Kicking Together: A Remote Football Experience With Interactive Feedback Using Augmented Reality

Qi XU
44191547-3

Master of Engineering

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

January 2022

Abstract

In the past two years, COVID-19 has been in a severe situation in most places. It is very difficult for us to go out to play or exercise with friends, which is very dangerous. Many people choose to exercise at home or give up exercise. As one of the most popular sports in the world, football is suitable for friends to play together. If it loses the company of friends, it will become very boring. With the development of technology, many sports applications that can be experienced without going out have brought new methods to people. And the existing football experience at home still faces some challenges, it lacks motivation, accompanying and intuitive feedback, which will have much influence about people football experience.

In this paper, we design an immersive remote football experience system in order to enable two remote people to play football with each other. At the same time, the system will give different feedback according to the user's actions when kicking the football. We mainly focused on using augmented reality to bring users a better remote football experience, including simulated football and interaction with the real environment. With the development of AR technology, people can see virtual objects wearing HMD and interact with them in the physical world. And we chose HoloLens2 as the main equipment for the research. It can display virtual football well and allow users to interact with it. The system provides personalized avatars for users to enhance the sense of company users get from friends. In order to make the virtual personalized avatar more realistic, it is based on the specific information of the user's face and body, so that each user has his own unique virtual avatar. We also use the kinect to capture the user's actions and analyzes them, thereby simulating the possible trajectory of the virtual football.

Keywords: augmented reality, personalized avatar, remote football, feedback

Acknowledgements

First of all, I want to express heartfelt thanks and gratitude to my supervisor, Prof. Jiro Tanaka. He gave me great help and guidance during my two years at IPLAB. He is very humorous but also has a rigorous and responsible academic attitude. He led me into HCI field and augmented reality field, and patiently guided my research and gave me a lot of valuable advice. I am very honor to be the student of Prof. Jiro Tanaka.

And I want to thank other members of our lab. We had a difficult time during COVID-19 epidemic. I am fortunate to have their companionship and encouragement. At the same time, we discuss with each other in research, give inspiration and help, so that I can complete my research well. I think we are not only comrades-in-arms who are fighting together in research now, but also friends in the future.

Finally, I would like to thank my family members. They gave me financial and spiritual support in my study and life. Their encouragement became the motivation for me to continue my research.

Contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Introduction

In the past two years, COVID-19 has been in a severe situation in most places. It is very difficult for us to go out to play or exercise with friends, which is very dangerous[1]. Many people choose to exercise at home or give up exercise. As one of the most popular sports in the world, football is suitable for friends to play together. If it loses the company of friends, it will become very boring. With the development of technology, many sports applications that can be experienced without going out have brought new methods to people. For example, in the video game industry, people can control virtual characters to play football through gamepads or computers. But this lacks a sense of real experience. People cannot play football through their bodies. However, they control virtual characters with their hands. There is also the emerging VR football, which allows people to physically control virtual characters immersively. However, the lack of contact with the real world and the inability to see familiar faces of friends can make users a little lonely.

In order not only to bring users a sense of immersion and to be able to connect with the physical world, but also to make physical contact with football, we tried to use AR technology[2]. Due to the versatility of AR, it can be seen in various fields, including education, art, remote communication, etc. First of all, AR technology can bring users a great sense of immersion and can interact with the real environment, which is very helpful for long-distance football systems[3].Secondly, AR technology can be combined with motion capture technology to track and analyze user actions through some additional devices[4], which can well realize the interaction between the user and the virtual football.

Microsoft's HoloLens2 is a head-mounted Mixed Reality[5] device that can be used as the core device of our system. On this basis, our remote football system mainly has the following functions:

- An interactive virtual football field is provided for users through space recognition.
- The system provides personalized full body avatars as virtual players to kick football. At the same time, it can also act as a virtual referee and audience, giving users feedback.
- Multi-user playing is provided so that users can play football with remote friends.
- By capturing the user's actions, the trajectory of the virtual football movement is simulated.

1.2 Organization of the Thesis

The rest of the thesis is organized as follows: Chapter 2 mainly introduces the background of this thesis. Chapter 3 introduces some related work. Chapter 4 will briefly describe our research goals and research methods. Chapter 5 will introduce in detail the content of the system design about our research. Chapter 6 shows the system implementation details, including some equipment and software. In Chapter 7, we will make a summary and talk about the future work.

Chapter 2

Background

2.1 Football Sports

Football has always been widely welcomed by people all over the world. It is not only a competitive sport, but also a good physical exercise. By cooperating with your partner's team to score goals, or to disrupt the opponent's offense, you can well promote the feelings between friends.However, due to the escalating COVID-19[6], many stadiums have been closed. People can't go out and regain their happiness on the court with friends.

People's lives are gradually moving towards informatization and digitization. The same is true for sports. For the 2.0 era of sports[7] , it is not only sports competition on the real world court, but also video game sports. The virtual reality equipment also greatly enhances the possibility of sports competition in the virtual world. Many football applications are also emerging.

2.1.1 Exsiting Football Application

People often play a football game on the table in their spare time. In recent years, there is an augmented reality table football[8]. Although it is augmented reality, it only projects some of the icons on the desktop. Players control the above dolls by controlling the club to pass and shoot. This is also a multiplayer football game that requires cooperation between players. It tests not only hand-brain coordination but also reaction speed. Although it is only at a table, it is very competitive.

Esports is becoming more and more popular, especially during this epidemic. Compared with outdoor sports, more and more people choose esports[9]. For example, FIFA2021[10], a video game in which users can control virtual players through a keyboard, mouse or gamepad, and compete with other players in the virtual stadium.It can give users as much space as possible, including passing, shooting, header, chest stop and other skills can be achieved. But it lacks the sense of immersion on the court. The 2D screen cannot give players the immersive fun of playing football, and can only operate virtual characters with their hands.

The increasing development of virtual reality equipment can greatly improve people's immersion when using it. VR football[11] is used by many players to get the experience of playing football alone at home. At the same time, it is also used by many professional football teams to train players' abilities and data analysis.But the disadvantage is that it can't connect with the real physical world, and playing football alone in the virtual world will feel lonely. At the same time, you can't get the company of your friends, and you can't feel the joy of multiplayer football. Although a small number of VR football games support multiple players, players still see virtual characters and cannot get a good impression of seeing friends' faces. The figure 2.1 shows the three types of football mentioned above.

Fig. 2.1 Exsiting football application

2.2 Augmented Reality

2.2.1 Augmented Reality

Although virtual reality technology has started a good practice in game applications. By increasing the interaction with users or device movements, applications have been robustly and continuously produced and reached more user experience goals. However, augmented reality [12]technology is also widely used in people's life, work and entertainment. For example, Google[13] provides many AR application services. For example, you can directly use the mobile phone camera to point at the text content you want to search in the real physical world to display related content. And it can directly place the 3D digital object in the real physical space where the user is. AR navigation can display 3D guidance arrows in the real world according to the destination the user needs to go to. The application of AR technology in games is also becoming increasingly popular, especially pokémon go[14].

2.2.2 Avatar in AR

Many augmented reality systems use 3D virtual models to improve user experience, including some applications that use virtual pets to give users a sense of companionship, such as virtual dogs[15] walking with users who are inconvenient to raise real dogs.

Similarly, avatars are also an important part of the virtual world. The virtual avatar is used to imitate the user's appearance and achieve interactive simulation behavior. People will see the captured 3D representation of their own body, and they will naturally control the real body movement in the context of the current real environment[16]. In the Augmented Reality Remote Collaboration system[17], avatars are often seen.. Many online meetings based on augmented reality use personalized avatars as participating members to participate in virtual meetings.

Chapter 3

Related Work

In this chapter we will introduce the related work of this research. First, we will introduce the research on football video sports, including playing football on the computer, using VR devices, and using projection experience. Then we will introduce the work related to body motion capture. Our research is mainly related to tracking the user's feet. Finally, we will introduce the related work of virtual avatar.

3.1 Virtual Sports Experience

In recent years, many sports have gradually developed into games, and users can experience sports through some equipment or technology indoors.

With the development of science and technology, more and more high-tech products have entered our lives, including sports. Rakuten[18] proposed a 2d video football experience. Through the projector, users can play football with virtual stars. this application allows the user to interact with virtual characters on the big screen, and the virtual football is also 2D. Its focus is to detect the user's foot and the 2d football position projected by the laser beam, as well as the use of pre-recorded video.

As VR technology matures, more and more fields introduce this technology. Fujun Wu[19]It introduces the application and development of "VR + sports".The main advantage is that it can give users an "immersive" experience, allowing users to experience the feeling of the sports field indoors. Deutsche Telekom[20] allows users to play penalty kicks with famous players. Other participants can watch from different angles, including playback, etc. through the headset. At the same time, the goal will also move with the number of goals scored.

AR (augmented reality) has become more and more popular in recent years, and many sports are integrated with AR, allowing users to experience virtual sports in a real environment, or to exercise better through AR. Tica Lin et al. [21] mainly proposed that immersive sports training can help athletes better by visualizing training content, which includes AR-based sports training systems. They also mainly analyze the shooting trajectory of different basketball actions through AR.

Tim Buckers et al. [22]proposed a multiplayer augmented reality sports game that uses mechanisms similar to video games to stimulate players to perform physical activities. It mainly realizes face-to-face dodgeball game through two Microsoft HoloLenses.Patrick Baudisch et al. [23] proposed an augmented reality game including basketball or football, where the ball is virtual. Markers are on the top of the user's head and the basket, and the system mainly determines position of users and the movement of virtual balls by identifying the markers.

3.2 AR/VR Remote System

In this digital world, more and more applications that allow people to connect at a distance include some online applications like video conferencing. But people lack an intuitive experience, 2D video lacks a traditional face-to-face communication feeling, and it is difficult to have some visual interactions.

Many AR or VR applications have been developed to give people a better remote experience. Sebastian Günther et al.[24] combines 3D printed objects with AR, which can be shared remotely through recognition on their surfaces. For example, the article presents a virtual remote chess example. In recent years, AR is also widely used in the video game field. Many capture the user's movements through a camera, thereby controlling virtual characters or objects. A Tele-Table system[25] which is an AR remote system allows users to interact with real objects remotely through this system.

Some applications that integrate the augmented interface iexisting remote communication systems such as Zoom.For example, in XRTeleBridge(XRTB)[26], users can control the avatar through gaze and gestures. With HMD, speakers and audiences can interact at a distance including virtual objects in a virtual space.

Christoph Leuze and Matthias Leuze[27] proposed an AR remote application about virtual flight. The system synchronizes the position of the virtual aircraft in the real physical world with data for multiple remote users. At the same time, the user can choose to use a headset or a mobile phone to see the virtual plane being operated by a user in the real world.

AVATAREX[28], the remote presence system of virtual avatars, provides a way to connect users who occupy the same space in the real world and their virtual copies in the virtual world. They implemented a simple remote game based on AVATAREX, and evaluated the performance of AVATAREX on different devices through user study.

3.2.1 Virtual Avatar

Avatars are often used to participate in remote communication, and there were many 2D avatars in the early days. Kori M. Inkpen[29] explored the effect of different avatars on the user's different work locations. Formal and lifelike avatars scored the highest on five categories of avatars. This means that more realistic avatar is more attractive.

3D Virtual avatar is a key component of AR/VR remote system. By using virtual avatars, users on both sides can communicate and interact more intuitively. The latest advances in reconstruction and tracking technology allow for easier live capture of humans as animated and realistic 3D avatars. Therefore, traditional 2D video-based teleconferencing systems are evolving into 3D-based immersive teleconferencing systems, for example, delivering avatars in interactive and shared enhanced or virtual spaces[30].

Harrison Jesse Smith and Michael Neff[31] explored that the user's actions and language are represented on a virtual avatar through VR. And through user study, the system can greatly improve the user's sense of immersion and the sense of presence when communicating.

In addition to remote systems, it is also very critical about the social dynamics in virtual environments and video game[32].

There are also many ways to construct this kind of virtual personalized avatar. For example, Alexandru Eugen Ichim et al.[33] described the process for creating a personalized 3D facial avatar through the phone camera. Based on recording the user's facial features, including the user's expressions, the system provides for creating and restoring the user's 3D face and expressions. The system is also deployed in real-time applications. And some methods were proposed for reconstruction using deep neural networks. PIFuHD[34] solves the limitation of low resolution for human model reconstruction. And provide users with a better way than other mannequin reconstruction techniques, thus helping users to easily build their own avatars.

3.3 Motion Tracking with AR

Many AR applications need to track people's behavior, and some use widely used surveillance cameras to track and recognize people in a timely manner based on vision[35]. By recording and capturing the user's daily behavior, CAPturAR[36] identifies and reconstructs the user's daily environment including the virtual identity of some items. The system uses HMD devices with multiple cameras to help users better handle daily affairs.

Sam Corbett-Davies et al.[37] explored and improved the real-time human-machine interface for augmented reality. By using motion capture as a technology, real people can influence concrete 3D objects in augmented reality. Motion capture usually involves tracking various parts of the body through an infrared camera to obtain streaming position data. We can achieve real human motion by applying this data to 3D models.

Microsoft kinect can provide excellent depth sensors and RGB cameras to track and capture various parts of the user's body, so it is gradually being used in combination with AR applications. For example, Lucía Vera et al.[38] proposed an AR system, which is based on kinect to capture the user's movements, as well as head movements, to control the virtual character and communicate with other users. Eduardo Souza Santos[4] propose a kinect-based tracking recognition to allow users to interact with virtual objects through gestures. Some augmented reality "virtual dressing rooms"[39] that simulate 3D clothing use a kinect to obtain the user's physical parameter measurement values and build the personalized model for users. And I-Jui Lee[40] proposed to control virtual 3D cartoon characters based on kinect bone tracking to improve the social interaction of ASD children.

Chapter 4

Research Goal and Approach

4.1 Goal

The research goal is to provide an immersive remote football experience system in order to enable two remote people to play football with each other. At the same time, the system will give different feedback according to the user's actions when kicking the football.

For a two-user remote system, how to better improve the user experience is very important. Most remote systems lack authenticity, immersion, and interaction between users, which we need to solve. In addition, as an AR football system, allowing users to better interact with virtual football is a key point. Based on these, we put forward the following research goals:

- The system will provide an immersive football experience between remote users.
- Personalized full body avatars will be provided as virtual players to kick football.
- Various feedback will be given to users to improve the experience including the interactive avatars.
- Simulation of virtual football trajectory based on user's actions.

4.2 Research Approach

Our research is mainly focused on using augmented reality to bring users a better remote football experience, including simulated football and interaction with the real environment, as well as rich feedback. With the development of AR technology, people can see virtual objects wearing HMD and interact with them in the physical world. Based on our goals, we chose HoloLens2 as the main equipment for the research. It can display virtual football well and allow users to interact with it.

We create personalized avatars for users to enhance the sense of company users get from friends. In order to make the virtual personalized avatar more real, it is based on the specific information of the user's face and body, so that each user has his own unique virtual avatar. By rigging the generated personalized avatar rig, we can preset a variety of animations for it, so as to be able to interact and feedback in real time with the user.

Through spatial recognition, the system store the recognized environment around the user in the system, and process it. According to the memorized information of the environment, the real environment information around the user is combined with the virtual football field. The collision improves the realism of the user experience.

For kicking football part, the kinect captures the user's actions and analyzes them, thereby simulating the possible trajectory of the virtual football. And the system will give feedback on user behavior, including foul action and excellent kicking action. The feedback also includes a yellow card warning or cheer given to the user by the avatar generated by the user as a virtual referee and the virtual audience. In order to solve the lack of companionship for users, we also add a multi-user element to the system. Users can play football with their remote partners and can see personalized avatars that are very similar to remote users.

4.3 Use Case

When users are bored at home alone, but want to play football with good friends, they can use our kicking together system. It allows two remote users to play football together as if they were face to face.

Firstly, the user needs to wear HoloLens2, connect the kinect to the computer and deploy it in front. After starting our system, user need to look around himself, including the ground, walls and ceiling. The system will automatically recognize the user's surrounding environment and generate a virtual football field overlaid on the real ground. The virtual football field includes grass and a virtual football. The user then needs to wait for the remote user to join.

After a remote user starts our system and joins, he needs to do the same so that the system can recognize its surroundings. Then a personalized avatar of the remote user on the virtual football field appears in front of the local user. The user can click the Generate Virtual Referee button and the Virtual Audience button on the UI interface to provide users with more intuitive feedback.

Then the user can play football with remote friends. By combining the capture of kinect and the recognition function of HoloLens2, the user can control the height and approximate direction of the virtual football. In addition, the virtual referee and the audience will give corresponding feedback according to the user's actions. When the user touches the football with his hand, it is a foul action, and the system will warn with text and voice. The virtual referee will also show a yellow card as a warning. The height of the virtual football we set to fly more than 2 meters will be considered as an excellent action. When the user makes a good kick, the virtual audience will cheer and applaud. After the user makes good moves many times, the virtual audience will imitate the user's actions. And if the user kicks the ball to the virtual referee, the virtual referee will kick the ball back to the user.

Chapter 5

System Design

In this chapter, we will introduce our system design. The first is an overview of the system, and the next is about every important part of our system approach.

5.1 System Overview

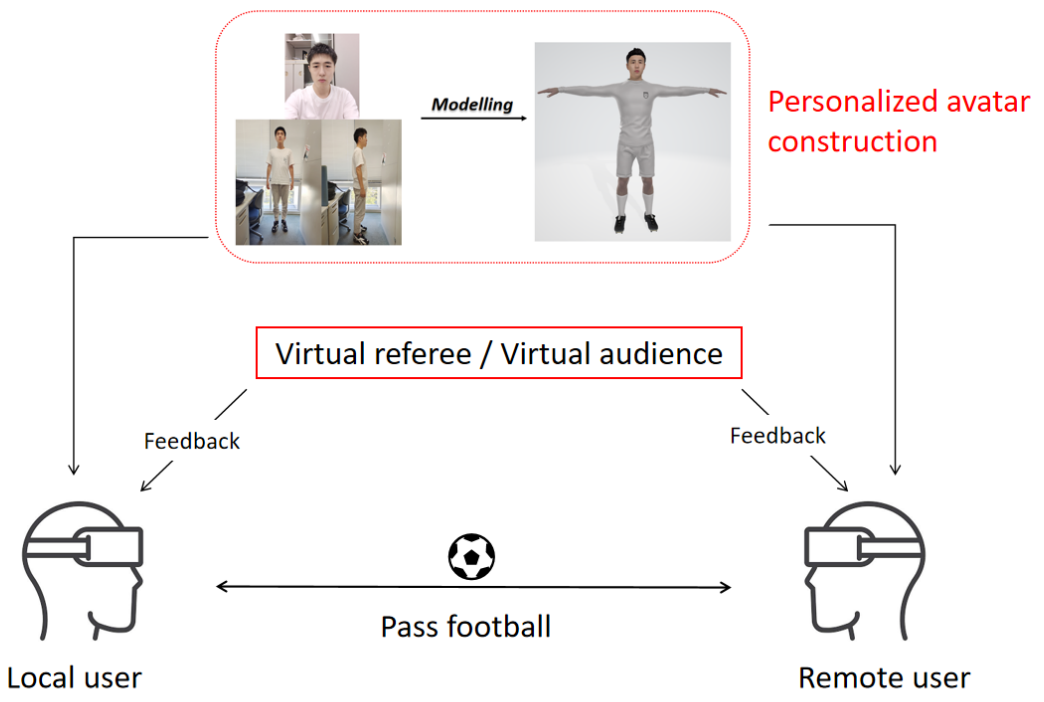

Fig. 5.1 System overview

Figure 5.1 shows the overall structure of our system. The system mainly contains two users, one is a local user and the other is a remote user. From the perspective of the local user, he can see the personalized avatar of the remote user and pass the ball smoothly with each other. In addition, users can see virtual referees and virtual audiences, and can give feedback when the user kicks the ball.

In my system, users can wear HoloLens2 to experience playing football with remote users. First of all, the system will provide two users with a personalized full-body avatar as what the other party sees in HoloLens2, as well as virtual referees and listeners. The system will then provide spatial awareness to identify the physical environment around the local user. According to the spatial information, the system will render a piece of green grass to simulate the ground on the football field. In addition, the football will collide with objects in the real physical scene. According to capturing the user's actions and the trajectory of the football, the user will be given corresponding feedback, which is mainly provided for the virtual referee and the audience.

As we have mentioned above, our research contains three key points and we will divide the system description into four parts:

- Part 1 is about personalized full-body avatar for players.
- Part 2 is related to construction of football environment.
- Part 3 is about kicking virtual football with remote users.
- Part 4 is about interactive feedback for users.

5.2 System Hardware

5.2.1 Mobile Display Device

Our remote football system needs to provide users with a realistic football playing environment and the avatar of remote users. So first of all, our system needs a mobile display device that allows users to move and do their hands freely. Secondly, the mobile device can be equipped with a camera to recognize the environment, so as to render a realistic football environment for the user. In order to bring users the greatest sense of immersion, we chose Microsoft HoloLens2 as the mixed reality HMD (Head-Mounted Display) device as shown in the figure 5.2.

Part of the input of our remote football system includes voice input, as well as receiving voice feedback. Microsoft HoloLens2 as a mobile display device also provides a microphone and headset, which can effectively support our system. At the same time, it has a depth camera and sensor, which can track and recognize the user's hand and accept the user's gesture input, as well as recognize and remember the user's surrounding environment, so as to render a better user kicking football environment for our system, which will greatly improve the user's immersive experience.

(a) Front (b) Side

Fig. 5.2 Microsoft HoloLens2

5.2.2 MoCap Devices

In order to be able to capture the user's actions and simulate the trajectory of the football based on the action information, we used the Microsoft Kinect Xbox one sensor as a depth sensor to record the RGBD information(include the color information and the depth information) and track the users' body movement.

As shown in the figure 5.3, it is Kinect Xbox one sensor including a Kinect adapter. By using this sensor, our system can well track and capture the user's body movements, thereby

Fig. 5.3 Kinect Xbox one sensor

simulating a more realistic football trajectory and giving corresponding feedback based on the user's movement information.

5.3 Personalized Full-body Avatar for Remote Player

For most sports, multiple people or teams playing together can feel the fun of sports. Football is such a sport. If the user can only experience playing football by himself, it will be very boring. Our system is online, and users usually want to play football with friends happily. It is difficult for users to have a pleasant experience if they are playing football with the face of a stranger. In order to provide players with a face-to-face experience of playing with remote friends anywhere, it is necessary to provide a personalized avatar of the user based on the characteristics of the user's body and face.

5.3.1 Personalize Embodied Avatar Construction

In this section, we discuss in detail the process of face model generation and body model generation. Figure 5.4 shows an overview of avatar model construction, which consists of two parts: face model generation and body model generation. It is mainly based on the user's 2D face photos and full-body photos to generate the user's 3D avatar model, which has a face and body very similar to the user.

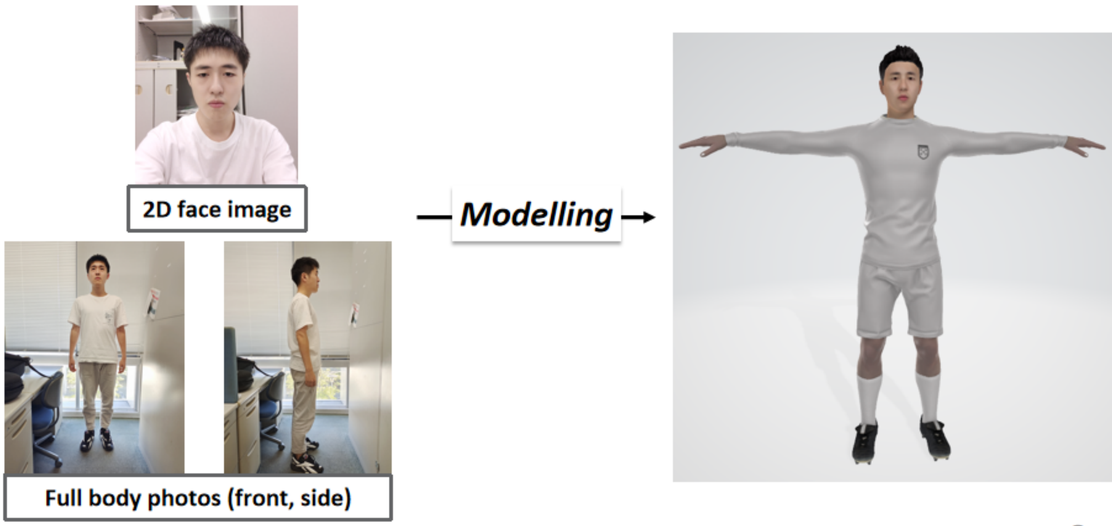

Fig. 5.4 Player's avatar construction

We use avatar SDK to generate the user's 3d face model, which is a technology to generate a realistic 3d face model based on the user's 2d face photo data.For generating the user's 3D body model, we use 3DLook software. Then we combine the 3D face model and body model into a complete full-body avatar model, which is the personalized avatar of the user.

- **Face model generation:** Figure 5.5 shows the transforming from the input image to the output 3D model file. This is the user's personalized face model generated by Avatar SDK, based on computer vision, deep learning and other technologies. This facial model includes the user's face, neck and hair. It is very realistic and very similar to the user's real face. We also changed a hairstyle for the head without covering the face.

- **Body Model Generation:** Figure 5.6 shows the conversion of a user's 2D body photo into a user's personalized 3D body model. We use 3DLook software to generate a 3D model through the user's two 2D body photos including the front and side, as well as basic information, including height, weight, etc. Because of its poor capture of facial features, we chose to use Avatar SDK to generate a facial model, and then combine the body model generated by 3DLook to construct a complete user personalized avatar model.

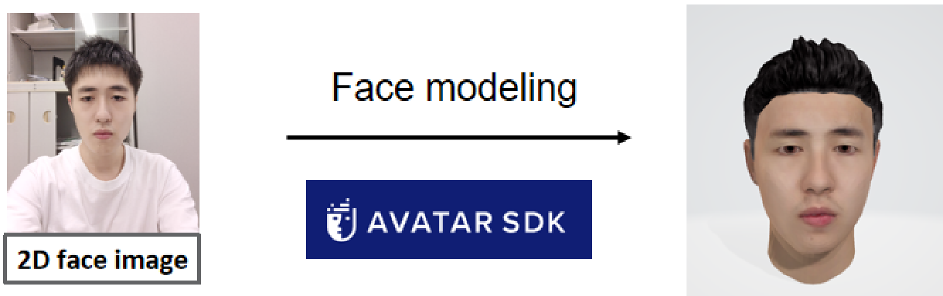

Fig. 5.5 Face model

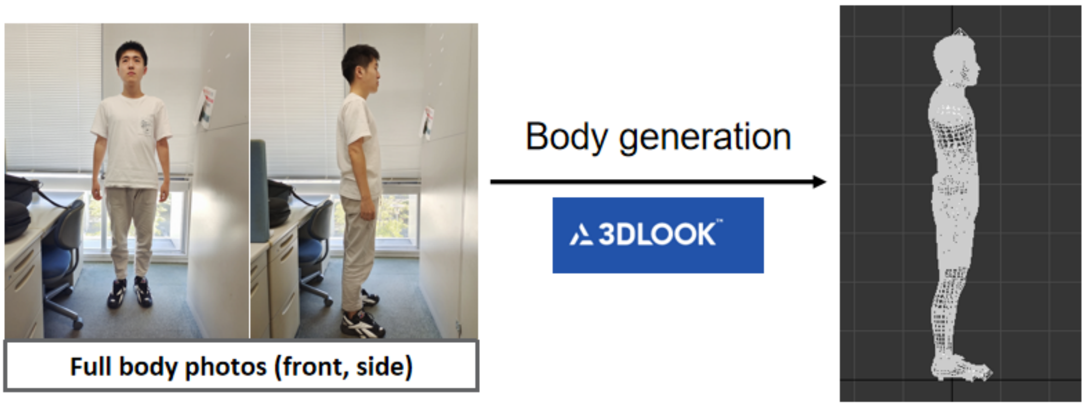

Fig. 5.6 Body model

5.3.2 Avatar Movement

After the modeling of the user's personalized avatar is completed, as shown in the figure 5.7, the user needs to be able to control the movement of the avatar. If the personalized avatar is not equipped with a skeleton, it will be a T-pose model and cannot be manipulated. For our remote football system, the avatar's movement needs to be consistent with the user and be able to make the same actions as ordinary people, such as kicking. So in order to ensure personalized avatar action, the skinning and skeleton rigging should be done in advance.

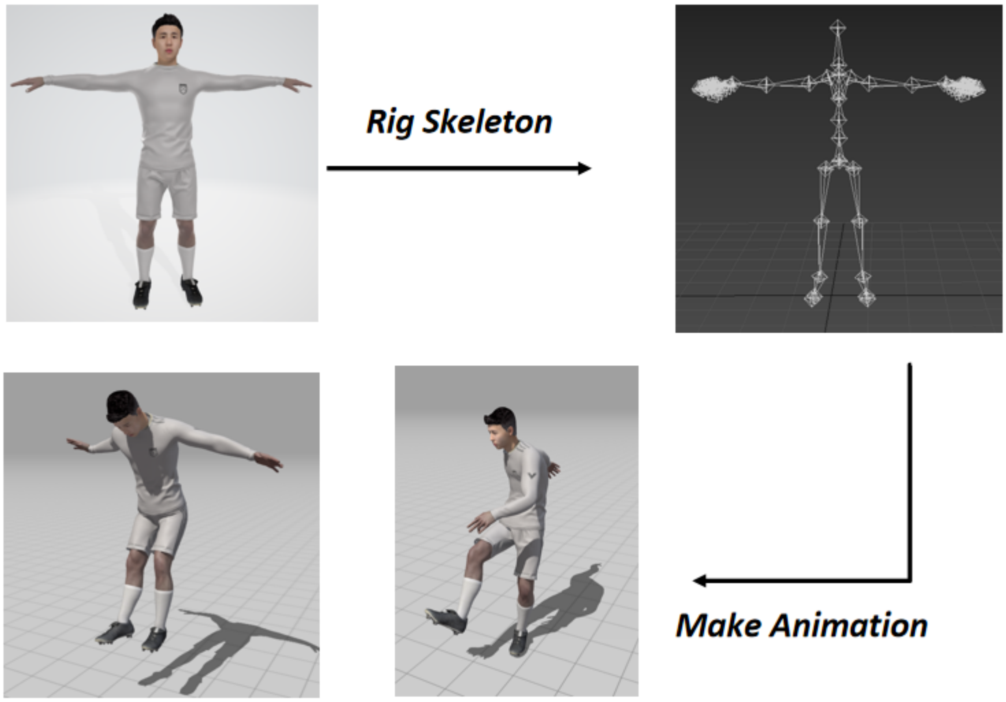

Fig. 5.7 avatar animation

The avatar's movement needs to be synchronized with the user's movement and shared with remote users. We use the Microsoft Azure service to upload the user's location to the server and update it in real time, and synchronize the location information to the user's personalized avatar. In this way, the avatar's location will be consistent with the user's location and shared with remote users. In this way, local users wearing HoloLens2 can see the avatars of remote users moving in real time.

Since there will be a great delay if you control the avatar's actions in real time by capturing the user's actions, we will pre-store an action library for the avatar's actions to facilitate the system's real-time call, such as the avatar's standing, walking, kicking, etc, as shown in the figure 5.8. This can make the virtual avatar more humane, so that users who use our system and remote friends can have a better experience when playing football.

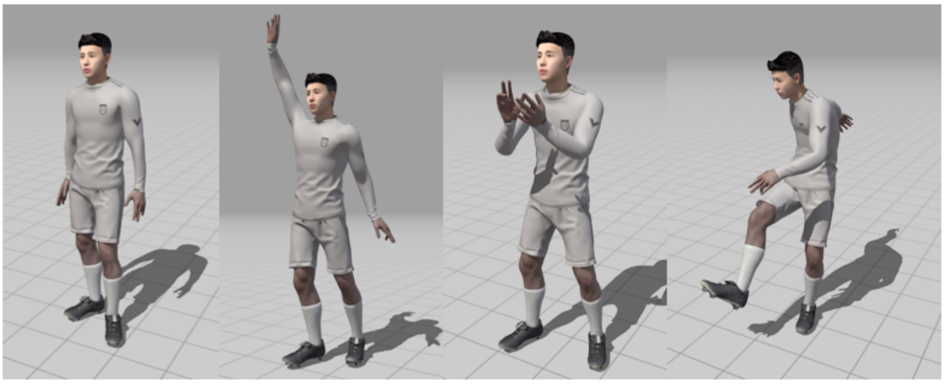

Fig. 5.8 Animation library

5.4 Interactive Football Field

In our system, we want to provide users with an experience of playing football with remote users. After completing the design and construction of the user's personalized avatar, the system needs to effectively recognize the user's environment, so as to provide a realistic football lawn and remote user avatars in front of local users. In order to make the environment more realistic, based on the information data of the environment where the user is located, objects in the real environment will collide with the virtual football.

To summarize the above, we will introduce two parts:

- Part 1 is about rendering football grass for local users and remote users.
- Part 2 is the interaction between football and the environment.

5.4.1 Football Field Generation

Figure 5.9 shows the virtual football field where local users and remote users are located. Our system will detect the user's environment, including the ground, walls, tables and chairs and other items, and then generate a piece of grass and cover the ground, and other objects in the user's surrounding environment will not be blocked by it. In this virtual football field, the personalized avatar of the remote user will be generated in front of the local user. In addition, a virtual football will be generated between them for them to kick and pass.

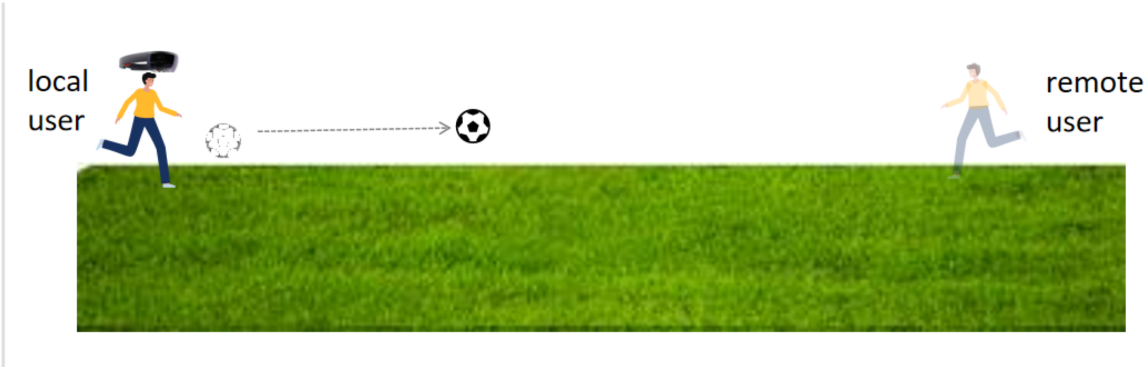

Fig. 5.9 Football field

5.4.2 Interaction Between Football and Real Environment

The system can perform spatial recognition, as shown in the figure 5.10, to store the recognized environment around the user in the system, and process it. According to the memorized information of the environment, the real environment information around the user is combined with the virtual football field. And give collider attributes to virtual football and real environment objects.

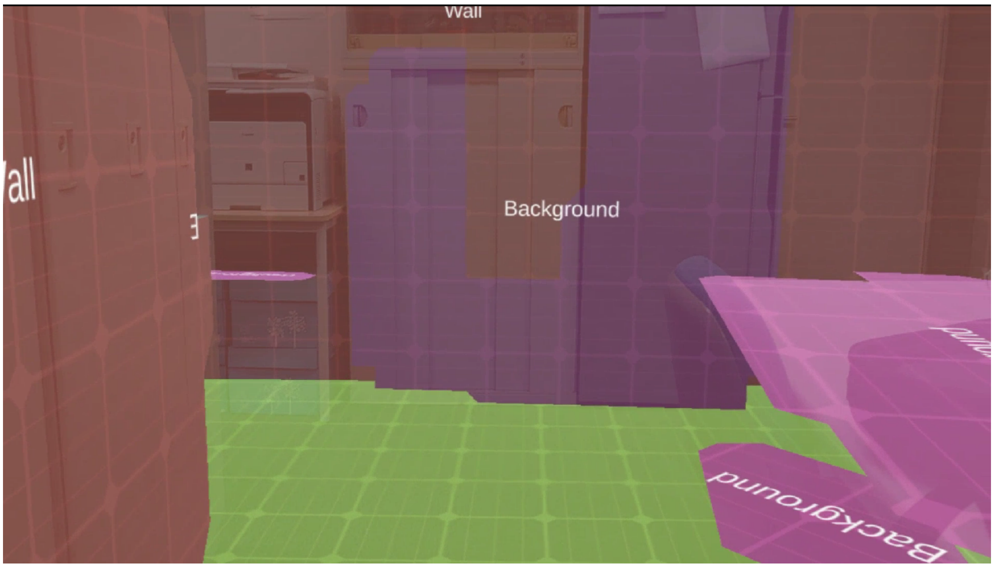

Fig. 5.10 Spatial understanding

In this way, the virtual football can collide with the real physical environment when it rolls. For example, the virtual football will stop rolling or roll to the left or right when it hits a wall, and the virtual football will be still and stuck after being kicked into the trash can.

Fig. 5.11 Rolling football blocked by wall

5.5 Kicking Virtual Football

In this section, we will introduce the detection of kicking motions and the simulation of virtual football. In our system, how to allow users to interact with virtual football is an important issue. Users with HoloLens2 can be recognized for gesture input, gaze input, and voice input, but the movement of the user's feet cannot be effectively recognized. So we use Kinect sensors to track and recognize the movements of the user's entire body.

As figure 5.12 shown, Kinect will first track and identify the user's actions, and through analysis and calculations, the data will be transmitted to HoloLens, thereby simulating the motion trajectory of the virtual football and showing it to the user. The trajectory of the virtual football mainly depends on the movement of the user's foot and the hitting point. However, due to limited technology and time, the point where the user's foot hits the football has not been considered. Based on the movement of the user's foot captured by the Kinect, we calculate the possible forces on the x-axis, y-axis, and z-axis of the football, so as to simulate the height of the football movement, the forward speed, the offset distance, etc.

Therefore, according to the calculated force, a force can be simulated in the system to give a virtual football. At the same time, we need to add appropriate physical attributes to

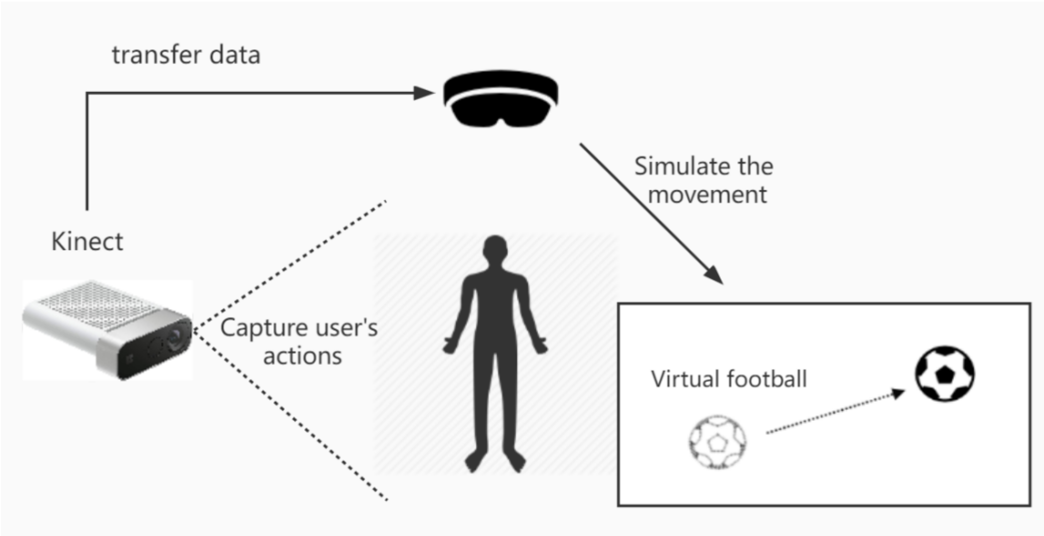

Fig. 5.12 Kicking Detection

the football, so that the virtual football can behave the same as the trajectory of the normal football in reality when it is subjected to force, such as the collision body attribute and the gravity attribute. In order to prevent the football from flying in the air all the time, adding the gravity attribute is necessary for the virtual football.

5.6 Two-player Remote Experience

We plan to implement multiplayer remote kicking in future work. So far we have mainly focused on the remote experience with 2 users. For the two-player remote kicking experience, users need to wear HoloLens2 and use Kinect to track and capture user movements. Firstly, based on the spatial recognition of the environment where each user is located, the football field is rendered for users at both ends. Secondly, the personalized avatar of the remote user is generated in front of him from the perspective of the local user. A football will be generated on the grass between the local user and the remote user avatar for the user to kick the ball and pass it.

After the start, the football will roll to the side of the user, who needs to catch the ball and kick it to the opponent. Kinect will capture the user's actions, and simulate the trajectory of the football movement based on the user's actions to display it in the perspective of both users. The opponent also needs to catch the ball and kick it back. This completes the mutual passing between the local user and the remote user.

The location of the remote user's avatar is mainly determined by the HoloLens worn by the user. We use Microsoft Azure service to share the user's location. At the same time, Photon networking is used for both users to connect. The virtual football also uploads and shares the spatial anchor point of each frame, so that the movement of the football can be synchronized from the perspective of both parties.

For the actions of the remote user's personalized avatar seen by the local user, the use of Kinect to control the avatar action is delayed and difficult to meet expectations, so it is more satisfactory to call the pre-made actions in the avatar action library.

In this experience, the physical scene in the virtual football field will also collide with the football. As shown in the figure 5.13 and figure 5.14, from the perspective of the local user, the football collides with the wall in front of him and will stop, while for the remote user's perspective, the football will stop unexpectedly, which may cause confusion for the remote user. Therefore, to experience multiplayer remote kicking, it is best to be in an open space or there is no barrier between the virtual avatars of the local user and the remote user, so that the experience will not be affected.

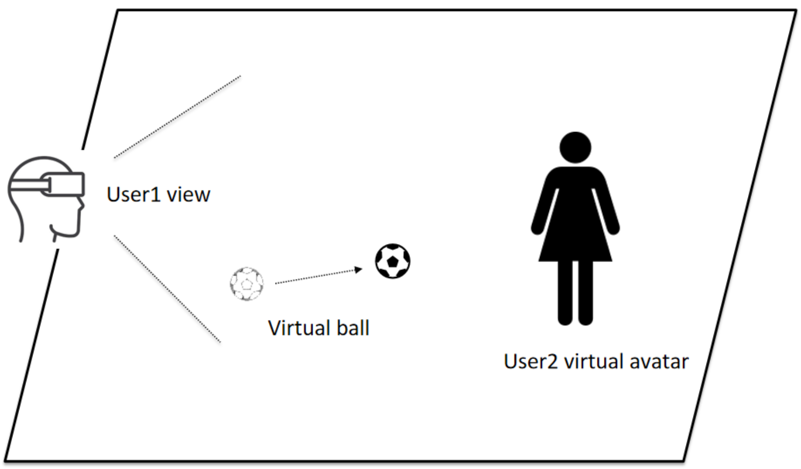

Fig. 5.13 User1 view

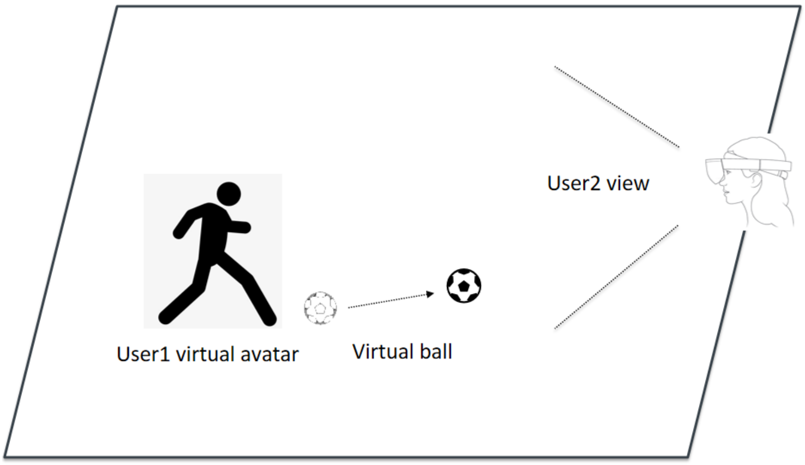

Fig. 5.14 User2 view

5.7 Feedback Design

In this section, a variety of different feedback for this system will be introduced. According to the user's actions and the tracking and recognition of the virtual football, the system will determine whether it is an excellent action or a foul action. For example, if the user touches the ball with his hand, it will be considered as a foul action by the system and will give relative feedback. As shown in the figure 5.15, we will introduce voice feedback, visual feedback, and avatar interactive feedback, which is also the core of this section.

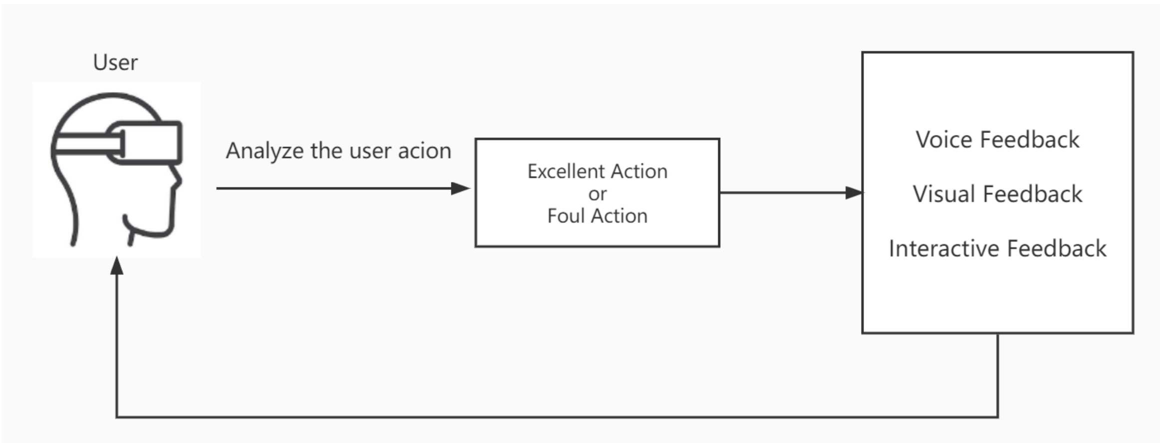

Fig. 5.15 Feedback design

5.7.1 Action Judgement

Based on the user's actions, that is, the data captured by the Kinect tracking, the trajectory of the football is simulated. If the maximum height of the football exceeds a certain value, we can think that the user has made a good kicking action. For example, the movement height of a virtual football exceeds 2 meters. At this time, the system will give voice encouragement, visual effects, and feedback from the virtual audience.

Recognize the user's hand and track the hand bones through HoloLens2. If the user's hand touches the virtual football, the system will recognize it as a foul action. At this time, the system will give voice warnings, visual text warnings, and feedback from the virtual referee.

5.7.2 Sound Feedback

For sound feedback, the system will give different voices for different actions of the user. For example, when the user makes an excellent action, the system will give voice encouragement, such as "nice kicking" and so on. If the user makes a foul action, such as touching the ball with his hand, the system will give a voice warning, such as "Do not touch football by hands".

In addition, when the user touches the football while playing the football, the system will give sound feedback, such as the sound of the football colliding with the foot. Similarly, when the football touches other collision objects, the system will also give sound feedback.

5.7.3 Visual Feedback

For visual feedback, one is text. A text warning is displayed in the middle of the user's field of vision. By binding the location of the text information to the user's camera, it can move with the user's perspective, that is, be fixed in the middle of the user's field of vision. Or it is bound to an object. For example, when the user touches a football, a text warning of foul will be given. This text is bound to the football and appears on top of the football. It will be displayed every time the user touches the football.

The other is visual special effects. For example, if the user makes an excellent action, the system will add special effects to the football, such as creating a ball of flame around the virtual football to show that the user has made an excellent football kicking action.

5.7.4 Interactive Feedback

In order to allow users to have a better intuitive experience, we also designed a personalized avatar in the feedback function for interacting with users. We designed two characters similar to NPC, virtual referee and virtual audience. The virtual referee mainly responds to the user's foul behavior, while the virtual audience responds to the user's foul behavior and excellent behavior.

The virtual referee gives different feedback according to the user's actions. As shown in the figure 5.16, if the user touches the football with his hand, the referee will give out a yellow card for warning. If the user kicks the ball towards the referee, the referee will kick the ball back to the user.

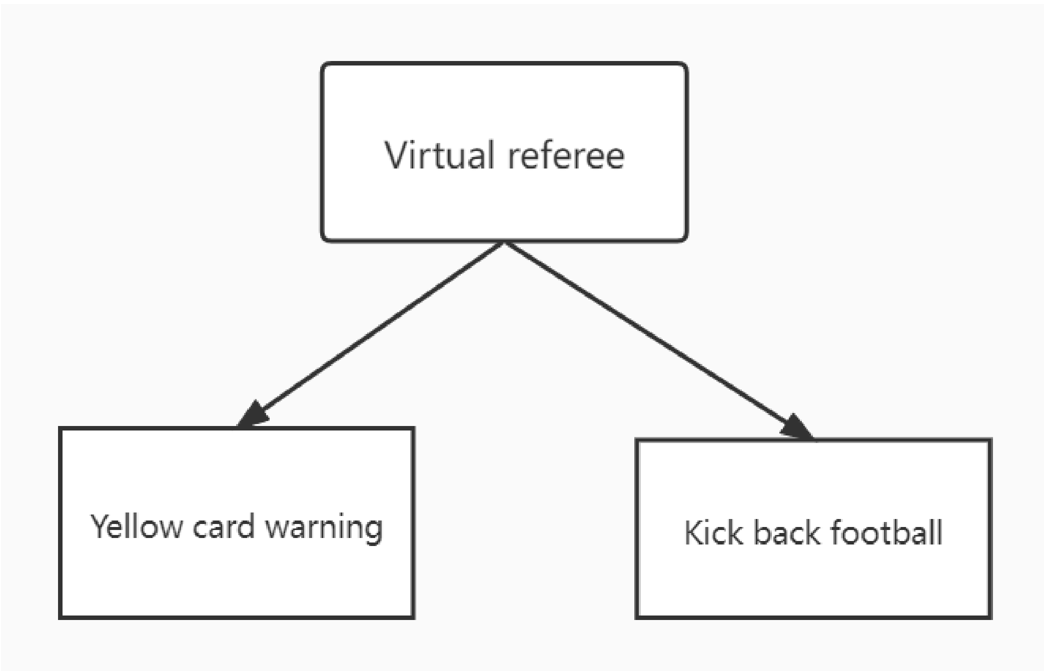

Fig. 5.16 Virtual referee feedback

For the yellow card warning feedback function, we use HoloLens2's hand tracking and add a hand touch receiver component to the object. After the touch is detected through the script, the corresponding function will be called, such as making the yellow card appear in the virtual With the referee's hand, let the virtual referee make a gesture of raising his hand in warning.

For the feedback about kicking back the football, We set up collision detection between soccer balls and virtual referees. When the football hits the virtual referee, the virtual referee will kick the ball back to the user.

The virtual audience will also give different feedback according to the user's kicking action. As shown in the figure 5.17, if the user makes an excellent kicking action, the virtual audience will applaud and cheer. If the user makes excellent kicks many times, we currently set it to 3 times, and the virtual audience will imitate the actions made by the user. This is mainly by tracking the user's joint data through Kinect and transmitting it to every joint point of the virtual audience. If the user makes a foul action, the virtual user will bow his head in frustration. This is mainly through calling different actions of the user in the preset action library, and the virtual audience will make feedback on different actions.

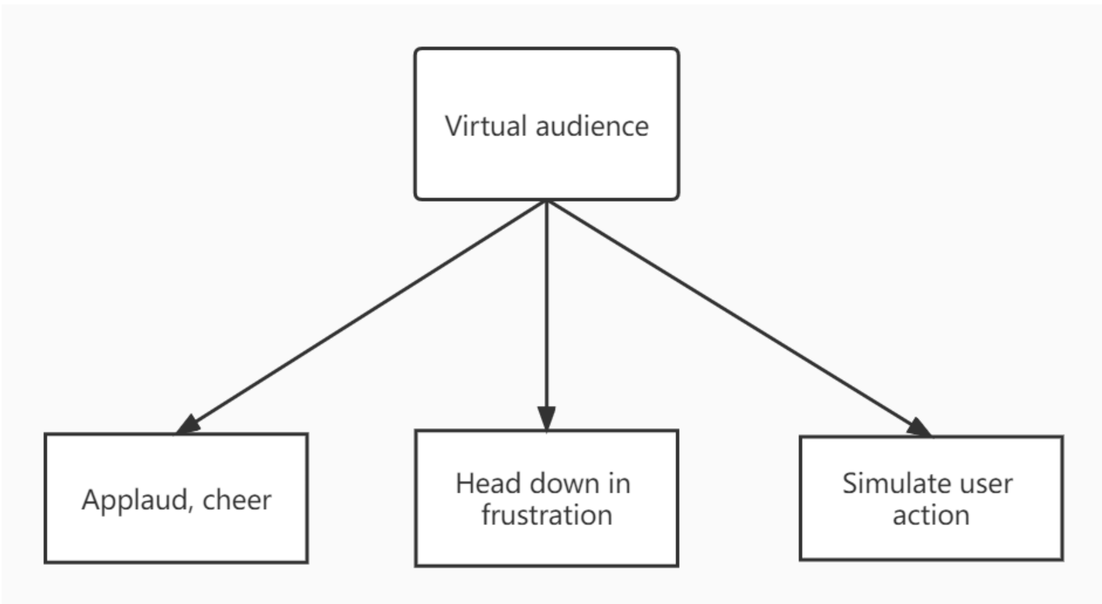

Fig. 5.17 Virtual audience feedback

5.7.5 Feedback in Two-player Experience

The feedback is not only unilateral, the system will be able to give feedback to both users when they play football remotely. The virtual referee and the virtual audience are shared by both parties, so when a user makes an action to trigger the corresponding feedback, both parties can see it.

For example, when the local user makes a good kick, both the remote user and the local user can watch the virtual audience applaud and cheer, or imitate the local user's excellent action. The same goes for remote users. When the local user touches the football with his hand, it is a foul action, and the virtual referee will show a yellow card as a warning, which can also be seen by the remote user. Because virtual referees and virtual spectators are also on the shared football field, the feedback made can be received by both remote and local users.

And for visual feedback, if a user commits a foul, the feedback he gets is as above. The remote user will also see visual feedback of the other user's foul in front of him, which will disappear after a few seconds. For audible feedback, such as the system's encouragement of good moves, this is shared and can be heard by both parties. The collision sound of the football is bound to the virtual football, and both sides are shared.

Therefore, by enriching the feedback mechanism, we can improve the user's experience as much as possible and eliminate the loneliness when playing football alone in a real environment.

Chapter 6

System Implementation

6.1 Deployment Environment

6.1.1 Development Devices

In order to deploy our system to HoloLens2, we need a windows10 operating system computer to develop and process. The configuration information of the computer we use is shown in the Table 6.1 shows.

Operation System	Microsoft Windows 10
CPU	AMD Ryzen 7 4800H with Radeon Graphics 2.90 GHz
Graphics Card	NVIDIA GeForce RTX 2060
Ram	16 GB

Table 6.1 Information of PC

6.1.2 Development Softwares

For the software, we used unity 2020.3.12f as the development platform and visual studio 2019 to write the code. In addition, 3DSMAX is used to build a personalized full body avatar for users.

6.1.3 Technical Support

The other technical supports are shown below:

- Avatar SDK: By combining hard-core computer vision, deep learning and computer graphics, Avatar SDK can quickly turn 2D photos into realistic virtual humans. With this technical support, we can generate a head model that is very similar to the user's face.
- 3DLook SDK: 3DLook could support generation of user body model. We use 3DLook to generate a 3D model through the user's two 2D body photos including the front and side, as well as basic information,
- Mixamo: It provides animations of many characters. At the same time, it can help us quickly rig the skeleton for the avatar.
- Mixed Reality Toolkit(MRTK) 2.7.1: MRTK is a cross-platform tool for building applications to HoloLens2. It can help us build a variety of interactions on HoloLens2 and provide the system with information about the user's surrounding environment by identifying the environment.

6.2 Personalized Full-body Avatar Generation

In this section, we will introduce about the generation of the user's personal avatar, including the detailed content of the generation about the face model and body model of the user avatar.

6.2.1 Face Model Generation

Through the Avatar SDK we can generate a realistic 3D model according to the user's face, including the texture of the face. After we input the 2D photo of the user's face, we can get the relevant model file. As shown in the figure 6.1 is the generation process of the user's personalized face.

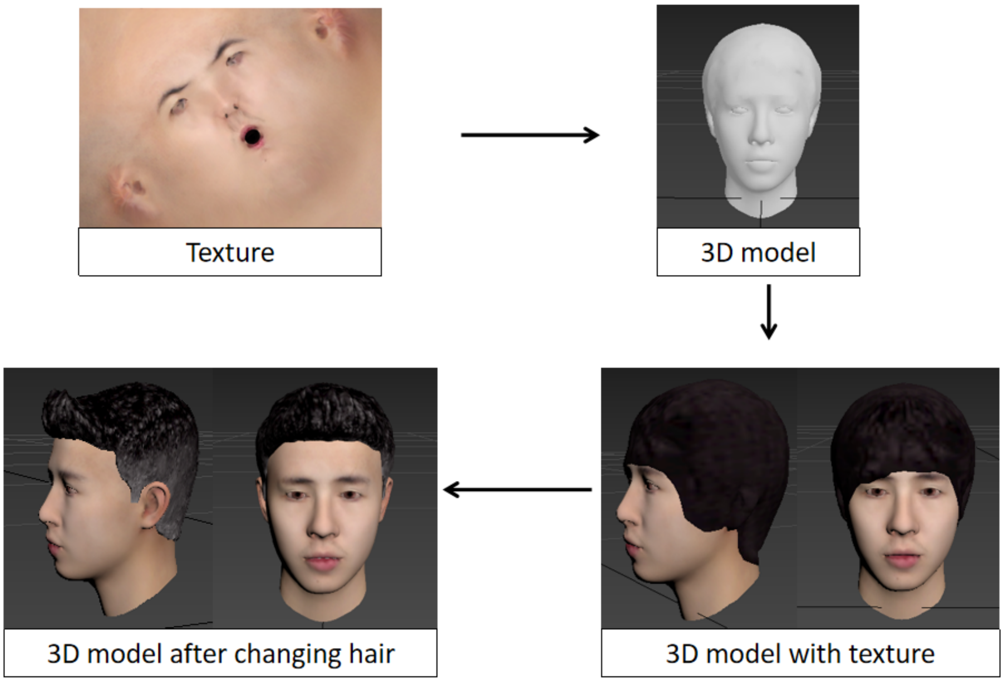

Fig. 6.1 Face model generation

1. **Generate 3D face model through 3D avatar:** First we can upload the user's 2D frontal photo to the Avatar SDK website. After a period of processing, we can get a 3D head model file and related texture map files that are very similar to the user's face.

2. **Mapping texture of face to the generated 3D face model:** We can mapping the texture of face to its corresponding 3D face model and rending the 3D face model with its texture. As a result, we can generate a realistic 3D face model based on a front face image of the user.

3. **Changing the hair of 3D face model:** Avatar SDK also provides users with a variety of hairstyles. Since the original hair style of the user's 3D face file is not suitable, we changed the original 3D model hair style to one that does not cover the face after processing according to the provided hair style file.

6.2.2 Body Model Generation

Through 3DLook, we can easily generate a personalized user body model based on the user's basic information and the user's 2D photos including the front and side. Since the processing of the face is not detailed enough and does not meet expectations, we only need the user body model generated by 3DLook. The figure 6.2 shows the generation process of the personalized avatar body model.

1. **3D body model generation through 3DLook** According to the user's basic information, including height and weight, combined with the input 2D body photos, 3DLook will help generate a 3D full-body model of the user.

2. **Modification of the body model** In our system, the user is a football player, so it is more appropriate to put on a football uniform for the user's avatar. And according to our expectations, some detailed modifications were made to the 3D body model.

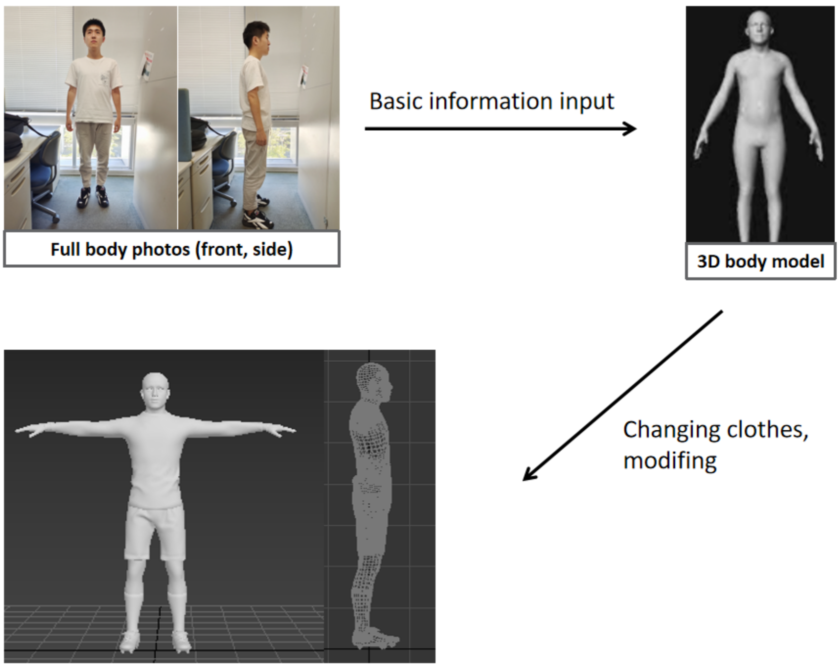

Fig. 6.2 Body model generation

6.2.3 Combination of Face Model and Body model

As mentioned above, after processing with 3DLook, the avatar of the user's whole body can be obtained. Although the overall structure of the body is very similar to the user, the recognition and reconstruction of the face is very rough. So we need to replace the head of the avatar obtained in this step with a more detailed face model generated using the Avatar SDK. We used 3DSMAX software to help us carry out a series of operations, including the cutting of the head model and the integration of the body model as shown int the figure 6.3. And then, we need to modify the texture file of the entire avatar model.

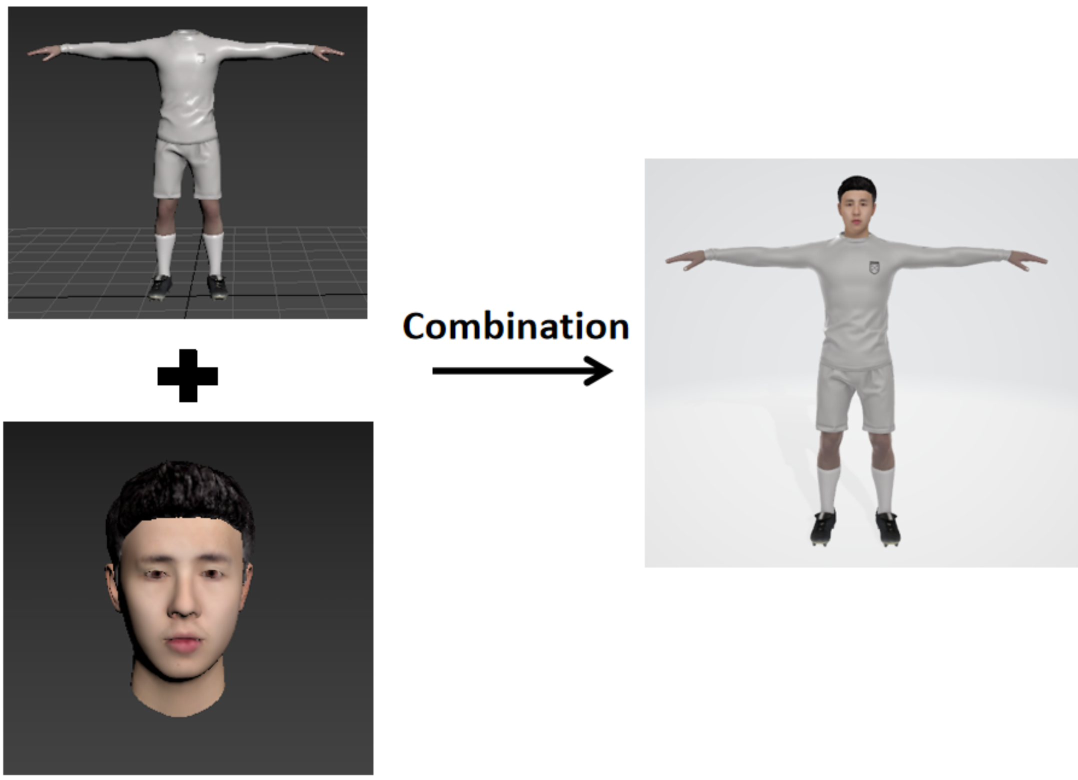

Fig. 6.3 The Combination of face and body model

6.3 Rigging and Animation

In the above section, we have completed the modeling of the virtual personalized avatar, but in order to enable us to make actions similar to humans, we need to bind the skeleton and pre-made animations for it.

In this section, we will introduce the rigging skeleton of the virtual personalized avatar that has been modeled, and make an animation for it, and make it do some actions.

6.3.1 Rigging for Avatar

Usually the model avatar model needs to be bound to the bones. After some pre-made animation data is transmitted, some actions such as walking and running can be well realized. By using 3DSMAX, a skeleton model that fits the user's virtual personalized avatar is constructed as shown in the figure 6.4.

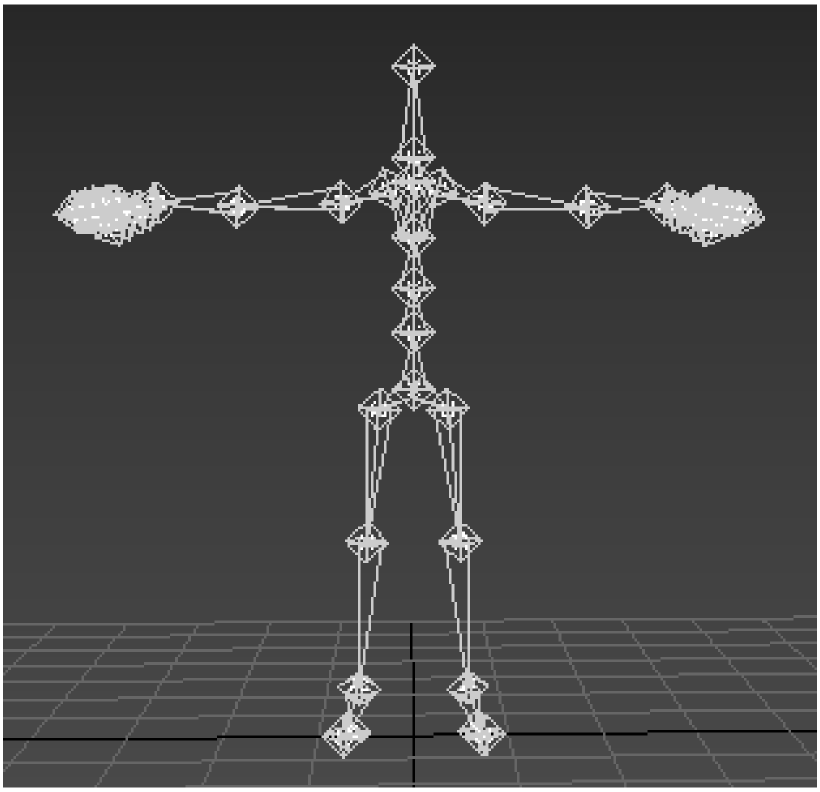

Fig. 6.4 Skeleton model

After the modeling, We need to adjust the nodes of the model to fit every skeleton node of the avatar model as shown in the figure 6.5.

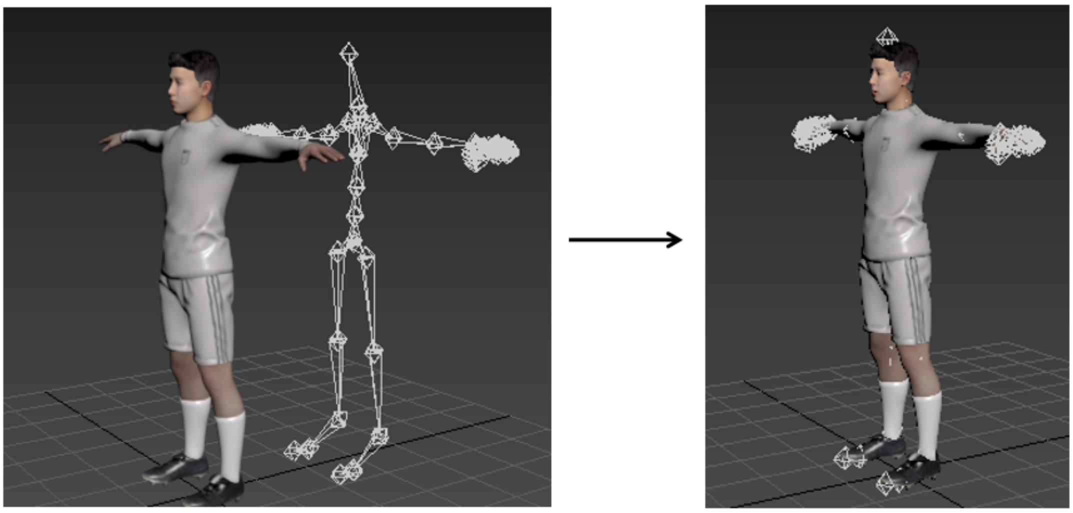

Fig. 6.5 Rigging the skeleton

6.3.2 Make Animation

After rigging the skeleton for the user's virtual personalized avatar, as shown in the figure 6.6, we need to input animation data for it. The Mixamo website provides a lot of different animations. We need to upload the user's personalized avatar model and input the animation data to our avatar model.

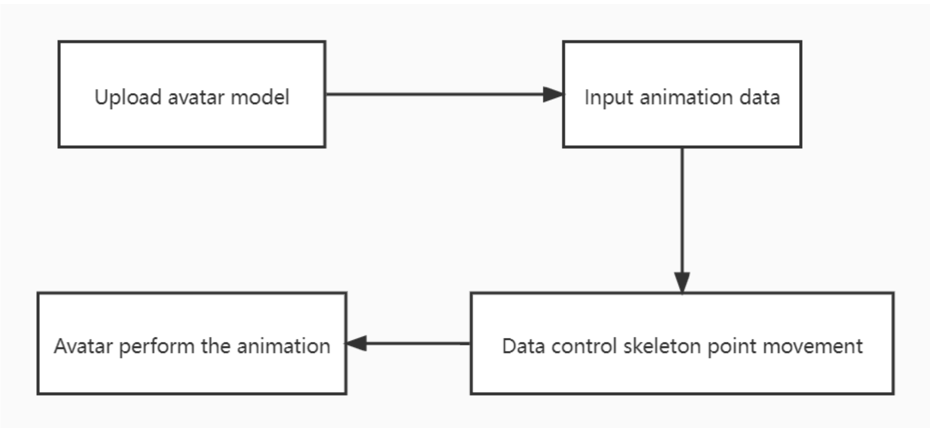

Fig. 6.6 Animation input flow

As shown in the figure 6.7, by inputting the animation data provided by mixamo to the personalized avatar of the rig good skeleton, the data is used to control the movement of the bones and then drive the movement of the entire body to perform the corresponding actions.

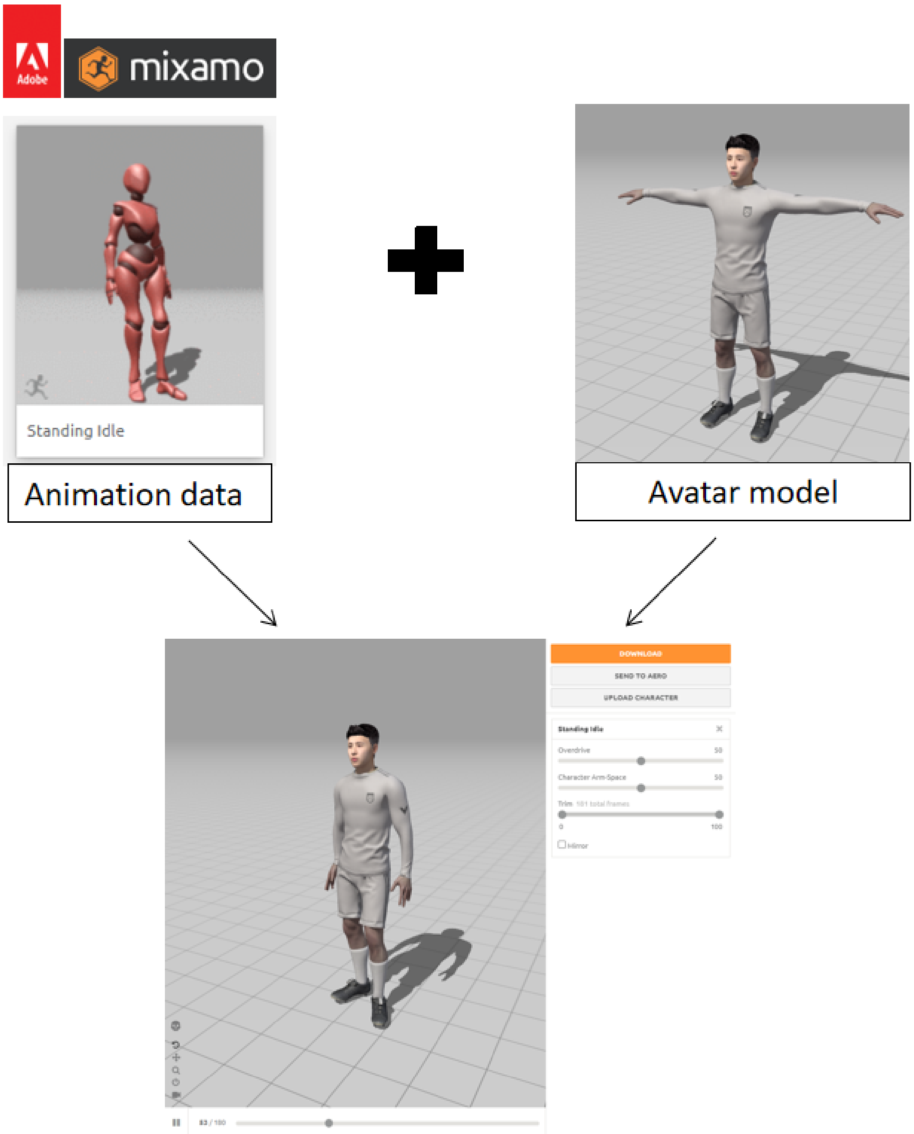

Fig. 6.7 Make animation for avatar

6.3.3 Animation Controller

We can download the animation made by the personalized avatar from Mixamo and import it into the unity project. As shown in the figure 6.8, we need to select "Humanoid" for the animation type, which is a humanoid animation that conforms to the human body skeleton points and performs actions.

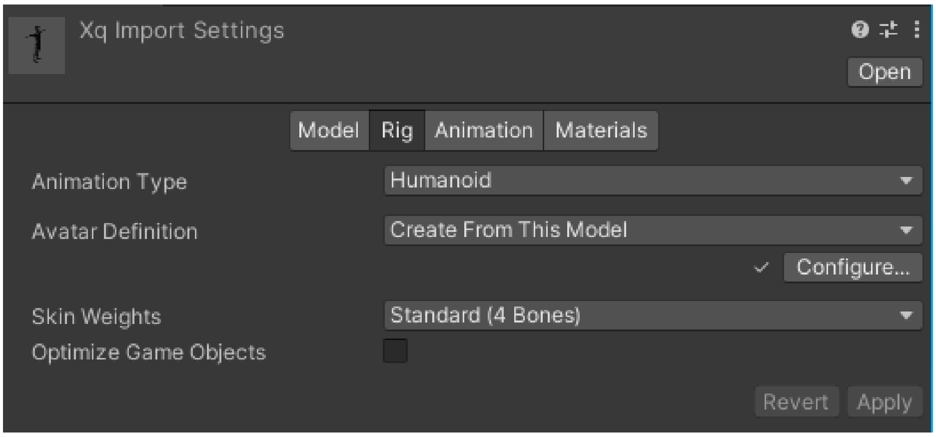

Fig. 6.8 Animation setting

After that, we need to create an animator component for our avatar object, as shown in the figure 6.9, where the controller is used to control the animation operation of our avatar, including the execution order. Avatar is the personalized avatar model of our users.

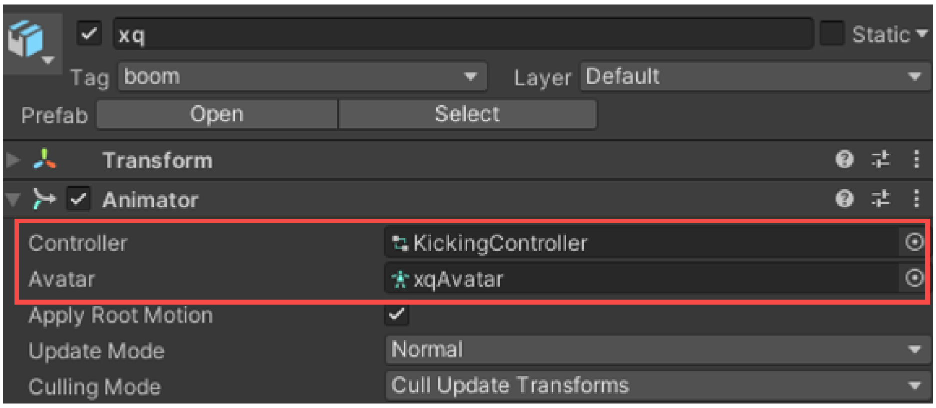

Fig. 6.9 Avatar controller

6.4 Virtual Football Field

6.4.1 Virtual Grass

For the virtual football field, our expectation is to cover the real ground with grass and add football elements. In order to achieve our expectations, that is, not to obscure the real environmental objects around the user, we try to use HoloLens2 for spatial awareness.

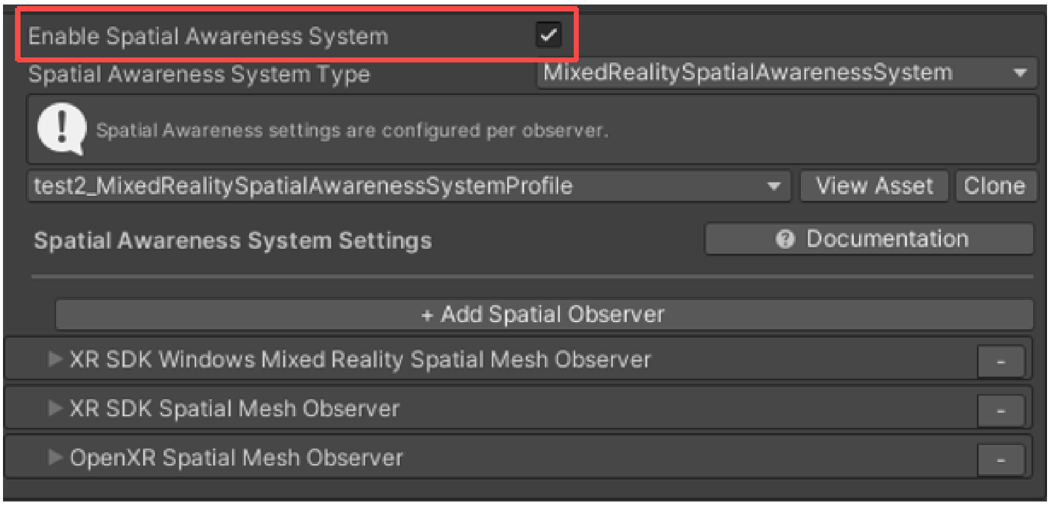

Fig. 6.10 Spatial awareness setting

As shown in the figure 6.10, we use the Mixed Reality Toolkit (MRTK) to set the specific details of spatial awareness, remove unnecessary grid occlusion and prevent real objects in the user's surrounding environment from being blocked by the virtual grass.

The display setting is about the relationship between virtual objects and real environment objects. there are three options. If you select none, the real space object will be occluded by the virtual object. If you select visible, the default mesh material will be used to cover the real environment objects recognized by the system, and the virtual objects will be occluded by the real environment objects. But such coverage will greatly affect the user's visual experience. So we choose occlusion, it will only show that objects in the real environment will not be occluded by virtual objects, and will not cover a layer of grid on the physical objects.

As shown in the figure 6.11, the virtual grass is well generated in the real physical environment and covered on the ground. Walls, tables, etc. will not be blocked and will act as collision objects to interact with the virtual football.

Fig. 6.11 Virtual grass

6.4.2 Football Construction

After spatial awareness is completed, a collision body will be automatically constructed for real objects in the user's surrounding environment, that is, real physical collisions can be generated with other objects. In order to achieve the desired effect of a very realistic football field, the virtual football can collide with real objects in the surrounding environment.

Firstly, we need to create a spherical object about the size of a football, and we need to add a Mesh Renderer component to it to modify the material map for it, including black spherical material and white spherical material as shown in the figure 6.12. In order to make the virtual football more like a real football, we need to add physical components to it, including Rigidbody and Sphere Collider as shown in the figure 6.13. Rigidbody is mainly to make the virtual football movement conform to the physical laws of the real world, add gravity to the virtual football, and make the speed of the football more realistic. At the same time, we have to add a spherical collider to the virtual football, which is to allow it to collide with other objects, such as real objects in the user's surrounding environment.

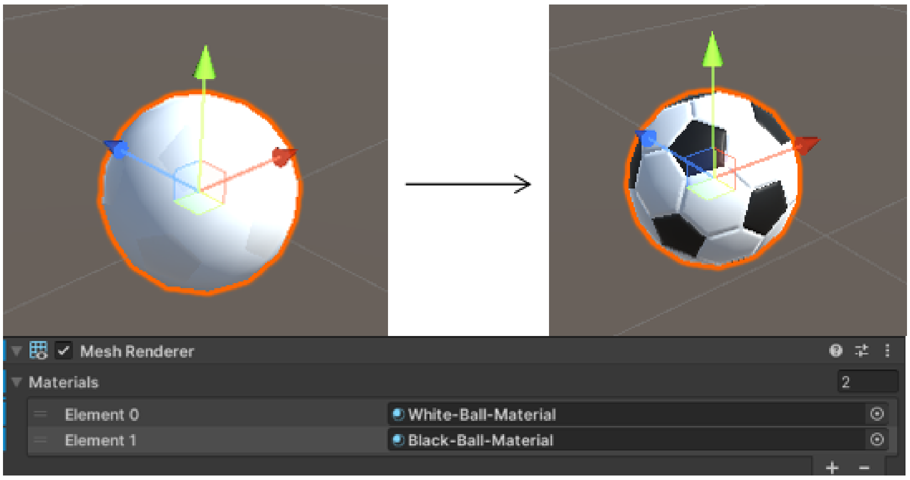

Fig. 6.12 Virtual football construction

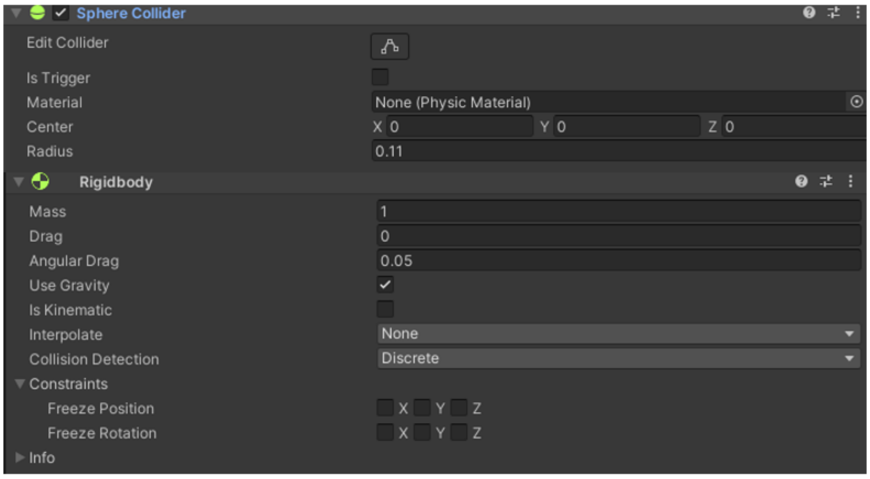

Fig. 6.13 "Rigidbody" and "Sphere Collider"

6.5 Multi-user Experience

6.5.1 Server

Football is a multiplayer sport, and it is always boring to play alone. So I added a multi-user kicking function to our system, mainly combining Photon and MRTK. As shown in the figure 6.14, it is about photon server, which is mainly for connecting in multi-player mode.

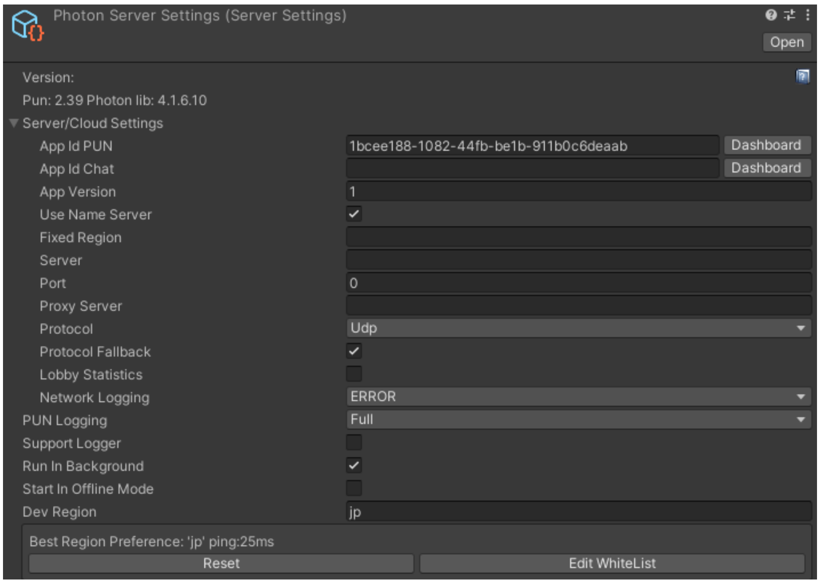

Fig. 6.14 Photon setting

6.5.2 Multi-player Creating

The server is necessary. For the multiplayer football system, we use the photon server. First, we need to register a server application on the photon official website to get the server id. As shown in the picture6.14is the photon server settings, we chose "jp" for the region, which can minimize the delay in connection and interaction between users.

After we have the server, we need to modify the client's information. First of all, we need to add user personalized avatars, and the personalized avatar models of the two users need to be set as prefabs to call at any time. As shown in the figure 6.15, we set host, which is the first user model to join the server, to player1, then the first created position is also the initial position of the first user's camera. The user model added later is set to player2, and the position of the model is consistent with the position of the camera, that is, the position of the user avatar seen by the other party is controlled by the position of HoloLens2.

We need to adjust the position of the camera near the eye position of the user's virtual avatar to achieve the best effect, that is, we can see other places well without being obstructed by our own model. The position of the virtual football is a fixed position and will be automatically generated between the two users.

```
private void StartGame()
{
    if (PhotonNetwork.IsMasterClient)
        CreatPlayer1();
    else
    {
        CreatPlayer2();
        return;
    }
    if (TableAnchor.Instance != null) CreateInteractableObjects();
}

private void CreatPlayer1()
{
    var player1 = PhotonNetwork.Instantiate(photonUserPrefab.name, Vector3.zero, Quaternion.identity);
}
private void CreatPlayer2()
{
    var player2 = PhotonNetwork.Instantiate(photonUserPrefab2.name, Vector3.zero, Quaternion.identity);
}
```

Fig. 6.15 Multi-player creating

6.5.3 Camera Position

For the HoloLens2 camera, it is fixed and cannot be changed with other objects. We usually keep the position of other objects consistent with the HoloLens2 camera instead of fixing the camera to other objects. Therefore, in our system, the user's personalized avatar model needs to be bound to the camera position of HoloLens2, and the initial position of the camera needs to be at the eye position of the model and a little forward to ensure that the field of view is not blocked by the model.

After aligning the position of the model with the camera, since the model will move with the position of the camera, it means the position of the HoloLens2 worn by the user. If the user rotates left and right, the model will also rotate left and right, which is no problem. But if the user raises his head or leaned over, the model will rotate up and down, which is not in line with the real person situation in the physical scene. As shown in the figure 6.16, we need to modify the script related to the camera and the avatar so that the user's model only rotates left and right, that is, with the y-axis rotation, the xz-axis needs to freeze the rotation of the model. Similarly, for the movement of the personalized avatar, what we expect is to move on the horizontal plane, that is, the xz direction. If it moves along the y-axis with the camera, it will fly, which is not in line with common sense. So we need to freeze the y-axis position of the avatar to keep it on a horizontal surface.

```
if (!photonView.IsMine)
{
    var trans = transform;
    trans.localPosition = networkLocalPosition;
    trans.localRotation = networkLocalRotation;
}

if (photonView.IsMine && isUser)
{
    var trans = transform;
    var mainCameraTransform = mainCamera.transform;

    trans.position = new Vector3(mainCameraTransform.position.x,
        trans.position.y, mainCameraTransform.position.z);

    Vector3 transR = trans.localEulerAngles;
    Vector3 maincCameraTansformR = mainCameraTransform.localEulerAngles;

    trans.localEulerAngles = new Vector3(transR.x, maincCameraTansformR.y, transR.z);
}
```

Fig. 6.16 Camera setting

6.6 Football Movement Simulation

In this section, we will mainly introduce how the virtual football is controlled by the player to kick out the desired trajectory.

First of all, for the data transmission between Xbox one Kinect and HoloLens2, we need to use a computer as the server at both ends of the connection. Because there is no direct connection between Kinect and HoloLens, direct data transmission cannot be realized.

We need to connect the computer and Kinect with a USB cable, and use the HoloLens2 simulation application on the computer and connect it with the HoloLens2 device using the same IP address, so that HoloLens2 can receive the data from Kinect. At the same time, as shown in the figure 6.17, in the unity project on the computer, we need to add data transmission components under the Kinect controller, including the server and the client.

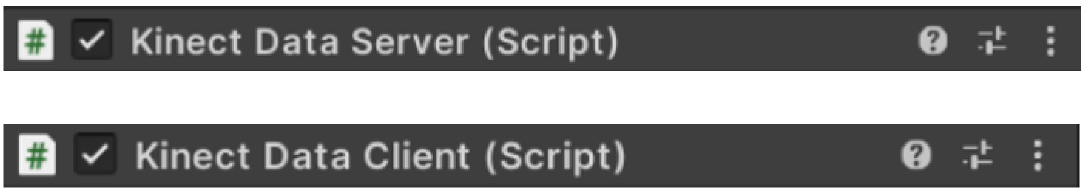

Fig. 6.17 Kinect components in unity project

We mainly calculate the possible trajectory of the simulated virtual football based on the coordinate change of the user's foot within one second. We first need to track the coordinates of the key points of the user's human body skeleton. As shown in the figure 6.18, it is the screen of tracking human skeleton displayed after Xbox one Kinect is connected to the computer during processing.

Because in the context of multiplayer football, the user's personalized avatar is used in the system and interacts with remote users. We try to obtain the user's skeleton key nodes in each frame captured by Kinect by tracking, and transfer them to the personalized avatars in the unity project. The humanoid we selected when importing the personalized avatar makes

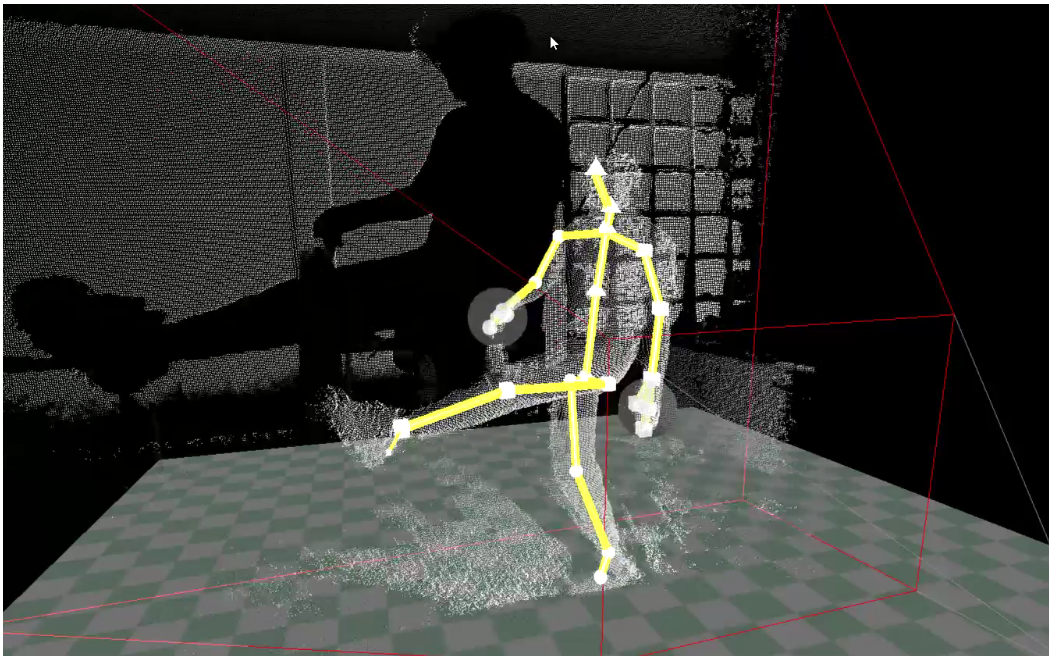

Fig. 6.18 Kinect body tracking

it have the same nodes as human bones. As shown in the figure 6.19, the user can control the actions of the personalized avatar under Kinect, such as waving and kicking.

Fig. 6.19 Control Avatar by Kinect

According to the skeleton node coordinates captured by Kinect, we can save the input to a csv file, read the joints coordinates in each file by calling as shown in the figure 6.20, and calculate the coordinates change of the user's foot on the x, y, z axis within one second to simulate possible trajectories of virtual football.

```
if(manager.IsJointTracked(userId, (int)joint))
{
    Vector3 jointPos = manager.GetJointPosition(userId, (int)joint);
    jointPosition = jointPos;

    if(isSaving)
    {
        if((secondsToSave == 0f) || ((Time.time - saveStartTime) <= secondsToSave))
        {

            using(StreamWriter writer = File.AppendText(saveFilePath))
            {
                string sLine = string.Format("{0:F3}, {1}, {2:F3}, {3:F3}, {4:F3}", Time.time,
                    ((KinectInterop.JointType)joint).ToString(), jointPos.x, jointPos.y, jointPos.z);
                writer.WriteLine(sLine);
            }

        }
    }
}
```

Fig. 6.20 Joints position writing

In order to better simulate the trajectory of football, we choose to use the method of instantaneous force and write it into the football C script. According to the calculated speed of the user's foot on the three axes of x, y, z in one second, a force is given to the football, so that it can obtain instantaneous acceleration in three directions at the same time.

6.7 Interactive Feedback for User

For feedback, we mainly detect the user's behavior, such as HoloLens2 tracking the user's hand movement. If the user's hand touches the virtual football, a series of foul feedback about the hand touch football will be triggered. For example, a line of warning will appear in front of the user's eyes, and a voice prompt not to touch the ball.

We add a hand touch receiver component to the virtual football in Unity, and associate the detected event with the called component through a C script. For the word with warning, its display is controlled by the setactive() function, and it disappears after setting a duration of five seconds. The voice component is called directly after receiving the hand touch event.

In order to better give users feedback when playing football, we have added virtual referees and virtual audiences. The preset actions are called at the right time to give users good feedback.

As shown in the figure 6.21, this is an Animator component on a virtual non-player character, which is used to control the virtual character through preset actions.

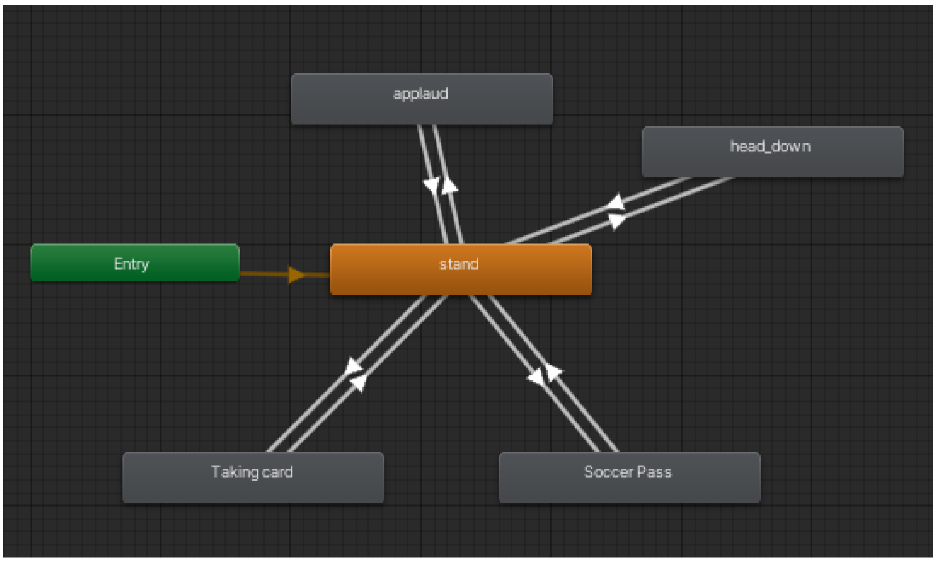

Fig. 6.21 Feedback controller

For example, when the system detects that the user's hand touches a football, the system will trigger a warning, and the virtual referee will show a yellow card to show the warning, as shown in the figure 6.22.

Fig. 6.22 Hand touching feedback

When the user kicks the ball and shows a good kicking action, the virtual audience will give cheering feedback. We preset that when a virtual football can be kicked to a certain height, it is considered a good kick. In our system, the user's foot reaches a coordinate change of one meter within one second on the y-axis, which is a good kick. As shown in the figure 6.23, when the user makes a good kick, the virtual audience gives cheering feedback.

Fig. 6.23 Cheering feedback

For simulating the user's excellent kicking action, we introduced in the previous section about using Kinect to recognize and track the user's joint nodes, and the user can control the personalized avatar. Therefore, in the same way, when an excellent action is detected, we mainly track the user's action in the next two seconds after touching the ball, and assign the received user's joint data to the virtual audience. The virtual audience can imitate the user's excellent kicking action.

For passing feedback, we added a kicking script to the avatar. Since both the football and the avatar have been added with collision components, we need to detect the collision between the football and the avatar and then call the preset kicking actions in the animator controller.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

In this research, we proposed an AR remote football system. In order for people to be able to play football with their distant friends anytime, anywhere, our system incorporates a multi-user element. And by providing the user's unique personalized avatar, the lack of companionship when the user is playing football can be reduced. We provide users with a virtual football field that is integrated with the real environment, allowing users to interact with real objects around them while playing. We added rich feedback to the system to provide users with a combination of personalized avatars as virtual referees and virtual spectators.

First, we create a very realistic virtual personalized avatar based on specific information such as the user's face and body. By rigging it, we can preset many animations to provide real-time interaction and feedback to the user. Second, through spatial recognition, we are able to build an interactive virtual football field. According to the environmental information around the user, the system will store the real object as a collider. So the user can kick the ball on a soccer field covered with virtual grass and collide with real objects. In addition to using HoloLens2 to provide users with virtual information and interactions, our system also uses kinect to capture and analyze user actions. According to the different actions of the user, different virtual football trajectories are simulated. And for excellent actions and foul actions, the system will give different feedback. By using the generated personalized avatars as virtual referee and virtual audience, interactive feedback can be given to users, such as the virtual referee taking out a yellow card when the user fouls the user, or the virtual audience cheering when the user makes a good kicking.

7.2 Future Work

In the future, we plan to increase the number of users, our system is currently two users, but playing football is generally a lot of players. So increasing the number of users playing together is what is going to be done. Secondly, the user's experience when playing football is not enough, and it is still unable to control the more realistic movement of virtual football, which is what we will solve. Finally, feedback to users will increase, and more feedback that can improve the user experience will be designed.

References

[1] Daniel Parnell, Alexander John Bond, Paul Widdop, and David Cockayne. Football worlds: Business and networks during covid-19. *Soccer & Society*, pages 19–26, 2021.

[2] Iuliana Andreea Sicaru, Ciprian Gabriel Ciocianu, and Costin-Anton Boiangiu. A survey on augmented reality. *Journal of Information Systems & Operations Management*, pages 263–279, 2017.

[3] Donghee Shin. How does immersion work in augmented reality games? a user-centric view of immersion and engagement. *Information, Communication & Society*, pages 1212–1229, 2019.

[4] Eduardo Souza Santos, Edgard A Lamounier, and Alexandre Cardoso. Interaction in augmented reality environments using kinect. In *2011 XIII Symposium on Virtual Reality*, pages 112–121. IEEE, 2011.

[5] Maximilian Speicher, Brian D Hall, and Michael Nebeling. What is mixed reality? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2019.

[6] Ashley Ying-Ying Wong, Samuel Ka-Kin Ling, Lobo Hung-Tak Louie, George Ying-Kan Law, Raymond Chi-Hung So, Daniel Chi-Wo Lee, Forrest Chung-Fai Yau, and Patrick Shu-Hang Yung. Impact of the covid-19 pandemic on sports and exercise. *Asia-Pacific journal of sports medicine, arthroscopy, rehabilitation and technology*, pages 39–44, 2020.

[7] Stefan Lawrence and Garry Crawford. Football 2.0. *Digital Football Cultures: Fandom, Identities and Resistance*, 2018.

[8] Katja Rogers, Mark Colley, David Lehr, Julian Frommel, Marcel Walch, Lennart E Nacke, and Michael Weber. Kickar: Exploring game balancing through boosts and handicaps in augmented reality table football. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2018.

[9] Jimena Brucki, Tobías Villecco, Iara Agustina Riera, and Martina Zanaboni Deleau. *Claytonclan: un formato integral de Esports*. PhD thesis, Universidad Argentina de la Empresa, 2021.

[10] A football video game from ea sports. https://www.ea.com/games/fifa/fifa-21, 2021.

[11] Virtual reality football training. https://rezzil.com/player22/, 2021.

[12] Mustafa Atalar and Mahmut Özcan. New augmented reality application in e-commerce and m-commerce. In *2017 International Conference on Computer Science and Engineering (UBMK)*, pages 332–336. IEEE, 2017.

[13] Google augmented reality application. https://arvr.google.com/ar/, 2021.

[14] Philipp A Rauschnabel, Alexander Rossmann, and M Claudia tom Dieck. An adoption framework for mobile augmented reality games: The case of pokémon go. *Computers in Human Behavior*, pages 276–286, 2017.

[15] Nahal Norouzi, Kangsoo Kim, Myungho Lee, Ryan Schubert, Austin Erickson, Jeremy Bailenson, Gerd Bruder, and Greg Welch. Walking your virtual dog: Analysis of awareness and proxemics with simulated support animals in augmented reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 157–168. IEEE, 2019.

[16] Chontira Nimcharoen, Stefanie Zollmann, Jonny Collins, and Holger Regenbrecht. Is that me?—embodiment and body perception with an augmented reality mirror. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 158–163. IEEE, 2018.

[17] Boram Yoon, Hyung-il Kim, Gun A Lee, Mark Billinghurst, and Woontack Woo. The effect of avatar appearance on social presence in an augmented reality remote collaboration. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 547–556. IEEE, 2019.

[18] Rakuten virtual football experience. https://corp.rakuten.co.jp/optimism2019/top.html, 2021.

[19] Fujun Wu. Construction of digital dynamic sports system platform based on vr technology. In *2021 2nd International Conference on Computers, Information Processing and Advanced Education*, pages 998–1002, 2021.

[20] A mixed reality football experience by publicis sapient. https://www.creativereview.co.uk/ar-sports-a-mixed-reality-football-experience-by-publicis-sapient/, 2021.

[21] Tica Lin, Rishi Singh, Yalong Yang, Carolina Nobre, Johanna Beyer, Maurice A Smith, and Hanspeter Pfister. Towards an understanding of situated ar visualization for basketball free-throw training. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2021.

[22] Tim Buckers, Boning Gong, Elmar Eisemann, and Stephan Lukosch. Vrabl: stimulating physical activities through a multiplayer augmented reality sports game. In *Proceedings of the First Superhuman Sports Design Challenge: First International Symposium on Amplifying Capabilities and Competing in Mixed Realities*, pages 1–5, 2018.

[23] Patrick Baudisch, Henning Pohl, Stefanie Reinicke, Emilia Wittmers, Patrick Lühne, Marius Knaust, Sven Köhler, Patrick Schmidt, and Christian Holz. Imaginary reality gaming: ball games without a ball. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pages 405–410, 2013.

[24] Sebastian Günther, Florian Müller, Martin Schmitz, Jan Riemann, Niloofar Dezfuli, Markus Funk, Dominik Schön, and Max Mühlhäuser. Checkmate: Exploring a tangible augmented reality interface for remote interaction. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–6, 2018.

[25] Daniel Chun-Ming Leung, Pak-Shing Au, Irwin King, and Edward Hon-Hei Yau. Remote augmented reality for multiple players over network. In *Proceedings of the international conference on Advances in computer entertainment technology*, pages 220–223, 2007.

[26] Prasanth Sasikumar, Max Collins, Huidong Bai, and Mark Billinghurst. Xrtb: A cross reality teleconference bridge to incorporate 3d interactivity to 2d teleconferencing. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–4, 2021.

[27] Christoph Leuze and Matthias Leuze. Augmented reality representation of virtual user avatars moving in a virtual representation of the real world at their respective real world locations. In *ACM SIGGRAPH 2021 Emerging Technologies*, pages 1–4. 2021.

[28] Timo Koskela, Mounib Mazouzi, Paula Alavesa, Minna Pakanen, Ilya Minyaev, Eero Paavola, and Jere Tuliniemi. Avatarex: Telexistence system based on virtual avatars. In *Proceedings of the 9th Augmented Human International Conference*, pages 1–8, 2018.

[29] Kori M Inkpen and Mara Sedlins. Me and my avatar: exploring users' comfort with avatars for workplace communication. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, pages 383–386, 2011.

[30] Dongsik Jo, Ki-Hong Kim, and Gerard Jounghyun Kim. Effects of avatar and background types on users' co-presence and trust for mixed reality-based teleconference systems. In *Proceedings the 30th Conference on Computer Animation and Social Agents*, pages 27–36, 2017.

[31] Harrison Jesse Smith and Michael Neff. Communication behavior in embodied virtual reality. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–12, 2018.

[32] Guo Freeman, Samaneh Zamanifard, Divine Maloney, and Alexandra Adkins. My body, my avatar: How people perceive their avatars in social virtual reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–8, 2020.

[33] Alexandru Eugen Ichim, Sofien Bouaziz, and Mark Pauly. Dynamic 3d avatar creation from hand-held video input. *ACM Transactions on Graphics (ToG)*, pages 1–14, 2015.

[34] Shunsuke Saito, Tomas Simon, Jason Saragih, and Hanbyul Joo. Pifuhd: Multi-level pixel-aligned implicit function for high-resolution 3d human digitization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 84–93, 2020.

[35] Christian AL Waechter, Daniel Pustka, and Gudrun J Klinker. Vision based people tracking for ubiquitous augmented reality applications. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, pages 221–222. IEEE, 2009.

[36] Tianyi Wang, Xun Qian, Fengming He, Xiyun Hu, Ke Huo, Yuanzhi Cao, and Karthik Ramani. Capturar: An augmented reality tool for authoring human-involved context-aware applications. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, pages 328–341, 2020.

[37] Hiromitsu Sato and Michael Cohen. Using motion capture for real-time augmented reality scenes. In *Humans and Computers*, pages 58–62, 2010.

[38] Lucía Vera, Jesús Gimeno, Inmaculada Coma, and Marcos Fernández. Augmented mirror: interactive augmented reality system based on kinect. In *IFIP Conference on Human-Computer Interaction*, pages 483–486. Springer, 2011.

[39] Sasadara B Adikari, Naleen C Ganegoda, Ravinda GN Meegama, and Indika L Wanniarachchi. Applicability of a single depth sensor in real-time 3d clothes simulation: Augmented reality virtual dressing room using kinect sensor. *Advances in Human-Computer Interaction*, pages 1–10, 2020.

[40] I-Jui Lee. Kinect-for-windows with augmented reality in an interactive roleplay system for children with an autism spectrum disorder. *Interactive Learning Environments*, pages 688–704, 2021.