Function Simulation and Instructional Preview Based on Augmented Reality

Kanghao Zhong
44201023-7

Master of Engineering

Supervisor: Prof. Jiro Tanaka
Graduate School of Information, Production and Systems
Waseda University

January 2022

Abstract

In recent years, with the development of Augmented Reality, people can preview products and understand products in advance intuitively before purchase them especially for e-commerce. However, it is also very hard for user to know functional information and how to use it in real scenes. For example, the usage of household appliances with new function, which could confuse some people especially the elder a lot.

Besides, as traditional way to guide people how to use products, textual manual has always been used and even online shopping also introduce products in form of text and picture. which need too much time to read and hard to understand directly. It also brings people inconvenience.

In this paper, we introduce AR technology to provide a new experience for preview to provide more intuitive simulation and easy-to-understand instructions. In order to make the preview experience more understandable and make users more familiar with usage, we design and accomplish this Augmented Reality system which provides experience corresponding function based on smartphone and an instructional preview. The system aims to provide users a more interactive experience which allow users to learn specific usage steps from avatar instruction and manipulate 3D models with their own hands. And by combination of smartphone and Leap-motion, user can be provided convenient and portable simulation experience.

Keywords: function simulation, augmented reality, 3d model, hand tracking

Acknowledgements

First of all, I would express my heartfelt thanks and gratefulness to my supervisor, Pro. Jiro Tanaka, in the term of two years study, he has helped and supported me a lot. In the study of AR and human-computer interaction, he also led me to gradually master it step by step. In terms of the determination and completion of the research topic, he also gave me a lot of constructive suggestions and proposed possibilities and solutions to the difficulties and problems I faced. I can say that he is the most grateful person for the two years.

Then, it is grateful to all IPLAB members. I enjoyed every day in the lab and was honored to be a part of it. Although we had only been together for about a year due to the COVID-19, I also feel the warmth like a family through mutual help, communication and learning. With the help and encouragement from them, I gradually integrated into the environment and made rapid progress. Thank you very much for your company.

Finally, I would like to express my gratitude to my parents and all family, who provided the opportunity to study abroad and supported me financially and mentally during this time .

Contents

List of figures vi

List of tables viii

1 Introduction 1

1.1 Introduction . . . 1

1.2 Organization of the thesis . . . 2

2 Background and Related Work 3

2.1 Using 3D model in AR . . . 3

2.2 Intuitive Free-hand Interaction with 3D Model . . . 4

2.2.1 Interaction in AR Glasses-based Augmented Reality . . . 4

2.2.2 Interaction in Smartphone-based Augmented Reality . . . 5

2.2.3 Adding Physical Properties to 3D Models for Realistic Interactions 5

2.3 Products Preview System . . . 6

2.3.1 3D Product Presentation in E-commerce . . . 6

2.3.2 Product Preview in AR . . . 7

2.3.3 Current AR Preview System and Research . . . 7

2.3.4 Interactive User Manual based AR . . . 8

3 Research Goal and Approach 10

3.1 Research Goal . . . 10

3.2 Research Approach . . . 10

3.3 Use Case . . . 12

4 System Design 15

4.1 System Overview . . . 15

4.1.1 System Hardware . . . 16

4.2 Module Design . . . 18

4.2.1 Function Simulation 18
4.2.2 Instructional Preview 23
4.3 Interaction Design 29
4.3.1 Interaction between People and Virtual Objects 30
4.3.2 Interaction between Real World and Virtual Objects 31

5 System Implementation 33
5.1 Development Environment 33
5.2 3D Model Preprocessing 34
5.2.1 Model Construction 34
5.2.2 Functional Trigger 36
5.3 Connect Smartphone and Leap Motion 37
5.3.1 Connection Framework 38
5.3.2 Wireless Data Transmission 39
5.4 Recognition Real Object and Real Scene to Realize Spatial Interaction . . . 41
5.4.1 Scanning of 3D Object and Using Object Target 41
5.4.2 Scanning Real Scene and Using Area Target 44
5.5 User Interface and Avatar 46

6 Conclusion and Future Work 49
6.1 Conclusion 49
6.2 Future Work 50

References 51

List of figures

3.1 Method to wireless data transmission 11
3.2 Spatial recognition and interaction 12
3.3 Use Cases 13
3.4 Use Flow 14

4.1 System overview 16
4.2 Smartphone Device & Leap Motion 17
4.3 Portable Devices 17
4.4 Flow Chart of Kettle Case 19
4.5 Construction of Each Components of Model 20
4.6 Construction of Collides of Model 21
4.7 Connection of Devices 22
4.8 Gesture Control with Mobile Devices 22
4.9 Structure Diagram of Instructional Preview 23
4.10 Avatar Instruction 24
4.11 Instruction Panel 25
4.12 Simulation Experience Based on Real Scene 26
4.13 3D Model of Scanned Area 27
4.14 Guidance Signals in Rooms 27
4.15 Virtual Water Based on Position of Faucet 28
4.16 The Result of Object Detection 29
4.17 Interaction between Users and Virtual Objects 30
4.18 Design of User Interface 31
4.19 Interaction between Real World and Virtual World 32

5.1 Construction of 3D model 35
5.2 Add key frames to components of model in Blender 35
5.3 Functional Trigger on Model 36

5.4 Functional Trigger on Hand 37
5.5 Connection Framework 38
5.6 Creation of Server 40
5.7 Build the Link 41
5.8 Object Scanning Target Image 42
5.9 Generated File & Test 42
5.10 Generate Unity Package File 43
5.11 Package Download & Import 44
5.12 Select Target Database 44
5.13 Official Recommended Devices 45
5.14 Scanning environment using Vuforia Area Target Creator 45
5.15 Creating Area Target 46
5.16 UI Buttons and Toggles 47
5.17 Construction of Avatar 48
5.18 Red Ray-casting from Avatar 48

List of tables

4.1 The PC Setup Information 18

5.1 Software and Information 34

Chapter 1

Introduction

1.1 Introduction

In conventional way of online shopping, if customers would like to know something about products, they have to view appearance in picture and read detail in text. In such case, complicated information and counter-intuitive experience will confuse users and reduce the enthusiasm of users.

As the usage of innovative design and technology is becoming more and more popular in e-commerce, many company have attempted to integrate augmented reality to their e-commerce platform, by means of which, consumers are able to preview products in advance [1]. Some researchers are also focusing on combining e-commerce and AR to display AR object, like design interior furniture [2]. So far, e-commerce with AR is developing rapidly in every sense.

As a result, with the development Augmented Reality in e-commerce, nowadays, people can preview products and understand products intuitively in lots of different platform. The most common examples are Amazon [3] and Ikea [4], integrated with AR technology, by which users are able to see 3d models of equal proportions of various products in real environment before deciding to purchase something.

However, even if current system has solved existing problems and make e-commerce more convenient, there are still two points that are able to be optimized. Firstly, it is still hard for user to know functional information and how to use it in real scenes, especially for some electric appliances. The second is that we can optimize the conventional instruction method by integrating AR elements, like avatar and 3d instruction board displayed instead of textual manual.

In this research, we implemented a function simulation and instructional preview AR system. According to Azuma [5], combining real world and virtual content and providing real-time interaction in 3D are necessary and sufficient condition for AR. Therefore, based on this principle, we have presented this research to simulate product function and optimize the instruction by enhancing the interaction between virtual content and virtual content. In addition, by combining smartphone and gesture tracking device, users can experience device in a portable and intuitive method.

1.2 Organization of the thesis

The rest of the thesis is organized as follows: Chapter 2 will introduce some background and related works about technology and previous system. Chapter 3, we will give the overall goal of this research and a brief introduction to the research method. Chapter 4 will introduce the design of our system in detail, including the overview of the system and the detail design of each part. Chapter 5 we will introduce the implementation part of our system, including the used software and the detail implementation methods of the functions. Chapter 6, we will briefly summarizes this work and introduce what we can optimize in future plan.

Chapter 2

Background and Related Work

In this chapter, I will introduce the background of our research about technique and practical design respectively. The technique almost includes 3D models in AR, interaction in AR and 3D models with properties for interaction. About practical design, there are previous preview system and instructional operating manual. In addition, in each part, I will also introduce some related works.

2.1 Using 3D model in AR

The definition of AR(augmented reality) can be explained as a technique that allows users to have an interactive experience with various digital perceptual information including digital visual elements, sound, or others which can be enhanced [6] in real environment.

In addition, to combine AR technology within a system, we can almost identify it into three characteristics [7]:

1. AR can integrate virtual images to real object(or environment) as a display technology
2. AR also is a real-time process which supports do several interactions simultaneously
3. AR need to be displayed in 3D and by means of using tracking technology to realize the 3D display

Therefore, as the first step of using augmented reality, undoubtedly, it is using 3d virtual images (or models) in system. Using 3D model in AR system can be seems as a critical component of long-term, well-thought-through digital strategies to ensure app longevity. So as to build 3D assets, there are two main method for augmented reality, including 3D modeling and 3D scanning. The main difference between them is the method of capturing the data of the object [8]. In order to enhance the immersion and vraisemblance, we used these two ways simultaneously in this research.

2.2 Intuitive Free-hand Interaction with 3D Model

Interaction can be one of important components in AR system, in many existing system, hand external devices is main method to realize the interaction. However, actually, intuitive free-hand interaction is a more immersive method for AR.

2.2.1 Interaction in AR Glasses-based Augmented Reality

In order to enhance the interactivity of AR system, the intuitive free-hand interaction is indispensable. Nowadays, mobile devices, especially wearable devices, have rapidly developed and made a vast improvement.

Yueting Weng, Chun Yu and Yingtian Shi [9] have proposed a novel sensing technique about hand gesture and facial information based on vision, which needs using AR glasses. They trained multiple different convolutional neural network (CNN) models and obtained hand or face detection interaction.

For more accurate detection, some developers used tracking devices to support AR glasses while parts of head-mounted devices has fundamental tracking function. Ikeuchi, Kohki and Otsuka [10], increase the immersive AR experiences by enhancing somatic sensation with the usage both of AR device and Kinect.

As a result, computing and wearable computing have incorporated several mature technology into a head-mounted package to provide free-hand interaction. However, sometimes the heavy equipment and limited input method make users inconvenient to have natural

3D interaction experience which plays a considerable significant roles in AR immersive experience [11].

2.2.2 Interaction in Smartphone-based Augmented Reality

As mentioned in last section, to avoid the shortcoming of head-mounted and improve the immersive experience, some researches have thought about new direction in developing portable devices like smartphone.

The one of method is using context-aware smartphone. Botev, Jean and Mayer, Joe and Rothkugel, Steffen [12] used smartphone with depth camera and developed Open-CV framework to realize multi-user interaction with virtual objects in smartphone platform. However, the drawback of this method is also obvious. First, the context-aware smartphone is necessary, which means the platform is limited. Additional, because it is based on trained data framework, developer need to maintain and update the sample database.

Another is using extra tracking device, such as Leap-Motion. Jing Qian, Jiaju Ma, Xiangyu Li, and Benjamin Attal [13] used Leap-Motion to track hand gesture and display AR content based on common smartphone. In this case, it enhance the immersion of AR experience and make it more portable and convenient.

2.2.3 Adding Physical Properties to 3D Models for Realistic Interactions

To enhance interaction, apart from tracking device, adding realistic properties to 3D models according to real corresponding object is considerable important. For achieving realistic experiences, AR interactions between the real and the virtual sometimes need to follow real physical rules. Imbert, Nicolas and Vignat [14] built and used 3D models with their own dimensions , volumes or weight. They tested their research by means of using eight different types of model with corresponding markers.

However, just using markers is not nature and convenient, besides in our case, in order to simulate functions, just adding law of physics is not enough. Thus we combined with last section work to interact with these 3D model with properties and added corresponding triggered events to realize function simulation

2.3 Products Preview System

In common sense, before people would like to know something about a product, there are multiple ways to obtain specific details of products. In traditional way, it is considerable cumbersome that people have to know the information through pictures and descriptions on paper or go directly to offline stores to check out products. However, as everyone knows, it is quite inconvenient.

2.3.1 3D Product Presentation in E-commerce

Nowadays, with the rapid development of the Internet and smart phones, there are more and more approaches to know products in advance. Current e-commerce make customers all around world be able to have an efficient shopping experience with using different e-commerce platforms online [15]. People can preview the appearance and read the relative information of products in the Internet.

Usual e-commerce mode is also through picture and text to indicate information. on account of its limitation, product displayed in 3D in e-commerce become a new and popular method to show more vivid detail of products. Hewawalpita, Supun and Perera, Indika [16] have done the research about the influence of product presentation shown in 3D for better customer experience in different e-commerce platforms.

For instance, while using online shopping platform, customers can check the product by 3D model view, instead of picture. In this case, customer can achieve more three-dimensional information and more intuitive preview experience. Moritz, Franka [17] also indicate that positive effect and claim the potential for future development.

2.3.2 Product Preview in AR

As extension of 3D model in product presentation, AR is the most novel form of technical expression integrated into e-commerce. By fusing the real environment and virtual object in 3D presentation, the understanding of products can be more intuitive and the interest of users in e-commerce will undoubtedly be further.

2.3.3 Current AR Preview System and Research

About products preview in AR, the interior design and furniture preview become a popular practical and research direction. Waraporn Viyanon, Songsuittipong, et al. [18] research and develop an augmented reality application which allows users preview how various furniture will fit in real environment like their home and details of products through AR display support customer to make decision more efficiently. In practical application field, “IKEA place” [4], a popular and famous android application ,which allows users to shop online and support virtually placing true-to-scale 3D models in your very own space(Markerless Augmented Reality). Alves, Carlos and Reis [19] indicate that "IKEA place" will bring more attraction and potential in the future. Apart form above works, many other application and researches are also developing. Similar to IKEA place, Amazon AR View [3], as one of biggest e-commerce platform, also support product preview in real environment.

In term of longitudinal development of product preview in AR, some AR research start focusing on improving the preview experience and make the AR experience more humane and convenient. As one method of markerless AR used in field of online shopping, Desai et al. [20] enhance the AR experience in real environment and add some more interaction in manipulation interface. For instance, the color and size can be changed by users and users can check the description of product in text and view the 3D model in real environment simultaneously. In addition, different types of objects can be displayed at the same time, providing a rich experience for the user. Ruichen, MA [21] present and implement a MR shopping system which provide hand scale preview and gesture interaction. In her case, the preview experience can be scaled making users more convenient and comparing with previous work ,by using hand gesture, more intuitive interactive experience can be provided to customers instead of interacting through smartphone screen.

In summary, current step of AR product preview systems have been developed fairly well, and some applications have already been used widely. Therefore, in case of this research, the main purpose is based on their research and improve the experience further. On account of lack of realistic interaction based on product features in previous works, briefly, in field of product preview, we aim to give more interaction and properties for 3D model to simulate the function of products.

2.3.4 Interactive User Manual based AR

In order to instruct users how to use products, as usual, there is a operation manual that explain the usage of product and what should be noticed according to different products. As well as last section, instructional manual is also provided to users more convenient through developing Internet techniques. People can read relative information and know the usage of products in advance by using smartphone or computer.

However, it is not intuitive even if current e-commerce can show information online. Therefore, if we can use AR to display the instructional information, it is no doubt that users can understand usage better and read content of manual in higher interest [22].

Fiorentino, Michele and Uva, et al. [23] research a system which display AR content by means of combining usage of a big screen and multiple cameras in different angle. Users can read the instructional labels, 3D virtual objects and animation which as AR content on big screen and manipulate objects in real environment.

However the drawback of using big screen is very obvious, heavy equipment will bring users inconvenience and reduce the passion to continue using this instructional system. In this case, it opposites to the initial purpose of system design. As we mentioned last section, Desai et al. [20] provided a new thought is combining the AR model preview and AR instructional content simultaneously through smartphone devices. In this case, the experience become portable and convenient.

Based on previous work, we designed instructional preview part in our system. We add avatar and combine with scene recognition to improve the instruction more intuitively and attractive.

Chapter 3

Research Goal and Approach

3.1 Research Goal

The goal of our research is to provide an instructional preview with Avatar and experience with function simulative model based on smartphone AR, which broaden the scope of application of the conventional system. By means of hand tracking and scenarios recognition, users can experience the function of target product and trigger series of interaction. By our instruction module, users can learn usage and detail information of product intuitively. Our goals can be divided into the following points:

1. Users can manipulate the 3D model with functionality and trigger different function events according to products using hands directly;
2. System can demonstrate and guide users how to use product by avatar, textual board and voice;
3. System can detect real-world objects and support different interaction in corresponding scenes.

3.2 Research Approach

To achieve above goal, we almost can divide the approaches into following parts:

1. Connect the smartphone and tracking device: Leap-Motion to make system both portable and operable. We can use c++ web-socket and hot-spot as bridge and set a library of recognizable gestures, in which case, we can transmit preset gesture data from Leap-Motion to smartphone. Then, we used a computer stick instead of laptop or desktop to make system portable. The method is as Figure 3.1 shown.

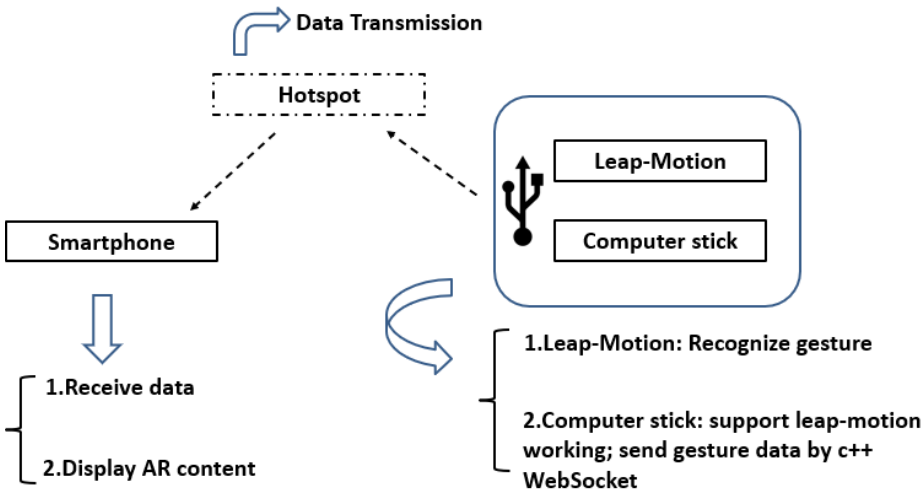


Fig. 3.1 Method to wireless data transmission

Then, in order to construct the model with functionality, we need build each part of product separately. After that, adding corresponding triggers and events on each part. In this way, the first goal can be achieved.

2. Search avatar model and corresponding appropriate motion animation to bind with avatar. Then setting animator to determine the sequence of motion. Meanwhile, adding a 3d textual instructional board and text related voice in system.

   In order to control above components, we develop multiple controller button in user interface, in addition, we also can select the model and instructed parts of product by controller. In this way, the second goal can be achieved.

3. In order to recognize real environment and add multiple interaction in scenes, we combined object target and area target to realize these module, as Figure 3.2 shows,

object target means detect specific real object and give events to it; area target means multiple AR contents can display and users can interact with them in recognized space. In this way, the third goal can be achieved.

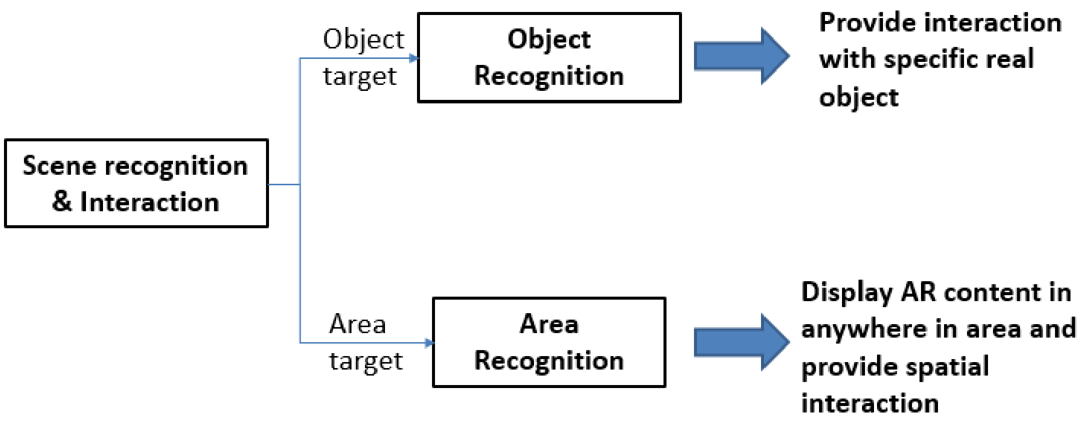


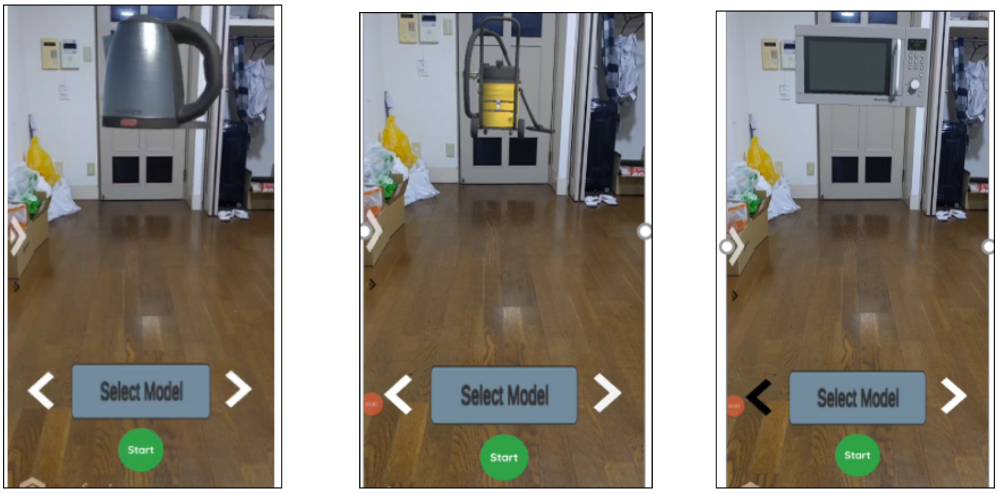
Fig. 3.2 Spatial recognition and interaction

On the basis of these methods, we designed a relatively complete system architecture to enable users to learn the information of product and decide to follow steps and experience the function of product.

3.3 Use Case

In use case part, we will introduce the main use flow of the system briefly, about which users can operate the system as following steps:

When users enter the system, there will be three cases can be selected, in our development, kettle is the main example to show function and other two cases also have been used as extended examples, which are microwave oven and dust cleaner. Users can browser three cases by clicking the button towards left or the button to right, after deciding, users need clicking buttons "select model" and "start", as Figure 3.3 shows.

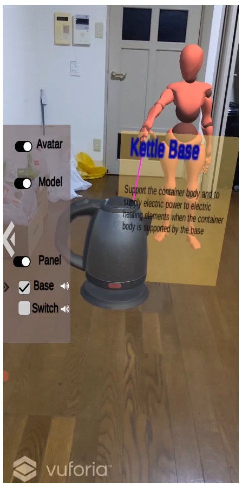


(a) Kettle (b) Dust Cleaner (c) Microwave Oven

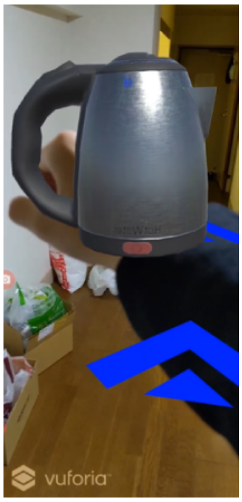
Fig. 3.3 Use Cases

Then selection interface will disappear and as Figure Figure 3.4a shows, the control user interface can display. This time, users can control the model by using hands to manipulate the model directly. Meanwhile, user interface also can interact with model or control the other object like avatar, text board or instructional voice. About these elements, we set red casting-ray to link the finger of avatar and components of model and it is related to the selected parts, meanwhile, the text board will show the corresponding information of model.

Besides, we provide instruction in specific space by recognizing related spatial information, as Figure 3.4b shows, users can manipulate the model following the guidance signal in prefab spatial position.



(a) UI and Controller



(b) Hand Control and Spatial Guidance

Fig. 3.4 Use Flow

In the case of kettle, users can experience the usage of pouring water, boiling water and pouring out water, during which, multiple instructional guidance based on spatial position will be shown to users.

About the various products example, apart from kettle that we will introduce how to use in detail, dust cleaner and microwave oven are also able to interact with user and simulate their functions in different scenarios. For instance, dust cleaner can detect plane like floor and if the real trash can are detected, dust effect will render the floor. When user grab the cleaner model to be close to floor, the effect will disappear and become " cleaner“Either, the microwave oven can detect different food, and recognize what can be heated or what can not.

Chapter 4

System Design

This chapter will introduce the design of the system. This chapter includes seven parts. The first part is the system design overview part. Then the middle part will introduce construction of 3D model, recognition in real environment gesture control and interaction in system design respectively. After that, we will introduce main function module design separately. The final part is about use case in which we will demonstrate our system in specific samples

4.1 System Overview

This system is mainly divided into two parts: function simulation and instructional preview. Figure 4.1 shows the overall design of this system.

The first part is **Function Simulation** . In this part, system provide users with different product preview options, which can be selected by users' own. After that, there will be corresponding 3D model display in real environment. The model is grabbable and on account of multiple triggers set on 3D model, the function of real products can be simulated. Besides, recording to characteristics of product, system can provide different function experience based on real environment including real object and real scenes.

The second part is **Instructional Preview**. In this part, the system can assist users to understand how to use the product. Apart from traditional method in text like operation manual which combined with AR board, the avatar also is used. The avatar can introduce the usage of product with pointing ray-cast, voice and operation demonstration in actions.

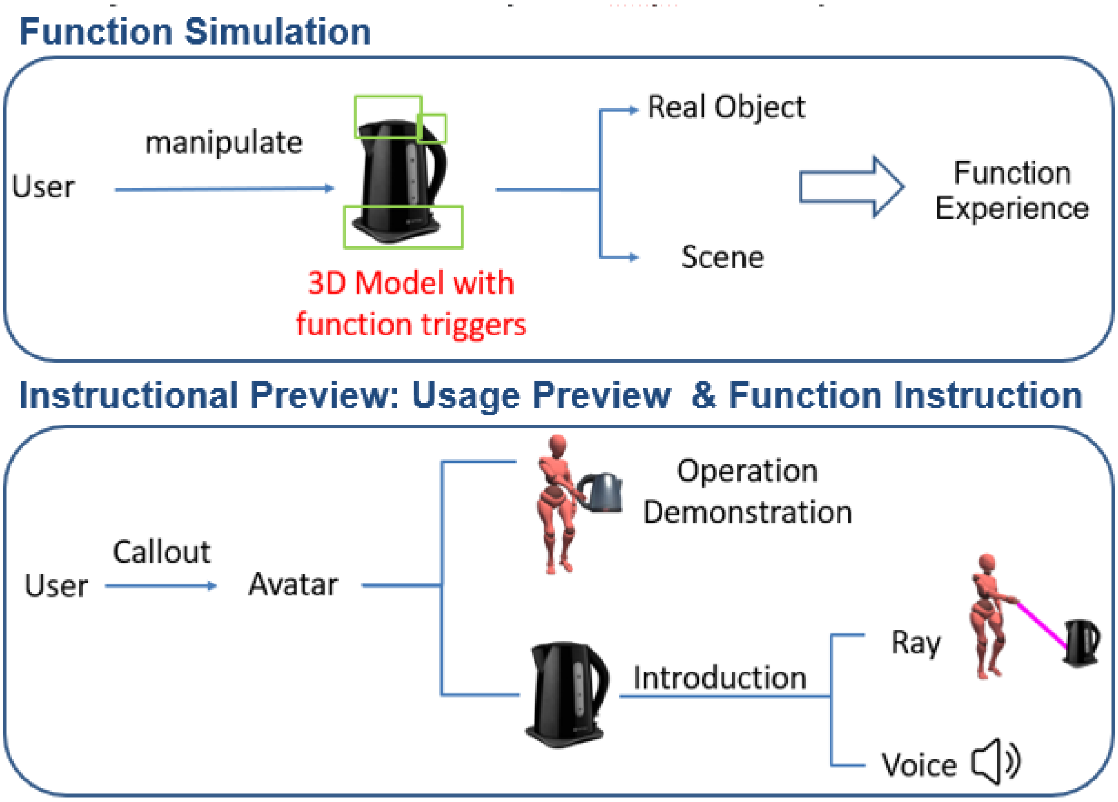


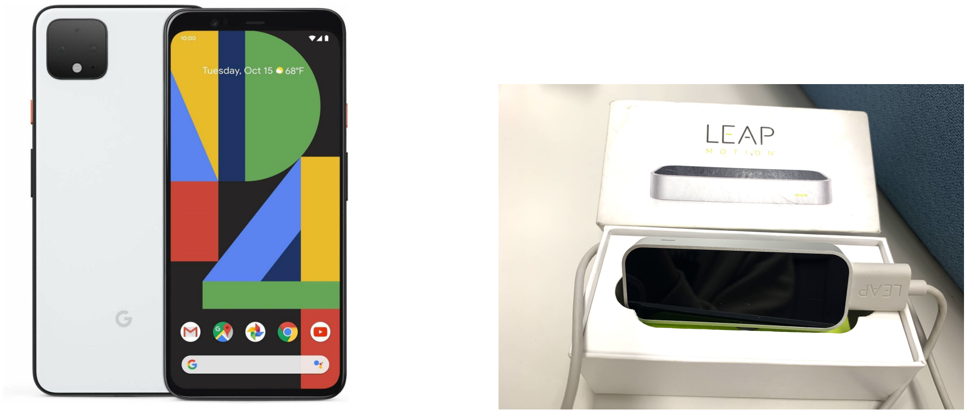
Fig. 4.1 System overview

4.1.1 System Hardware

In order to enhance the portability and convenience, instead of Head Mounted Device(HMD), we use smartphone as our device.

Here, we select Google Pixel 4, through which as Figure 4.2a shows, we can display virtual things in the midair the then users can interact with them by devices.

In addition, in order to address the lack of immersion and authenticity in usage of smartphone, we use Leap Motion[24] motion capture device to capture hand gesture, in which case, users can manipulate 3D model directly. Figure 4.2b is Leap Motion we use in research.



(a) Google Pixel 4 (b) Leap Motion

Fig. 4.2 Smartphone Device & Leap Motion

We use portable power source and computer stick to make Leap Motion portable and set a compute unit as a data relay, as Figure 4.3 shows, computer stick is Intel Atom x5-Z8300 and power source is Xiaomi Charger Pal2(10000mAh)

Fig. 4.3 Portable Devices

In order to use SDK and software to build the system and process programs, we use a PC with Windows 10 operating system. The configuration of the laptop, as development platform we use is shown in the Table 4.1 shows

Operation System	Windows10
CPU	AMD Ryzen 7 4800H CPU @ 2.90GHz
Graphics Card	NVIDIA GeForce GTX 1660 Ti
RAM	16.00GB

Table 4.1 The PC Setup Information

4.2 Module Design

For better understanding of this research, in this part, we will introduce the specific meaning of two main part: **Function Simulation** and **Instructional Preview** respectively. Additional, we will use the case of **Kettle** as example to explain in detail.

4.2.1 Function Simulation

In current stage, as we know, several existing AR system can preview the appearance of products in advance. However, as we mentioned in previous chapter, they are limited and not intuitive especially for some products with multiple function. For instance, household electrical appliances with control buttons, it is not easy for the elder to know how to use them and what should be noticed.

In this case, so as to provide more intuitive system which make users comprehend the usage, we design this system with module function simulation. The first step we need to do is construct **3D Model with Functionality**. Compared with traditional AR preview system which used static model, 3D model in this research can be manipulated which can trigger dynamic animation by interacting with surroundings. The main idea is using trigger, and we added multiple triggers on 3D model which can lead to interaction. As trigger method we used, they are almost divided into following parts:

1. Collision between user's hands and virtual button on 3D model, which means users can use hand to contact button on model to trigger function animation;

2. Collision between 3D models, which means the events can be triggered when specific trigger conditions are generated between the two models;

3. Real object detection and area detection, and this is based on different products used in specific usage scene.

The introduction of 3 method in detail will be explained in following session and so as to make it easier to understand, here we list one example of kettle first.

As we know, the usage of kettle can be divided into pour water, boil water and pour out water into cup/ bottle, therefore, in this system,we simulate this three process as Figure 4.4 shows.

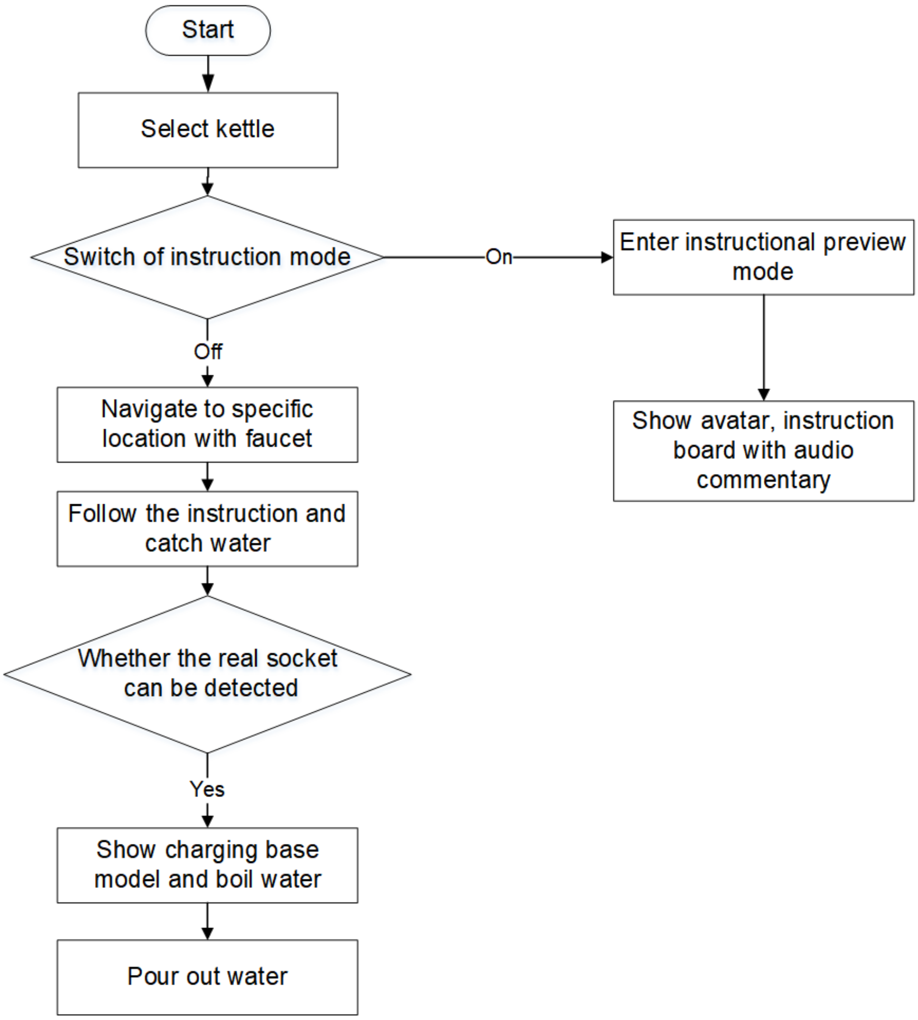


Fig. 4.4 Flow Chart of Kettle Case

After we choose the kettle case, we can switch into two module by open or close the instructional board and avatar. When we enter simulation mode, we can manipulate the kettle model as the real one. The method design and interaction part will be introduced in following section.

3D Model with Functionality

In this part, the main idea of construction of function simulation will be introduced. That is using 3D model with functionality.

As we mentioned in last section, in order to simulate some specific electrical appliance, just using static model is not enough. In this case, the first thing which we need to do is building each important components of product model respectively and certainly, components are able to fit together, as Figure 4.5 shows. After that, according to real products, we can polish the model, for example adding a effect of light reflection so that the model looks more realistic.

What is more, how to give the model functionality and how to trigger it play a more important role in the construction of model. Here we use collide module in Unity, which can trigger preset events when collision occurs with surrounding target.

Fig. 4.5 Construction of Each Components of Model

In the same way, we list kettle as the example, according to the characteristic, there are three triggers has been set on the kettle model: trigger on the top of kettle(open or close cap); trigger on the handle of kettle(switch of boiling water); trigger on the bottom of kettle body and charging base(another trigger of boiling). The framework is like Figure 4.6 shows.

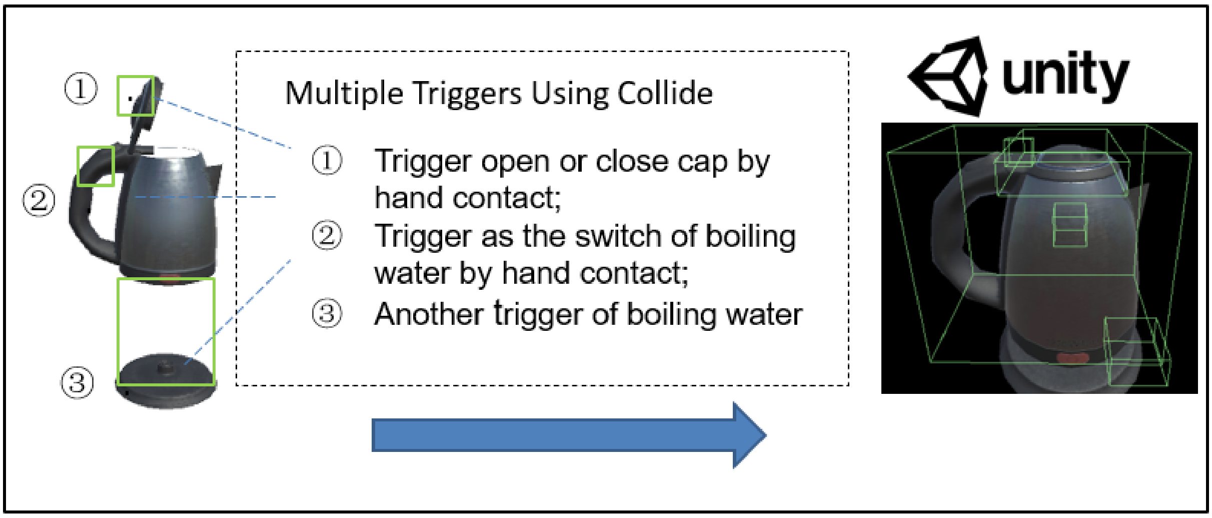


Fig. 4.6 Construction of Collides of Model

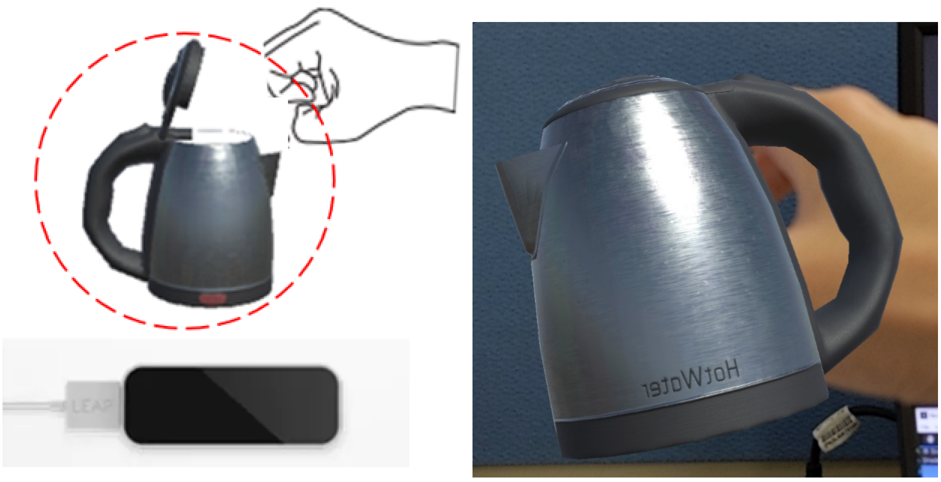
Gesture Control with Mobile Devices

In this system, as we mentioned in previous part, we use smartphone and Leap Motion as basic device, in this case, it brings more portable experience comparing with Head Mounted Devices. By means of compute stick running a C++ Web-Socket server, as Figure 4.7 we can connect Leap Motion to smartphone faultlessly. The specific method and step will be introduced in next chapter, implementation part.

Fig. 4.7 Connection of Devices

Leap Motion capture the gesture of users, for example, when user would like to grab the object, the hand will be close to any parts above the surface of this object which lead to collision between hand and model, then, if the gesture of grabbing is detected,as Figure 4.8a, the position of model will follow the hand.

After the connection of devices, the gesture data from Leap Motion can be transmitted to smartphone, in this case, we can just use our hand to interact with 3D model through smartphone, as Figure 4.8b.



(a) Hand Gesture Detection (b) Transmit Data to Smartphone

Fig. 4.8 Gesture Control with Mobile Devices

4.2.2 Instructional Preview

In general AR preview system, users can observe the appearance of products, however, limitation as we mentioned are weak interactive experience and lack of information about usage. The first shortcoming we has solved by our function simulation module, and the second can be resolved by this part.

In this research, in order to make the acquisition of information more intuitive and enhance the instruction of product usage, based on traditional method, we add avatar and instruction UI(User Interface) as controller. The structure of instructional preview module almost like Figure 4.9 shows.

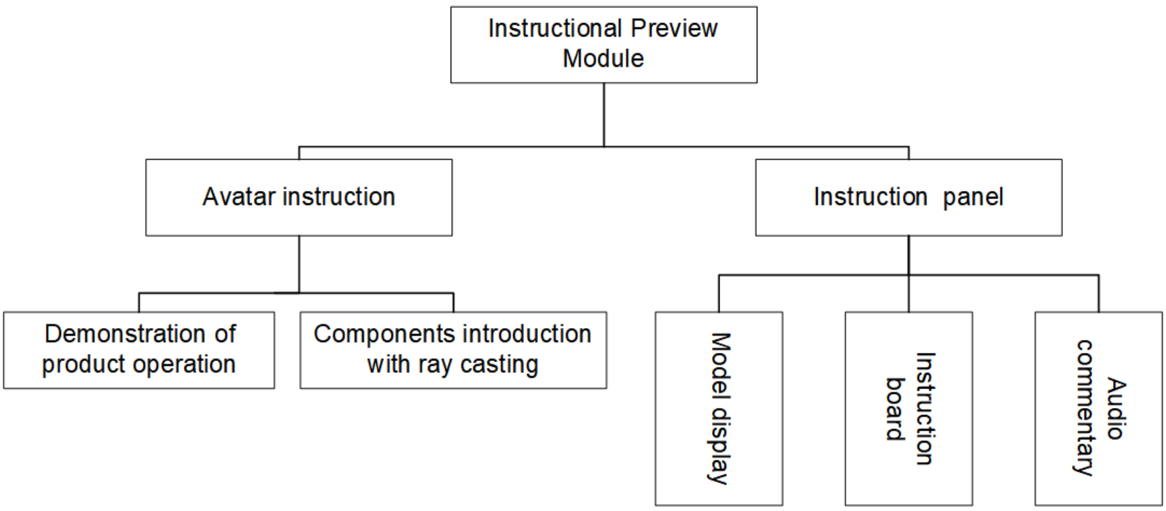
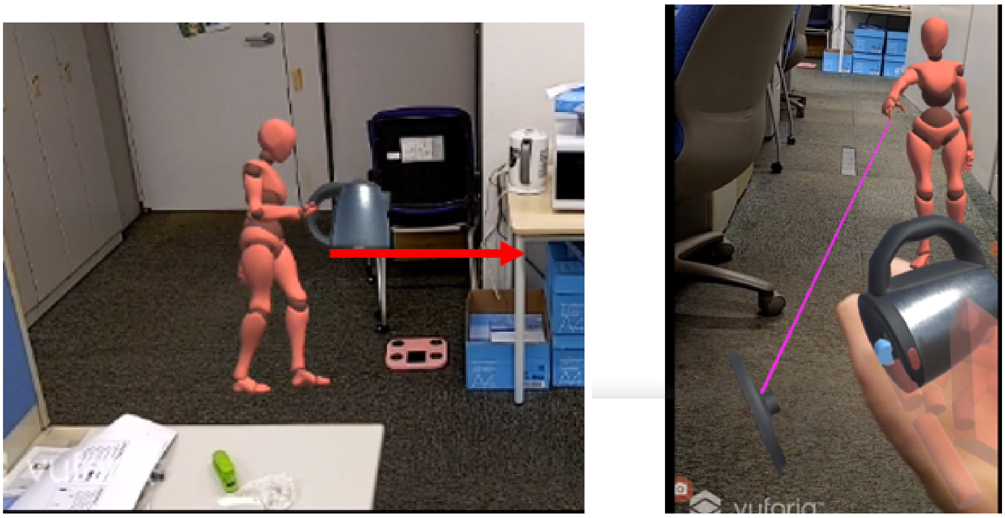


Fig. 4.9 Structure Diagram of Instructional Preview

Same as last section, we explain this using kettle case. After selecting the kettle as model, we need to click start button to enter case. It consists of following parts:

1. Avatar instruction. Avatar can demonstrate the usage and operation of products by animation, on the other hand, it can also show the position of triggers and introduce different components with ray casting.They are shown as Figure 4.10.



(a) Avatar grab the virtual kettle to other scenes and show how to boil water step by step

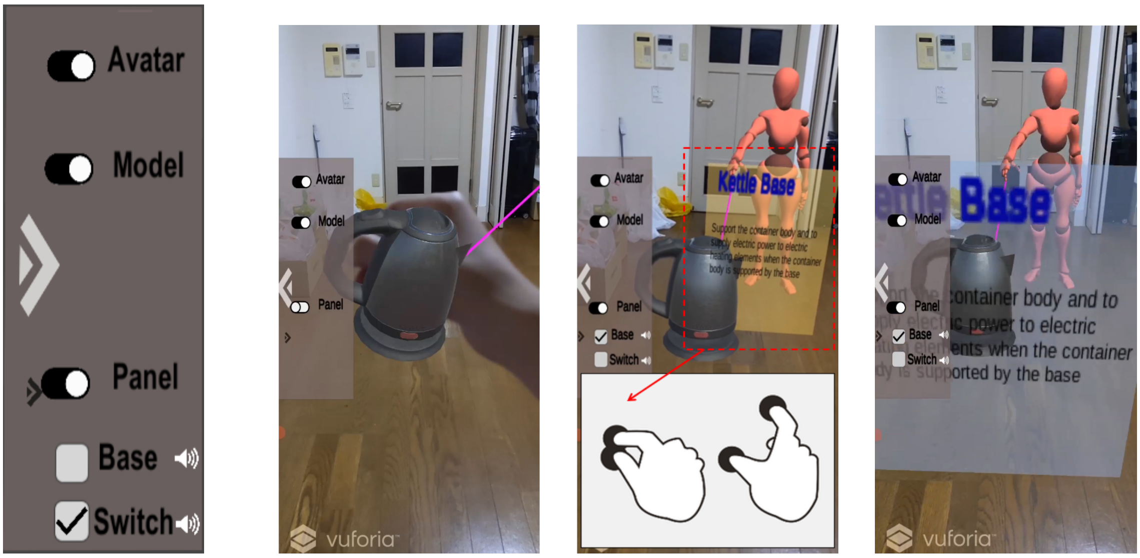
(b) Using ray casting to introduce

Fig. 4.10 Avatar Instruction

2. Instruction panel. It can be divided into two parts in general.

   First is instruction board, which is a text board about information corresponding to position/components pointed by avatar, meanwhile, if user think the display of avatar is redundant, they can move the position of avatar and even let it disappear, as Figure 4.11b . The size of board can be zoomed by swiping the screen with two fingers, as Figure 4.11a.

   The second is using UI panel, as Figure 4.11d. We can see hided window on the lift of screen, by opening this window, the control panel will display. In this panel, users can control the display of kettle, the introduction board, avatar and audio commentary(which is also related to position/components pointed by avatar.



(a) Hided Control Window (b) Hide the display of avatar and manipulate the kettle (c) Swiping the screen with two fingers (d) The size of board has be zoomed in

Fig. 4.11 Instruction Panel

In our research, the main purpose is to simulate the use process of some products and experience this process through AR, which is a technology that combines virtual information with reality. By means of AR, superimposing the daily life environment and virtual objects into the same space,we can experience functions with virtual products in real environments as if we were using actual products.

In order to apply more realistic usage experience, we should simulate the usage in real scenes according to specific product, for example, if user would like to experience the usage of kettle in real scene, obviously, it is not enough that system just use 3D model with animation. Therefore, we combine two methods to enhance the simulation immersion in real environment . The first is scene recognition which uses Vuforia "Area Target" function and object detection which uses Vuforia " Object Target". The framework is shown as Figure 4.12, and two parts play different roles respectively.

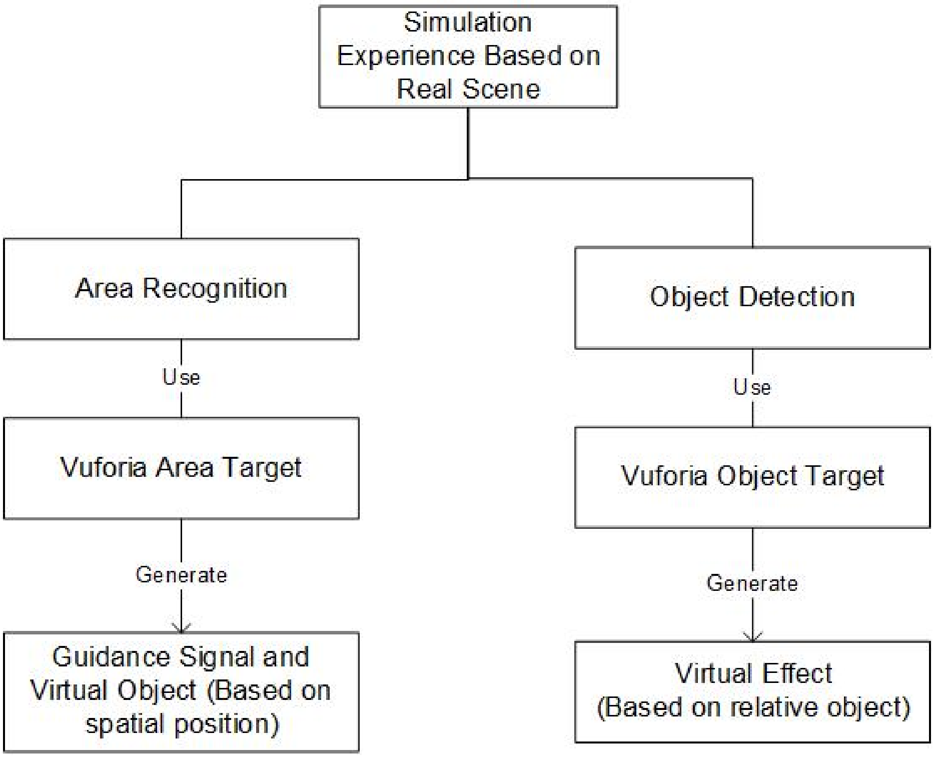


Fig. 4.12 Simulation Experience Based on Real Scene

Area Recognition

In this part, we merge real environment and virtual object by Vuforia Area Target which can identifying spatial information and display the virtual object in specific spatial position. To realize the function, first step is to use 3D spatial scanner which is provided by the official and create corresponding area database,as Figure 4.13 shows. After we install the corresponding version of Vuforia and import the model database package into project, the system will enable us to track and augment areas and spaces, adding virtual elements into real area.

Fig. 4.13 3D Model of Scanned Area

With the method of area target, users are able to use our system in a wider space and have more types of interactive experience not only between a single model and user themselves. The purpose can be divided into two parts:

- Generate guidance signal: As Figure 4.14 shows, guidance signal can reminder user where to go an what should be done while using product, in this case, it provides more intuitional instruction than voice or just text.

(a) Room A

(b) Room B

Fig. 4.14 Guidance Signals in Rooms

- Display virtual object related to product: Varieties of products need to be used in specific place, for example, we have to be close to somewhere with a faucet while using kettle to catch water. If everything including faucet is showed in the form of virtual model, lack of immersion will weaken usage experience and lower the interest in use. Therefore, as , we add virtual object into 3D area model based on real environment.

Fig. 4.15 Virtual Water Based on Position of Faucet

Object Detection

In this system, apart from area recognition, object detection is also necessary. Generally, area target is stationary, which means each things in scanned space need keeping the same position and not moving. Therefore, area target cannot detect one separate object in the room. However, in practical use, it is no doubt one product is always to be used with other separate objects inevitably. In this case, using object detection is a good substitute.

To detect object, steps what we should do is similar to achieve area target. First, we created a 3D model database for each object through Vuforia Object Scanner. Second, when scanned object in it is detected, the system will display related function simulation result, including corresponding virtual objects can be generated based on the detected object, or displaying relative effect around the detected object. Then, users can experience function process as using real one.

For instance, when we would like to boil water using kettle after catching water, the first thing what should be done is find socket and then connect kettle to power base. In this case, we scan a real socket and set it as target, then add virtual power base on it and vapor and the lighting red signal effect which means water has been boiled. The result almost like Figure 4.16.

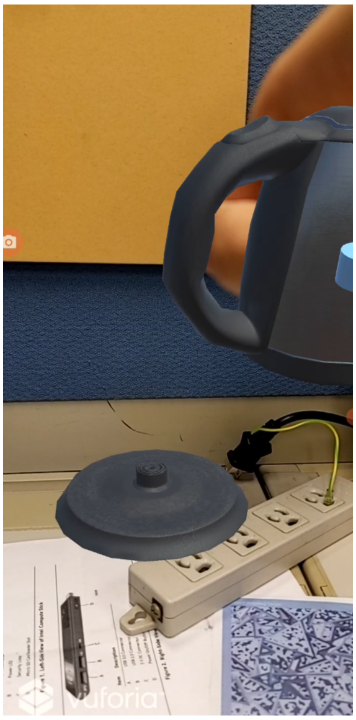


Fig. 4.16 The Result of Object Detection

4.3 Interaction Design

In order to enhance the experience of function experience with 3D model, abundant interaction is necessary for whole system. Therefore, how to design more realistic and direct interaction become one important core point.

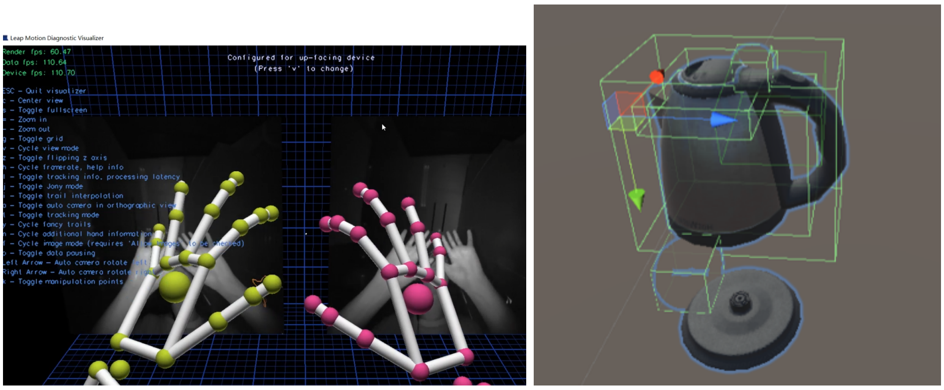
In this system , the type of interaction almost can be divided into two categories. The first is interaction between people and virtual object and the second is interaction between real environment and virtual object, through these two methods, corresponding function can be triggered or relative instruction can be shown to users. Following, I will introduce them respectively in detail.

4.3.1 Interaction between People and Virtual Objects

In this part, user can interact with virtual things by two ways. One is using hand to operate the virtual object or trigger some results directly and another is through the screen-touch of smartphone, by means of which, users can manipulate virtual elements with UI(User Interface) of our system and control events.

1. Trigger interactive events by hand.

   The core idea of this part is to use the "Hand Tracking" function of Leap Motion to directly obtain user's hand gesture and coordinates which we have mentioned in previous page. Generally, the hand tracking of Leap Motion and its interaction is based on banding hand skeleton(Figure 4.17a) which has collision volume in space. Then, we add relative Box Collide in Unity on 3D model(Figure 4.17b, position according to products) and give them corresponding events which are able to be triggered by collision. After that, we set hand skeleton as prerequisite of collision events. As the result, if the collision between them is detected, the events will be triggered.



(a) The Skeleton of Hand

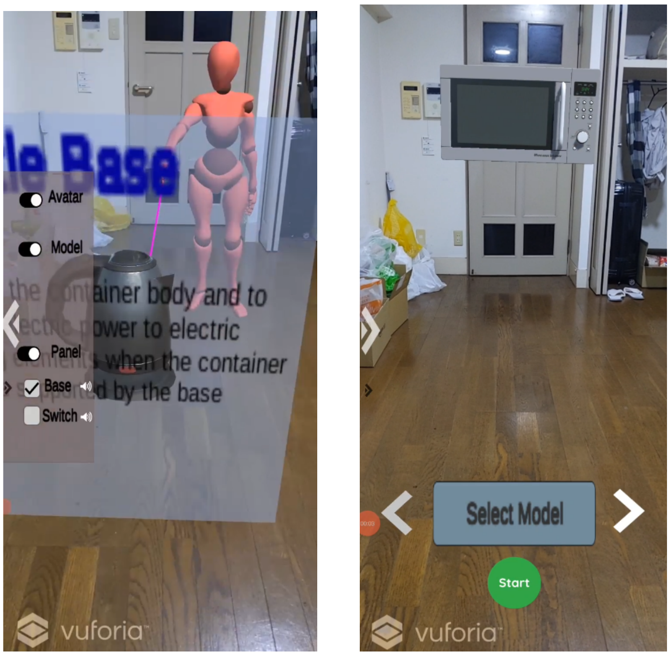
(b) 3D Model with Box Collide(Trigger preset events)

Fig. 4.17 Interaction between Users and Virtual Objects

For different products and in different scenes, we can add Box Collide to virtual object like object target, area target or even image target. Besides, about the relative event, it also can be in various types(effect, voice, etc.).

2. Trigger interactive events by UI.

   Comparing with Head Mounted Devices, one of advantage of using smartphone is can manipulate things by interface easily, and sometimes it can be a more convenient way to interact. In the case of this research, we also design interface to control not only interaction with various virtual objects, but also instructional voice and mode switching(Figure 4.18a & Figure 4.18b).



(a) Control Interface (b) Mode Switch Interface

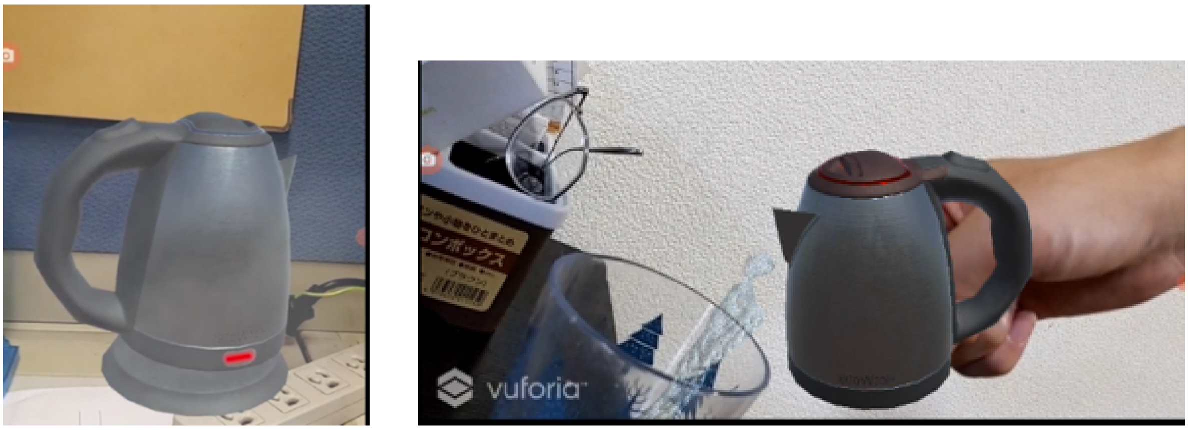
Fig. 4.18 Design of User Interface

4.3.2 Interaction between Real World and Virtual Objects

In order to simulate some electric appliances, just using one interaction is limited, in this case , we need to the second method, interaction between real world and virtual objects.

The interaction between real world and virtual object means the real object can be detected by system, at the same time, the virtual object can be displayed by AR. When they come into contact or detect each other, some events like new virtual object or effect can be triggered.

In the system, we have designed multiple this type of interaction, for instance, for simulating the process of boiling water and pouring hot water into a glass, system should detect the real object first. If the system detect the socket, the power base will display. While putting kettle on the base, the vapor effect, sound of boiling and red indicator light will be shown to user, as Figure 4.19a show. After that, if system detect the real glass cup, the user can pour hot water as Figure 4.19b show.



(a) Boil Water(with vapor effect, sound and red signal light)

(b) Pour Out Water

Fig. 4.19 Interaction between Real World and Virtual World

Chapter 5

System Implementation

5.1 Development Environment

This system uses the following interfaces and software for development:The Software Environment as following shown:

1. Unity 2020.3.22f1, it provides application development platform based on universal Windows platform and then system will be run on Pixel 4, a smartphone with depth camera, in which we can debug and test our functional modules.

2. Blender 3.0, a free open source 3D model generation software, in our research, we used Blender to create each components of model and add different animation or effect based on corresponding function requirement.

3. Vuforia 10.1.4, a software development kit which can be used in unity. Developers can easily add various AR content and add different interaction by using it. In addition, it can realize spatial recognition, object tracking and series of function effectively. We enhanced experience with Vurforia to combine virtual content and real world better.[25].

4. Android Studio, which is an Google's Android Integrated development tools. In our research, we used it as the basic environment for function development and debugging.

5. Orion 4.0, which is the framework provided by Ultraleaf company and can be used for Leap motion [13]. It can support 3D gesture tracking and basic data transmission. Here we combine it with trained hand model with a series of gesture to recognize user's hands and

track trained gesture, and then we transmit the captured hand gesture data by a prefab C++ WebSocket based on hot spot.

Additionally, we have used Visual Studio 2019 as program compiler. It is mainly about C scripting of unity, so as to debug and development. The versions of some of the software and services used are shown in Table 5.1.

Table 5.1 Software and Information

Software	Information
Unity	2020.3.22f1
Blender	Blender 2.93.4
Microsoft Visual Studio	Microsoft Visual Studio Community 2019 V16.8.5
Vuforia	Vuforia 10.1.4

5.2 3D Model Preprocessing

In this system, as we mentioned in last chapter, the 3D model we used should be dynamic, therefore, it is important that we need think about how to construct dynamic 3D model with functionality. The main step can be divided to following two part: model construction and function trigger.

5.2.1 Model Construction

In order to construct 3D model, here we used Blender and build each part respectively, and according to different products' function, model have different components which are dynamic.

For instance, as the Figure 5.1 shows, this is the kettle case, which includes six parts: base, bick, body, button, cap and handle. Each components can move separately on account of different usage, and we can make connection between two parts, like button and cap, when we click the button using our hand, the cap will do a rotation animation to simulate the cap opening used in catching water scenario.

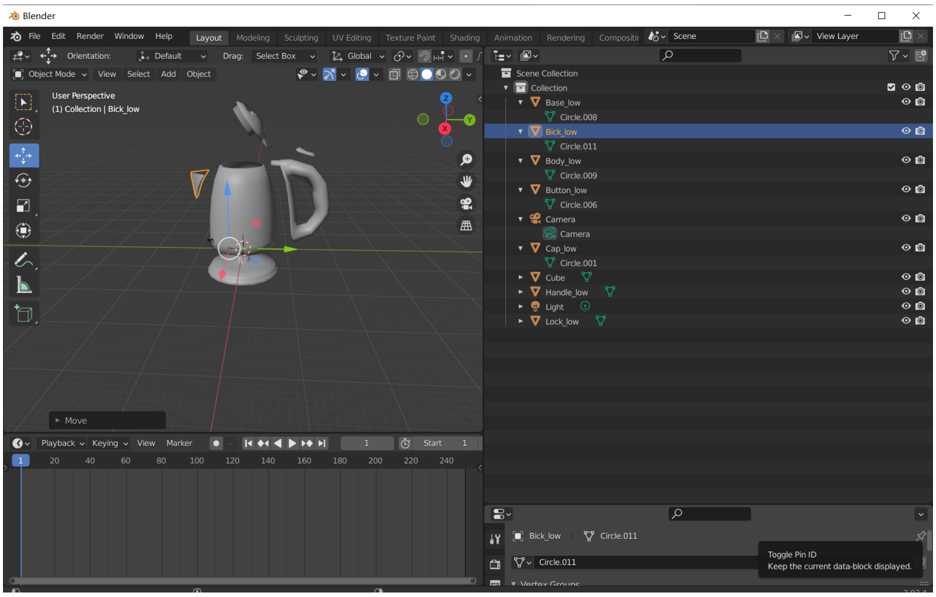


Fig. 5.1 Construction of 3D model

After constructing, to simulate function, we need add animation to each relative components of model, for instance, as the Figure 5.2 shows, it means add key frames for the door component of microwave. As the result, it seems almost the same as usage of the real one.

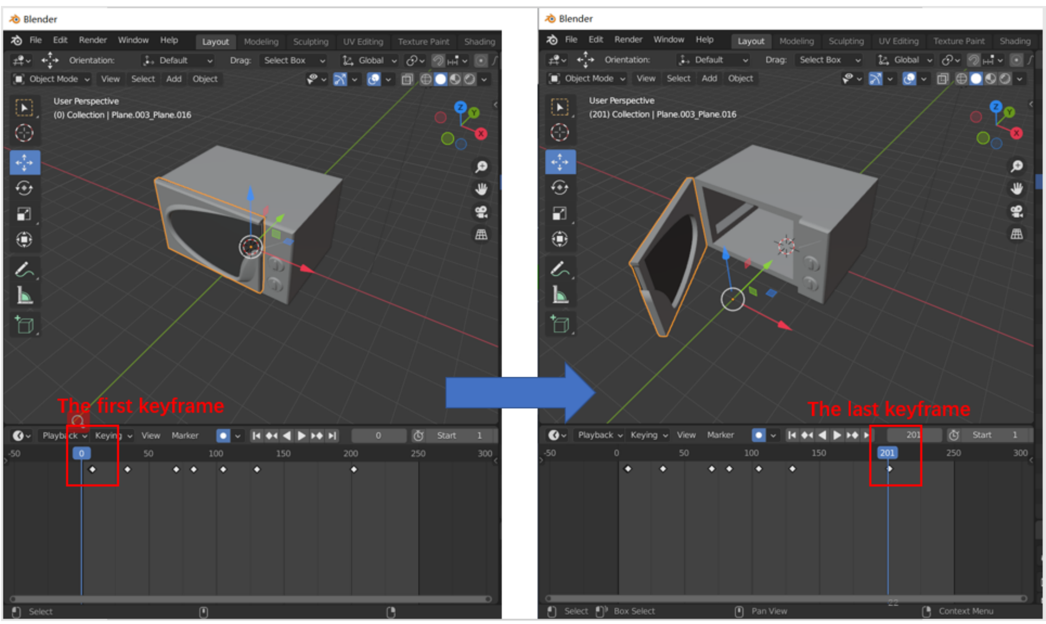


Fig. 5.2 Add key frames to components of model in Blender

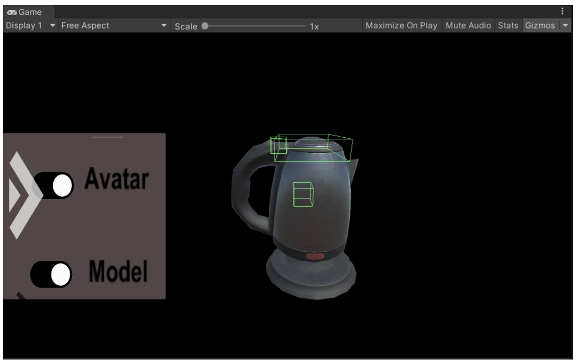
After animations have been added to models and export the model in form of ".fbx" which can be used directly in Unity and turn into next step.

5.2.2 Functional Trigger

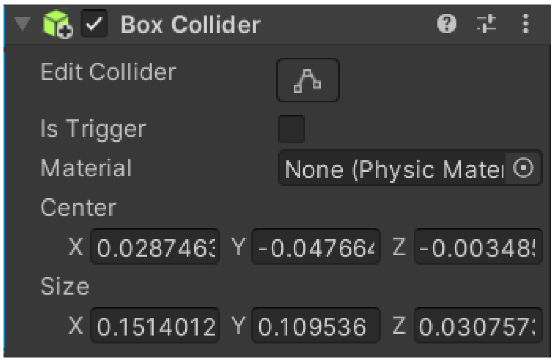
The first is to import the 3D model we have constructed in previous step into Unity, and in order to give them functionality, we should combine them with Collide component in Unity. As we introduced in last chapter, this is the most important process for our construction.

As the Figure 5.3 shows, there is a default function component called Box Collider, which give the model or the parts of model collision box framework. As the Figure 5.3a, we add the cap part, button part and base part with this component. In this case, when another 3D model with Box Collider contact with the kettle model, the collision can be detected.

The size and center of Box Collider are able to be adjusted, as the Figure 5.3b shows, if we would like to improve the sensitivity of collision detection, we can edit the size more expansively, on the contrary, we can narrow the range of box and change the center of it.



(a) Structure of Box Collider

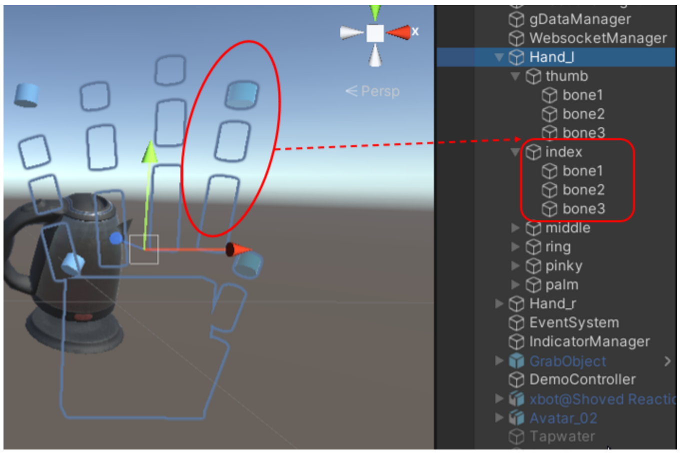


(b) Box Collider Component in Unity

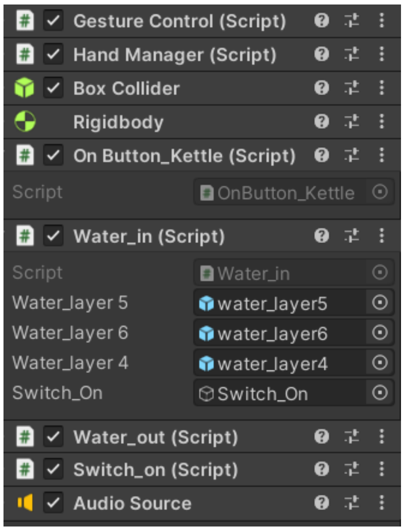
Fig. 5.3 Functional Trigger on Model

While the model has box collider already, we need use our hand as trigger and build relative events in scripts, which means when we touch the model with collider, the specific events can be triggered.

To demonstrate the processing of triggers, here we use the construction of kettle model in Unity as example. The Figure 5.4a shows hand model, which is related to the real-time movement of our hands. Same as kettle model, hand model also bind with both box collider and rigid body, and can detect the collision. Here, we set the index finger of lift hand as trigger. As Figure 5.4b shows, we added multiple functional scripts on hand model, when collision between hand and specific parts of model are detected, corresponding events can be shown to users. Besides, by the means of position calibration and give transparent material to hand(the display of virtual hand can be concealed), users can trigger the events directly by using their own hands rather than virtual hand, in which case the simulation experience must be immersive and emulational.



(a) Structure of Hand



(b) Different Functional Events using Scripts

Fig. 5.4 Functional Trigger on Hand

5.3 Connect Smartphone and Leap Motion

As usual, when users want to manipulate virtual through AR, they are always using AR headset devices like Google Hololens, however, to combine with more convenient and portable AR experience to simulate the function, we would like to find a method which make a connection between smartphone and Leap Motion.

Here, after finding related work which enable free hand-interactive smartphone AR, we decided to use their method as initial prototype: connect Leap motion, a tracking device, to a smartphone, in this way, we can interact with AR content by smartphone and hand directly instead of heavy desktop or even laptop. Meanwhile, the specific connection method is based on wireless network built in their system and hand tracking data can be transmitted by a built-in web socket.

5.3.1 Connection Framework

We begin with an Android smartphone with depth camera, in our case, using Google Pixel 4. The application has been developed and researched in unity, for display of AR content, it is the usage of Vuforia SDK to provide basic augmented reality asset parts.

As we know, the Leap Motion Controller is always used in computer device which preset with specific USB interface and sometimes it also be connected with AR or VR headset which also integrated with needed USB. However, it cannot connect to smartphone, to connect them, the structure can be generalized as Figure 5.5, treating a computer stick replace the laptop and make connection using normal universal serial bus(USB), and realize data transmitting using C++ WebSocket as server and hotspot as wireless method and about the wireless method, we will introduce in next section.

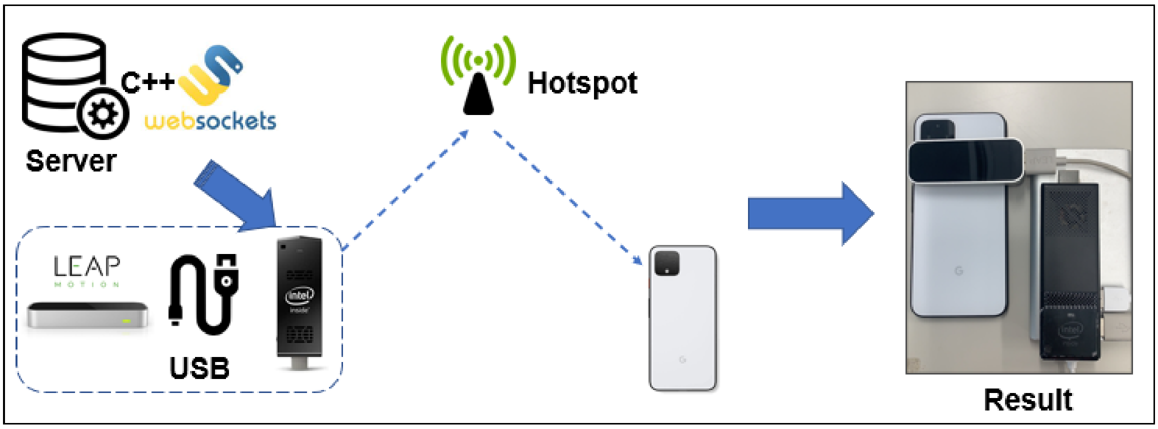


Fig. 5.5 Connection Framework

In order to get real-time position and movements of hands, we connect the Leap Motion to our Google pixel 4 and it is significant to calibrate the relative position between virtual 3D hand and real hand step by step. With this calibration, we can move our hand to control virtual object or interact with them simultaneously and fluently. Furthermore, in order to improve the accuracy of calibration, we have confirmed and ensured the relative position of smartphone camera and leap motion, in which way, motion of 3D hand can be closer to target object.

5.3.2 Wireless Data Transmission

As we mentioned, we transmit the captured hand tracking data by meansof utilizing C++ Web Socket as server, which includes series of trained hand gesture in advance. And then, by connecting with the compute unit, data can be sent to smartphone using hot spot, in the case of smartphone, a relative client will receive and stores this hand tracking data. The implementation of creation of server and setup of starting leap service almost as the Figure 5.6 shows which means the programming of creating server and the figure 5.7a shows the executable file of server which also includes detect the hand data and transmission of data. Every time we want to use this system, the first thing we should do is opening this server and starting the service.

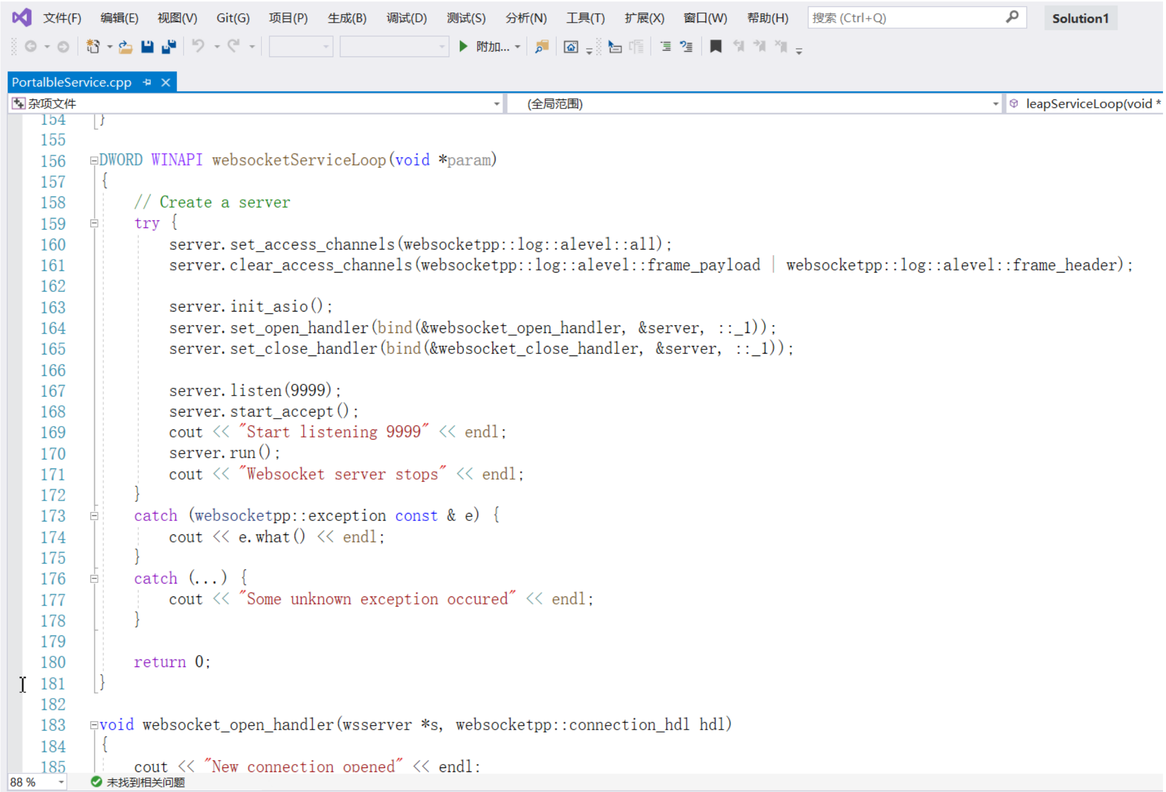
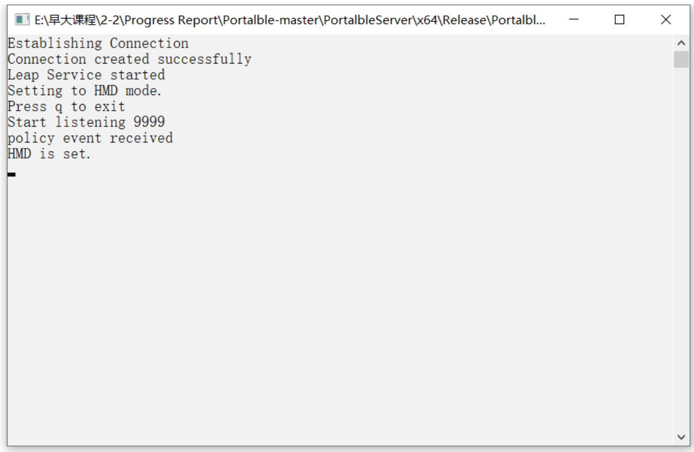
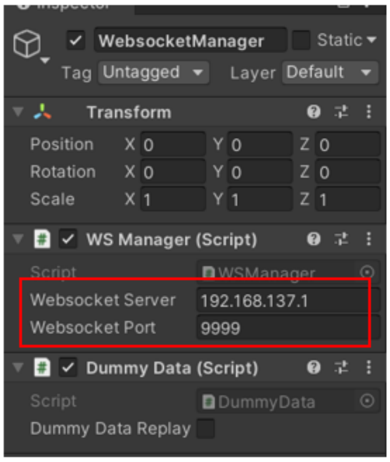


Fig. 5.6 Creation of Server

At the same time, we create a data transmission scripts in Unity, named "WS Manager" under WebsocketManager. As Figure 5.7 shows, after we open the executable file and start the leap service as Figure 5.7a, the smartphone can receive the data sent. Therefore, the WS Manager scripts is to provide the hand tracking data. The Figure 5.7b demonstrate that the gateway address write in WebSocket should be the IP address of hotspot, in this case, the data transmission from the compute stick to smartphone can be realized.



(a) Starting the Leap Service



(b) WS Manager in Unity

Fig. 5.7 Build the Link

5.4 Recognition Real Object and Real Scene to Realize Spatial Interaction

In our system design, we are committed to build a simulation system that users can pre-experience the usage and function of various products by manipulating 3D model and what is more, whether users themselves or 3D model can generate different interaction in real spatial scene. In order to this module, we used Vuforia object target and Vuforia area target to realize spatial interaction in different situation.

5.4.1 Scanning of 3D Object and Using Object Target

As we explained in system design part, the 3D object recognition is to realize the interaction with flexible gadgets, therefore we need to use Vuforia object target function module. By using the official scanner app and scan different object target, we can achieve needed database.

First, we should install and use the Vuforia object scanner app [25] , which can scan one object in tiny size and generate relative data files. The app can be downloaded from Vuforia developer library in the official website.

Then, in order to scan the target object, we should download the Object Scanning Target Image, as Figure 5.8 .Then, print it in appropriate size, in which the area of non-patterned fit to the size of target object. In this case, we can put the object that we would like to set as the target on the Object Scanning Target Image and use Vuforia object scanner to scan the target object around. Then we can get the file in form of ".od". as the Figure 5.9a shows. If we want to test whether it is successful to recognize the target object, the app also can provide the test function, as Figure 5.9b shows, the green column means success.

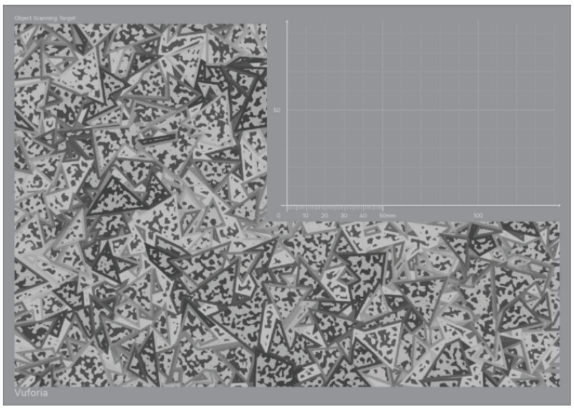
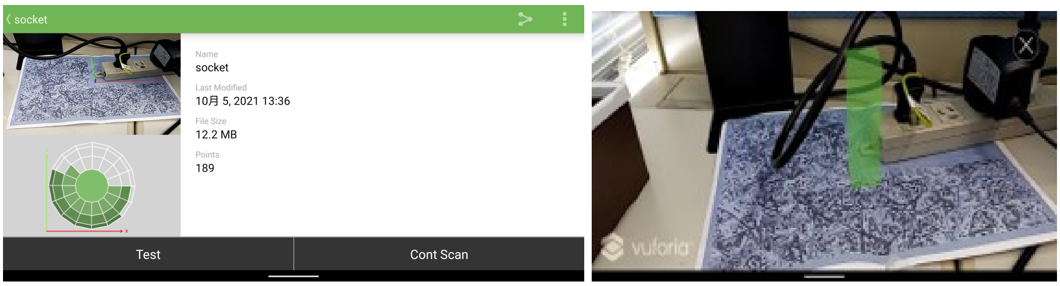


Fig. 5.8 Object Scanning Target Image



(a) ".od" File of Object Target

(b) Vuforia Test for Target Object

Fig. 5.9 Generated File & Test

After that, what we have done is using the object target in Unity. By the way, the file in form of ".od" cannot be used in Unity directly, in this case, what we should do is almost like following steps:

1. Transmit the ".od" file to computer terminal and upload it to the Target Manager of Vuforia official website, as the Figure 5.10 shows. What should be noticed is the target type must be 3D Object.

2. Achieve the new Unity package file in Vuforia website, as the Figure 5.11a shows, we need select the development platform in Unity Editor form. Then download and import it into Unity, the file like the Figure 5.11b shows.

3. Add the Object Target in Hierarchy(Have installed the Vuforia SDK in advanced) and select the database of object target named "socket" which is we generated previously, as the Figure 5.12.

4. Add the events under the Object Target, and this is based on the type of model, in the case of socket, we give it function collider to trigger boiling water(specific operation has introduced in 3D model preprocessing part)

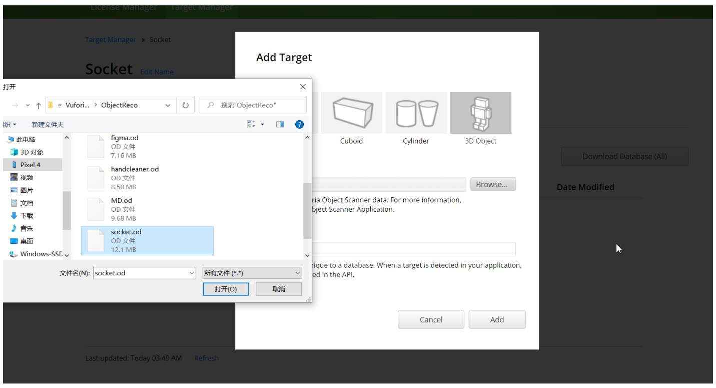
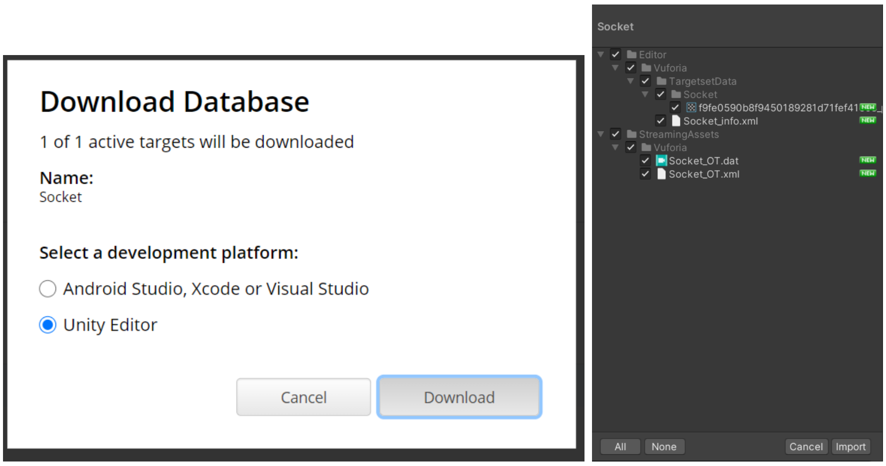


Fig. 5.10 Generate Unity Package File



(a) Download Unity Package File (b) Import Package into Unity

Fig. 5.11 Package Download & Import

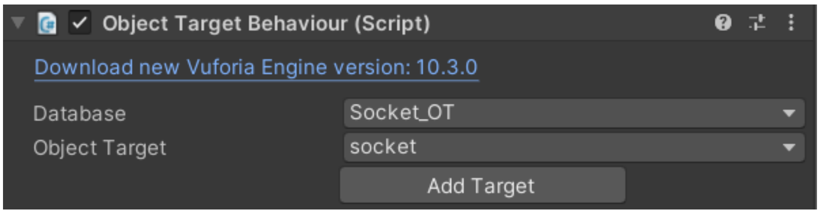


Fig. 5.12 Select Target Database

5.4.2 Scanning Real Scene and Using Area Target

Apart from object target, we also used area target. As we mentioned in last part, the 3D object recognition is to realize the interaction with flexible gadgets, but it cannot support the interaction with both fixed or large object and environment. Thus, solving this problem is the purpose of recognizing the real scene.

As prerequisites, something need to be noticed. Unity Editor had better be up to version 2020.3.0+ and Area Targets need be generated later than version 9.7 of Vuforia Engine. Otherwise, it may occur some display errors or even cannot show AR objects.

We need to use Vuforia Area Target which enable users to custom different AR contents and interaction in the range of area as augmented space. To implement it, the processing we should do like following shows:

1. Scanning the environment. First we should prepare scanner, there are several official recommendations as Figure 5.13 shows, among them, we choose use the first one, iPad with inbuilt LiDAR sensors which enabled ARKit.

Vuforia Area Targets supports scans made with:

- **ARKit enabled devices with inbuilt LiDAR sensors,**
- **Matterport™ Pro2 3D camera,**
- **NavVis M6 and VLX scanners, and**
- **Leica BLK360 and RTC360 scanners.**

Fig. 5.13 Official Recommended Devices

After that, we need to install the Vuforia Area Target Creator App[26], which supports creating small Area Targets on iOS devices with LiDAR sensor. We used this app scan different rooms and generate the area target file in iPad, as Figure 5.14 shows.

Fig. 5.14 Scanning environment using Vuforia Area Target Creator

2. Creating Area Target. After scanning, we imported the file of room model into our laptop, differ from object target, this file can be directly used by Unity. In unity, we can see 3D model of scanned room and in this space, we can add multiple components.

In our research, we added directing signal and prompt tag in as area target, as Figure 5.15 shows. Besides, in order to simulate the flowing water, we added audio source with gradual change in location of faucet. When we go around in real scene, we can see these signals and if we wall to faucet, the sound of flowing water can be louder and louder.

3. Develop with Area Target. For developing more interaction, we also add collider in area target, for instance, we can give a transparent cube on faucet, in this case, users can turn on or turn off it by triggering collision.

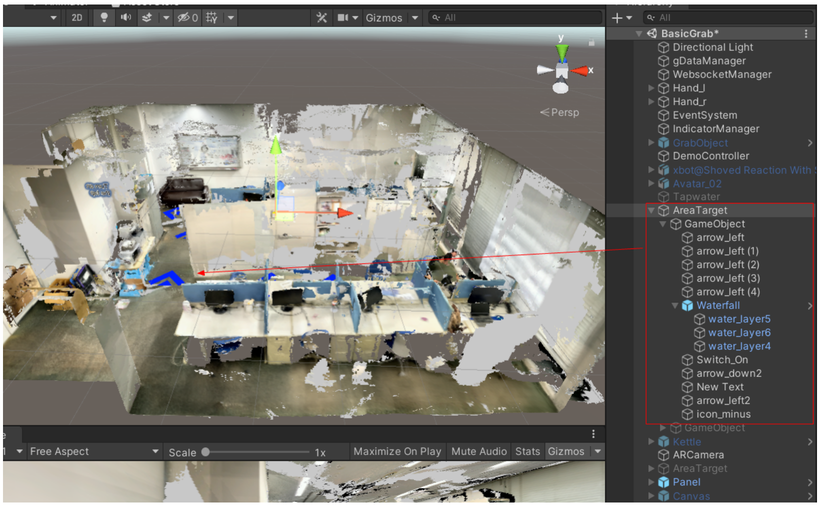


Fig. 5.15 Creating Area Target

5.5 User Interface and Avatar

About instructional preview module, the main implementation method is through combination of UI (User Interface) and Avatar. UI provides a controller of most AR object including product, instructional board, avatar and so on.

As Figure 5.16 shows, we almost used two UI components: buttons and toggles. Buttons are used to control display of object and decide certain events, like "next" and "prev" object can select which model display. Toggles are used to select just one situation from juxtaposed options, like "bgroup" and "bgroup_s" are two toggles, but we just can select one of them to show different text in instructional text board.

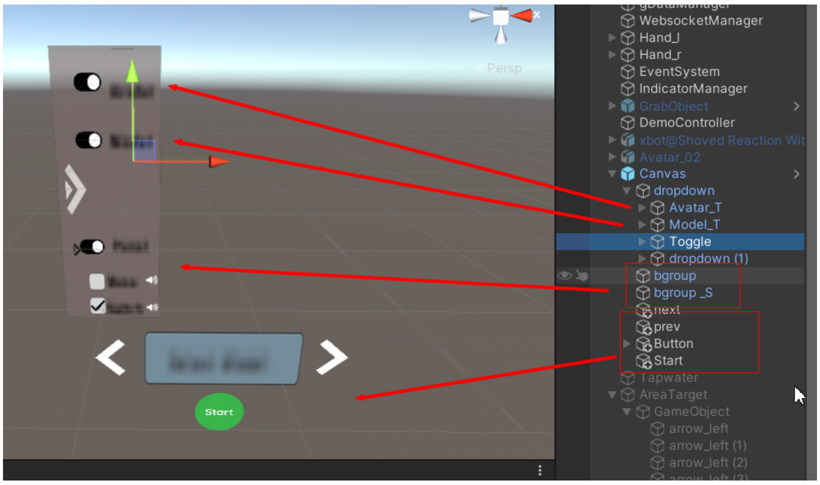
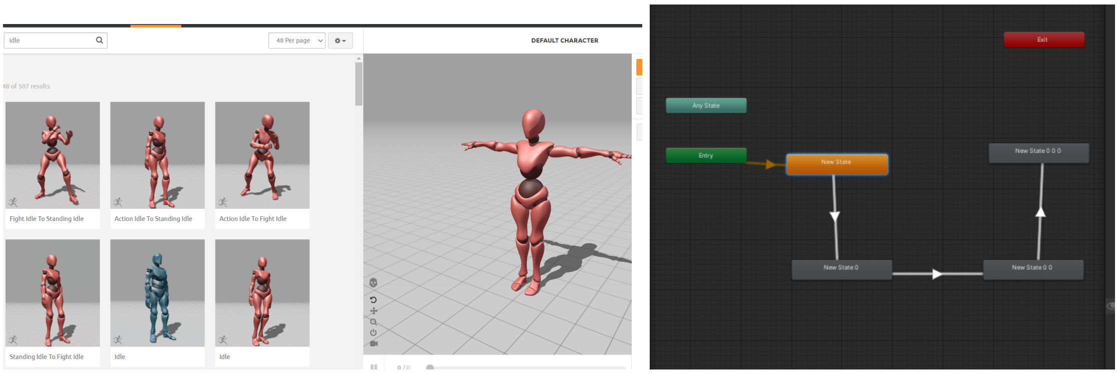


Fig. 5.16 UI Buttons and Toggles

On the other hand, avatar supports instruction is set as instructor, users can obtain usage demonstration and guide of products components from it. The implementation of avatar can be divided into two parts: construction of avatar and the red ray-casting from avatar.

- Construction of avatar

  Here we used Mixamo, a website including thousands of full-body 3D model with series of body action animation, where you can freely combine different actions and assign them to the avatar. To simulate the manipulation of real person, we select the action the nearest to actual use of the case. As the Figure 5.17a shows, there are multiple action can be downloaded and the Figure 5.17b shows that we used animator in Unity to combine a series of actions and build a new group to give certain action we would like to use.



(a) Obtain Avatar from Mixamo (b) Animator of Avatar

Fig. 5.17 Construction of Avatar

- Red Ray-casting from Avatar

  To instruct users usage of each part of products, we used red ray to point the specific position of model.To implement this, we Line Rinder in Unity and realize ray-casting function used scripts as the Figure 5.18.

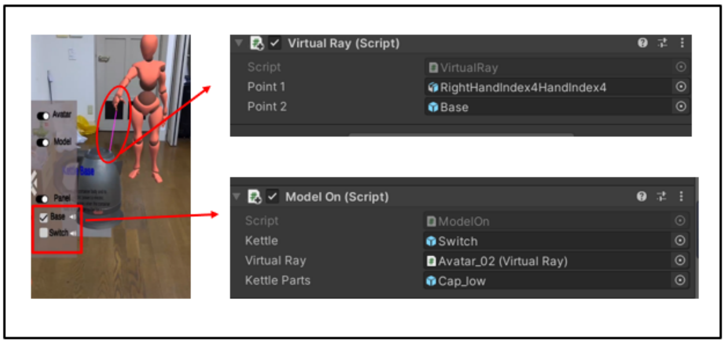


Fig. 5.18 Red Ray-casting from Avatar

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this research, we developed a AR product preview system which allows users to experience function simulation process and provides users with intuitive instruction simultaneously.

We designed our system module based on conventional AR e-commerce system, apart from basic preview function, users can experience different function according to corresponding products using their own hands, at the same time they can view the instruction and follow steps. Totally, we divided our system into two parts: function simulation and instructional preview.

In function simulation part, in order to provide immersive and emulational function simulation experience, we used smartphone as AR elements display device and combined with Leap-Motion as tracking device. We constructed 3d model with functionality by adding animation triggers on related components of product model. Then, We combine with various scene information, in which users not only can interact with 3d model of product, but also other elements according to different products. We used Vuforia to recognize real object and detect spatial information in different scenes. In addition, in recognized real environment, we set multiple triggers to enhance the interaction among users, virtual content and real world.

Besides, about instructional preview part, we enhance the comprehensive information in three dimension environment and make instruction more intuitive and attractive by incorporating avatar and presenting the traditional picture and text mode in AR. In order to instruct how to use the product and provide detailed information of product, avatar has been bound with series of animation about usage step. Then, we designed red-ray to build connection between avatar and specific part of product model, and add 3d instruction board in mid air and voice explanation. Through adding multiple control button in user interface, users can control all above elements, in such case, they can manipulate the product and check the instructional information at the same time.

In summary, we designed and implemented a function simulation preview system with instruction based on augmented reality. We proposed 3d model with functionality which can trigger function events, by means of connection of smartphone and tracking device. Meanwhile, users can be provided instruction information and follow the instruction to interact with something in real scene.

6.2 Future Work

The main idea of our system is to realize the function simulation , so as to develop and comprehend, we used kettle as use case to explain what have done in this research. We have added other two cases: vacuum cleaners and microwave, in future stage, we can perfect these two cases and plan to develop more product case in different scenarios.

Also, cause of connecting smartphone and leap-motion by wireless hot-spot, we plan to optimize the network condition to solve the video freezing or delay.

References

[1] Minjung Park and Jungmin Yoo. Effects of perceived interactivity of augmented reality on consumer responses: A mental imagery perspective. *Journal of Retailing and Consumer Services*, page 101912, 2020.

[2] Jeff KT Tang, Wan-Man Lau, Kwun-Kit Chan, and Kwok-Ho To. Ar interior designer: Automatic furniture arrangement using spatial and functional relationships. In *2014 International Conference on Virtual Systems & Multimedia (VSMM)*, pages 345–352. IEEE, 2014.

[3] amazon App. webpage-amazonarview. https://www.amazon.com/adlp/arview, 2021.

[4] ikeaApp. webpage-say hej to ikea place. https://www.ikea.com/au/en/customer-service/mobile-apps/say-hej-to-ikea-place-pub1f8af050, 2021.

[5] Ronald T Azuma. A survey of augmented reality. *Presence: teleoperators & virtual environments*, pages 355–385, 1997.

[6] Steve Mann and Feiner. Wearable computing, 3d augmented reality, photographic/videographic gesture sensing, and veillance. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, pages 497–500, 2015.

[7] Ronald Azuma, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon Julier, and Blair MacIntyre. Recent advances in augmented reality. *IEEE computer graphics and applications*, pages 34–47, 2001.

[8] wikitude. blog-3d-content-for-augmented-reality. https://www.wikitude.com/blog-3d-content-for-augmented-reality/, 2021.

[9] Yueting Weng, Chun Yu, Yingtian Shi, Yuhang Zhao, Yukang Yan, and Yuanchun Shi. Facesight: Enabling hand-to-face gesture interaction on ar glasses with a downward-facing camera vision. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021.

[10] Kohki Ikeuchi, Tomoaki Otsuka, Akihito Yoshii, Mizuki Sakamoto, and Tatsuo Nakajima. Kinecdrone: enhancing somatic sensation to fly in the sky with kinect and ar. drone. In *Proceedings of the 5th Augmented Human International Conference*, pages 1–2, 2014.

[11] Huidong Bai, Gun Lee, and Mark Billinghurst. Using 3d hand gestures and touch input for wearable ar interaction. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, pages 1321–1326. 2014.

[12] Jean Botev, Joe Mayer, and Steffen Rothkugel. Immersive mixed reality object interaction for collaborative context-aware mobile training and exploration. In *Proceedings of the 11th ACM Workshop on Immersive Mixed and Virtual Environment Systems*, pages 4–9, 2019.

[13] Jing Qian, Jiaju Ma, Xiangyu Li, Benjamin Attal, Haoming Lai, James Tompkin, John F Hughes, and Jeff Huang. Portal-ble: Intuitive free-hand manipulation in unbounded smartphone-based augmented reality. In *UIST'19 Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pages 133–145, 2019.

[14] Nicolas Imbert, Frederic Vignat, Charlee Kaewrat, and Poonpong Boonbrahm. Adding physical properties to 3d models in augmented reality for realistic interactions experiments. *Procedia Computer Science*, pages 364–369, 2013.

[15] Ann Marie Fiore, Jihyun Kim, and Hyun-Hwa Lee. Effect of image interactivity technology on consumer responses toward the online retailer. *Journal of Interactive Marketing*, pages 38–53, 2005.

[16] Supun Hewawalpita and Indika Perera. Effect of 3d product presentation on consumer preference in e-commerce. In *2017 Moratuwa Engineering Research Conference (MERCon)*, pages 485–490, 2017.

[17] Franka Moritz. Potentials of 3d-web-applications in e-commerce-study about the impact of 3d-product-presentations. In *2010 IEEE/ACIS 9th International Conference on Computer and Information Science*, pages 307–314. IEEE, 2010.

[18] Waraporn Viyanon, Thanadon Songsuittipong, Phattarika Piyapaisarn, and Suwanun Sudchid. Ar furniture: Integrating augmented reality technology to enhance interior design using marker and markerless tracking. In *Proceedings of the 2nd International Conference on Intelligent Information Processing*, IIP'17, pages 1–7, New York, NY, USA, 2017. Association for Computing Machinery.

[19] Carlos Alves and José Luís Reis. The intention to use e-commerce using augmented reality-the case of ikea place. In *International Conference on Information Technology & Systems*, pages 114–123. Springer, 2020.

[20] Yashvi Desai, Naisha Shah, Vrushali Shah, Prasenjit Bhavathankar, and Kaisar Katchi. Markerless augmented reality based application for e-commerce to visualise 3d content. In *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, pages 756–760, 2021.

[21] MA Ruichen. A mixed reality shopping system providing hand scale preview and gesture interaction. pages 3–29, 2020.

[22] Raffaele De Amicis, Alessandro Ceruti, Daniela Francia, Leonardo Frizziero, and Bruno Simões. Augmented reality for virtual user manual. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, pages 689–697, 2018.

[23] Michele Fiorentino, Antonio E Uva, Michele Gattullo, Saverio Debernardis, and Giuseppe Monno. Augmented reality on large screen for interactive maintenance instructions. *Computers in Industry*, pages 270–278, 2014.

[24] ultraleap. webpage-ultraleap. https://www.ultraleap.com/, 2021.

[25] Vuforia. webpage-vuforia. https://library.vuforia.com/articles/Training/Vuforia-Object-Scanner-Users-Guide.html, 2021.

[26] VuforiaArea. webpage-vuforiaarea. https://library.vuforia.com/features/environments/area-targets.html, 2021.