HandPoseMenu: Hand Posture-Based Virtual Menus for Changing Interaction Mode in 3D Space

Chanho Park
Electronics and
Telecommunications Research
Institute
Daejeon, South Korea
chanho@etri.re.kr

Young-Suk Yoon
Electronics and
Telecommunications Research
Institute
Daejeon, South Korea
ys.yoon@etri.re.kr

Hyunwoo Cho
Electronics and
Telecommunications Research
Institute
Daejeon, South Korea
nonoriri@etri.re.kr

Sung-Uk Jung
Electronics and
Telecommunications Research
Institute
Daejeon, South Korea
brcastle@etri.re.kr

Sangheon Park
Electronics and
Telecommunications Research
Institute
Daejeon, South Korea
shpark12@etri.re.kr

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

ACM ISBN 978-1-4503-6891-9/19/11.
http://dx.doi.org/10.1145/3343055.3360752

Abstract
A system control technique (changing the interaction mode) is needed to provide various spatial interaction techniques for virtual, augmented and mixed reality (VR/AR/MR) applications. We propose HandPoseMenu, a hand posture-based virtual menus system in 3D space. HandPoseMenu recognizes the non-dominant hand's posture using a depth sensor that is attached to a wearable display and presents the virtual menu corresponding to the hand posture beside the non-dominant hand. When the user selects a desired function among the presented menu as a dominant hand, the interaction mode changes. The combination of the two system controls (graphical menu and hand posture) makes it possible to display the menu that the user wants to use in 3D space. This paper describes the implementation of HandPoseMenu and its application.

CCS Concepts
•Human-centered computing → Interaction techniques; Mixed / augmented reality; Virtual reality;

Author Keywords
mixed reality; virtual reality; graphical menu; hand posture; gesture recognition; head-mounted display.

Introduction and Related Work

With the advent of the latest wearable devices and spatial interaction techniques, virtual, augmented and mixed reality (VR/AR/MR) can be applied to various applications in three-dimensional (3D) space. Users can experience the 3D space through wearable displays (head-mounted-display (HMD), HoloLens [2], etc.) and perform spatial interaction with the controller or hand gestures [10]. Because it requires various interaction functions depending on the type of application, it is necessary to be able to change the user's interaction mode during the main task.

The graphical menu is a common method for switching the interaction mode in 3D space [1, 5, 6]. If the graphical menu is fixed at a specific location [1], or the user cannot view the main task and the menu at the same time [6], the user must exit from the main task and switch the interaction techniques. It is also possible to provide a hierarchical menu to occupy a minimal amount of the field of view, but it can take a long time to select a target when the target is deeply hidden [5].

Gesture interface [12, 14, 18, 19] which involves a change of position and/or orientation of the body can be a powerful system control technique. Generally, one gesture interface provides one interaction function by the user-defined approach. Although it has the advantage of quickly changing the interaction mode, all gestures must be learned by the users to provide various kinds of interaction functions. On the other hand, users can use a physical tool's corresponding grasp as a gesture [20]. This approach provides tool-related interaction functions and is easy for the users to memorize, but similar grasping gestures (due to similar shapes of the objects) can confuse the users.

We propose HandPoseMenu (see Figure 1), a hand posture-based graphical menus system for changing the user interaction mode in 3D space. It uses a depth sensor to recognize the non-dominant hand posture and displays a virtual menu corresponding to a user-defined hand posture beside the user's non-dominant hand. The user can switch the interaction function by selecting the menu target with the dominant hand while performing the main task.

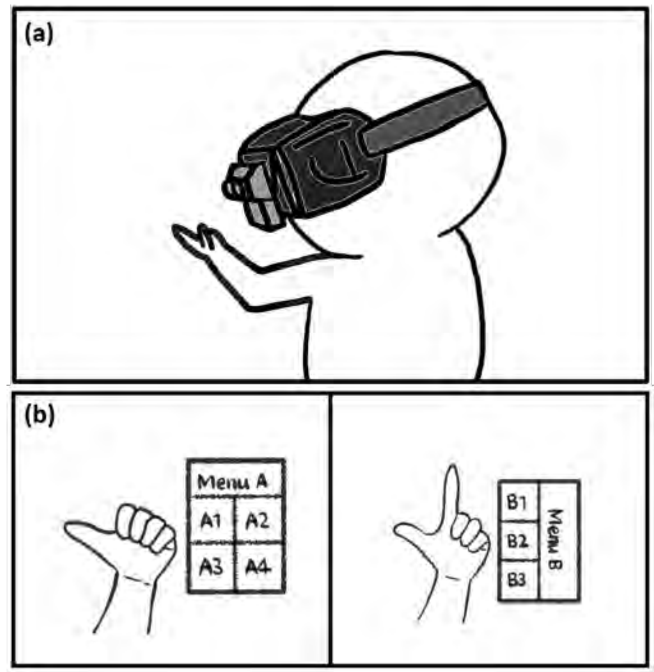

Figure 1: The concept of HandPoseMenu. (a) Example use of HandPoseMenu (b) Various menus can be presented according to the hand posture

Figure 2: Implementation of HandPoseMenu

Figure 3: Set of candidate non-dominant hand postures

HandPoseMenu

HandPoseMenu is displayed in 3D space around the user's hand [4, 7] providing a wider display area compared to the body parts (forearm [6], hand [11]) and real objects [13]. To avoid confusion with the user's natural hand movement and to improve the hand recognition rate of our system, a menu is displayed on the side of the little finger [4] as shown in Figure 1(b). Our system can present the menu reliably by using the hand posture which contains only static movement, and it is easy for the users to learn. For the touch screen, a system control method that combines a hand posture and a graphical menu [8, 16] was proposed, but we move to a 3D environment (VR and MR). Similar to [16], our system recognizes the non-dominant hand's posture, presents a menu corresponding to the posture, and selects the target menu to change the interaction mode with the dominant hand.

Implementation

We implemented HandPoseMenu (see Figure 2) using Oculus Rift DK2 (960 x 1080 per eye at a 75Hz, 100° FOV diagonally) for VR space. We also considered MR space with Ovrvision Pro, a front-facing stereo camera (960 x 950 per eye at a 60Hz, 99° FOV diagonally). Leap Motion was attached to HMD for hand tracking and hand posture recognition. We used Unity 3D to render graphics.

Hand Posture Recognition

HandPoseMenu recognizes hand posture by Leap Motion attached to the front of the wearable display. We selected 16 hand postures as candidates (see Figure 3) to create a machine learning model by referring to previous studies [17, 21]. We recruited 6 right-handed participants, ages 28 to 34, and collected non-dominant hand (left hand) posture data. In this task, the subject did not wear a display because the subject had to take the same hand posture as the posture in the image presented in the experiment. In addition, our system was intended to present a virtual menu based on egocentric hand posture, but we collected the data by placing the Leap Motion on the desk in order to easily acquire 3D skeleton data based on the top center of the sensor.

Before the experiment, the subjects had preliminarily practiced a set of hand postures and measured the 3D distance (H_{length}) from tip of thumb to tip of little finger at the P2 posture. When the experiment started, the subject took each hand posture that appeared on the screen and recorded hand skeleton data obtained from the Leap Motion SDK at a 50Hz frame rate for 5 seconds. The extracted 29 features were as follows (feature values related to distance were divided by H_{length} to make our system robust to different hand sizes [15]):

- The distances between the fingertips and palm center (5 features)
- The distances of the fingertips from the plane corresponding to the palm region (5 features)
- The distances between fingertips (4 features)
- The angles corresponding to the orientation of the fingertips with respect to the hand orientation (5 features)
- The angles between adjacent fingertips (10 features)

Each hand posture was repeated five times, and a total of 120000 data frames (6 people x 16 hand postures x 5 times x 5 seconds x 50Hz) were collected. We used LIBSVM [9], a library for support vector machines, to classify the 16 hand postures. The results of a 6-fold cross validation showed a mean accuracy rate of 94.49%.

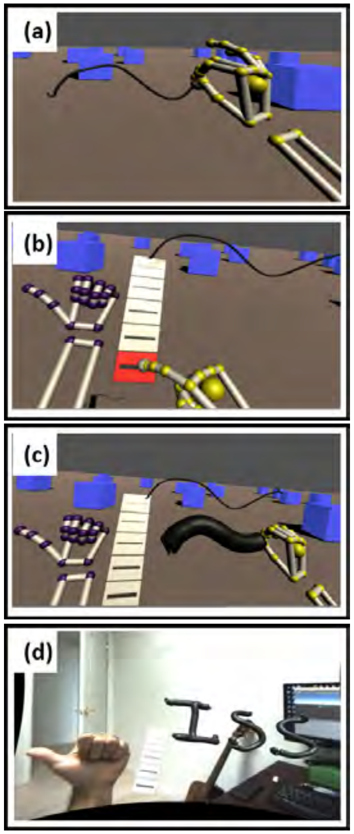

Figure 4: (a) 3D drawing task in VR (b) Select the thickest pen (changing the interaction mode) (c) Perform the 3D drawing task with the thickest pen (d) 3D drawing task in MR

Controlling Graphical Menu
While the user is performing the main task in 3D space, we set the following trigger to display the graphical menu:

- The angle between the vertical vector at the top of the Leap Motion and the normal vector of the back of the hand is within θ $(= 60°)$

If the hand tracking condition is satisfied, the hand posture is classified in real-time based on the SVM model. Our system determines the final hand posture in the current frame as the most counted posture for t $(= \ 0.5s)$. Then, a virtual menu corresponding to a predetermined hand posture is displayed. When a graphical menu appears, the user changes the interaction mode with a dominant hand. The target menu is selected when the 3D distance between the index fingertip of the dominant hand and the target menu is within D_{thr} $(= 1.5cm)$.

Application

We demonstrated a 3D drawing and modeling task application using the HandPoseMenu. We applied a learning model using 6 hand postures (P1∼P6) to the application for real-time hand posture classification and configured the corresponding menu as shown in Figure 5. First, the user controls the activation of the interaction mode of the dominant hand by the P1 posture (Figure 5(a)). Then, the user creates a pre-made 3D model or perform a 3D drawing using the thickness, color, material menus (Figure 5(b)∼(e)). Also, the user can delete the 3D objects or drawing contents (Figure 5(f)).

For example, Figure 4(a) shows a 3D drawing task using the pinch gesture of the dominant hand [3] in VR. When the user wants to change the thickness of the pen, the user can take the P3 posture to display a thickness menu, select the thickest pen, and finally perform a 3D drawing with

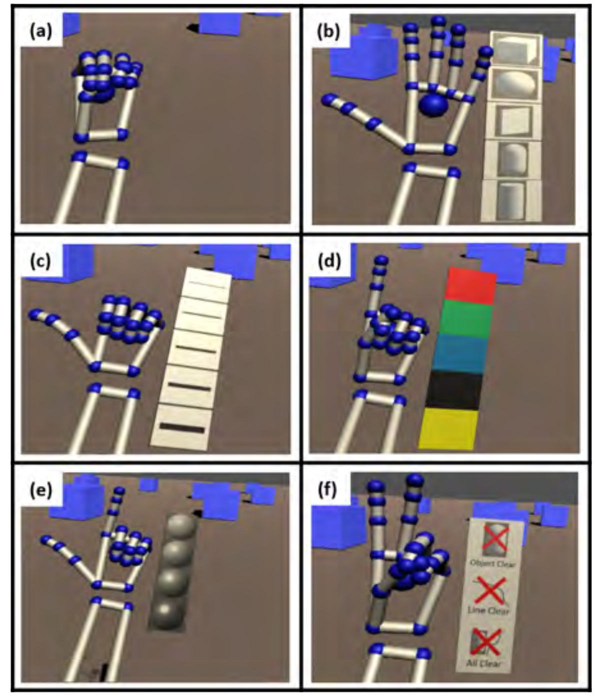

Figure 5: An example of HandPoseMenu in 3D modeling and drawing (a) P1: activate/deactivate the interaction mode (dominant hand) (b) P2: create a 3d model (c) P3: thickness of the pen (d) P4: color of the pen (e) P5: material of the pen (f) P6: delete the 3D objects or drawing contents

the thickest pen (Figure4(b)∼(c)). Also, our system can be used in MR environment (Figure 4(d)).

Conclusion and Future Work

In this study, we proposed a hand posture based graphical menus system: HandPoseMenu. The user can control the virtual menus using hand postures corresponding to the purpose of the main task, and can change the interaction mode. In future work, we will compare the usability of our proposed method, which combines a graphical menu and hand posture, with the existing system control methods (only the graphical menu or the gesture interface).

Acknowledgments

This research was supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program. [1375026932, Development of participational AR platform for large-scale cultural space]

REFERENCES

1. 2019. Holostudio. (2019). Retrieved June 30, 2019 from https://www.youtube.com/watch?v=BRIJG0x_We8.
2. 2019. Microsoft HoloLens. (2019). Retrieved June 30, 2019 from https://www.microsoft.com/en-CY/hololens.
3. 2019. Pinch Draw Module. (2019). Retrieved June 30, 2019 from https://gallery.leapmotion.com/pinch-draw/.
4. 2019. Project North Star. (2019). Retrieved June 30, 2019 from https://developer.leapmotion.com/northstar.
5. 2019. Steam VR Environment: Destinations. (2019). Retrieved June 30, 2019 from https://developer.valvesoftware.com/wiki/Destinations/SteamVR_Environment.
6. Takumi Azai, Shuhei Ogawa, Mai Otsuki, Fumihisa Shibata, and Asako Kimura. 2017. Selection and Manipulation Methods for a Menu Widget on the Human Forearm. In *Proceedings of the CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, 357–360. DOI:http://dx.doi.org/10.1145/3027063.3052959
7. Takumi Azai, Mai Otsuki, Fumihisa Shibata, and Asako Kimura. 2018. Open Palm Menu: A Virtual Menu Placed in Front of the Palm. In *Proceedings of the 9th Augmented Human International Conference (AH '18)*. ACM, Article 17, 5 pages. DOI: http://dx.doi.org/10.1145/3174910.3174929
8. Drini Cami, Fabrice Matulic, Richard G. Calland, Brian Vogel, and Daniel Vogel. 2018. Unimanual Pen+Touch Input Using Variations of Precision Grip Postures. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. ACM, 825–837. DOI: http://dx.doi.org/10.1145/3242587.3242652
9. Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* (2011), 27:1–27:27. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.
10. Barrett Ens, Ahmad Byagowi, Teng Han, Juan David Hincapié-Ramos, and Pourang Irani. 2016. Combining Ring Input with Hand Tracking for Precise, Natural Interaction with Spatial Analytic Interfaces. In *Proceedings of the Symposium on Spatial User Interaction (SUI '16)*. ACM, 99–102. DOI: http://dx.doi.org/10.1145/2983310.2985757
11. Zhenyi He and Xubo Yang. 2014. Hand-based Interaction for Object Manipulation with Augmented Reality Glasses. In *Proceedings of the 13th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry (VRCAI*

'14). ACM, 227–230. DOI: http://dx.doi.org/10.1145/2670473.2670505

12. Yangho Ji, Sunmok Kim, Young-Joo Kim, and Ki-Baek Lee. 2018. Human-like sign-language learning method using deep learning. *ETRI Journal* 40, 4 (August 2018), 435–445. DOI: http://dx.doi.org/10.4218/etrij.2018-0066
13. Hyeongmook Lee and Woontack Woo. 2010. Tangible Spin Cube for 3D Ring Menu in Real Space. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems (CHI EA '10)*. ACM, 4147–4152. DOI: http://dx.doi.org/10.1145/1753846.1754117
14. Shahzad Malik, Abhishek Ranjan, and Ravin Balakrishnan. 2017. Interacting with Large Displays from a Distance with Vision-Tracked Multi-Finger Gestureal Input. In *Proceedings of the ACM symposium on User interface software and technology (UIST '05)*. ACM, 43–52. DOI: http://dx.doi.org/10.1145/1095034.1095042
15. Giulio Marin, Fabio Dominio, and Pietro Zanuttigh. 2014. Hand gesture recognition with leap motion and kinect devices. In *2014 IEEE International Conference on Image Processing (ICIP)*. IEEE, 1565–1569. DOI: http://dx.doi.org/10.1109/ICIP.2014.7025313
16. Fabrice Matulic, Daniel Vogel, and Raimund Dachselt. 2017. Hand Contact Shape Recognition for Posture-Based Tabletop Widgets and Interaction. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17)*. ACM, 3–11. DOI: http://dx.doi.org/10.1145/3132272.3134126
17. Ludovico Minto, Giulio Marin, and Pietro Zanuttigh. 2015. 3D hand shape analysis for palm and fingers identification. In *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, Vol. 07. IEEE, 1–6. DOI: http://dx.doi.org/10.1109/FG.2015.7284879
18. Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. 2013. User-Defined Gestures for Augmented Reality. In *Human-Computer Interaction - INTERACT 2013 (INTERACT '13)*, Vol. 8118. Lecture Notes in Computer Science, Springer, 282–299. DOI: http://dx.doi.org/10.1007/978-3-642-40480-1_18
19. Hemant Bhaskar Surale, Fabrice Matulic, and Daniel Vogel. 2019. Experimental Analysis of Barehand Mid-air Mode-Switching Techniques in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, Article 196, 14 pages. DOI: http://dx.doi.org/10.1145/3290605.3300426
20. Yukang Yan, Chun Yu, Xiaojuan Ma, Xin Yi, Sun Ke, and Yuanchun Shi. 2018. VirtualGrasp: Leveraging Experience of Interacting with Physical Objects to Facilitate Digital Object Retrieval. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, 1–13. DOI: http://dx.doi.org/10.1145/3173574.3173652
21. Yifan Zhang, Congqi Cao, Jian Cheng, and Hanqing Lu. 2018. EgoGesture: A New Dataset and Benchmark for Egocentric Hand Gesture Recognition. *IEEE Transactions on Multimedia* 20, 5 (May 2018), 1038–1050. DOI: http://dx.doi.org/10.1109/TMM.2018.2808769