Implementing Hand Gestures for a Room-scale Virtual Reality Shopping System

Chunmeng Lu
Waseda University
Kitakyushu, Japan
Email: lcm0113@163.com

Boyang Liu
Waseda University
Kitakyushu, Japan
Email: waseda-liuboyang@moegi.waseda.jp

Jiro Tanaka
Waseda University
Kitakyushu, Japan
Email: jiro@aoni.waseda.jp

***Abstract*—In the virtual reality (VR) environment, the user is required to input information and achieve interaction with virtual objects. At present, most VR systems provide some input devices, such as keyboard and controller. However, utilizing such devices is not intuitive, especially in the case of a VR shopping system. In real applications, we use our hands to handle objects. In virtual applications, using hand gestures to interact with a VR shopping store will provide us a more intuitive VR shopping experience. Following the needs of the room-scale VR shopping activities, we have introduced a new gesture classification for the gesture set, which has three levels to classify hand gestures based on the characteristic of gestures. We have focused on the gestures in level 3. We have built a room-scale VR shopping system and applied the new hand gesture set for the interaction in the VR shopping system. We conducted experiments to evaluate the accuracy of gestures in a VR shopping environment. The classification and set of gestures together when evaluated showed that these specific gestures were recognized with high accuracy.**

Keywords- Room-scale Virtual Reality Shopping; Gesture set; Gesture classification; Gesture interaction.

I. INTRODUCTION

Currently, people can roam in the virtual environment through Head-Mounted Display (HMD). As shopping is one of the most important activities in the real world, a virtual shopping environment can be a part of the virtual environment. We are familiar with e-commerce or online shopping. We can extend online shopping to the virtual environment [1] because the virtual reality (VR) shopping experience has the potential to allow users to surpass geography and other restrictions.

With the improvement of VR technology, many researchers and companies attempt to apply VR technology in the e-commerce field to find profitable economic value. Alibaba is a famous IT company and is known for its online shopping services. Tianmao is one of its online shopping services. Tianmao presented a VR shopping application, called Tianmao buy+ [2], for smartphones. The Tianmao buy+ attempted to combine the convenience of online shopping and the facticity of physical store shopping. From simple and cheap VR devices and smartphones, people in China can view stores around the American Times Square and pay for orders online. The VR application creates in people a feeling of the shopping experience at the American Times Square.

Some companies employ VR technology to create virtual stores. The furniture company IKEA presented a room-scale VR kitchen to show its design [3]. In the room-scale VR kitchen, a user can utilize HTC vive to view the equal proportion VR kitchen, and even interact with the VR environment compared to physical furniture stores, the VR environment can provide more functions and interactions. Users can view the kitchen freely in the comfort of their room. In the VR kitchen, a user can easily change the color of the furniture, an impossible task in a physical store.

The VR technology company, inVRsion, presents a VR supermarket system based on room-scale VR [4]. Their retail space, products, and shopping VR experience solutions provide an immersive shopping environment. In the VR shopping environment, a manager can analyze customer behavior through eye-tracking for market research insights. The system can help the seller to test his category projects, new packaging, and communication instore before implementation. A user can search his target products more easily than in a physical supermarket. This system tries to provide a method for users to view a big virtual supermarket in a room.

This article is organized as follows. In Section I, the Introduction is presented. Section II specifies the problem to be solved. Section III presents the research purposes and approaches. Section IV presents related works. Section V introduces the design of the system. Section VI introduces the implementation of the designed system. Section VII specifies the detailed process of gesture recognition. In Section VIII, we discuss how to apply gestures set in a room-scale VR shopping environment. In Section IX, we present evaluation results. In Section X, we conclude and present future works.

II. PROBLEM

The VR shopping environment provides an emulated environment in which the virtual objects are similar to physical objects. We normally use our hands to touch and grab objects around us. Thus, the use of controllers in a VR shopping environment is not sufficiently immersive (see Figure 1). The user lacks the feeling sensations to hold a virtual object when utilizing controllers.

The number and function of buttons in the controllers are limited (see Figure 2), thus restricting the interaction with products when they are used in a room-scale VR environment. Conversely, the human hand can perform various gestures for human-computer interaction. Compared to controllers, hand gestures can improve the VR immersion shopping experience and allow a greater dictionary of commands for the VR shopping system.

Figure 1. Using controllers in the virtual environment

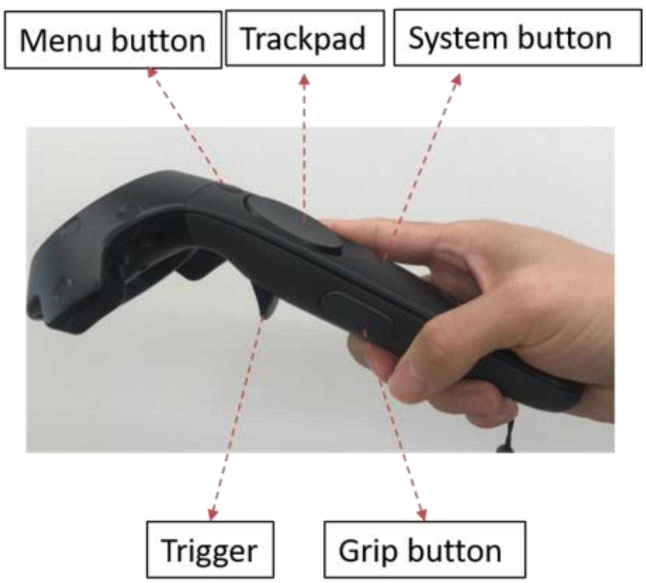


Figure 2. The buttons of the controller

III. Goal And Approach

The main features of VR include immersion, plausibility, interaction, and fidelity. In this research, we aim to present a new hand gesture set suitable for room-scale VR shopping activity to replace the controllers and improve immersion experience. We introduce a new gesture classification functionality to achieve a more structured gesture set. We apply a room-scale shopping system to provide an immersive virtual shopping environment to simulate a physical shopping store. In the room-scale VR shopping environment, the user is able to walk in it and to view the virtual shopping environment through the HMD. In the designed new gestures for the room-scale VR shopping system, users can interact with the VR environment by natural hand gestures instead of with controllers. We introduce a gesture classification functionality with three levels to classify hand gestures based on gesture characteristics. By summarizing the hand gestures, we obtain a new hand gesture set for room-scale VR shopping activity. The hand gesture set is expected to improve user convenience and immersion experience within a room-scale VR shopping system.

We utilize the VR devices to build the room-scale VR shopping system with two sensor stations installed in the room. These sensor stations create a walking area for the user. When moving in this area, user motion information is captured by the sensor stations. The system receives rotation and three-dimensional data coordinates from the HMD worn by the user. The sight vision in the virtual environment moves synchronously with HMD. We employ a depth sensor to recognize the hand gestures. The virtual environment allows the user to experience his virtual and physical hands moving synchronously. The system can distinguish the gestures and allow for interaction between the user and with the virtual environment.

IV. Related Work

A. Virtual Reality

Virtual Reality (VR) allows for users to interact in a virtual environment as one is in the physical world. It provides an illusion of "being there" [5].

In previous years, the development of computer graphics, 3D technology, and electrical engineering, permitted improvements in HMD of VR. Recently, VR devices have gained space outside of the laboratory. Some companies have introduced simple and easy-to-use VR devices for the consumer market, such as HTC Vive and Oculus Rift.

B. VR Shopping

People can navigate in the virtual environment through HMD. In the last decades, many VR shopping environments have been presented. Some works aim at improving VR shopping experience, while others research on the interaction in virtual shopping environments. Bhatt presented a theoretical framework to attract customers through a website based on these three features: interactivity, immersion, and connectivity [6]. Chen et al. presented a VRML-Based virtual shopping mall. They analyzed the behavior of a customer in a Virtual Shopping Mall System. They also explored the application of intelligent agents in shopping guidance [7]. Lee et al. designed a virtual interactive shopping environment and analyzed if the virtual interface had positive effects [8]. Verhulst et al. presented a VR user study. They applied the VR store as a tool to determine if the user in the store wished to buy food [9]. Speicher et al. introduced a VR Shopping Experience model. Their model considered three aspects: customer satisfaction, task performance, and user preference [10].

The previous researches presented good features of VR shopping that raised interests from retail and online shopping companies for VR shopping. The company IKEA presented a room-scale VR environment in which the user can view a virtual kitchen and interact with the furniture [3]. In another example, inVRsion provides a virtual supermarket shopping system, the Shelfzone VR [4]. In the future, more applications are expected for VR shopping.

C. Room-scale VR and Hand Gesture

When the user is moving his gaze in the virtual world with HMD, he cannot move his physical body in the real world. Thus, the absence of sensory movement experience between the virtual and physical environments reduces immersion experience in VR. This absence also causes motion sickness in some users [10][11]. Nevertheless, if walking is synchronized both in virtual and physical environments, this provides an improved sensory experience.

Room-scale VR shopping environment provides an emulated environment in which virtual objects are similar to real objects. Commonly, we use our hands to touch and hold objects around us. Thus, the use of controllers in VR shopping environment causes it not to be sufficiently immersive. The user lacks the sensation of holding virtual objects similar to physical ones when using controllers. Additionally, the number and function of buttons in the controllers are limited, which restricts the interaction when utilizing them in a room-scale VR environment. There are three buttons on a controller. To achieve VR shopping activities, these three buttons are designed for the interaction method in the VR shopping system. Compared with controllers, the use of hand gestures can improve the VR immersion shopping experience and provide a rich interaction dictionary of commands for the VR shopping systems.

Hand gestures have widely been utilized in human-computer interfaces. Gesture-based interaction provides a natural intuitive communication between people and devices. People use 2D multi-touch gesture to interact with devices such as smartphones and computers in daily life. The 3D hand gesture can be employed for some devices equipped with a camera or depth sensor. The most important point in hand gesture interaction is how to provide computers the ability to recognize commands from hand gestures [12]. Wachs et al. summarized the requirements of hand-gesture interfaces and the challenges when applying hand gestures in different applications [13]. Yves et al. presented a framework for 3D visualization and manipulation in an immersive space. Their work can be used in AR and VR systems [14]. Karam et al. employed a depth camera and presented a two-hand interactive menu to improve efficiency [15]. These previous studies show that there is potential for hand gestures in the human-computer interaction field.

In a previous work, we extended 2D multi-touch interaction to 3D space and introduced a universal multi-touch gesture for 3D space [16]. We called these midair gestures in 3D as 3D multi-touch-like gestures (see Figure 3).

To recognize 3D multi-touch-like gestures, we present a method using machine learning. However, the use of machine learning alone is not sufficient to perform the task accordingly. Although machine learning techniques can recognize the hand shape, it fails to address the begin and end timing of gestures movement. If we cannot precisely recognize the gesture timeframe, it is difficult to provide a fast response to the performed gesture. Therefore, proper timing of events is required. We use a depth camera to detect the state of fingers to discover the begin and end timeframe of the gestures.

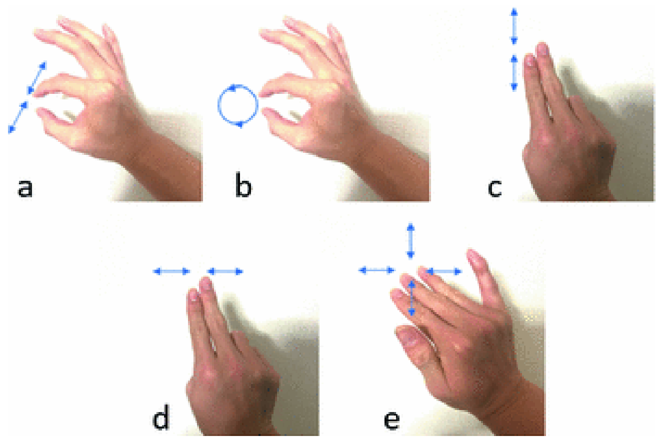


Figure 3. Five 3D multi-touch-like gestures: (a) zoom in/out, (b) rotate, (c) scroll, (d) swipe, and (e) drag

The previous related works elucidated the broad application prospects for room-scale VR shopping and gestures. The proposed work on gesture set design for room-scale VR can be used in these systems to provide immersive VR shopping experience.

V. SYSTEM DESIGN

A. Room-scale VR Shopping Environment

To achieve VR shopping activity, we design a VR shopping store as the shopping environment, similar to physical stores. We arrange desks, shelves, and goods in the VR shopping store as shown in Figure 4.

Figure 4. A VR store

The room is provided with an empty area in a real room, which is included in a 3D space. We use a tracking sensor to capture the motion and rotation of HMD in the 3D space for the VR shopping system use. As shown in Figure 5, the length, width, and height of the 3D space is 4 m, 3 m, and 2 m, respectively. The 3D space contains walking area of the room.

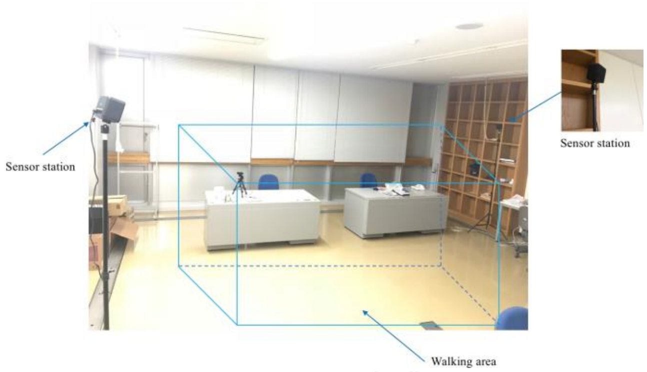


Figure 5. 3D space and walking area

In the room-scale VR shopping environment, there is also a virtual walking area, as shown in Figure 6. The virtual walking area is the same as the area in the actual room. Because the VR shopping store is larger than the real room, the user can change the virtual walking area while looking at the entire VR shopping store.

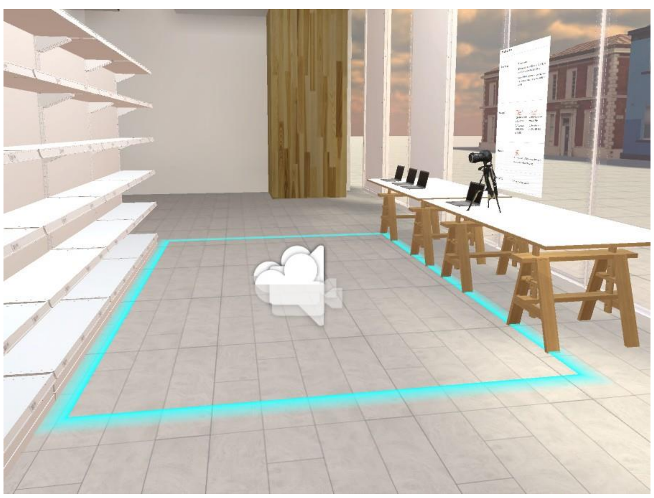


Figure 6. Walking area in VR environment

B. Gesture Set

In our research, we designed a series of hand gestures specially for the room-scale VR shopping activity. The hand gestures must provide natural and suitable interaction between the user and system. Based on the specific activities in room-scale VR shopping system, we design these 14 gestures in the system according to previous work [16]:

(1) OK gesture;
(2) No gesture;
(3) push/pull gesture;
(4) waving gesture;
(5) pointing gesture;
(6) grab gesture;
(7) holding gesture;
(8) drag gesture;
(9) rotation gesture;
(10) zoom in/out gesture;
(11) click gesture;
(12) two-fingers scroll/swipe gesture;
(13) opening/closing gesture; and
(14) changing area gesture.

These gestures create a new gesture set for a room-level VR shopping system.

C. Gesture Classification to define levels

We propose a new type of gesture classification that classifies the hand gestures according to different characteristics of gestures. There are three levels of gesture classification:

Level 1: Core static hand positions are classified into level 1. In level 1, gestures englobe hand shape without hand movements. The classic example is pointing gestures.

Level 2: Dynamic palm motions are classified into level 2. Level 2 only considers the palm movement regardless of finger shape. The classic example is pulling and pushing.

Level 3: Combined hand gestures are classified into level 3; they combine the features of level 1 and level 2 gestures. Level 3 considers the finger movement and shape and palm movement.

In the VR shopping environment, the hand gestures are divided into different levels. In Figures 8 and 9, the red arrow indicates the finger movement direction.

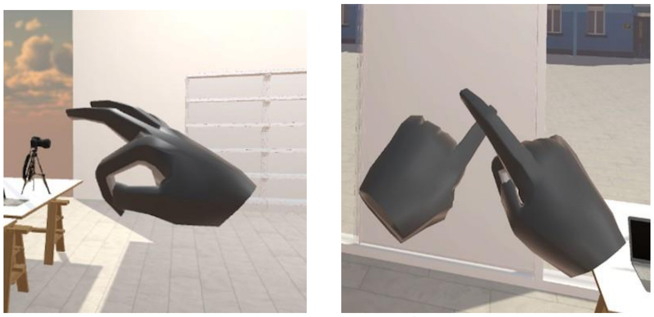
The gesture classification method classifies the gesture set in a structured way and provides a structure that can be used in other gesture sets in different VR systems. Based on the system structure, researchers can design appropriate gestures for their own VR systems.

1) Level 1 Gestures:

In our system, we use level 1 gestures to send feedback to the system. These gestures are static signals of the shopping system, and the system only needs to detect the hand shape. Table I shows the functions of level 1 gestures, and Figure 7 depicts two level 1 gestures.

TABLE I. FUNCTION OF GESTURES IN LEVEL 1

Level	Gesture	Function
1	OK	Inform a positive feedback to the system.
1	NO	Inform a negative feedback to system.



(a) OK gesture (b) NO gesture

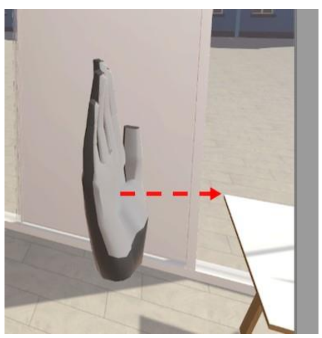
Figure 7. OK and NO gestures positive or negative feedback to system

2) Level 2 Gestures:

Level 2 Gestures are palm movements. After selecting a virtual object, the user can control or interact with level 2 gestures. Table II shows the functions of level 2 gestures, and Figure 8 shows two level 2 gestures.

TABLE II. FUNCTION OF GESTURES IN LEVEL 2

Level	Gesture	Function
2	Push/pull	Push or pull a virtual object with a hand.
2	Waving	Make virtual object return to the original position.



(a-1) Push gesture (a-2) Pull gesture

(b) Waving gesture

Figure 8. Level 2 gestures: push/pull and waving

3) Level 3 Gestures:

Gestures in level 3 combine hand gestures of finger shapes and hand movements. They combine the features of Level 1 and Level 2 gestures.

The design of a suitable and convenient gesture set for the user determines an immersive VR shopping experience. In levels 1 and 2, gestures are simple. Because level 3 gestures are more complex, they are the focus of this research.

In level 3, the system identifies hand shapes and detects finger and hands movement simultaneously.

As the gestures indicate different messages they are divided into three classification categories:

- The core gesture: pointing gesture
- Gestures for interacting with a virtual object: (1) grab gesture, (2) holding gesture, (3) drag gesture, (4) rotation gesture, and (5) zoom in/out gesture;
- Gestures for interacting with menu: (1) click gesture, (2) scroll/swipe gesture, and (3) opening/closing gesture;
- Gesture for interacting with space: change area gesture

In the gesture set, the pointing gesture is the most important gesture, because it is used to select a target object or button before any interaction, as shown in Table III and Figure 9.

TABLE III. FUNCTION OF POINTING GESTURE

Level	Gesture	Function
3	Pointing	Point a virtual object with the index finger.

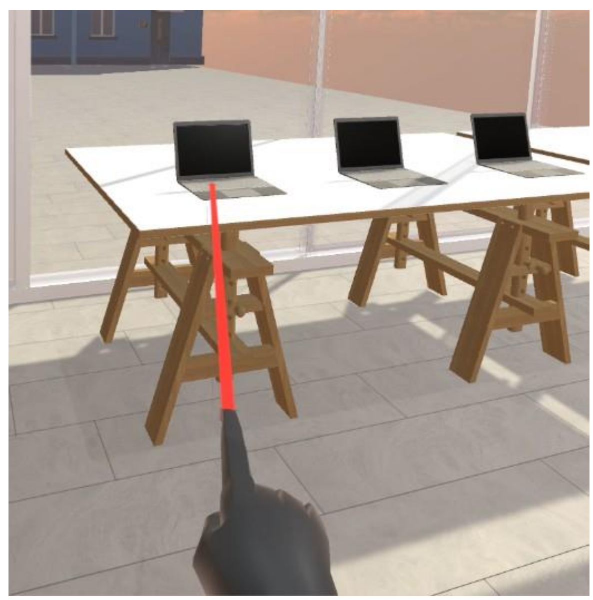


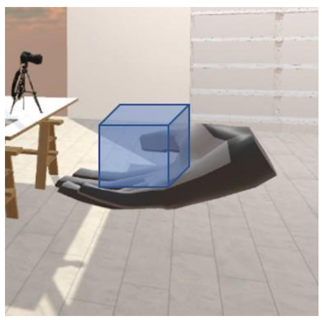
Figure 9. Pointing gesture

Some gestures are mainly used to interact with virtual objects in a VR shopping store, such as moving a virtual object. We design these gestures to achieve the following simple functions: grab, hold, drag, rotation, and zoom in/out, as shown in Table IV and Figure 10.

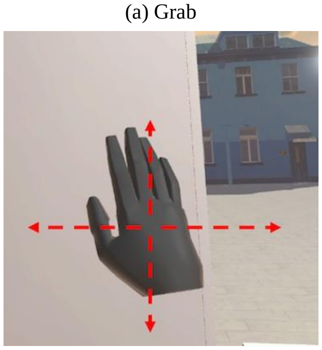
TABLE IV. FUNCTIONS OF GESTURES FOR INTERACTING WITH OBJECT

Level	Gesture	Function
3	Grab	The object is approximated to the hand that grabs it.
3	Hold	Holds a virtual object with one hand.
3	Drag	Move virtual object freely with a hand.
3	Rotation	Rotate a virtual object when observing it with a hand.
3	Zoom in/out	Zooms in or out of a virtual object using the relative motion of thumb and index finger.

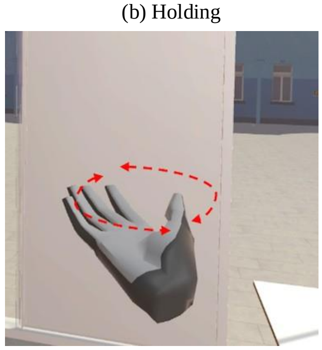
(a) Grab



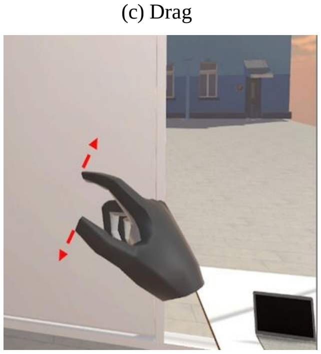
(b) Holding



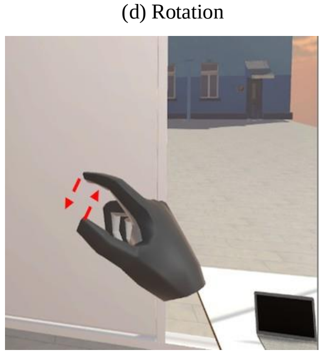
(c) Drag



(d) Rotation



(e-1) Zoom in



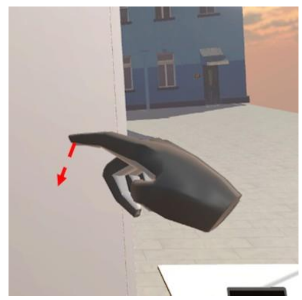
(e-2) Zoom out

Figure 10. Gestures for interacting with object

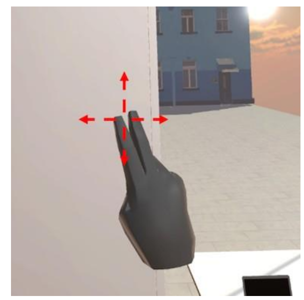
In some cases, the user requires to interact with the menu to perform shopping activities. To achieve that, we design the following gestures: click, scroll/swipe, opening/closing, as shown in Table V and Figure 11.

TABLE V. FUNCTIONS OF THE GESTURES FOR INTERACTING WITH MENU

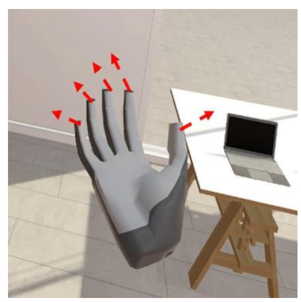
Level	Gesture	Function
3	Click	Click the buttons with index finger.
3	Scroll/swipe	Use two fingers gestures to control menus in user interface.
3	Opening/closing	Spread the fingers to open the dashboard or close the hand to close the dashboard.



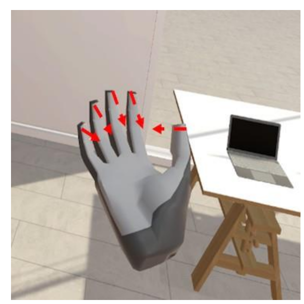
(a) Click



(b) Scroll/swipe



(c-1) Opening



(c-2) Closing

Figure 11. Gestures for interacting with the menu

In a room-scale VR shopping system, the user can walk in the physical walking area in the room. However, owing to the size of the room, the pedestrian zone may sometimes be smaller than the VR shopping store. Thus, we design a gesture for the user to change the area in the room-scale VR shopping store to meet this need. Table VI and Figure 12 show the change in area gesture. To accomplish the area change gesture, one must stretch the index finger and thumb. In the system, when the user intends to change the position of the VR shopping store, one can point the position on the virtual floor with this special hand gesture and use the thumb to click, and the index finger to inform the system where one intends to go.

TABLE VI. FUNCTIONS OF THE GESTURES FOR CHANGING AREA

Level	Gesture	Function
3	Changing area	Use index finger to point at a new position on the floor and make thumb click index finger, then view in VR will move to the new position.

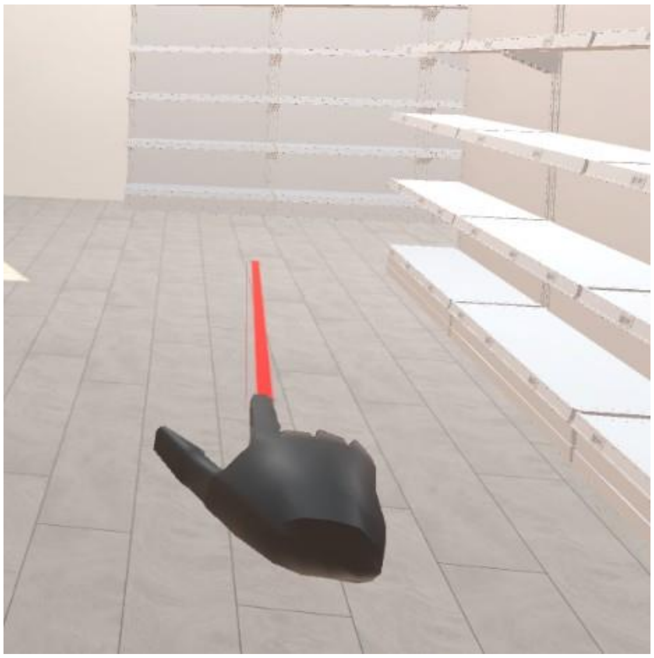


Figure 12. Gesture for changing area in VR environment

VI. SYSTEM IMPLEMENTATION

In the room-scale VR shopping system, we use HTC Vive as the VR device. The HTC Vive has a head-mounted display (HMD), two controllers, and two base stations. The HMD offers dual 3.6-inch screens with 1080 x 1200 pixels per eye. The refresh rate is 90 Hz, and the field of view is 110 degrees. The controllers have five buttons: multifunction trackpad, grip buttons, dual-stage trigger, system button, and menu button. The two base stations achieve room-scale tracking of HMD and controllers.

We use Leap Motion [17] as the depth sensor to recognize the hand gesture. Leap Motion can track the coordinate and the fingertip and center of palm rotations and transfer the data to the VR system. In the VR system, Leap Motion is mounted on the HMD, as shown in Figure 13.

The VR device and depth sensor require a PC to work accordingly. We use the Unity 3D as the software to build the room-scale VR system. With the Unity 3D, data is processed from Leap Motion and design the virtual shopping environment.

In the room-scale VR shopping system, the user can walk in his own room. Therefore, it is necessary to track the head movement of the user in the 3D space of the room. HTC Vive provides the base stations tracking sensor. We place the two base stations in the two corners of the room.

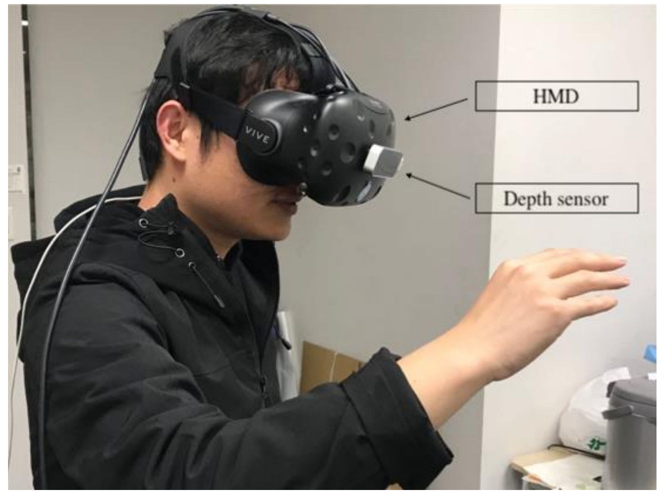


Figure 13. HMD and depth sensor

VII. GESTURE RECOGNITION

We use Leap Motion as the depth sensor to track the hand. Leap Motion tracks the joints, fingertips, and palm center of the hand of the user. In addition, Leap Motion records the positions of these important points in the hand of the user in every frame.

With the original position data, we use Machine Learning methods to recognize the hand shapes. By combining the hand shapes and motion, we achieve recognition of the gestures that we design for the room-scale VR shopping system.

A. Hand Shape Recognition

First, we need to confirm how many hand shapes need to be identified because different hand gestures have the same handshape. For example, drag, holding, and rotation gestures share the same handshape. Their differences lie in the palm movement. As we have introduced all 14 gestures, we summarize the following hand shapes that we recognize, also as shown in Figure 14.

The system should be able to distinguish natural hand movements from interaction movements. Therefore, we need to recognize the natural hand shape. There are nine hand shapes that the system needs to realize.

Below is the relationship between the 14 hand gestures and the nine hand shapes:

a. OK handshape (including one gesture): OK gesture.

b. Pointing handshape (including three gestures): pointing gesture, NO gesture, click gesture.

c. Extending the handshape (including five gestures): push/pull, waving, holding, drag, and rotation gestures.

d. Grab handshape (including one gesture): grab gesture.

e. Zoom handshape (including one gesture): zoom in/out gesture.

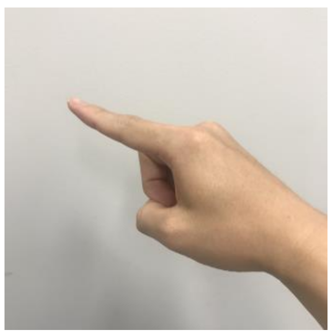
f. Scroll/swipe handshape (including one gesture): scroll/swipe gesture.

g. Opening/closing handshape (including one gesture): opening/closing gesture.

h. Changing area handshape (including one gesture): changing area gesture.

i. Natural hand shape: the hand shape when the user move the hands without interaction intention.

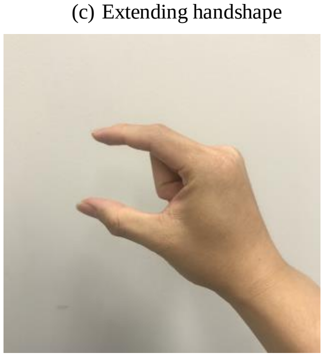
(a) OK hand shape



(b) Pointing hand shape

(c) Extending handshape

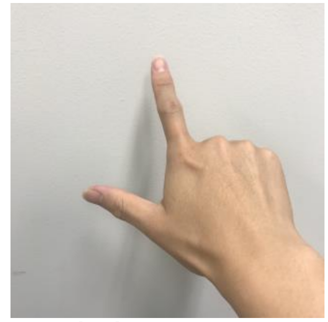
(d) Grab handshape



(e) Zoom handshape

(f) Scroll/swipe

(g) Opening/closing



(h) Changing area

(i) Natural hand shape

Figure 14. The nine hand shapes

Figure 15. Key Points: two blue points represent palm center and wrist joint; red points represent the endpoints of bones in the hand

The support vector machine (SVM) method is employed to learn the nine hand shapes and to accomplish the multi-label classification method of the system. We utilize the open-source software, libsvm-3.22, in the VR store [18]. There are four steps for multi-label classification:

1. data collection
2. data normalization and scale
3. model training
4. predicting

(1) Data Collection

From a hand, we capture the endpoints of bones and palm center as "feature points" to describe hand structure [16]. As shown in Figure 15, we track each feature point position data in every frame of the VR environment.

(2) Data Normalization and Scale

We calculate other positions of feature points relative to the palm center. Data normalization follows these steps:

1. Move positions to make palm center on the origin coordinate.

2. Rotate the points to make palm parallel to the x-z axis plane.

3. Rotate the points around the y coordinate axis to make the palm point the z-axis.

Then we scale the data to [-1, 1].

(3) Model Training

The third step of handshape recognition is model training. The objective is from the feature points to recognize the nine hand shapes. Various SVM models have supervised learning strategies with associated learning algorithms that analyze data for classification and regression analysis. We selected the Classification SVM Type 1 (i.e., C-SVM classification) to the system. There are three steps to train with the data and obtain a classifier model:

1. Capture 50 group coordinates of the key points for every hand shape.

2. Normalize and scale the data and obtain nine group training sets.

3. Through the training sets, obtain the multi-label classifier model.

The model training part is preparation work for the system. The multi-label classifier model is employed to realize handshapes in time for the system.

(4) Prediction

After obtaining the multi-label classifier, we utilize it in the VR shopping system to recognize the nine hand shapes. When a user observes the room-scale VR shopping store, the user can freely move the hands. The depth sensor tracks the hands and obtains a group of original data set in every frame. Subsequently, the multi-label classifier predicts the result. The predicted result informs the system which handshape the user is performing. If the hand shape is a natural handshape, the system will disregard it. Conversely, the system acknowledges the command and interacts accordingly.

We put a hand above the Leap Motion and perform the nine handshapes 100 times, respectively. We record the data and test the accuracy of the method, shown in Table VII.

TABLE VII. THE ACCURACY OF RECOGNITION THE NINE HAND SHAPES

Handshape	**OK**	**Pointing**	**Extending**
Accuracy	92%	94%.	94%
Handshape	**Grab**	**Zoom**	**Scroll/swipe**
Accuracy	92%	93%.	93%
Handshape	**Opening/closing**	**Changing area**	**Natural**
Accuracy	94%	93%.	95%

B. *Motion Detection*

The system is also required to detect motion because gestures are defined by both hand shape and motion together.

Once getting the hand shape, the system begins the motion detection. The system calculates the hand data in each frame. For different hand shapes prediction, the system detects hand center or different fingertips to realize gestures.

Based on the labeled handshapes from 1 to 9, there are nine situations to recognize the movement.

Situation 1: for OK hand shape, a level 1 category gesture, no need for motion recognition is required.

Situation 2: for the pointing hand shape, if the system detects two hands are in pointing hand shape, the system requires to detect the positions of two index fingertips. If the two index fingertips are close to each other, it indicates the user is performing NO gesture. Conversely, if one hand is in pointing hand shape, the system detects the direction and motion of the index finger to select a target or click a button.

Situation 3: for extending hand shape, the strategy is more complex. (a) if it is detected that the center of the palm is face oriented and gradually approaches the face, it is recognized as a pull gesture; (b) if the palm center is detected as forward-oriented and to move forward, it is recognized as push gesture; (c) if the palm is detected as left-orient and it moves to the left, it is recognized as waving gesture; (d) if the palm is detected as top-oriented, it is recognized as holding gesture; (e) if the palm is detected as forward-oriented and moves on a vertical plane, it is recognized as drag gesture; and (f) if the palm is detected as top-orient and rotating around the palm center, it is recognized as rotation gesture.

Situation 4: for grab hand shape, the system detects the movement of the center of the palm, and the target object follows the movement of the center of the palm.

Situation 5: for zoom hand shape, the system detects the movement of index and thumb fingertips. If the fingertips move away from each other, it is recognized as a zoom-in gesture; if the fingertips move towards each other, it is recognized as a zoom-out gesture. The movement distance is employed to change the size of the target object.

Situation 6: scroll/swipe handshape; the system detects the movement of the index finger. The movement distance employed to control the menu.

Situation 7: for opening/closing hand shape, the system detects the movement of the index, middle, and thumb fingertips. If they are moving towards each other, it is recognized as the closing gesture; if they are moving away from each other, it is recognized as the opening gesture.

Situation 8: for changing area hand shape, the system detects the direction of the index finger and movement of thumb fingertip. If the index finger points to a position on the floor and thumb fingertip click the index finger, it is recognized as the changing area gesture, and the user moves to the target position.

Situation 9: for natural hand shape, the system does not need to detect any motion because the user moves his hands freely in 3D space and does not wish to interact with the system in this situation.

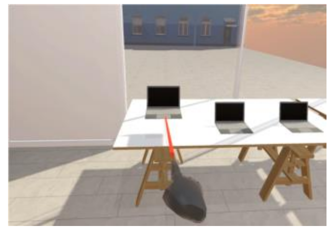
VIII. APPLY GESTURE SET IN ROOM-SCALE VR SHOPPING ENVIRONMENT

We apply the gesture set in VR environment to build an interactive system. We design a typical shopping activity as an example: viewing and buying a laptop.

Firstly, the user can move to the desk with changing area gesture where the laptops are displayed in the room-scale VR shopping environment, as shown in Figure 16(g). In this situation, the desk is distant from the user.

The user then selects a laptop with the pointing gesture. Once selected, the laptop performs a bounce animation, as shown in Figure 16(a). The user can manipulate the laptop with the hold gesture, as shown in Figure 16(b). After that, he can view the detailed information of the laptop with zoom in/out gesture, open/close gesture, and scroll/swipe gesture, as shown in Figure 16(c), Figure 16(d), and Figure 16(e). Finally, the user can perform an OK gesture to inform the system of the decision to choose it, as shown in Figure 16(f).

By using the gestures we designed, we emulate typical shopping activities in a room-scale VR shopping environment. The system provides the shopping activity process similar to daily shopping activities with physical hand gestures.



(a) Pointing a product

(b) Hold the product

(c) Zoom in/out the product

(d) Use Open/close gesture

(e) Use scroll/swipe gesture

(f) Use OK gesture

(g) Use changing area gesture

Figure 16. The gestures used in shopping activities

IX. EVALUATION

Five students were invited to use the room-scale VR shopping system. The ages of the students were 19 - 27. They performed the nine handshapes 50 times respectively. We collected the data and used the SVM method to find the error of the handshape classification.

TABLE VIII. ERRORS OF CLASSIFICATION OF NINE HAND SHAPES FOR EVERY USER FOR 50 TIMES

Handshape	User1	User2	User3	User4	User5
OK	1	0	0	2	1
Pointing	0	0	1	2	1
Extending	3	2	4	5	4
Grab	1	2	2	3	3
Zoom	0	1	1	3	1
Scroll/swipe	1	1	1	2	1
Opening/closing	0	0	1	1	0
Changing area	1	1	1	4	2
Natural	4	3	4	5	4

TABLE IX. THE ACCURACY RATE OF HANDSHAPE RECOGNITION OF 5 USER

User	Amount	Accuracy	Accuracy Rate
1	450	439	97.56%
2	450	440	97.78%
3	450	435	95.56%
4	450	423	94.00%
5	450	433	96.22%

Table VIII shows the classification errors of every handshape for every user. From Table VIII, we can see that extending hand shape and natural hand shape have relatively higher error rates because these two hand shapes are relatively similar.

Then we can obtain the accuracy rate when a user performs handshapes in the system, as shown in Table IX.

X. LIMITATION

Although we did some experiments to examine the accuracy rate of gestures, the method, including selection gestures and SVMs, still requires improvements. The quality of pattern recognition requires to be thoroughly checked by statistical methods. The number of users involved in the experiment is insufficient to obtain robust experimental results.

In addition, although the method is a solution for immersion problem, it requires training to recognize user gestures. Implementing AI systems that recognize standard human gestures may help eliminate training needs. The problem of the lack of haptic feedback in VR gesture interaction remains.

XI. CONCLUSION AND FUTURE WORK

In this research, we built a room-scale VR shopping system and proposed a new hand gesture set for the room-scale VR shopping system. We employed the gesture set as an alternative to VR device controllers to solve its limitations in VR shopping activities. We introduced a new gesture classification in the gesture set. Three levels for the classification methods are designed. The gestures in level 1 are static hand posture. The gestures in level 2 are dynamic gestures with motion, and the level 3 gestures are the combination of level 1 and level 2 gestures.

For level 3 gestures, we introduced three categories to classify them: core gestures, gestures for interaction with virtual objects, gestures for interaction with menu and interaction with space. The classification helps us understand the gesture set in the room-scale VR shopping system. In addition, the gesture set, and 3-level classification method can be easily transferred to other VR or AR systems.

To achieve complex gestures recognition, we applied the SVM method in the proposed VR shopping system. In the end, the user could walk around in his room to view the VR shopping store and interact with the system with natural hand gestures. We evaluated the accuracy of the gesture set. The results show that gestures have a high recognition accuracy. This suggests that using gestures to replace controllers in a VR environment has a promising prospect.

In the future, we plan to further improve the room-scale VR shopping system such as by implementing an AI system for recognizing gestures. We will improve the accuracy of gesture recognition and the convenience of interaction with system by looking into the recognition failures. The gesture set can be extended to have more specific features in the system and we need to study more complex context before extending the gesture set such as in a more crowded environment with multiple rows of merchandise, where one might partially obstruct the other. We would like to conduct some experiments to evaluate the efficiency of the proposed gesture set by comparing the proposed system with the system using controllers.

REFERENCES

[1] Chunmeng Lu and Jiro Tanaka. “A Virtual Shopping System Based on Room-scale Virtual Reality,” The Twelfth International Conference on Advances in Computer-Human Interactions (ACHI 2013) IARIA, Feb. 2019, pp. 191-198, ISSN: 2308-4138, ISBN: 978-1-61208-686-6

[2] Buy+ the first complete VR shopping experience. https://www.alizila.com/video/buy-first-complete-vr-shopping-experience/. Accessed November 29, 2019.

[3] IKEA highlights 2016 virtual reality. https://www.ikea.com/ms/en_US/this-is-ikea/ikea-highlights/Virtual-reality/. Accessed November 22, 2019.

[4] ShelfZone VR shopping experience by inVRsion. http://www.arvrmagazine.com/shelfzone-vr-shopping-experience-by-invrsion/. Accessed November 22, 2019.

[5] Serrano, Berenice, Rosa M. Baños, and Cristina Botella. "Virtual reality and stimulation of touch and smell for inducing relaxation: A randomized controlled trial." Computers in Human Behavior, vol. 55, pp. 1-8, 2016, doi: https://doi.org/10.1016/j.chb.2015.08.007. Accessed November 22, 2019.

[6] Bhatt, Ganesh. "Bringing virtual reality for commercial Web sites." International Journal of Human-Computer Studies, vol. 60, pp. 1-15, 2004, doi: http://dx.doi.org/10.1016/j.ijhcs.2003.07.002. Accessed November 22, 2019.

[7] Chen, Tian, Zhi-geng Pan, and Jian-ming Zheng. "EasyMall-An Interactive Virtual Shopping System." 2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery, vol. 4. IEEE, 2008, doi: http://dx.doi.org/10.1109/FSKD.2008.124.

[8] Lee, Kun Chang, and Namho Chung. "Empirical analysis of consumer reaction to the virtual reality shopping mall." Computers in Human Behavior, vol.24, pp.88-104, 2008, doi: http://dx.doi.org/10.1016/j.chb.2007.01.018.

[9] Verhulst A, Normand J M, Lombart C, et al. "A study on the use of an immersive Virtual Reality store to investigate consumer perceptions and purchase behavior toward non-standard fruits and vegetables." 2017 IEEE Virtual Reality (VR). IEEE, pp.55-63, 2017, doi: 10.1109/VR.2017.7892231. Accessed December 7, 2019.

[10] Speicher, Marco, Sebastian Cucerca, and Antonio Krüger. "Vrshop: A mobile interactive virtual reality shopping environment combining the benefits of on-and offline shopping." Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol 1, No. 3, Article 102, doi: http://dx.doi.org/10.1145/3130967. Accessed November 22, 2019.

[11] Lin, JJ-W., et al. "Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment." Proceedings ieee virtual reality 2002, pp. 164–171, doi: http://dx.doi.org/10.1109/VR.2002.996519.

[12] Garg, Pragati, Naveen Aggarwal, and Sanjeev Sofat. "Vision based hand gesture recognition." World Academy of Science, Engineering and Technology, vol 49, pp. 972-977, 2009. http://waset.org/Publications?p=25. Accessed November 22, 2019.

[13] Wachs, Juan Pablo, et al. "Vision-based hand-gesture applications." Communications of the ACM, vol 54, pp. 60-71, 2011, doi: http://dx.doi.org/10.1145/1897816.1897838. Accessed November 22, 2019.

[14] Boussemart, Yves, et al. "A framework for 3D visualisation and manipulation in an immersive space using an untethered bimanual gestural interface." Proceedings of the ACM symposium on Virtual reality software and technology. ACM, pp. 162-165, 2004, doi: http://dx.doi.org/10.1145/1077534.1077566. Accessed November 22, 2019.

[15] Karam, Hani, and Jiro Tanaka. "Two-handed interactive menu: An application of asymmetric bimanual gestures and depth based selection techniques." International Conference on Human Interface and the Management of Information. Springer, Cham, pp. 187-198, 2014, doi:http://dx.doi.org/10.1007/978-3-319-07731-4_19. Accessed November 22, 2019.

[16] Lu, Chunmeng, Li Zhou, and Jiro Tanaka. "Realizing Multi-Touch-Like Gestures in 3D Space." International Conference on Human Interface and the Management of Information. Springer, Cham, pp.227-239, 2018, doi: http://dx.doi.org/10.1007/978-3-319-92043-6_20. Accessed November 22, 2019.

[17] Leap Motion, 2018. https://www.leapmotion.com/. Accessed November 22, 2019.

[18] LIBSVM -- A Library for Support Vector Machines, 2016. https://www.csie.ntu.edu.tw/~cjlin/libsvm/. Accessed November 22, 2019.