# A virtual shopping system based on room-scale Virtual Reality

**44161694-9 LU CHUNMENG**

Supervisor: Prof. Jiro TANAKA

Graduate School of Information, Production and Systems
Waseda University

A thesis submitted for the degree of
*Master of Engineering*

July 2018

# Abstract

Virtual Reality (VR) has three important features: interactivity, immersion and connectivity. Applying Virtual Reality in online shopping field will provide an immersive shopping experience, in which user views a virtual shopping store. However, present VR shopping systems are chair-scale VR. It means user need to sit or stand at a position when using VR. In this case, user's view in a VR store is limited. He just could view and get object around him. User needs to change his view and position frequently to approach targets. It is not natural and convenient. Room-scale VR is a way to solve this problem, which will allow user to walk in his room when viewing in a VR store. Thus, the property, room-scale, will make the activity in virtual shopping store similar to that in physical world.

In the virtual reality environment, user needs to input information and achieve interaction with user interface and virtual object. At present, most virtual reality systems provide some devices, such as keyboard and controller, for user to interact with the virtual reality environment. However, using devices is not a natural and immersive experience, especially in virtual reality shopping experience. In the real world,we use our hands to get object. Thus in virtual world, using our hand gestures to interact with VR shopping store will give a more immersive VR shopping experience.

There are many gestures could be used in VR system. According to the needs of the room-scale VR shopping activities, we present a new gesture set for interaction. We introduce a new gesture classification for the gesture set which has three levels to classify hand gestures based on the characteristic of gestures. We focus on researching the gestures in level 3, because the level 3 gestures have more interaction languages.

In our research, we build a room-scale virtual reality shopping system and apply the new hand gesture set for the interaction in the VR shopping system.

**Keywords:** room-scale Virtual Reality, gesture set, gesture classification

# Acknowledgements

In the past two years, I have gotten lots of help from Professor Tanaka. Thanks to his patient guidance, I learned a lot on both knowledge and interpersonal communication. I remember the first time I wrote conference paper. Professor Tanaka explained the research points patiently again and again until I understood it. He helped me to improve my communication skill and academic ability. Without professor's guidance, I would not have these meaningful two years in Waseda University.

# Table of contents

# List of figures

# List of tables

# Chapter 1

# Background

## 1.1 Virtual Reality

Virtual Reality (VR) allows the simulation of real-life situations in a tridimensional computer-generated environment where the user can interact with the environment as if he/she were in the real world. In VR it is possible to interact in real time with different objects and experience a sense of presence. It is a kind of the illusion of "being there" [1].

Virtual Reality is not new concept and it has been improved for more than 50 years. VR technology was firstly presented in 1950s. In 1957, Morton Heilig, a cinematographer, who invented a multi-sensory motorcycle simulator known as the Sensorama. The simulator enabled users to watch 3D films, which was later patented in 1962. The Sensorama combined all sensory channels by generating visual scenes of the city driven, engine and city sounds for auditory stimulation, smell of exhaust and food and seat vibration in its interaction concept in order to give its users the illusion of being part of the virtual environment. However, the Sensorama was is very huge and not easy to use. People need a light, convenient and comfortable device. A Head-Mounted Display (HMD) was introduced by Ivan Sutherland in 1965 [2]. The concept was later developed to the device,the Sword of Damocles, which is claimed by some researchers to be the world first HMD to track the head in real time. Starting with it and with the development with computer graphics, 3D technology and electrical engineering, HMD gradually improved in the last years. Virtual reality also was studied and developed by scientists. In recent years, virtual reality devices come out of laboratory. Some some technology companies have introduced simple and easy-to-use VR devices for the consumer market, such as HTC Vive and Oculus Rift.

Interactivity, immersion and connectivity are important features of virtual reality. The most important feature is immersion [3]. Immersion refers to the feeling of being deeply engaged in a virtual world as if it were the real one [4]. Using the VR devices, people view the virtual world through display. The computer graphic and 3D technology render all the object in the virtual world. The user's view in virtual world will move synchronously with user's head. This makes user feel that he is view another parallel world. Because of this good feature, in this research, I apply VR technology to build a virtual reality shopping environment, in which user feels he is browsing a real store.

## 1.2 Room-scale VR

When user is moving their view in virtual world with a HMD, he could not move his physical body in real world. Thus, there is a great gap between sensorial moving and physical moving. It will reduce the immersion of VR greatly. This gap will also cause motion sickness for some people [5, 6]. If user's walking is synchronous both in virtual world and real world, it is a more great experience feeling. In order to achieve this target, we need to track user's walking in real world.

In 1990, Wang et al. from The University of North Carolina at Chapel Hill presented a room-sized head-tracking system for VR HMD [7]. They used three cameras on head to capture infrared LEDs mounted on the room ceiling. Then, processing the infrared images, they could get the position and rotation of user's head. Their system offers good accuracy in tracking head motion (0.1 degree in rotation and 2 mm in translation) at long range (3 meters between source and sensor). This project continued for more than 10 years. In 1992, other group researchers extended Wang's work. They built a scalable head-tracking system for HMDs [8]. In this work, measurements of head position and orientation were produced at a rate of 20-100 Hz with 20-60 ms of delay. The system's accuracy had not been measured precisely, but the resolution of position and orientation is 2 mm and 0.2 degrees. Then in 1999, Welch et al. presented a better tracker for virtual and augmented environment. Their work was a vastly improved version of that 1992 system in The University of North Carolina. Their new device, Hiball, is lighter, more fast and more convenient [9]. Some researchers used depth sensor, like Kinect, to capture user's walking in room [10].

In recent years, researchers aim on improving room-scale using experience and applying them in some fields. Ruddle et al. explain the importance of using navigation when walking in virtual environment.[11]. In their experience, the behavioral data indicates that both translational and rotational body-based information are required to accurately update one's position during navigation, and participants who walked tended to avoid obstacles, even though collision detection was not implemented and feedback not provided. The experience shows that a walking interface would bring immediate benefits to a number of virtual environment applications. Suma et al. introduce "impossible spaces," a new design mechanic for virtual environments that wish to maximize the size of the virtual environment that can be explored with natural locomotion. Such environments make use of self-overlapping architectural layouts, effectively compressing comparatively large interior environments into smaller physical areas. Their results demonstrate that users perceive distances to objects in adjacent overlapping rooms as if the overall space was uncompressed, even at overlap levels that were overtly noticeable [12]. As user walks in a real room, some researchers try to add some useful physical information in virtual world. Simeone et al. find that experiencing Virtual Reality in domestic and other uncontrolled settings is challenging due to the presence of physical objects and furniture that are not usually defined in the Virtual Environment. To solve this challenge, we explore the concept of Substitutional Reality in the context of Virtual Reality: a class of Virtual Environments where every physical object surrounding a user is paired, with some degree of discrepancy, to a virtual counterpart. They present a set of guidelines for the design of future Reality experiences [13]. Sra M. et al. present a novel system for automatically generating immersive and interactive virtual reality environments using the real world as a template. They apply depth sensors in the system. They capture indoor scenes in 3D, detect obstacles like furniture and walls, and maps walkable areas to enable real-walking in the generated virtual environment. Depth data is additionally used for recognizing and tracking objects during the VR experience. The detected objects are paired with virtual counterparts to leverage the physicality of the real world for a tactile experience [14]. Some researchers analyze the relationship between virtual path and physical walking [15]. In the past few years, some companies provide VR devices for customer market. Some of them are equipped sensors for tracking user's walking in room. Oculus Rift and HTC Vive are two famous VR devices. Oculus Rift provides infrared sensors to track user's motion in room. HTC Vive and its sensor stations are showed in Fig. 1.1. The tracking sensors of HTC Vive could track user's motion in a area with 3 meters length and 4 meters width.


(a) HMD and controllers (b) Sensor stations

Fig. 1.1 HTC Vive

## 1.3 VR shopping

Now people could roam in the virtual environment through HMD. As shopping is one of the important activities in our daily real world, a virtual shopping environment could be a part of the virtual environment. We are familiar with e-commerce or online shopping. So we could extend online shopping to the virtual environment.

In the last decades, many VR shopping environments have been presented. Some works aim on improving VR shopping experience and some works research on the interaction in virtual shopping environments. Basing on arguments about virtual space, Bhatt presented a theoretical framework that highlights the relative importance of interactivity, immersion, and connectivity for attracting customers through a Web site. They argue that in order to attract customers through Web sites, e-commerce companies are required to balance a trade-off between interactivity, immersion, and connectively, depending on their business objectives [3]. Chen et al. think that most e-Commerce platforms only provide users simple, 2D image-based and text-based interfaces to access the products. Such kind of tedious environment neither makes consumers enjoy fun during purchase, nor provides real-time personalized services according to different personal attributes and actual shopping behaviors of consumers on base of assisting both of buyers and vendors. They present a VRML-Based virtual shopping mall. They stress comprehensively analyzing customers' personal information, special demands, shopping behaviors and related historical records recorded by the Virtual Shopping Mall System. They also present how to employ Intelligent Agent technology based on above marketing decision model to interactively provide consumers a good purchase guidance service [16]. Lee et al. design a virtual interactive shopping environment and investigate whether the user interface of the VR shopping mall positively affects customer satisfaction in comparison with the ordinary shopping mall [17]. Kerrebroeck et al. examine the potential of a relaxing Virtual Reality (VR) experience in a shopping mall. They try to find whether VR could provide a relaxing shopping experience for people. Comparing with crowed shopping mall, VR shopping could give a relaxing shopping environment. Their experience shows that consumers reported more positive responses on all measured outcome variables after being exposed to the Virtual Reality experience. Besides, the effect on mall attitudes, satisfaction and loyalty is more pronounced when crowding is perceived to be high. [18]. Verhulst et al. present an immerse virtual reality user study. In the study, they aim on investigating how customers perceive and if they would purchase non-standard (i.e. misshaped) fruits and vegetables (FaVs) in supermarkets and hypermarkets [19]. Speicher et al. presented a mobile interactive virtual reality shopping environment combining the benefits of on- and offline shopping. They summarized previous researching results and introduced a Virtual Reality Shopping Experience (VRSE) model based on customer satisfaction, task performance and user preference. Besides, they research on the problems about searching for a product in a WebVR online shop. They found that using speech input in combination with VR output proved to be the best regarding user performance (speed, error rate) and preference (usability, user experience, immersion, motion sickness) [5].

The previous researches show some good features of VR shopping. Thus some retail companies and online shopping companies have become interested in VR shopping. IKEA company presented room-scale VR environment, in which user could view a virtual kitchen and interact with the furnitures [20]. Another example comes from inVRsion, which provides a virtual supermarket shopping system, Shelfzone VR [21]. In the future, there will be more application on VR shopping.

From these previous research, we find that the property, room-scale, could improve immersion of virtual reality. So it also could optimize the VR shopping experience. In our research, we will room-scale VR technology to build the virtual environment for shopping experience. User will be able to walk in a room-sized area to view a virtual reality shopping store.

## 1.4 Organization of the Thesis

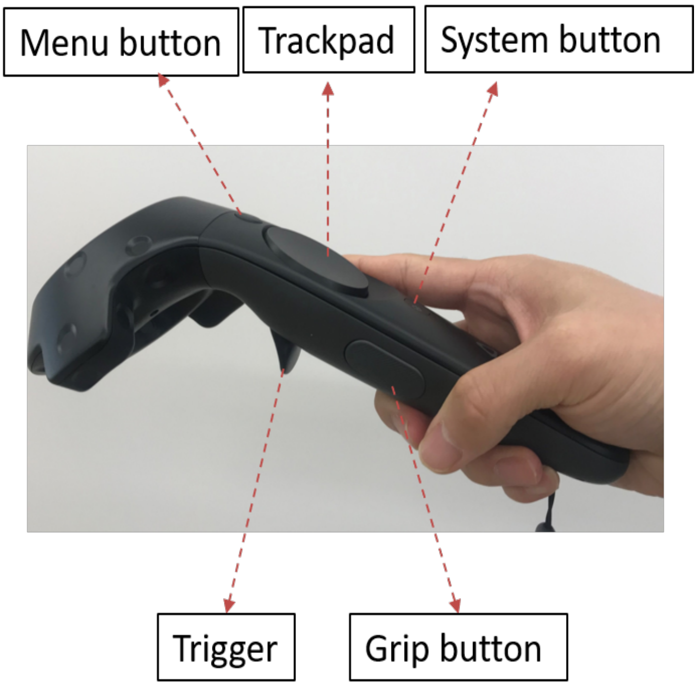
The rest of this thesis is organized as follows: In chapter 2, we will describe the problem that we want to solve and introduce our research purpose and the approaches. In chapter 3, we will introduce the system design. In chapter 4, we will introduce the implements of our system. In chapter 5, we introduce some related works. And in chapter 6, we will make conclusion and introduce the future works.

# Chapter 2

# Research Goal and Approach

## 2.1 Problem

VR shopping environment tries to provide a emulational environment, in which the virtual objects are similar to our real world. In our real world, we use our hands to touch and catch the objects around us. Thus, using controllers in VR shopping environment is not immersive enough. User will feel a gap when using controllers to catch the virtual but emulational objects. Besides, the amount and functions of buttons in the controllers are limited, which limits the interaction languages when using controllers in room-scale VR environment. As showed in Fig. 2.1a, there are only three kind of buttons on the controllers. To achieve VR shopping activities, we could use these three buttons to design interaction method in the VR shopping system. However, as showed in Fig. 2.1b, human hand has a complex structure. And human hand could perform many gestures for human-computer interaction. Comparing with controllers, using hand gestures could not only improve the immersion of the VR shopping experience, but also achieve more interaction languages for the VR shopping system.




(a) The buttons of controller

(b) Hand structure

Fig. 2.1 Controller and hand

## 2.2 Research Goal

In this research, we aim to present a new hand gesture set which is suitable for room-scale VR shopping activity to replace the controllers. Researching on the hand gesture set, we introduce a new gesture classification to make the gesture set more structural. We apply room-scale shopping system to provide a immersive virtual shopping environment, which is simulation of physical shopping store. In the room-scale VR shopping environment, user could walk around in his room to view the virtual shopping environment through an HMD. We design the new gestures for the room-scale VR shopping system. User could interact with VR environment by the natural hand gestures but not the controllers. We introduce a gesture classification for gestures, which has three levels to classify hand gestures basing on the characteristic of gestures. Summarizing the hand gestures, we get a new hand gesture set especially for room-scale VR shopping activity. Comparing with controllers, hand gestures have rich interaction language and are more convenience. The hand gesture set will improve the convenience and immersion of room-scale VR shopping system.

## 2.3 Research Approach

Hand gestures have been widely used in human computer interface. Gesture-based interaction provides a nature, intuitive communication between people and devices. People use 2D multi-touch gesture to interact with devices like smart phone and computer in the daily life. 3D hand gesture could be used for some devices equipped with camera or depth sensor. The most important problem in hand gesture interaction is how to make hand gestures to be understood by computers [23]. Wachs et al. summarizes the requirements of hand-gesture interfaces and the challenges when applying hand gestures in different application [24]. Yves et al. present a framework for 3D visualization and manipulation in an immersive space. Their work can be used in AR and VR system [25]. Karam et al. use depth camera and present a two-hand interactive menu to improve efficiency [26]. Lu et al. shows an example [22] that using depth camera to recognize their multi-touch-like gestures in 3D space. These previous researches shows that hand gestures have many possibilities for human-computer interaction field.

Different gestures are given different meanings in different system. Thus, we need a gestures set defined by VR shopping system.

In this research, we introduce a new gesture set specially designed for room-scale VR shopping experience. In the gesture set, we introduce a gesture classification, which has three levels to classify hand gestures basing on the characteristic of gestures. The systematic classification method could help user to understand the interaction in the room-scale VR shopping system. We use VR device to built the room-scale VR shopping system. In the system, there are two sensor stations installed in the room. The two sensor stations create a walking area for user. When moving in the walking area, user' motion will be captured by the sensor stations. System will get rotation and three-dimensional coordinates of the HMD worn by user. The view in virtual environment will move synchronously with HMD. Then we use depth sensor to recognize the hand gestures. In the virtual environment, user could see his virtual hands moving synchronously with his physical hands. The system could realize the special gestures when user performing near depth sensor, and achieve the interaction with virtual environment.

# Chapter 3

# System Design

In the chapter, we introduce the room-scale VR shopping system, the gesture set and the gesture classification.

## 3.1 VR shopping environment

In order to achieve VR shopping activity, we design a VR shopping store as the shopping environment, which is similar to the stores in real world. We place some desks and shelves in the VR shopping store, as shown in Fig. 3.1a. For a shopping store, goods are a must. We placed some virtual 3D models in the VR shopping store. Here we put four laptop models and a camera with tripod on the desks, as shown in Fig. 3.1b.

In the VR shopping environment, user could view the store and interact with the system. The virtual goods are interactive. User could pick up the 3D models and scan them. User also could call out information of the goods and view the information freely.

## 3.2 Room-scale walking area

In the room-scale VR shopping system, we want to make it possible that user could walk around in his own room to view the virtual shopping environment. So we need to prepare an empty area in real room. The empty area is included in a 3D space. We will use tracking

(a) A VR store

(b) Indoor scene

Fig. 3.1 Overview of VR shopping environment sensor to capture the motion and rotation of HMD in the 3D space when user using the VR shopping system. As shown in Fig. 3.2, the length of 3D space is 4 meters, the width of 3D space is 3 meters and the height of 3D space is 2 meters. The 3D space contains walking area of the room.

In the VR shopping environment, there is also a virtual walking area, as shown in Fig. 3.3. The virtual walking area is same with the area in real room. As the VR shopping store is larger than our real room, user could change the virtual walking area when view the whole VR shopping store.

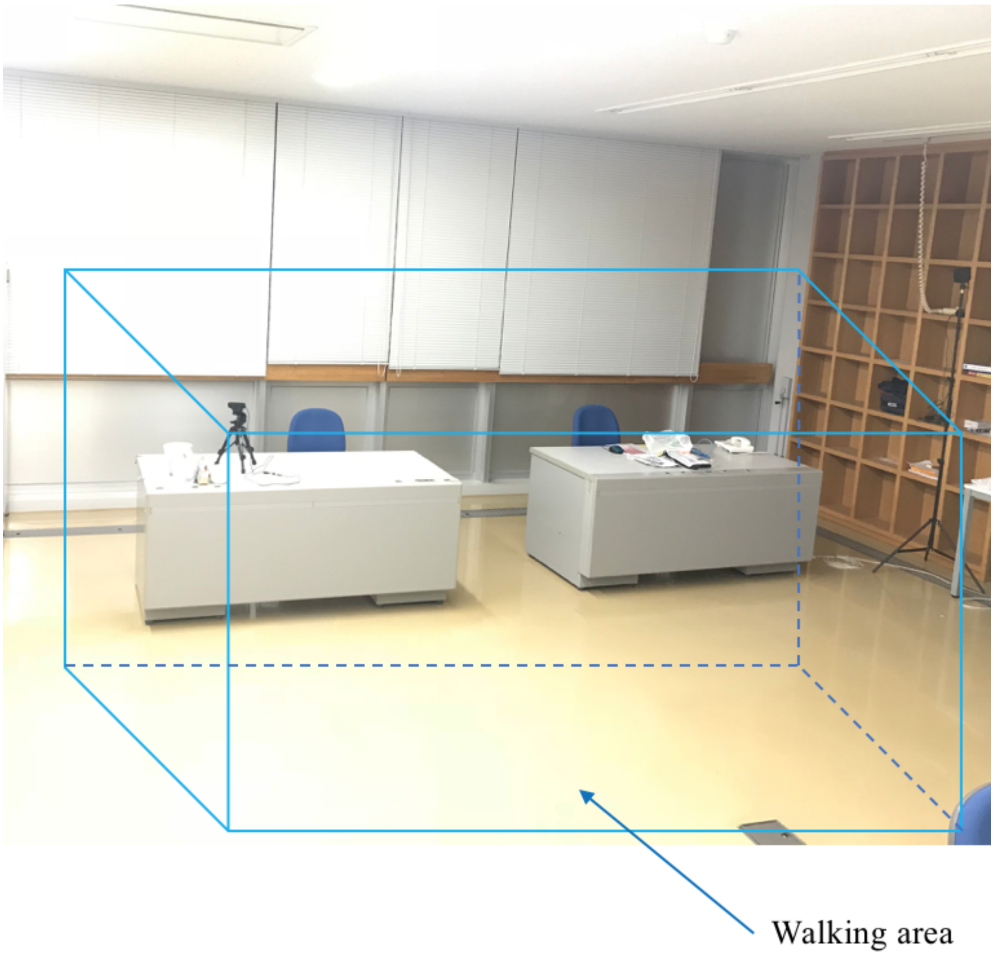



Fig. 3.2 3D space and walking area in real room

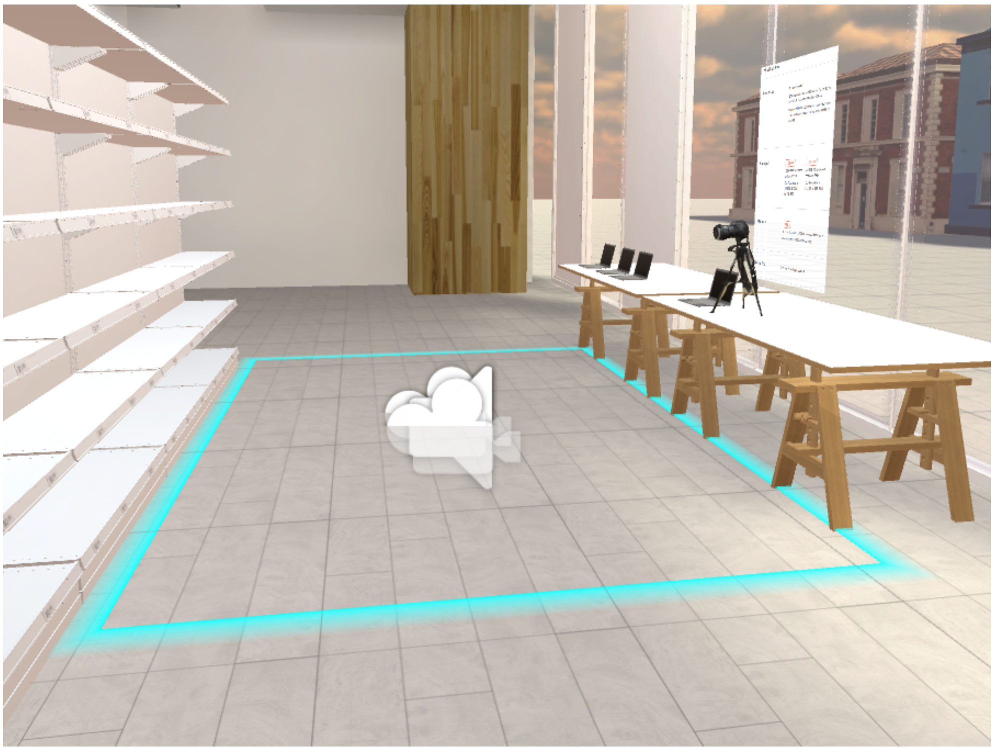


Fig. 3.3 Virtual walking area in the VR shopping environment

## 3.3 Gesture Set and Gesture Classification

### 3.3.1 Gesture Set

There are so many gestures that could be used in VR system. While, in our research, we need to design a series of hand gestures specially for the room-scale VR shopping activity. The hand gestures must provide a natural and suitable interaction for user and system. According to the particular activities in room-scale VR shopping system, we design these 14 gestures in our system:

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

These gestures combine a new gesture set for room-scale VR shopping system. We want to design the functions of these gestures, as shown in Table 3.1. Comparing with controllers, the gestures have rich interaction language and could achieve more complex functions.

| Number | Gesture | Description and Function |
|---|---|---|
| 1 | Pointing | Point a virtual object with index finger. |
| 2 | Holding | Hold a virtual object on one hand. |
| 3 | OK | Give positive feedback to system. |
| 4 | NO | Give negative feedback to system. |
| 5 | Push/pull | Push or pull a virtual object with a hand. |
| 6 | Rotation | Rotate a virtual object when viewing it with a hand. |
| 7 | Drag | Move virtual object freely with a hand. |
| 8 | Waving | Make virtual object return to the original position; Change interface with quickly waving hand |
| 9 | Click | Choose a virtual object or button with index finger. |
| 10 | Zoom in/out | Make a virtual object show a larger or smaller size using relative motion of thumb and index finger. |
| 11 | Opening/closing | Open or close five fingers to open or close a box. |
| 12 | Grab | Make object move close to hand and grab it with hand. |
| 13 | Scroll/swipe | Using two fingers to view information in system. |
| 14 | Changing area | Use index finger to point a new position on the floor and make thumb click index finger, then view in VR will move to the new position. |

Table 3.1 The gesture set for room-scale VR shopping system

### 3.3.2 Gesture Classification

We present a new gesture classification to classify the hand gestures basing on their different characteristic. The gesture classification has three levels:

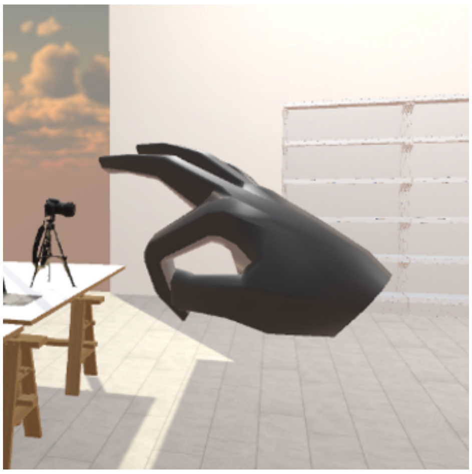
- Level 1: Core static hand postures are divided into the level 1, as shown in Fig. 3.4. In level 1, gestures are just hand shape without hand motion. The classic example is pointing gesture.
- Level 2: Dynamic palm motions are divided into the level 2, as shown in Fig. 3.5.. In level 2, we just care about the palm movement. We don't care about the shapes of fingers. The classic examples are pull and push.
- Level 3: Combination hand gestures are divided into the level 3, as shown in Fig. 3.6, Fig. 3.7, Fig. 3.8 and Fig. 3.10. Combination hand gestures combine the features of level 1 and level 2 gestures. In level 3, we care about the motion and shapes of fingers and the motion of palm.

In VR shopping environment, the hand gestures are divided into different levels. In the figures of level 2 and level 3, the red arrows mean the fingers movement trends.

The classification method will make the gesture set more structural. This classification method provides a structure that could also be used in other gesture sets in different VR system. Basing on the systematic structure, researchers could design suitable gestures for their VR systems.

#### 3.3.2.1 Level 1 Gestures

In our system, we use level 1 gestures to give feedback to system. Level 1 gestures are static signals for shopping system. We do not need care about the motion of fingers or hands. System just needs to detect the hand shapes. Fig. 3.4 shows two level 1 gestures and Table. 3.2 shows the descriptions and functions of level 1 gestures.



(a) OK gesture

(b) NO gesture

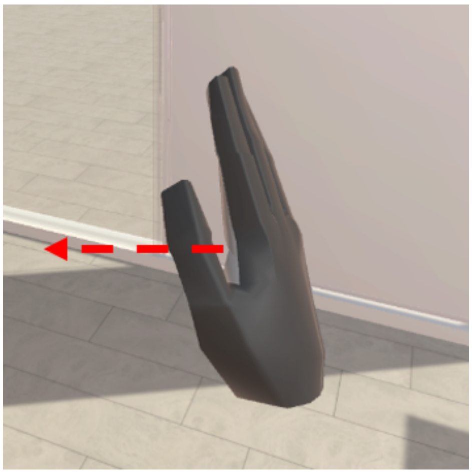
Fig. 3.4 Level 1: Core static hand postures

| Level | Gesture | Description and Function |
|---|---|---|
| 1 | OK | Give positive feedback to system. |
| 1 | NO | Give negative feedback to system. |

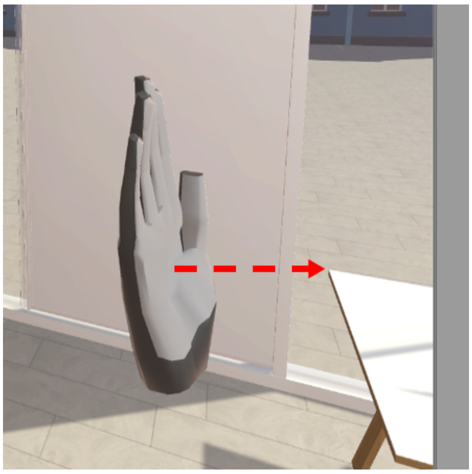
Table 3.2 Description and Function of gestures in Level 1

#### 3.3.2.2 Level 2 Gestures

Level 2 gestures are palm motions. After choosing a virtual object, user could use level 2 gestures control or interact with it. Fig. 3.5 shows two level 2 gestures and Table. 3.3 shows the descriptions and functions of level 2 gestures.



(a) Push gesture



(b) Pull gesture

(c) Waving gesture

Fig. 3.5 Level 2: Palm movement

| Level | Gesture | Description and Function |
|---|---|---|
| 2 | Push/pull | Push or pull a virtual object with a hand. |
| 2 | Waving | Make virtual object return to the original position;<br>Change interface with quickly waving hand |

Table 3.3 Description and Function of gestures in Level 2

#### 3.3.2.3 Level 3 Gestures

Gestures in level 3 are hand gestures that combines finger shapes and hand motion. These gestures is complex and combine the features of level 1 and level 2 gestures.

Designing a suitable and convenient gesture set for user determines whether user could feel a immersive VR shopping experience. Level 1 and level 2 gestures are simple and a little week. Thus, level 3 gesture set is the focus of our research.

In level 3, we need to recognize the hand shape and detect motion of the fingers and hands at the same time. We design the gesture set, and their functions are shown in Table. 3.1.

In the gesture set, pointing gesture is the most important gesture, because we need to choose a target object or button with pointing gesture before any interaction. Usually, user moves his hands freely in 3D space with a natural hand shape and system does not recognize the hand shapes. The pointing gesture can be a signal of interaction intention from user to system. System will start realizing hand shapes. In the VR environment, there will be an arrow at the index fingertip to help user confirm the pointing direction.

In level 3, the gestures have different usage. Thus, we need introduce a classification for level 3 gestures.

Some gestures are mainly used to interact with virtual products in VR shopping store, such as moving a virtual object. We design these gesture to achieve it: grab gesture, hold gesture, drag gesture, rotation gesture, and zoom in/out gesture.

In some cases, we need to interact with menu to achieve shopping activities. We design these gestures: click gesture, scroll/swipe gesture, opening/closing gesture.

In the system, we will change area in the VR environment frequently because the walking space in real world is usually smaller than VR shopping store. There is the classification:

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

#### 3.3.2.4 The Core Gesture: Pointing Gesture

In a physical shopping store, when customer wants to view or buy a product, he need to walk to close to the product and grab it with hand. While in VR system, VR environment provides more possibilities for user to interacting with products models. In our room-scale VR shopping system, user could use pointing gesture to choose an object on the table or shelf which are a little far away from user. Then the object gets into a mode in which it could be interactive. So, choosing the target object with pointing gesture is the first and most important step in the shopping activities. The pointing gesture is shown in Fig. 3.6 and Table. 3.4.

| Level | Gesture | Description and Function |
|---|---|---|
| 3 | Pointing gesture | Point a virtual object with index finger. |

Table 3.4 Description and function of the core gesture

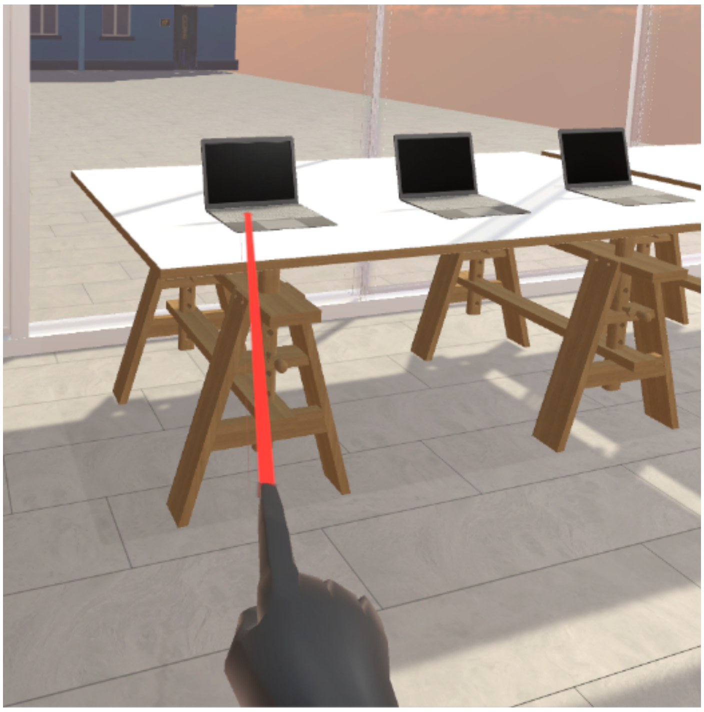
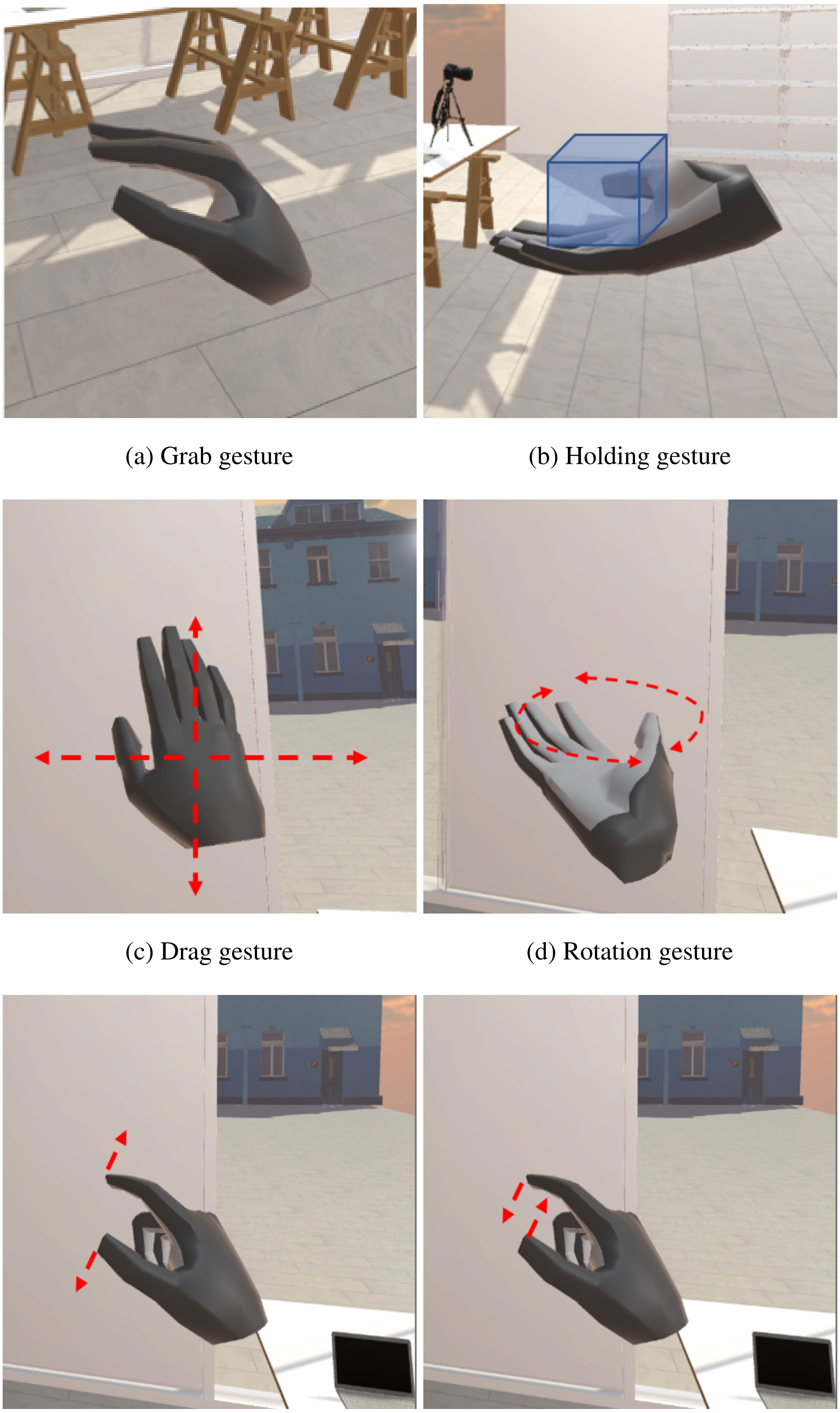


Fig. 3.6 Pointing gesture

#### 3.3.2.5 Gesture for Interacting with Virtual Object

Interacting with virtual objects in our system is also important. We not only try to simulate the interaction of physical shopping store in VR environment, but also want to design more useful gestures for the the room-scale VR shopping environment. We design the following gestures for interacting with virtual objects: grab gesture, hold gesture, drag gesture, rotation gesture and zoom in/out gesture. The gestures are shown in in 3.7. And their descriptions and functions are shown in Table. 3.5.



(a) Grab gesture (b) Holding gesture

(c) Drag gesture (d) Rotation gesture

(e) Zoom in (f) Zoom out

Fig. 3.7 The gestures for interacting with virtual objects

| Level | Gesture | Description and Function |
|---|---|---|
| 3 | Grab gesture | Make object move close to hand and grab it with hand. |
| 3 | Holding | Hold a virtual object on one hand. |
| 3 | Drag | Move virtual object freely with a hand. |
| 3 | Rotation | Rotate a virtual object when viewing it with a hand. |
| 3 | Zoom in/out | Make a virtual object show a larger or smaller size using relative motion of thumb and index finger. |

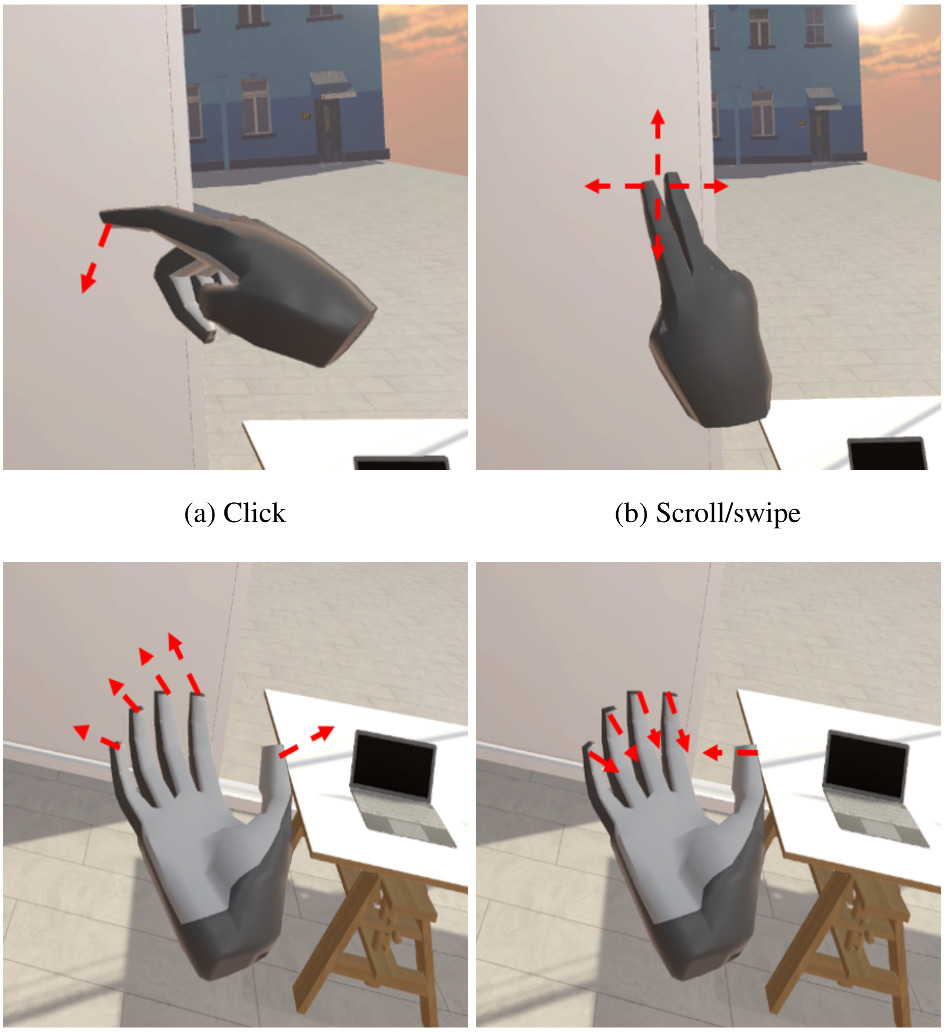
Table 3.5 Description and Function of gestures for interacting with virtual objects

#### 3.3.2.6 Gestures for Interacting with Menu

In many cases, user need to interact with menu in the user interface, such as clicking a button. We design these gestures for interacting with menu: click gesture, scroll/swipe gesture and opening/closing gesture. The gestures are shown in 3.8. And their descriptions and functions are shown in Table. 3.6.

| Level | Gesture | Description and Function |
|---|---|---|
| 3 | Click gesture | Click the buttons with index finger. |
| 3 | Scroll/swipe | Using two fingers gestures to control menus in user interface. |
| 3 | Opening/closing | Open or close five fingers to open or close the dashboard. |

Table 3.6 Description and Function of gestures for interacting with virtual objects



(a) Click (b) Scroll/swipe

(c) Opening (d) Closing

Fig. 3.8 Gestures for interacting with menu

#### 3.3.2.7 Gestures for Interacting with Space

In room-scale VR shopping system, user could walk in the real walking area in his own room. However, the room-scale walking area is alway smaller than the VR shopping store. As shown in Fig. 3.9, the blue rectangle in the overhead view of the VR shopping store shows the walking area in VR environment which simulates the walking area in physical room. The walking area in both VR environment and physical room have the same area. Thus, we need to design a gesture for user to changing area in the room-scale VR shopping store, as explained in Table. 3.7. Fig. 3.10 shows the changing area gesture. When performing changing area gesture, user needs to extend his index finger and thumb finger. In the system, when user wants to change area in the VR shopping store, he could point at position on the virtual floor with the special hand gesture and use thumb to click the index finger to tell system that he wants to move to there.

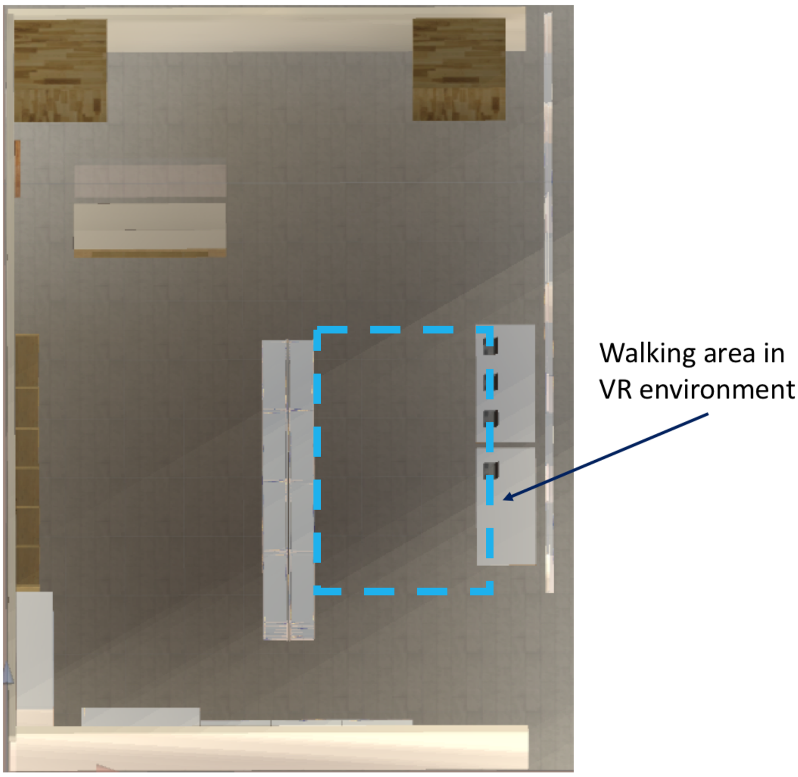



Fig. 3.9 Walking area in VR environment

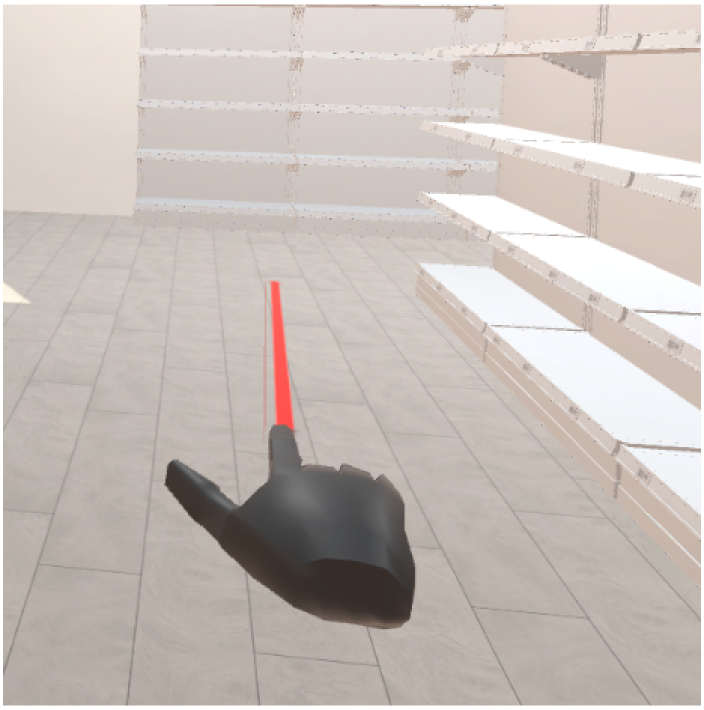


Fig. 3.10 Change area gesture

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

Table 3.7 Description and Function of the gesture for interacting with space

### 3.3.3 Summary of The Gesture Set

After designing and analyzing the gestures that we need in our system, we summarize the 14 gestures in the gesture set. There are two gestures in level 1, two gestures in level 2, and ten gestures in level 3, as shown in Table. 3.8.

| level | Gesture | Description and Function | |
|---|---|---|---|
| 1 | OK | Give positive feedback to system. | |
| 1 | NO | Give negative feedback to system. | |
| 2 | Push/pull | Push or pull a virtual object with a hand. | |
| 2 | Waving | Make virtual object return to the original position; Change interface with quickly waving hand. | |
| 3 | Pointing | Core gesture | Point a virtual object with index finger. |
| 3 | Grab | Interact with virtual object | Make object move close to hand and grab it with hand. |
| 3 | Holding | | Hold a virtual object on one hand. |
| 3 | Drag | | Move virtual object freely with a hand. |
| 3 | Rotation | | Rotate a virtual object when viewing it with a hand. |
| 3 | Zoom in/out | | Make a virtual object show a larger or smaller size using relative motion of thumb and index finger. |
| 3 | Click | Interact with menu | Choose a virtual object or button with index finger. |
| 3 | Scroll/swipe | | Using two fingers to view information in system. |
| 3 | Opening/closing | | Open or close five fingers to open or close a box. |
| 3 | Changing area | Interact with Space | Use index finger to point a new position on the floor and make thumb click index finger, then view in VR will move to the new position. |

Table 3.8 Summary of the gesture set for room-scale VR shopping system

# Chapter 4

# Implementation

## 4.1 System Hardware

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

We use Leap Motion [27] as the depth sensor to recognize the hand gesture. Leap Motion could track the coordinate and rotation of the fingertips and center of palm and transfer the data to VR system. In the VR system, Leap Motion is stuck on the HMD, as showed in Fig. 4.1.

The VR device and depth sensor need to work with a PC. Table 4.1 shows the information of the PC we use. In our research, we use Unity 3D as the software to build the room-scale VR system. With the Unity 3D, we could precess the data from Leap Motion and design the virtual shopping environment.

| | |
|---|---|
| Operation System | Microsoft Windows 10 |
| CPU | Intel Core i7-7700K @4.2 GHz |
| Graphics Card | NVIDIA GeFource GTX 1080 |
| Ram | 16 GB |
| Software | Unity 2017.3.0f3(64-bit) |

Table 4.1 The information of PC

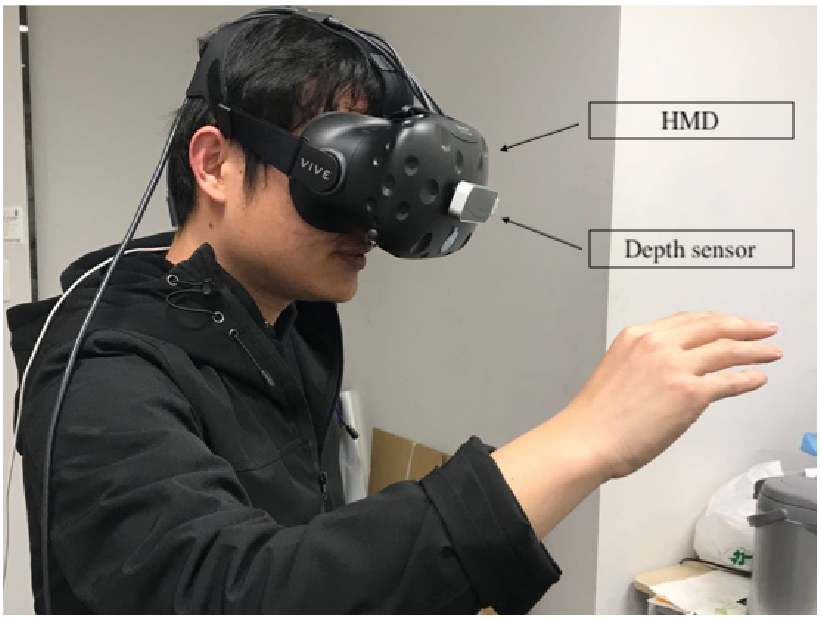



Fig. 4.1 HMD and depth sensor

## 4.2 Room-scale VR Setup

In the room-scale VR shopping system, user could walk around in his own room. So we need to track the motion of user's head in the 3D space of room. HTC Vive provides the tracking sensor: base station. As showed in Fig. 4.2, we place the two base station in the two corner of the room. The distance between two base station is 5 meters. The height of base stations is 2 meters. Thus the base stations create a 3D space with 2 meters height, 4 meters length and 3 meters width. The base stations could capture the coordinate and rotation of HMD. The data will be transfer wirelessly to PC. Using the data, user's viewing in virtual will move synchronously with user's head in physical world.

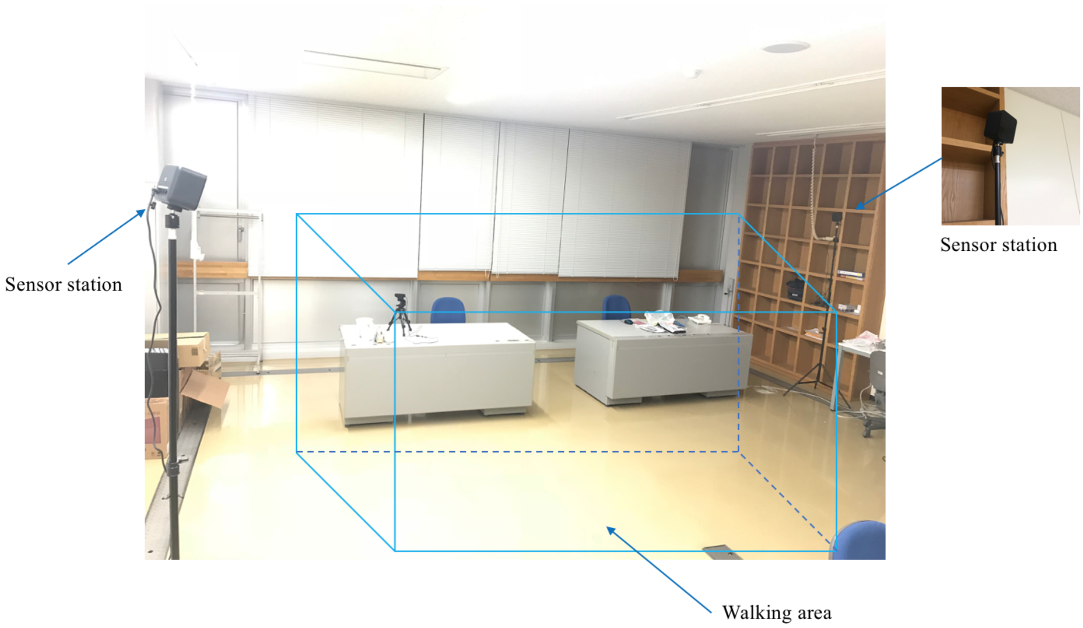



Fig. 4.2 Base stations, 3D space and walking area in physical room

## 4.3 Gesture Recognition

In our system, we use Leap Motion as the depth sensor to track hand. Leap Motion could track the joints, fingertips and palm center of user's hands. Meantime, Leap Motion could record the positions of these important point in user's hand in every frame.

With the original position data, we could use Machine Learning method to recognize the hand shapes.Then combining the hand shapes and motion, we will achieve recognizing the gestures that we design for the room-scale VR shopping system.

### 4.3.1 Hand Shape Recognition

At first, we need to confirm how many hand shapes that we need to recognize. In some cases, several hand gestures have the same hand hand shape. For example, drag gesture, holding gesture and rotation gesture have the same hand shape. Their difference is the motion of palm. As we have introduced all the 14 gestures we design in chapter 3, we summary the following hand shapes that we will recognize, also as showed in Fig. 4.3.

Besides, system also needs to realize whether user's hands just move naturally without interaction intention. So we need to recognize the natural hand shape. Thus there are 9 hand shapes that we need to realize.

Here is relationship between 14 hand gestures and 9 hand shapes:

(1) OK hand shape (including 1 gestures): OK gesture.

(2) Pointing hand shape (including 3 gestures): pointing gesture, NO gesture, click gesture.

(3) Extending hand shape(including 5 gestures): pull/push gesture, wave gesture, holding gesture, drag gesture, rotation gestures gesture.

(4) Grab hand shape (including 1 gestures): grab gesture.

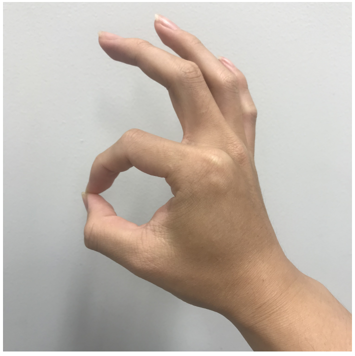
(5) Zoom hand shape (including 1 gestures): zoom in/out gesture.

(6) Scroll/swipe hand shape (including 1 gestures): scroll/swipe gesture.

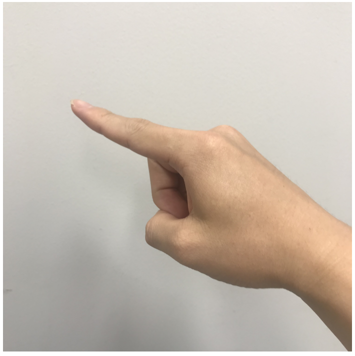
(7) Opening/closing hand shape (including 1 gestures): opening/closing gesture.

(8) Changing area hand shape (including 1 gestures): changing area gesture

(9) Natural hand shape: the hand shape when user move hands without interaction intention



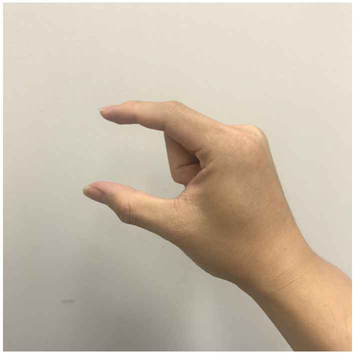
(a) OK hand shape



(b) Pointing hand shape

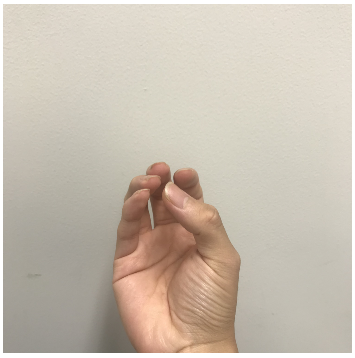
(c) Extending hand shape

(d) Grab hand shape

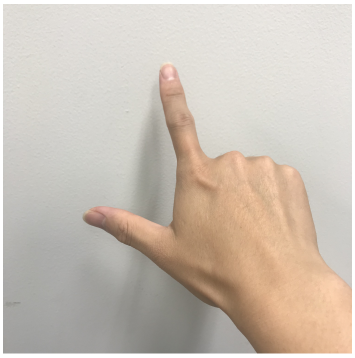


(e) Zoom hand shape

(f) Scroll/swipe hand shape



(g) Opening/closing



(h) Changing area

(i) Natural hand shape

Fig. 4.3 The nine hand shapes

Then we adapt SVM method to our system to realize these nine hand shapes. So we need the multi-label classification method in our system. We apply open source software, libsvm-3.22, in our system [28]. There are four steps for multi-label classification:

(1) data collection

(2) data normalization and scale

(3) model training

(4) predicting

#### 4.3.1.1 Data Collection

In the first step of multi-label classification, we need to confirm what kind of position data we need. In a hand of human, there are nineteen bones related to five fingers. We could use the endpoints of these bones and palm center as “key points” to describe these hand shapes. As showed in Fig. 4.4, the coloer points show key points of hands. The key points could be tracked by Leap Motion and we could get all the position data in every frame. With the position data, we could describe the hand structure the hand shapes.

Fig. 4.4 Key Points: two blue points represent palm center and wrist joint; red points represent the endpoints of bones in hand

#### 4.3.1.2 Data Normalization and Scale

We can make the palm center as the origin coordinate. Then we calculate other key points positions relative to palm center.

Data normalization follows these steps:

(1) Translate all the points until the palm center is on the origin coordinate.

(2) Rotate the points around the palm center until the palm parallel to the x-z axis plane.

(3) Rotate again the points around the y coordinate axis until the palm points the – z axis.

Then we scale all the data to [-1, 1]. This step will eliminate noise and improve accuracy.

#### 4.3.1.3 Model Training

The third step of hand shape recognition is model training. Out goal is to recognize the nine hand shapes. There are many SVM models and we choose to use Classification SVM Type 1 (also known as C-SVM classification) in our system.

There are three steps to train with data and get a classifier model:

(1) Capture 50 groups coordinates of the key points for every hand shape.

(2) Normalize and scale the data and get 9 groups training sets.

(3) Through training sets, get the multi-label classifier.

The model training part is preparation work for system. The multi-label classifier model will be used to realize hand shapes in time in our system.

#### 4.3.1.4 Predicting

After getting the multi-label classifier, we will use it in VR shopping system to recognize the nine hand shapes. When user viewing the room-scale VR shopping store, user moves his hands freely. Depth sensor tracks the hands and gets a group of original data set in every frame. Then with the multi-label classifier, we will get the predicting result. The predicting result will tell system which hand shape that user is performing. If the hand shape is natural hand shape, system will not give feedback; if getting other hand shape, system will respond to user's interaction intention.

We put a hand above the Leap Motion and perform the the nine hand shapes for 100 times respectively. We record the data and test the accuracy of the method. In the system, the accuracy of classifier is showed in Table. 4.2.

| Hand shape | OK | pointing | extending |
|---|---|---|---|
| Accuracy | 92% | 94% | 94% |
| Hand shape | grab | zoom | scroll/swipe |
| Accuracy | 92% | 93% | 93% |
| Hand shape | opening/closing | changing area | natural |
| Accuracy | 94% | 93% | 95% |

Table 4.2 The accuracy of recognition the nine hand shapes

### 4.3.2 Motion Detection

After knowing the hand shape, the system needs motion detection because gestures are defined by both hand shape and motion together.

Once getting the hand shape, system starts motion recognition step. After the frames when hand shape recognized, system will continuously calculate and record the positions of key points in every frame captured by depth camera. For different hand shapes, system detects motion of different fingertips or hand to realize gestures.

Basing on the hand shapes from label 1 to label 9, there are nine situations when recognize motion.

**Situation 1**: for OK hand shape, OK gesture is the level 1 gesture and system do not need to recognize the motion.

**Situation 2**: for pointing hand shape, if system find that two hands are in pointing hand shape, system need to detect the positions of two index fingertips. If the two index fingertips is close, it shows that user is performing NO gesture. If only one hand is in pointing hand shape, system need to detect the direction and motion of index finger, because we need to use pointing gesture to choose a target or click a button.

**Situation 3**: for extending hand shape, it is a little complexed. (a) if detecting that the palm center orients to face and moving toward to face, it is pull gesture; (b) if detecting that the palm center orients forward and moving forward, it is push gesture; (c) if detecting that the palm center orients left and moving to left, it is waving; (d) if detecting that the palm center orients to sky, it is holding gesture; (e) if detecting that the palm center orients forward and moving on a vertical plane, it is drag gesture; (f) if detecting that the palm center orients to sky and rotating around the palm center, it is rotation gesture.

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

**Situation 5**: for zoom hand shape, system detects the motion of index and thumb fingertips. If the fingertips move away from each other, it is zoom in gesture; if the fingertips move to close, it is zoom out gesture. The movement distance will be use change the size of target object.

**Situation 6**: scroll/swipe hand shape, system detects the motion of index finger. The movement distance will be used to control menu.

**Situation 7**: for opening/closing hand shape, system detects the motion of index, middle and thumb fingertips. If their motion is moving close to each other, it is closing gesture; if their motion is moving away from each other, it is opening gesture.

**Situation 8**: for changing area hand shape, system detects the direction of index finger and motion of thumb fingertip. If index finger points to a position on the floor and thumb fingertip clicks the index finger, it is changing area gesture and user will move to the target position.

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

## 4.4 Apply Gesture Set in Room-scale VR Shopping Environment

Once finishing the works on gesture set designing and gesture recognition, we apply the gesture set in our VR environment to build the interaction system. In order to test the hand gesture interaction, We design a typical shopping activity: viewing and buying a laptop.

Firstly, user could move to the desk with change area gesture where the laptops are placed in the room-scale VR shopping environment, as showed in Fig. 4.5. In this situation, the desk is a little far away from user and is out of the original walking area of user.

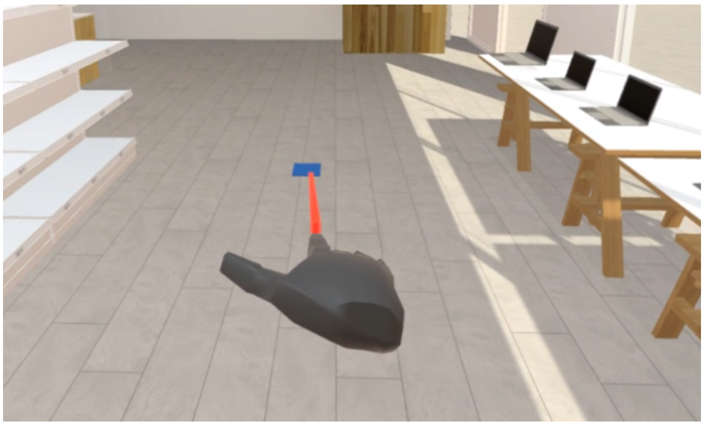
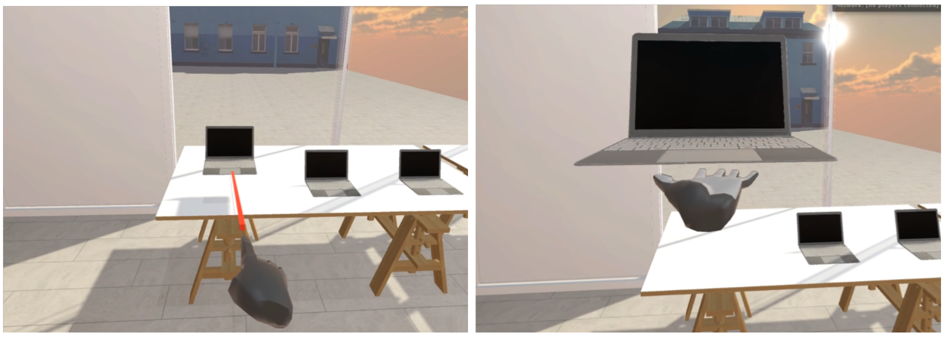


Fig. 4.5 Using change area gesture

Then user selects one of the laptops with pointing gesture. Once being selected, the laptop will show a bounce animation, as showed in Fig. 4.6a. User could make the laptop move to his hand with hold gesture, as showed in Fig. 4.6b.



(a) Using pointing gesture (b) Using hold gesture

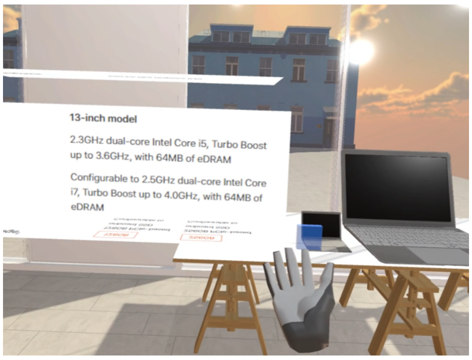
Fig. 4.6 Select and hold the target product

After that, user could view the details of the laptop with zoom in/out gesture, as showed in Fig. 4.7a. Besides, user could call out menu to check more information with open/close gesture and scroll/swipe gesture Fig. 4.7b and Fig. 4.7c.

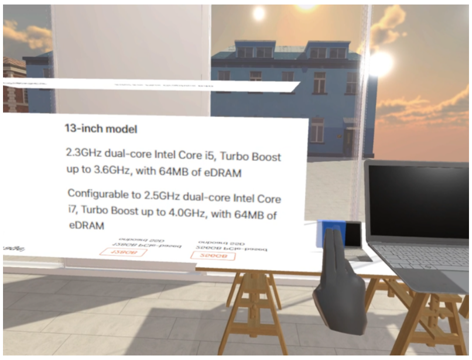
Finally, he could perform the OK gesture to tell the system that he decides to it, as showed in Fig. 4.8.

Using the gestures that we design, we could achieve a typical shopping activity in the room-scale VR shopping environment. Our system provides the shopping activity flow which is similar to our daily shopping activity with natural hand gestures.

(a) Using zoom in/out gesture




(b) Using open/close gesture




(c) Using scroll/swipe gesture

Fig. 4.7 Using zoom in/out gesture, open/close gesture and scroll gesture


Fig. 4.8 Using OK gesture

# Chapter 5

# Related Work

With the perfection of VR technology, many researchers and companies try to apply VR technology in e-commerce field and want to find a way to generate economic value. Alibaba is a famous IT company and is known for its great online shopping services. Tianmao is one of its online shopping services. Tianmao presented a VR shopping application, called Tianmao buy+, running on smartphone. The tianmao buy+ tried to combine the convenience of online shopping and the facticity of physical store shopping. With the simple and cheap VR device and smartphone, people in China could view the stores around American Times Square and pay for orders online. The VR application gives people a feeling that they are shopping at American Times Square. The VR shopping experience is really fantastic.

Some companies use VR technology to create virtual store. IKEA is a famouse furniture company. It presented a room-scale VR kitchen to show its beautiful design [20]. In the room-scale VR kitchen, user could use HTC vive to view the equal proportion VR kitchen, even could interact with the VR environment like opening range hood. Comparing with physical furniture stores, VR environment could provide more function and interactions. User could view the kitchen freely in his own room without warring about crowd. In the VR kitchen, user also could change the color of furnitures easily, which is impossible in a physical store.

A VR technology company, inVRsion, presents a VR supermarket system based on room-scale VR [21]. Their retail space, products and shopping experience VR solutions provide an immersive shopping environment. In the VR shopping environment, businessman could analyze shopper behavior through eye-tracking for extremely powerful market research insights. The system could help seller to test his category projects, new packaging and communication in store before implementation. User could search his target products more easily than physical supermarket. This system tries to provide a method for people to view a big virtual supermarket in his own room.

The previous related works prove the broad application prospect of room-scale VR shopping. Our work about gesture set for room-scale VR could be used in these system to provide better VR shopping experience.

# Chapter 6

# Conclusion and Future Work

## 6.1 Conclusion

In this research, we build a room-scale VR shopping system and design a new hand gesture set for the room-scale VR shopping system. Hand gestures have more interaction languages and we use the gesture set to replace the controllers of VR device to improve the limitation of controllers in VR shopping activities. Researching on the gesture set, we introduce a new gesture classification. We design a three levels for classification method. The gestures in level 1 are static hand shape, the gestures in level 2 are movement palm and the gestures in level 3 combine features if gestures in level 1 and level 2. The gestures in level 1 and level 2 are simple and is not enough for room-scale VR shopping activities. Thus, we focus our research on the level 3 gestures.

For the gestures in level 3, we introduce three categories to classify them: core gesture, gestures for interaction with virtual object, gestures for interaction with menu and gesture for interaction with space. The classifications help us to understand the gesture set in room-scale VR shopping system. Also the gesture set and 3-level classification method could be transplanted to other VR or AR system conveniently. In order to achieve the complex gestures, we apply SVM method in our VR shopping system. In the end, user could walk around in his room to view the VR shopping store and interact with system with natural hand gestures.

## 6.2 Future Work

In the future, we plan to improve the room-scale VR shopping system. We will improve the currency of gesture recognition and the convenience of interaction of system. The gesture set could be extended to add more interesting functions in system. We will also design some experiments to evaluate how much our gesture set improves the efficiency comparing with controllers in room-scale VR shopping environment.

# References


[1] Berenice Serrano, 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*, 55:1 – 8, 2016.

[2] Quinate Chioma Ihemedu-Steinke, Rainer Erbach, Prashanth Halady, Gerrit Meixner, and Michael Weber. Virtual reality driving simulator based on head-mounted displays. In *Automotive User Interfaces: Creating Interactive Experiences in the Car*, pages 401–428. Springer International Publishing, 2017.

[3] Ganesh Datt Bhatt. Bringing virtual reality for commercial web sites. *International Journal of Human-Computer Studies*, 60(1):1 – 15, 2004.

[4] Bob G. Witmer and Michael J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998.

[5] Marco Speicher, Sebastian Cucerca, and Antonio Krüger. Vrshop: A mobile interactive virtual reality shopping environment combining the benefits of on- and offline shopping. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(3):102:1–102:31, sep 2017.

[6] James Jeng-Weei Lin, Henry B. L. Duh, Donald E. Parker, Habib Abi-Rached, and Thomas A. Furness. Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Virtual Reality, 2002. Proceedings. IEEE*, pages 164–171. IEEE, 2002.

[7] Jih-fang Wang, Ronald T. Azuma, Gary Bishop, Vernon Chi, John Eyles, and Henry Fuchs. Tracking a head-mounted display in a room-sized environment with head-mounted cameras. In *Helmet-Mounted Displays II*, volume 1290, pages 47–58. International Society for Optics and Photonics, 1990.

[8] Mark Ward, Ronald Azuma, Robert Bennett, Stefan Gottschalk, and Henry Fuchs. A demonstrated optical tracker with scalable work area for head-mounted display systems. In *Proceedings of the 1992 symposium on Interactive 3D graphics*, pages 43–52. ACM, 1992.

[9] Greg Welch, Gary Bishop, Leandra Vicci, Stephen Brumback, Kurtis Keller, et al. The hiball tracker: High-performance wide-area tracking for virtual and augmented environments. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 1–10. ACM, 1999.

[10] Vix Kemanji Ketoma, Philip Schäfer, and Gerrit Meixner. Development and evaluation of a virtual reality grocery shopping application using a multi-kinect walking-in-place approach. In *Intelligent Human Systems Integration*, pages 368–374. Springer International Publishing, 2018.

[11] Roy A. Ruddle and Simon Lessels. The benefits of using a walking interface to navigate virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 16(1):5, 2009.

[12] Evan A. Suma, Zachary Lipps, Samantha Finkelstein, David M. Krum, and Mark Bolas. Impossible spaces: Maximizing natural walking in virtual environments with self-overlapping architecture. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):555–564, 2012.

[13] Adalberto L. Simeone, Eduardo Velloso, and Hans Gellersen. Substitutional reality: Using the physical environment to design virtual reality experiences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3307–3316. ACM, 2015.

[14] Misha Sra, Sergio Garrido-Jurado, Chris Schmandt, and Pattie Maes. Procedurally generated virtual reality from 3d reconstructed physical space. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, pages 191–200. ACM, 2016.

[15] Eike Langbehn, Paul Lubos, Gerd Bruder, and Frank Steinicke. Bending the curve: Sensitivity to bending of curved paths and application in room-scale vr. *IEEE transactions on visualization and computer graphics*, 23(4):1389–1398, 2017.

[16] Tian Chen, Zhi-geng Pan, and Jian-ming Zheng. Easymall-an interactive virtual shopping system. In *Fuzzy Systems and Knowledge Discovery, 2008. FSKD'08. Fifth International Conference on*, volume 4, pages 669–673. IEEE, 2008.

[17] Kun Chang Lee and Namho Chung. Empirical analysis of consumer reaction to the virtual reality shopping mall. *Computers in Human Behavior*, 24(1):88 – 104, 2008.

[18] Helena Van Kerrebroeck, Malaika Brengman, and Kim Willems. Escaping the crowd: An experimental study on the impact of a virtual reality experience in a shopping mall. *Computers in Human Behavior*, 77:437 – 450, 2017.

[19] Adrien Verhulst, Jean-Marie Normand, Cindy Lombard, and Guillaume Moreau. A study on the use of an immersive virtual reality store to investigate consumer perceptions and purchase behavior toward non-standard fruits and vegetables. In *Virtual Reality (VR), 2017 IEEE*, pages 55–63. IEEE, 2017.

[20] Ikea vr experience, 2016. https://www.ikea.com/ms/en_US/this-is-ikea/ikea-highlights/Virtual-reality/index.html.

[21] Shelfzone vr, 2016. https://invrsion.com/shelfzone.

[22] Chunmeng Lu, Li Zhou, and Jiro Tanaka. Realizing multi-touch-like gestures in 3d spaces. *Human-Computer Interaction International Conference 2018*, July, 2018 (accepted, to appear).

[23] Pragati Garg, Naveen Aggarwal, and Sanjeev Sofat. Vision based hand gesture recognition. *World Academy of Science, Engineering and Technology*, 49(1):972–977, 2009.

[24] Juan Pablo Wachs, Mathias Kölsch, Helman Stern, and Yael Edan. Vision-based hand-gesture applications. *Commun. ACM*, 54(2):60–71, February 2011.

[25] Yves Boussemart, François Rioux, Frank Rudzicz, Michael Wozniewski, and Jeremy R. Cooperstock. A framework for 3d visualisation and manipulation in an immersive space using an untethered bimanual gestural interface. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '04, pages 162–165. ACM, 2004.

[26] Hani Karam and Jiro Tanaka. Two-handed interactive menu: An application of asymmetric bimanual gestures and depth based selection techniques. In *Human Interface and the Management of Information. Information and Knowledge Design and Evaluation*, pages 187–198. Springer International Publishing, 2014.

[27] Leap motion, 2018. https://www.leapmotion.com/.

[28] libsvm-3.22, 2016. https://www.csie.ntu.edu.tw/~cjlin/libsvm/.