George Bebis, Mircea Nicolescu, Richard D. Boyle, and Xander Twombly. 2007. Vision-based hand pose estimation: A review. *Computer Vision and Image Understanding* 108, 1 (2007), 52 – 73. DOI: http://dx.doi.org/https://doi.org/10.1016/j.cviu.2006.10.012 Special Issue on Vision for Human-Computer Interaction.
5. W. Lu, Z. Tong, and J. Chu. 2016. Dynamic Hand Gesture Recognition With Leap Motion Controller. *IEEE Signal Processing Letters* 23, 9 (Sept 2016), 1188–1192. DOI: http://dx.doi.org/10.1109/LSP.2016.2590470
6. Zhihan Lv, Alaa Halawani, Shengzhong Feng, Shafiq Ur Réhman, and Haibo Li. 2015. Touch-less Interactive Augmented Reality Game on Vision-based Wearable Device. *Personal Ubiquitous Comput.* 19, 3-4 (July 2015), 551–567. DOI: http://dx.doi.org/10.1007/s00779-015-0844-1
7. Christiane Moser and Manfred Tscheligi. 2015. Physics-based Gaming: Exploring Touch vs. Mid-air Gesture Input. In *Proceedings of the 14th International Conference on Interaction Design and Children (IDC '15)*. ACM, New York, NY, USA, 291–294. DOI: http://dx.doi.org/10.1145/2771839.2771899
8. Florian 'Floyd' Mueller, Richard Byrne, Josh Andres, and Rakesh Patibanda. 2018. Experiencing the Body As Play. In *Proceedings of the 2018 CHI Conference on*

Human Factors in Computing Systems (CHI '18). ACM, New York, NY, USA, Article 210, 13 pages. DOI: http://dx.doi.org/10.1145/3173574.3173784

9. Paulo Peixoto and Joao Carreira. 2005. A Natural Hand Gesture Human Computer Interface using Contour Signatures. (2005).

10. Alexandros Pino, Evangelos Tzemis, Nikolaos Ioannou, and Georgios Kouroupetroglou. 2013. Using Kinect for 2D and 3D Pointing Tasks: Performance Evaluation. In *Human-Computer Interaction. Interaction Modalities and Techniques*, Masaaki Kurosu (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 358–367.

11. Lawrence Sambrooks and Brett Wilkinson. 2013. Comparison of Gestural, Touch, and Mouse Interaction with Fitts' Law. In *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration (OzCHI '13)*. ACM, New York, NY, USA, 119–122. DOI: http://dx.doi.org/10.1145/2541016.2541066

12. M. C. Bessa Seixas, J. C. S. Cardoso, and M. T. G. Dias. 2015. The Leap Motion movement for 2D pointing tasks: Characterisation and comparison to other devices. In *2015 International Conference on Pervasive and Embedded Computing and Communication Systems (PECCS)*. 15–24.

13. J. Svoboda, M. M. Bronstein, and M. Drahansky. 2015. Contactless biometric hand geometry recognition using a low-cost 3D camera. In *2015 International Conference on Biometrics (ICB)*. 452–457. DOI: http://dx.doi.org/10.1109/ICB.2015.7139109

14. Aaron Tabor, Scott Bateman, Erik Scheme, David R. Flatla, and Kathrin Gerling. 2017. Designing Game-Based Myoelectric Prosthesis Training. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1352–1363. DOI: http://dx.doi.org/10.1145/3025453.3025676

15. L. Vokorokos, J. Mihal'ov, and E. Chovancová. 2016. Motion sensors: Gesticulation efficiency across multiple platforms. In *2016 IEEE 20th Jubilee International Conference on Intelligent Engineering Systems (INES)*. 293–298. DOI: http://dx.doi.org/10.1109/INES.2016.7555139

16. I. Vrellis, A. Moutsioulis, and T. A. Mikropoulos. 2014. Primary School Students' Attitude towards Gesture Based Interaction: A Comparison between Microsoft Kinect and Mouse. In *2014 IEEE 14th International Conference on Advanced Learning Technologies*. 678–682. DOI: http://dx.doi.org/10.1109/ICALT.2014.199

17. Juan Pablo Wachs, Mathias Kölsch, Helman Stern, and Yael Edan. 2011. Vision-based Hand-gesture Applications. *Commun. ACM* 54, 2 (Feb. 2011), 60–71. DOI:http://dx.doi.org/10.1145/1897816.1897838

18. Frank Weichert, Daniel Bachmann, Bartholomäus Rudak, and Denis Fisseler. 2013. Analysis of the Accuracy and Robustness of the Leap Motion Controller. *Sensors* 13, 5 (2013), 6380–6393. DOI: http://dx.doi.org/10.3390/s130506380

19. Yuan Yao, Po-Tsung Chiu, and Wai-Tat Fu. 2017. A Gestural Interface for Practicing Children's Spatial Skills. In *Proceedings of the 22Nd International Conference on Intelligent User Interfaces Companion (IUI '17 Companion)*. ACM, New York, NY, USA, 43–47. DOI:http://dx.doi.org/10.1145/3030024.3038265