Context-Aware AR Scheduling Assistant Using Personalized Avatars

ZHAO Ruizhi

44191572-9

Master of Engineering

Supervisor: Prof. Jiro Tanaka

Graduate School of Information, Production and Systems

Waseda University

September 2021

Abstract

With the hustle and bustle of daily life today, it can be difficult to keep up with everything happening in life. A scheduling assistant can be of great help. Traditionally, the physical scheduling book and mobile application can offer some solutions. With the development of augmented reality technology and the research on the motivational role of virtual humans, we can introduce virtual assistants into the scheduling system.

In this research, we propose and implement an augmented reality scheduling system that aims to provide users with intuitive and effective scheduling and guidance based on context. The system mainly consists of personalized avatar generation and context-aware assistant.

The system generates the personalized human avatar from a single full body image. After the 3D human shape reconstruction, texture mapping operation we obtained a human avatar. In order to make the avatar have a more natural performance, we added movement and voice to avatars.

The system obtains contextual information about the user's current time, location, weather, objects in front of them and their status from device sensors and networks. By matching this information with pre-defined task information, a schedule can be derived from this. Then, according to the schedule, the system can call out the user's pre-generated personalized human avatars as a virtual assistant to achieve schedule reminder and guidance, and achieve intuitive interaction with the user. When a user is not familiar with the indoor environment, the system can give route guidance with intuitive signs.

In our study, we applied the Context-aware Scheduling Assistant system to several typical scenarios, such as medication reminders, meeting reminders and indoor navigation.

Keywords: Augmented Reality, Scheduling Assistant, Indoor Navigation, Context-aware

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Jiro Tanaka. He has given me a lot of guidance and help during these two years of study and research at IPS. He is a responsible, rigorous and humorous person. He led me into the field of HCI and AR. He mentors me on how to express my idea and content in presentations and seminars. He gave me advice on how to improve my research and write thesis. It is my great honor to be a student of Prof. Jiro Tanaka.

Secondly, I would like to thank all the members of the IPLAB. The senior members gave me guidance on study and research, especially my coach Ma Ruichen. It is grateful to all my lab-mates. We discussed a lot on research topic and they gives me lots of inspiration. We are the partners in research, and we will be friends for a long time in our future work and life.

Finally, I want to give my gratitude to my family members. Thanks for giving me the financial support and mental support in my study and life. Without your support, I cannot be in Japan to keep on my study and research.

Contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Introduction

People have many different things to do in life and work, but sometimes they will omit and forget something. In view of this situation, people usually write down what they are going to do in advance for memory and reminder. Schedule books and Post-it notes are traditional physical records. However, only when users take the initiative to read them, they can get information from them, instead of active reminders. With the popularity of mobile phones, a variety of calendar assistant software also began to appear in large numbers, but there are limitations that can only provide two-dimensional graphic information, unable to provide vivid and intuitive hints like the human assistant in reality.

In order to allow the schedule assistant to provide vivid interaction and close to real human experience, we can consider introducing Augmented Reality (AR) and virtual avatars. Virtual images have long been used in the online world such as games and social software. Many games today seek to provide players with increasingly anthropomorphic avatars as game characters [1]. When the avatar has a more realistic appearance, close to the user's appearance, and has its own body movements and voice output, the user can get a better sense of reality and immersion. Augmented reality can directly display virtual information in the real world. Microsoft's HoloLens is such a head-mounted Mixed Reality (MR) [2] device that also provides spatial awareness.

In this research, we propose a method to apply augmented reality technology to the schedule assistant system. We generate anthropomorphic, realistic and personalized avatars and use them as schedule assistants to interact with users in the AR world. At the same time, our system makes full use of the perceptual capabilities of mixed reality devices to schedule schedules based on preset tasks and perceived contextual information, and then calls out an avatar as a virtual assistant to provide schedule reminders, guidance, and even lead. The user arrives at the destination. During this period, the user can interact with the avatar in voice and motion.

Other than traditional physical notebook or mobile applications, our idea is to make a AR Schedule Assistant which can provide interactions between the user and the system. Instead of just showing the static information, we think to give verbal and nonverbal interactions between user and virtual assistant.

Different from any other solutions that exist, our proposal can achieve the following functions.

- System can generate personalized avatar as virtual assistant;
- System can provide intuitive remind and guidance instead of just texts;
- System can provide interactions between user and avatar;
- System can support intuitive indoor navigation.

1.2 Organization of the Thesis

The rest of the thesis is organized as follows: Chapter 2 introduces the background of the thesis. Chapter 3 is about the related works which show how the other researchers work for giving better human computer interactions and some related technology. Chapter 4 will talk about the research goal and also the approaches in brief. Chapter 5 is the system design part, where we will introduce the design concept, ideas and typical scenarios in detail. Chapter 6 will be the system implementation part where the detailed environment and implementation will be talked. The last part, Chapter 7, will be the conclusion and future work part, where we will conclude the previous content and talk about future possibilities.

Chapter 2

Background

2.1 Schedule Reminder in Daily Life

Schedules and reminders are unavoidable things in people's lives. What people do next is restricted by their current context.

In life, people usually have their own schedule reminders to prevent forgetting certain tasks. Of course, if you are absolutely confident in your memory and scheduling ability, you can just keep things in your mind. Under normal circumstances, people can record schedules or tasks in paper notebooks, or due to the development of mobile phones today, they can also record tasks in mobile phone software [3][4]. The figure 2.1 gives an example on physical and digital application.

(a) Physical notebook (b) Mobile application: MS To Do

Fig. 2.1 Traditional schedule reminder

Criteria of Better Schedule Reminder

As mentioned above, people usually use paper notebooks and mobile phone software. However, they usually have some flaws.

- Paper notebooks can only serve as reminders when people take the initiative to read them, and they cannot be proactively reminded.
- Although the mobile phone software can alarm at an appropriate time [5], it only has text content or voice messages. Compared with real human assistants, these interactions are weak and not vivid enough to give reminders in the real environment, so they will affect the user's acceptance of the reminders to a certain extent.
- For the lack of contextual information judgment, paper notes or sticky notes cannot directly perceive contextual information; mobile phones can obtain a lot of information, but they cannot see where the user is at any time like human eyes, and cannot make full use of the location and space of the current environment Information about objects in.

In summary, we need a schedule assistant that can perceive contextual information and can give vivid reminders and guidance to users in the right time and space, just like human assistants in reality.

For these requirements, the augmented reality technology that is booming today can meet our needs.

2.2 Augmented Reality

Augmented reality technology is now developing rapidly and is increasingly being used in people's lives and work [6]. Augmented reality technology can accurately locate in the real environment, superimpose the virtual information generated by the computer on the real world scene seen by the user, realize the expansion of the real world information, and enable the user to see a richer world. The result of the integration of various technologies." Augmented reality technology has been widely used in industrial production, education, entertainment and other fields. Under ideal circumstances, virtual scenes and real scenes are perfectly integrated, so AR technology has three characteristics: virtual and real combination; real-time Interactive; three-dimensional registration [7].

Virtual reality technology is a technology closely related to augmented reality technology. Augmented reality technology can be regarded as an extension of virtual reality technology. Virtual reality technology is to conceive a virtual scene in which the user is immersed and interacts with virtual objects in a natural way. Milgram's Reality-Virtuality Continuum describes the mixed reality relationship diagram that transitions from a real environment to a virtual environment [8].

According to the different devices and forms, we can divide AR into Project AR that places cameras and projection devices in a wide area; mobile AR that uses mobile phones; and HMD AR that uses head-mounted devices. Microsoft's HoloLens is a popular HMD. Microsoft developed the Mixed Reality Toolkit for mixed reality for developers to use; at the same time, it also provides some service support on Azure.

2.3 Virtual Avatar

A virtual avatar or virtual agent can represent a character in the virtual world or connect the user and digital world [9]. This character can be the users themselves or another person assigned or random.

Virtual avatars have been widely used in the game field. The virtual character in the game can represent the player, the image can be an animal, plant, etc., or it can be a human. Nowadays, with the increase of platform computing power, allowing users to customize a more anthropomorphic role can be seen in many places.

In the field of virtual reality and augmented reality, an anthropomorphic and personalized avatar can be regarded as a digital mapping of real people in the real world, which can bring a better sense of reality to users [10] [11].

In order to pursue a person who is more suitable for the physical world, we can restore human body characteristics and appearance characteristics as much as possible. The more fully considered, the more immersive experience and realism can be brought to users.

Chapter 3

Related Work

3.1 Personalized Human Avatar

Virtual avatars are digital recreations of various human characteristics, which include geometric shapes, movements, behaviors, etc. As an agent approach, avatar technology is widely used in games, media, engineering, medicine, and other fields. For the research level of avatars, they can be divided into four aspects: geometric appearance modeling, generation and control of motion, behavior expression, and cognitive expression, and most of the current research belongs to geometric appearance modeling and generation and control of motion [12].

Compared with standardized and uniform virtual characters, customized avatars with the characteristics of the target character can bring a stronger sense of realism and immersion to users. Waltemate investigated if and how virtual body presence, and emotional response are influenced depending on the specific look of users' avatars, which varied between (1) a generic hand-modeled version, (2) a generic scanned version, and (3) an individualized scanned version [13]. Their work prove the effect of personalized avatar.

The geometric appearance modeling of the virtual human is mainly the digital reproduction of the skin, muscles, hair, and clothes of the character, combined with computer graphics to realistically present the virtual character in the virtual 3D space. There are several typical methods exist today to generate avatars:

- adjusting the body parameters (body type, height, hair style, gender, etc.) of preset models;
- point cloud reconstruction techniques based on 3D scans [14];
- using machine learning methods to reconstruct the human body from a single photograph [15].

3.2 Context-Awareness in Augmented Reality

Context and context-awareness have been intensively studied in various fields such as ubiquitous computing, intelligent user interfaces or recommender systems [16]. Context awareness has been applied to many areas, such as health care [17], pervasive games [18] and mobile multimedia devices.

Augmented reality is now making its way into work and life. Although most of the current applications are usually used only occasionally for a certain purpose, for continuity, multi-purpose scenarios are slowly being considered. So, when using AR applications, we have to increasingly consider the user's current context.

Grubert presented the concept of Pervasive Augmented Reality, aiming to provide such an experience by sensing the user's current context and adapting the AR system based on the changing requirements and constraints [19]. They also present a taxonomy for Pervasive Augmented Reality and context-aware Augmented Reality, which classifies context sources and context targets relevant for implementing such a context-aware, continuous Augmented Reality experience.

Some researchers also applied this concept into application field [20]. Zhu proposed an authoring for context-aware AR system [21]. The system provides concepts and techniques for creating AR content for context-aware AR applications. Using this system, users can add and arrange various contents such as text, images, and computer-aided design (CAD) models spatially, and specify logical relationships between AR contents and maintenance contexts.

3.3 AR Indoor Navigation

Where am I? How do I get from my starting point to my destination? This is a question that people have to face regularly. The development of navigation technology has provided a great convenience to people who face the problem of finding their way around. Thanks to the popularity of smart mobile devices, people can more easily use GPS technology for outdoor navigation. Map applications such as Google Maps have become an important aid for people to travel. However, due to technical limitations such as signal strength, as mentioned by Nirjon et al [22], GPS cannot be effectively utilized in indoor scenarios.

Indoor Navigation Technologies

For a navigation system, we usually need to understand: what data the system needs, how the system generates a navigation path based on the acquired data and what data the system will provide to the user. Both indoor and outdoor navigation systems need to follow these basic principles to obtain information from the user. A navigation system needs to include the following features: the user's current location, destination information, generated paths and real-time navigation.

Indoor navigation starts with localizing the user's position in the environment, i.e., knowing where the person is. Relevant researchers have summarized and compared the commonly used indoor positioning techniques [23]. Current researchers have used different methods for indoor spatial localization, mainly the following.

- Receiving signal strength indication: This technique uses signals from different devices to determine the user's location. Radio signals from unknown locations, such as Wi-Fi signals [24], Zigbee signals [25], Bluetooth signals [26], etc., can be received by the user's device, and then the user's location and target location are calculated based on the signal strength and the location of the transmitting device.

- IMU sensor technology: This technology uses inertial measurement unit sensors to obtain the relative motion of the user. Using sensors that can detect the user's motion in terms of distance, height and rotation, the current position of the user can be calculated as long as the initial position is known [27].

- Computer vision based: with the development of computer vision and simultaneous localization and mapping (SLAM) techniques, marker-based algorithms are widely used for localization. Gimeno et al[28], Hüibner et al [29]. use markers such as QR codes attached to walls or floors, combined with preset virtual building models, to determine the user's current location.

Augmented Reality for Navigation

With the development of AR technology, it has also been widely used in indoor navigation. By overlaying virtual information on the real environment, users can get more intuitive route hints and the additional information makes the navigation process more efficient [30].

There are two main types of AR-based indoor navigation: handheld mobile devices and head-mounted devices. The former has Android [31] or Apple phones with AR functions, while the latter has Microsoft HoloLens series, etc. HoloLens has some advantages in positioning accuracy due to better depth sensors. At the same time, since the device is worn on the head like glasses, it frees the user's hands and also alleviates the sense of disconnection between the human sight in the real environment and the content displayed on the phone. HoloLens is a good carrier for indoor navigation due to its good spatial perception and inference ability. Many studies have also confirmed its usability in indoor navigation [32][33].

Chapter 4

Research Goal and Approach

4.1 Goal

Our research aims to provide users with intuitive and effective schedule instructions and guidance based on their current context.

This research concentrated on using mixed reality HMD to give user reminders and guidance with a virtual human avatar as assistant. Therefore, we think our goal is to accomplish a system that meets the following requirements.

- Generate intuitive personalized human avatar as assistant;
- Arrange the schedule and assistant according to the context;
- Allow interaction between the avatar and the user;
- Guide user to destination with indoor navigation.

4.2 Approach

To achieve our research goal, we can divide our work into two main parts: one is how to generate intuitive personalized human avatars; and the other is how to use a suitable avatar to act as an assistant who can interact with users in a schedule arranged according to the context.

Personalized Avatars

A personalized human avatar can provide users with a better sense of reality and can also play a better role in motivating. We have designed a set of procedures that can easily generate avatars.

Appearance: In order to make the avatar fit the body shape of the user or target model, we adopted PIFuHD method to achieve 3D body reconstruction. Only by uploading a full-length photo, we can get a 3D reconstruction of the person in the photo with no texture. After that, we can use Blender software to add texture information to the model, and then get a model that looks very similar to a static person.

Body Movement and Speech: In order to make the avatar more vivid, we decided to make him move and talk. In order to allow the avatar to have actions, we bound bones to the avatar on Mixamo, and added actions that fit the scene. In order to allow the avatar to speak, instead of just displaying text prompts, we use the text-to-speech service, which can convert text into words that the avatar can speak.

Context-Aware AR Scheduling Assistant

As a virtual schedule assistant, the virtual avatar can remind and guide the user's schedule, as well as interact and feedback with the user.

Context-Awareness: We first need to obtain the user's current context information in order to compare with the context of the task preset by the user. In order to obtain time information, we can call the system time. In order to obtain the current user's location information, we use the GPS positioning system to obtain the current longitude and latitude information of the device, and the spatial perception capability of HoloLens 2 can determine the user's location in the indoor space. Weather conditions are an important context. We use the weather data and forecast service from OpenWeather to obtain weather data at the target location. In order to be able to perceive and recognize the target object, we use Vuforia's model recognition function to make it clear what appears in the user's field of vision.

After obtaining this contextual information, we can compare this information with preset task data at a certain frequency to arrange different schedules, such as reminding users to leave in advance on rainy days.

User-Avatar Interactions: In order to detect the user's behavior, we detect the contact state of the user's hand with the target object. When the user comes into contact with the target object within a certain period of time, we can regard it as an event. In order to detect the contact event, we add a collision body for the user's hand and the target object. When the two contact, a collision event will occur, and the system will get the feedback immediately.

AR Indoor Navigation: The system also provides a service to bring users who do not know the route to the destination. We mainly solve the problem of indoor route navigation. The system needs to place anchor points in the target space in advance. Here we use Microsoft's Azure Spatial Anchor service, which can store anchor data in the cloud, so the anchor point map can be shared with users who have not yet reached the location. The system stores the id and adjacency information of the spatial anchor in Azure Table Storage, which can be retrieved and queried. For the user, when the user sends a navigation request to the system, the virtual assistant can query the anchor point of the current location from Azure, and generate the shortest path route, and use visual clues to prompt and navigate. When the user arrives at the destination, the navigation ends and feedback is given.

In summary, we have adopted several different technologies to achieve our goals and requirements. We use Holographic display to obtain an augmented reality experience; use 3D reconstruction and motion binding, and voice services to obtain anthropomorphic avatars; use sensors and online services in the device to obtain current contextual information; at the same time we use spatial anchor technology to achieve Indoor navigation for users.

Chapter 5

System Design

In this chapter, we will firstly introduce the overview of the system, and then explain each components of it.

5.1 System Overview

Figure 5.1 shows the overview the system design. This system covers two main stages in the order of existence: personalized avatars and context-aware schedule assistant.

The first part is **Personalized Avatars Generation** . In this stage, the system needs to create a personalized and customized virtual avatar in order to act as a schedule assistant and task instructor in the system.

The second part is **Context-Aware Scheduling Assistant**. At this stage, the need to wear an MR HMD as the entrance to the AR world and this part can be divided into several main fields as follows:

- Context recognition and scheduling matching;
- Task reminders and guidance from the virtual schedule assistant;
- Indoor navigation when necessary;
- User task completion status detection and feedback

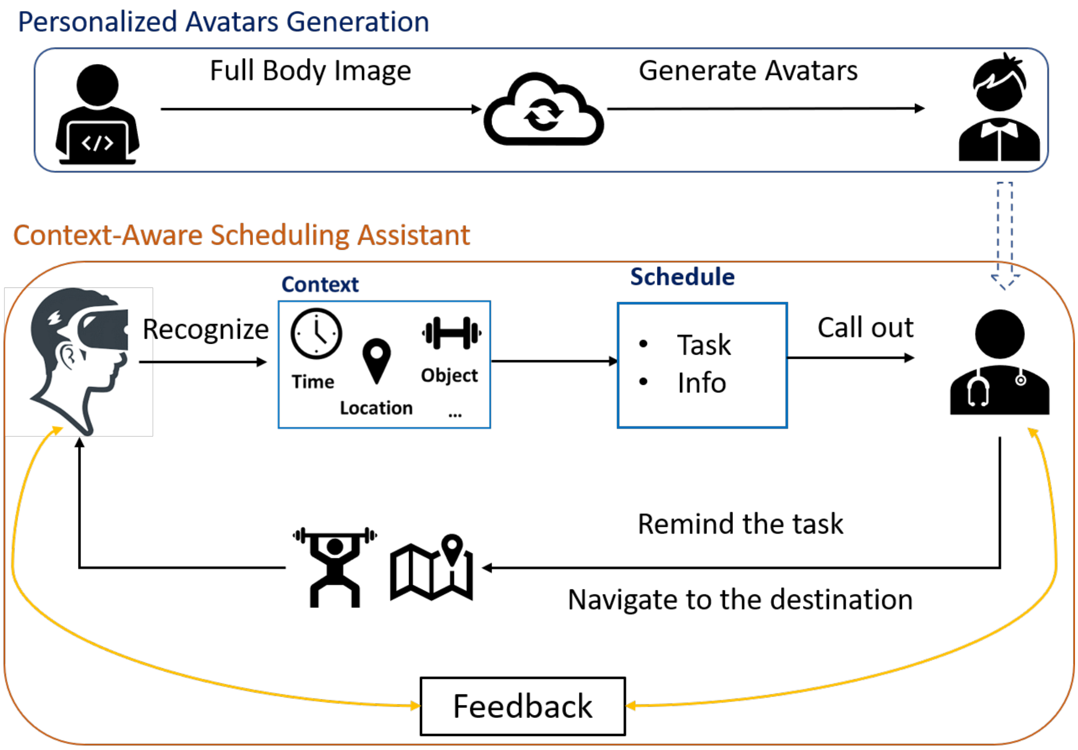


Fig. 5.1 Overview of system design

The next few sections of this chapter are detailed explanations of the key parts of this research.

5.2 Personalized Avatars Generation

Personalized avatars can be used as virtual assistants to remind the user of the schedule and give some guidance on the task or how to get to the destination. Compared to traditional text messages or the voice assistant, the virtual avatar as an assistant could bring more incentive effect and visual information.

In the first step, we need to introduce how to generate 3D virtual avatars which look like a real human, for example the user himself or herself.

5.2.1 Human Shape Reconstruction

Body shape is an important factor in character modeling, we need to create a human body model similar to the target model based on the human body information provided by the user.

There are two general approaches to generating avatars that are similar in size to the user's goals. One is to adjust various parameters on a general model body shape, such as gender, height, fatness, etc. The other method is based on photos or videos uploaded by users that contain the body shape data of the user or his target model. By extracting the target information in the picture or video, the system can generate a 3D reconstruction model that meets the requirements.

We can perform 3D human body reconstruction based on photos uploaded by users. The user should upload a full-body frontal photo of the model, which is the basis for 3D body reconstruction.

In this study, as shown in the figure, we can choose a full-body T-pose photo. Of course, the pose does not have to be this way, as long as it is a full-body photo of the model. After that, with the help of PIFuHD [15], we can quickly generate a human body model. The figure 5.2 shows the input and output result. The user uploads the full body image to the system, the PIFuHD function can then output the 3D reconstruction model in .obj format.

(a) Input (b) Front (c) Side (d) Back

Fig. 5.2 3D human shape reconstruction: (a)Input: a full body frontal image; (b,c,d)Front, side and back view of the generated 3D reconstruction shape model

5.2.2 Texture Mapping

Texture is a very critical part in the process of generating realistic models. If there is no texture, the generated model is like a plaster statue, with only shape information and lack of color information that is important for vision. In order to generate a suitable texture map, the color of the image can be projected onto the vertices of the body model. In this way, when we specify the texture for the human body model obtained from the previous step, we can get a complete and more visually realistic avatar.

(a) Without texture (b) With texture

Fig. 5.3 Add texture to human model

The figure 5.3b is the result that the texture is added to the model without texture in figure 5.3a. We can clearly see that by adding textures, the system can generate a more visually realistic avatar, instead of just a plaster statue.

5.3 Context-Aware System

This system is a context-aware system. In this design, the system needs to identify the context and give the corresponding schedule according to the context. Then, according to the schedule, the pre-created avatar acts as a virtual schedule assistant to prompt and guide the user.

5.3.1 Context Recognition

Context can be divided into two parts: human factors and environmental factors in our system. The human factors include the identity of the user and the related personnel, such as who will participate in this scheduled task. The environmental factors include the location, time, some objects and the state of the object.

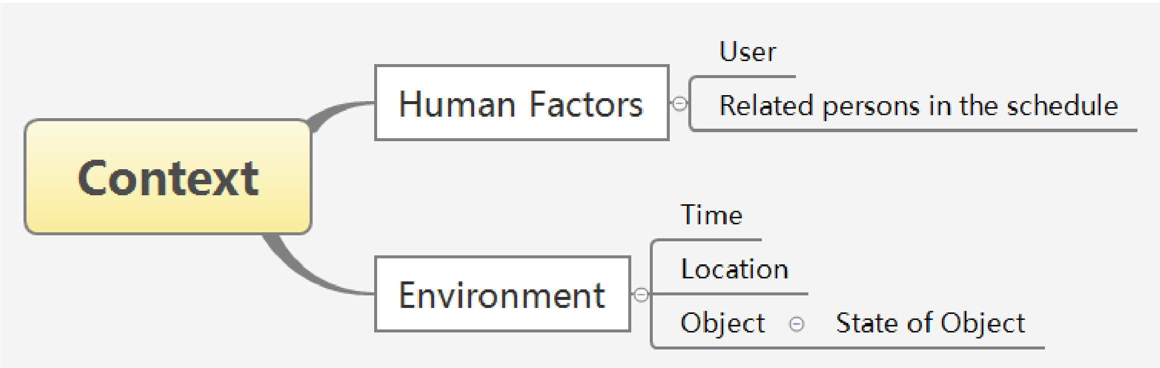


Fig. 5.4 Context factors

The current time data can be obtained from the system time, and the current position information can be obtained using GPS data. We mainly use the latitude and longitude data. After obtaining the time and location data, the system can determine the user's current space-time environment. Different times, different locations, or a combination of the two can be regarded as different contexts.

The object and its state are also an important part of the context, for example the state of weather. Weather information is more common among them, and can affect the factors of schedule arrangement. To obtain and forecast weather information, we can use third-party weather services, such as openWeather. The system can automatically query the current location and the current location of the destination, as well as the weather conditions when the task occurs, for subsequent adjustments to the schedule.

The system also needs to recognize objects in the environment, because some schedules are triggered by objects or are required to complete tasks. For the object recognition method, we can use Vuforia's object recognition method or Microsoft's custom vision service.

5.3.2 Context-Aware Scheduling Arrangement

When the context changes, the schedule arrangement sometimes changes accordingly. For example, the distance to the destination and the remaining time can affect the way traffic travels. For another example, whether the food is sufficient or not can affect whether the user decides to go to the supermarket to purchase.

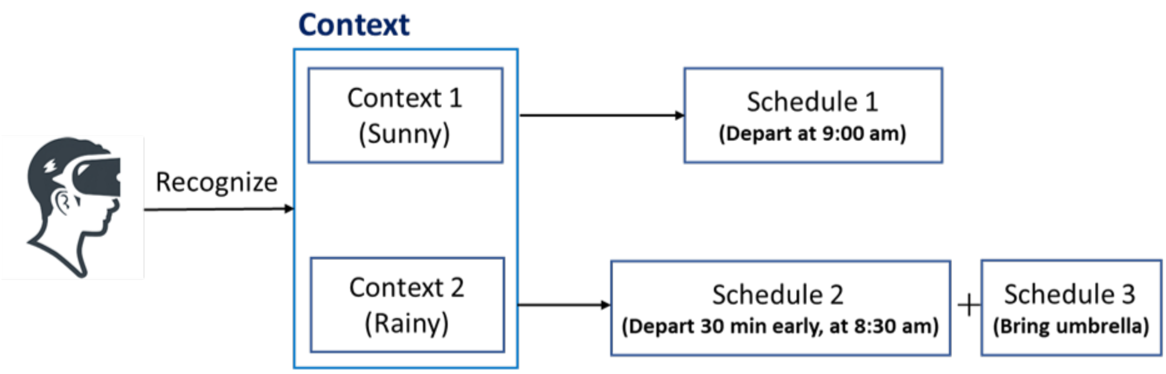


Fig. 5.5 Context-aware scheduling arrangement

The figure 5.5 is an example that how the context influence the schedule arrangement. For schedules that require travel, when the weather is about to rain, the schedule assistant needs to remind users to leave in advance to avoid time delay. At the same time, the system also needs to remind users to collect the clothes that are hanging out and close doors and windows to avoid raining clothes and rooms.

5.4 Avatar as a Schedule Assistant: Remind Task and Give Guidance

After determining the schedule, the system will evoke the virtual avatar as the user's schedule assistant. A virtual schedule assistant will appear near the user, remind the user what to do next, tell the user task information, and take the user to the destination when the user needs it.

This section will explain how the virtual avatar reminds the user of task information as a schedule assistant, focusing on the interaction between the user and the avatar. How to navigate will be explained in the next section.

5.4.1 Add Voice and Movement to Avatar

After creating a realistic avatar as a schedule assistant, how to make it more like a real human schedule assistant is a problem worth considering. In order to solve this problem, we can consider adding more interactivity to the avatar. First of all, we can add actions to the created avatar as body language, and at the same time, we can add voice functions to it, which makes the avatar like a real human. He can say the schedule with corresponding body actions beside the user.

Usually, it is more natural for the avatar to use voice for task prompts and interact with users compared to only text display. This makes the system more like face-to-face communication between people. This can also increase user acceptance of the system. Figure 5.6 shows an example that the avatar speeches to user to remind taking medicine.

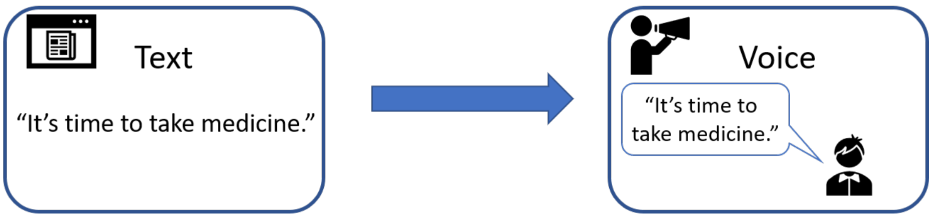


Fig. 5.6 Voice from avatar

The original avatar lacks skeletons, so we first need to bind skeletons to the avatar. After binding the skeletons, we can then design related actions, such as standing, nodding, etc., according to the needs of the task situation. The figure 5.7 is an example. We design the movement of drink to remind the user to drink something, such as warm water when taking medicine.

Fig. 5.7 Movement of Drink As an Example

5.4.2 Interactions Between User and Avatar

In the system design, it is our goal to make the interaction between the avatar and the user more natural. It is our goal to be able to guide the user while improving the user's work efficiency.

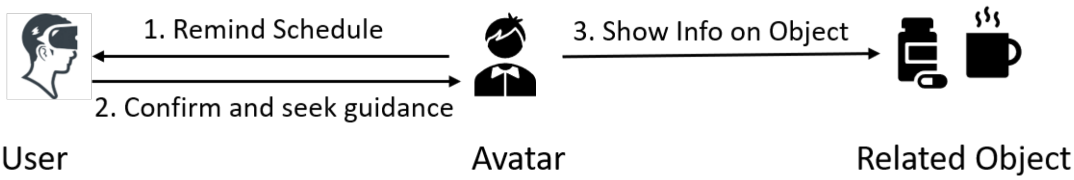


Fig. 5.8 User avatar interaction example

The figure 5.8 is a typical way of interaction. First, the avatar reminds the user of the schedule task information, accompanied by the avatar's voice and action language. After being reminded, the user confirms the task information, and if there is any doubt, he can seek guidance. Then, avatar receives feedback from users and gives guidance. The guidance in figure 5.8 is to display relevant information and guidance next to the target object.

5.5 Route Guidance

After the virtual schedule assistant is evoked, the avatar will remind the user where to go and what to do next. If the user does not know how to reach the destination, the system will also give instructions to navigate the user to the destination. The figure 5.9 shows the pipeline to navigate the user to the destination if the users is now familiar with that place.

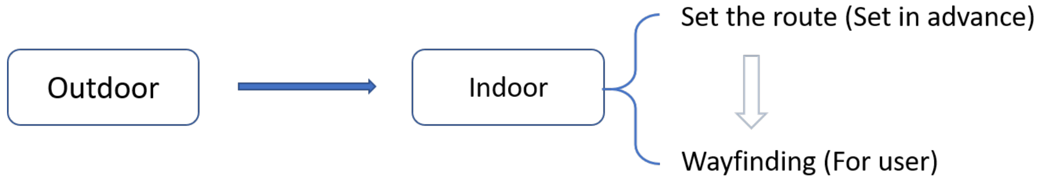


Fig. 5.9 Method of navigating user to destination

For long-distance outdoor navigation, current map service providers, such as Google Maps, can already provide sufficient services to provide location and route inquiries. Therefore, we don't need to spend too much energy on this, we just need to use the services they provide. The research in this paper focuses on indoor route navigation.

For indoor navigation, we can divide the research into two main parts: route setting and navigation. These two parts will be introduced in detail below.

5.5.1 Set the Route

For outdoor navigation, you need to have a map first, and there will be various positioning points on the map, such as buildings, and then route planning and navigation can be carried out based on the positioning points. For indoor navigation, it is also necessary to provide the type of positioning point first. At present, there are different methods for how to set an anchor point, such as placing a Bluetooth beacon as an anchor point in an indoor space, and then after the user's device is connected to the Bluetooth network, the current location of the user is calculated based on the signal strength. In this study, we use spatially aware mixed reality technology to realize our positioning and navigation.

This article uses Azure Spatial Anchor service [34] to set the spatial anchor point. By using this service and using the space perception capabilities of the device HoloLens, we can place anchors in the place of interest. These anchors can be regarded as landmarks on the map in this study. When we place anchors, we can add descriptions to them to illustrate the location or number information of these anchors.

(a) Set anchors at the corridor (b) Set anchors in the room

Fig. 5.10 Set anchors in spatial

The figure 5.10 shows the anchors which were set in spatial, the figure 5.10a are the anchors set at the corridor and the figure 5.10b are the anchors set in the room. With the spatial understanding and environment recognition provided by Azure and HoloLens, the system can remember where we put the anchors in spatial.

At the same time, in order to be able to use these anchors across devices, we upload the location and environment information of these anchors to the Azure cloud, so that other people who have never been in this space can also get the anchor information, which is equivalent to sharing Indoor map.

Because environmental information is sometimes not sufficient, only recording the spatial location data of the anchor may cause position drift, which may lead to errors in the way-finding stage, so we added the adjacent information of the anchor, which can improve the accuracy of the anchor's position in the environment.

5.5.2 Way-finding Method

Before introducing the wayfinding method of the system, we will further briefly explain the role of the spatial anchor in the wayfinding system. In the wayfinding navigation system, the spatial anchor can be regarded as a node. The side length between adjacent nodes is the distance between two spatial anchors. If there is an edge between two nodes, it means that the two points are directly connected, which means that you can go directly from one spatial anchor to another spatial anchor. At this time, the pre-set nodes in the space can form a weighted network, and each edge is a walking channel.

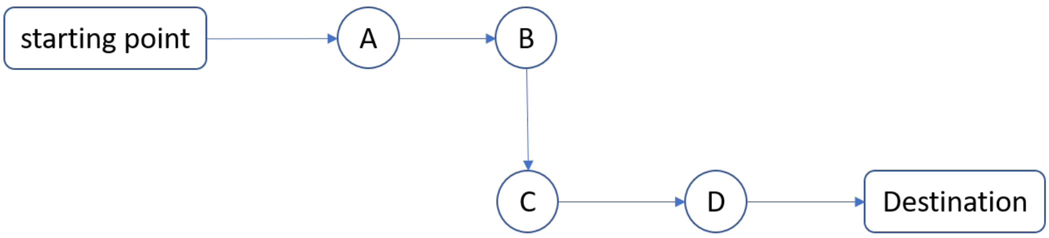


Fig. 5.11 Wayfinding using the pre-set anchors

As the figure 5.11 shows, the user can choose the origin and destinations from the created anchors. Once the starting point and the destination are selected, the system can then start to find the shortest path between the two end points. Since the edge length between two nodes represents the distance between the two spatial anchors, there is no negative weight edge. We can choose the Dijkstra's algorithm to realize this pathfinding process.

5.5.3 Navigating with AR Assistant

After arranging the route, how to visually show the route to the user and give guidance to the user as they move forward is a content worthy of consideration and research.

The traditional paper texture map statically displays a two-dimensional map of an area, and the user needs to judge his current location and find a route to the destination. Today's electronic map applications can display the user's current location and highlight the route to the destination. Some applications, such as Google Maps, can even provide AR map functions outdoors. But currently they cannot provide indoor navigation services. We can absorb their advantages and use them in our own system.

In real life, when a person enters a building for the first time, people who are familiar with the building can guide and give directions. Human guides can provide good instructions and companionship, but providing guidance services to everyone will lead to huge human resources costs, which is not very realistic. We can combine the advantages of the current electronic map and the actual manual guide: the electronic map can highlight the forward route on the map; the human guide can walk in front of the user to play a leading role.

The system can display the anchor point in the forward path according to the user's forward route, and change the state of the anchor point after passing the anchor point. First, the system will determine the user's current location based on environmental information. Then, after determining the location, the system uses a color to highlight all the anchor points on the way to the destination location. These anchor points can play a role of visual guidance. At the same time, we adopt an arrow that follows the user as a visualization Navigation clues. The figure 5.12a lists the navigation clues. The user keeps moving forward following the navigation and will pass through one anchor point. These anchor points will change state after the user passes by, indicating that the position has been passed. Comparing figure 5.12b to figure 5.12c, we can find that the color of some anchors are changed, which means user has passed these anchors.

Our AR assistant mainly acts as the entrance and guide of the navigation system in navigation tasks. When users arrive in a new indoor environment and need destination navigation services, they can interact with the AR assistant by voice to enable route navigation.

(a) Navigation clues

(b) Start Navigation

(c) Pass several anchors

Fig. 5.12 AR indoor navigation

At the same time, the virtual avatar representing the guide can be combined with a directional arrow to guide the user to the end. The figure 5.13 is the example of using virtual avatar as guide when the system navigating the user to destination. When the user reaches the destination, the virtual avatar can give voice and body language feedback, indicating that the user has successfully reached the destination.

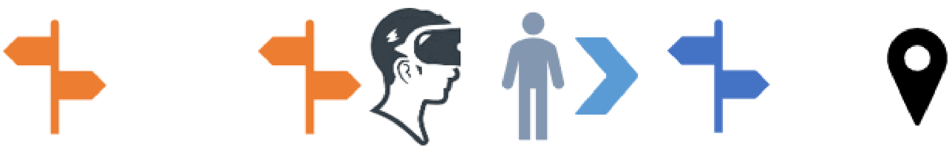


Fig. 5.13 Using avatar when navigating

Since the user is not familiar with the environment in which navigation is required, if the user's own avatar is used as a guide to provide route navigation services, it does not fit the actual situation. Therefore, we can choose to use an avatar of a person familiar with this route as an assistant at this stage.

5.6 Feedback Design

After giving the user guidance, the system can recognize the user's task completion and give corresponding voice and body language feedback to play a better motivational effect.

5.6.1 Task Completion Detection

Detecting the user's task completion is an important prerequisite for providing targeted feedback. We have designed two main information collection methods.

The first way is to monitor user actions and determine whether a series of actions have occurred during this period of time. The system detects whether a touch event occurs between the user's hand and the target object.

If the touch events occur sequentially in a certain order within a short period of time, it can be considered that the user completed the task on time. The figure 5.14 shows an example. Action 1 and Action 2 are within two dots, indicating that these two actions are within a period of time.

Fig. 5.14 A series of actions occurred during a period of time

In the second way, users actively provide feedback to the system by speech. In some cases, it may be difficult for the system to monitor or judge whether the behavior has occurred and whether the task is completed. At this time, the user can actively feedback the progress to the system. This way can be regarded as a supplement to the first way.

5.6.2 Feedback Interaction Design

After the system knows the completion of the user's task, the system will also give targeted feedback, mainly voice feedback and body language feedback.

Fig. 5.15 User actively reports the completion of the task

Feedback can be mainly divided into positive feedback after the user completes the task, and negative feedback after the user missed the task or failed. In the positive feedback, the virtual avatar can say some words of encouragement and praise to the user, accompanied by positive body movements. In the negative feedback, the avatar will criticize or spur users. Regardless of positive or negative feedback, the purpose is to motivate users through the behavior of the avatar, hoping to successfully complete the schedule task next time.

5.7 Scenarios

In figure 5.1 we show the framework of the system, which is mainly the main component of the system. In order to show our system better and more intuitively, we have designed a set of typical application scenarios. In this set of scenarios, we considered environmental factors such as time, location, objects and their states, as well as human factors such as mission partners.

The things people have to do next and their schedules will differ depending on the time and place. In the example scenario, we will show how the system functions in the user's day's work and life. Here, the user is a freshman who has never entered IPLAB. In order to use this system, task scheduling and schedule planning in different situations need to be set in advance. This system focuses on identifying context and providing schedule assistant services accordingly.

5.7.1 Take Medicine

Pre-set task: take medicine in warm water at home in the morning

During the user's use of the system, the system will detect the current context in real time and query the preset tasks that match its status.

In the morning, the user puts on the device HoloLens in his room, and the system detects that the current context matches the requirements of the pre-set medication task, so it triggers a reminder to take the medication schedule.

After confirming the arrangement, the system will call the avatar to start reminding and guiding services, and interact with users. The figure 5.16 shows the main processed in this schedule:

- AR assistant appears with the voice of reminding user to take medicine.
- After getting the remind message, when the user sees the medicine, the system shows information on how to take the medicine.
- After taking medicine, the system detects the behavior and gives feedback: AR assistant pose first pump and praise user's task completion

The figure 5.17 shows the scene that display guidance on taking medicine with pictures and avatar actions. The picture in the mid-air provide the message on how to take medicine and the appearance of the medicine bottle, and the avatar will give a motion on taking medicine with drinking movement.

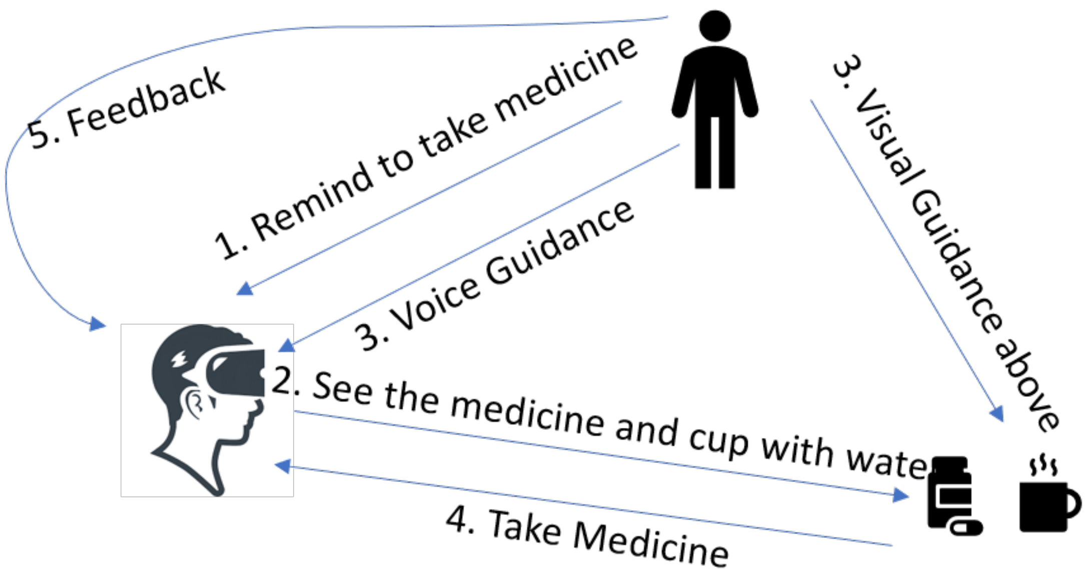


Fig. 5.16 Schedule: remind and guide to take medicine

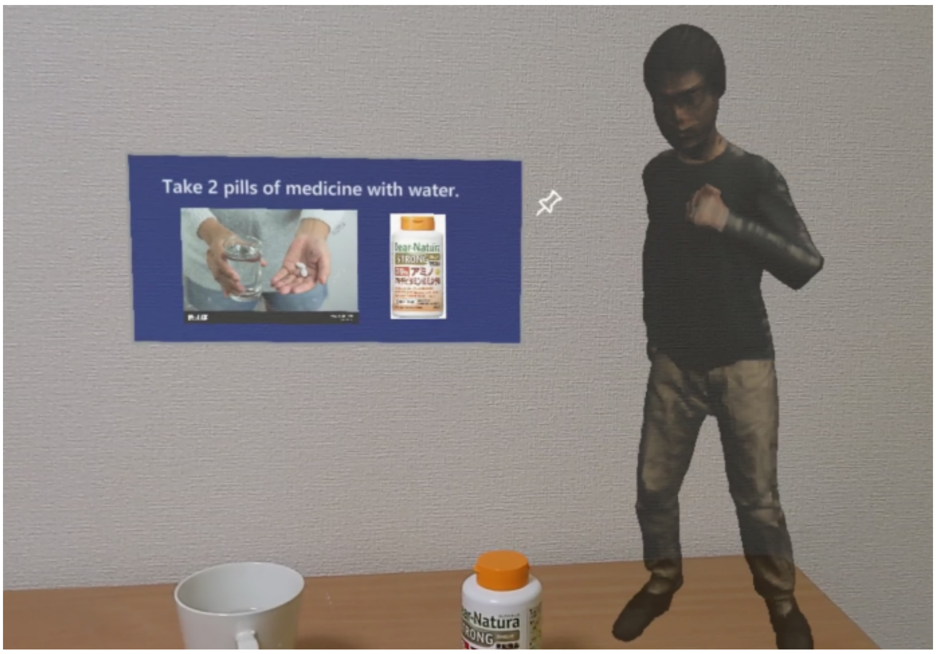


Fig. 5.17 Guidance on taking medicine

5.7.2 Participate in a Meeting

Pre-set task: Participate in a meeting in IPLAB this afternoon. If it rains, depart half an hour earlier and bring umbrella.

Adjustment of Schedule According to Context

In this scenario, the system first needs to adjust and determine the schedule based on weather state information.

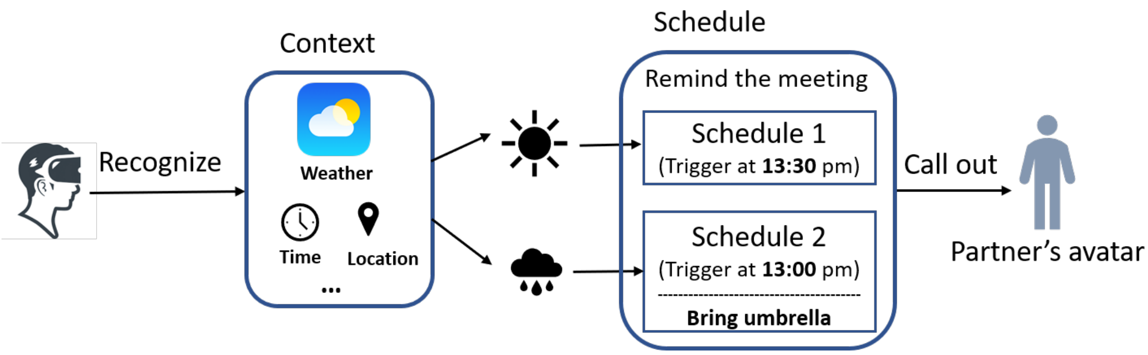


Fig. 5.18 Context-Aware Schedule: remind meeting

When the weather is fine, you only need to remind the user that there will be a meeting in the afternoon at the preset time and display relevant information about the meeting. When it rains now or in the future, the system needs to adjust the schedule, remind users to leave in advance, and also remind users not to forget to bring an umbrella.

(a) Schedule 1 (b) Schedule 2

Fig. 5.19 Remind task in two context

Fig. 5.20 2D interface about the meeting information

At the same time, in order to better play the reminder role of the schedule assistant, the system uses the avatar of another participant to remind tasks. Compared with the user's own avatar, the avatar of the other attendees can play a stronger role in motivating and supervising. The figure 5.18 shows the scene which happens at home of the user. The figure 5.19 shows the result in real environment. The figure 5.19b shows when the weather will be rainy, the avatar will remind not only the meeting information which is just like listed in figure 5.20, but also remind user to bring umbrella.

Indoor Navigation to IPLAB

The user arrived at the Graduate School of IPS college, but he didn't know where IPLAB. At this time, the system continues to serve as the role of the assistant, and it will take the user to the destination.

The figure 5.21 shows shows the main processed in this schedule:

- User ask the virtual assistant the route to IPLAB;
- The system recognize where the user is and query the route utilizing Azure spatial anchor information
- System get route and the avatar guide user to destination

Since the user does not know how to get to the destination, it would be a bit out of place to use the user's own avatar at this time, so the avatars of other participants who are familiar with the route continue to be used as schedule assistants to provide guidance services.

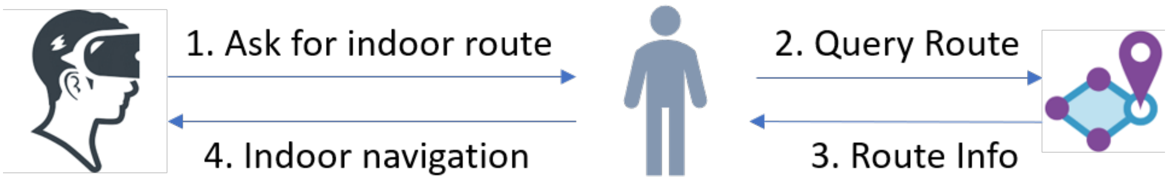


Fig. 5.21 Navigate user to destination

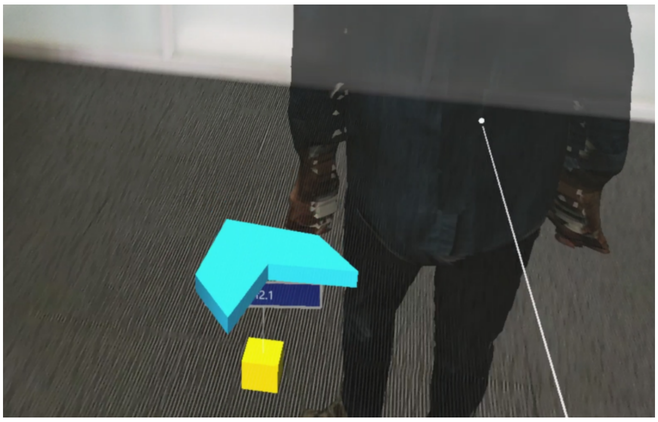


Fig. 5.22 Avatar guide user to destination

Chapter 6

System Implementation

6.1 Hardware Setup

To achieve the AR Scheduling Assistant system, we use Microsoft HoloLens 2 as our device, which can understand the surrounding environment with mixed reality display.

Through the Microsoft HoloLens 2 as figure 6.1 shows, we can display the avatar as virtual assistant in the midair and then the user with this Head Mounted Device (HMD) can interact with the virtual avatar and virtual objects.

(a) Front

(b) Side

Fig. 6.1 Microsoft HoloLens 2

To develop our system, we need a computer with the Windows 10 operating system to support HoloLens development. The configuration of the computer we use is shown in the Table 6.1.

Table 6.1 The information of PC

Category	**Information**
Operation System	Microsoft Windows 10
CPU	Intel(R) Core(TM) i7-6500U CPU @2.5GHz 2.59GHz
RAM	8 GB

6.2 Develop Environment

We established the system and built the 3D module on Unity 2019 and use the C sharp as the development language in Visual Studio 2019 Version 16.9.4. We also use the PIFuHD and blender Version 2.83 to generate the virtual model as the avatar. And some other software environment supports are shown as follows:

- Mixed Reality Toolkit (MRTK) v2.5.4: MRTK is a cross-platform development kit for mixed reality applications provided by Microsoft. We use its basic input methods, spatial awareness function to get environment information.

- Vuforia SDK: Vuforia can enhance the ability to recognize the 3D object and environment. We use Vuforia Engine to recognize the ground plane and build AR experience for mobile devices.

- Azure Spatial Anchors (ASA) v2.7.0: With the help of ASA, we can build mixed reality applications with spatial context. And the spatial anchors can be stored in Azure Cloud.

- Azure Storage: The Azure Storage platform is Microsoft's cloud storage solution for data storage scenarios. In order to access the context data easily, we choose Azure Storage to store the information of pre-set tasks and the spatial anchors.

6.3 Personalized Avatars Generation

The generation of avatars is an important part of the system. The human avatar will act as a virtual schedule assistant in the subsequent sessions, just like a real person, reminding, guiding, and interacting with users.

In order to generate a realistic and anthropomorphic human avatar, the method we adopted is mainly divided into two parts:

- first, generate a static avatar of the whole body from a single full body image;
- then, add some common and life-like actions and movements to the avatar, making it more than just a "static statue".

6.3.1 Static Avatar Generation

In order to generate an personalized human avatar, we must firstly reconstruct the outline of the person. Our goal is to quickly rebuild a suitable model which can be used. After comparison and try among several different technologies which have been mentioned in related works, we adopted Facebook's PIFuHD method to reconstruct the 3D human shape.

Using this method, we can complete the body shape reconstruction work with only one photo with human's full body. The figure 6.2 shows the process, and the right three images are different views of the generated model.

Next, we need to add texture to the reconstructed model. Since the model is generated from a single image, we can directly map the texture in the image to the appearance of the model. We use Blender to complete this operation: the front of the model directly uses the texture mapping in the picture, and the back of the model can be adjusted according to the actual situation. The figure 6.3 shows the process and the final result we can get.

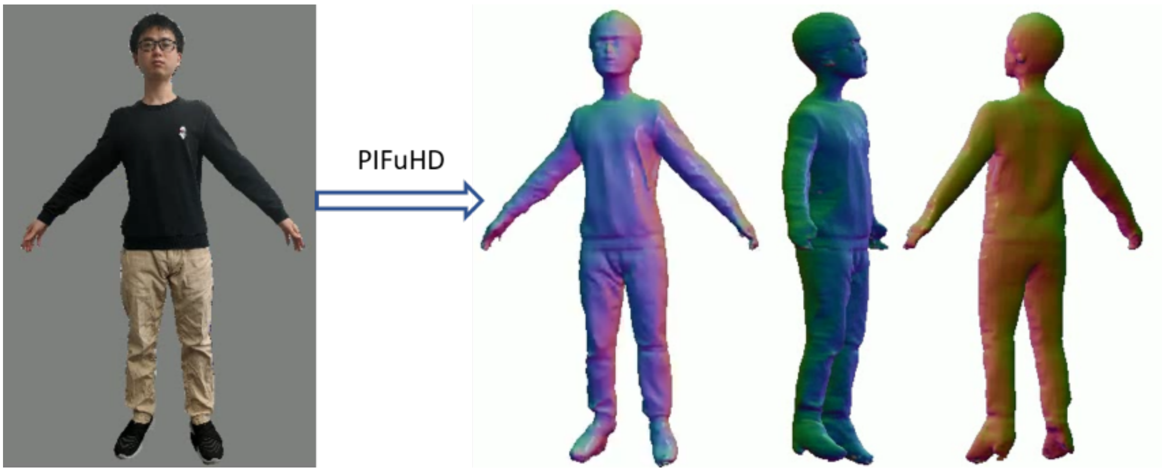


Fig. 6.2 3D human shape reconstruction through PIFuHD

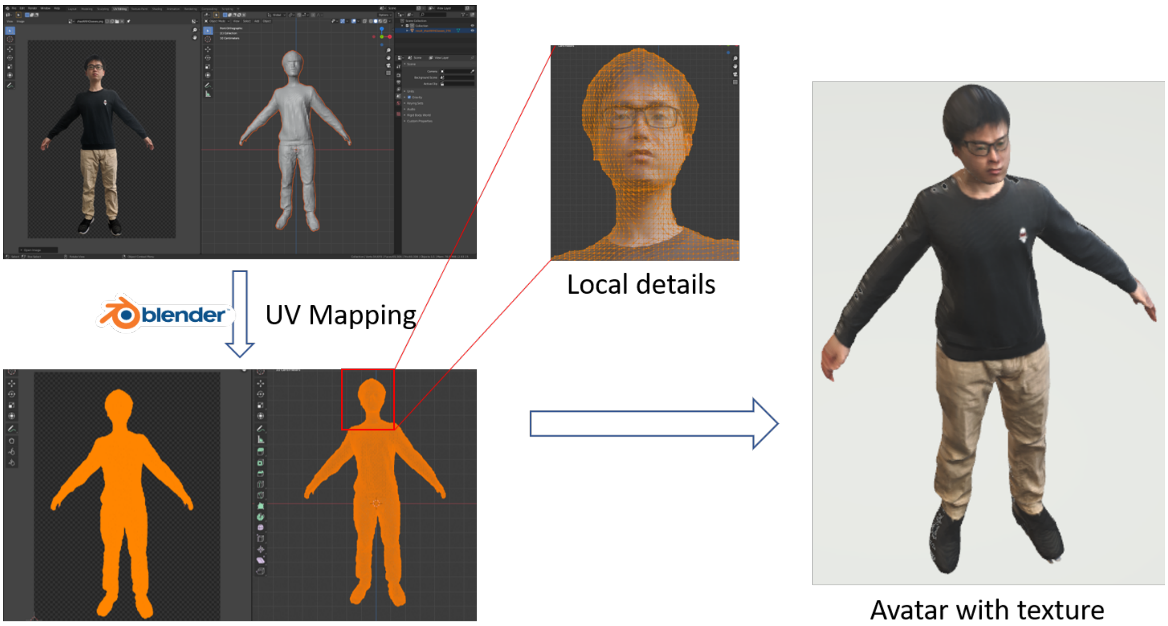


Fig. 6.3 Add texture from image using UV mapping

6.3.2 Rigging Avatar and Making Animation

A static avatar cannot give feedback on body movements when interacting with the user, which will also reduce the realism of the avatar.We use Mixamo from Adobe to bind bones and create animations, which is a web-based service for creating 3D character animations.

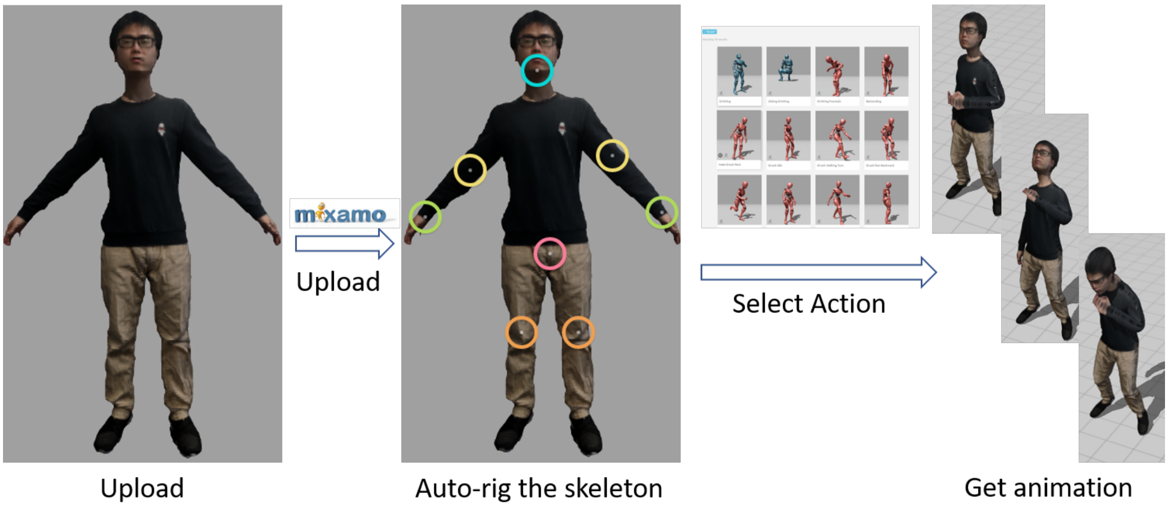


Fig. 6.4 Generate animation

We uploaded the textured avatar to Mixamo [35] and then rigger the skeleton to this avatar. Mixamo's technology uses machine learning methods to automate the steps of character animation processing, from 3D modeling to binding and 3D animation. After that, Mixamo will automatically bind the skeleton. We can select actions suitable for the scene from Mixamo's rich action library and generate animations. The figure 6.4 shows the process.

6.4 Context Recognition and Matching

In figure 5.1, we divide the system into two parts. In the second part, we said that the system can recognize the user's current context, arrange schedules and tasks according to different contextual information, and call a suitable avatar as a virtual assistant.

Recognizing the context is the starting point and focus of the second part. In the system design chapter, we explained several points of the system, and now we explain the technical details of its implementation.

6.4.1 Context Recognition

Context recognition is the first step of context awareness. Only when the current environment is known can targeted measures be taken.

Location and Time

Time and location are the basic information in the context and the most important information in the schedule. Here we explain how to identify the temporal and spatial environment the user is currently in.

For the user's current time context, we can query and use the current system time of the device. Since the system is networked, it is also network time.

For the location context information, in order to determine the user's location in the world environment, we use the GPS in the device to obtain latitude and longitude data.

For the specific location of the user indoors, we can use the spatial perception function of HoloLens 2. Due to the inclusion of an RGB camera and a depth camera, HoloLens 2 can more accurately know the relative position of the user in the indoor environment. The figure 6.5 is an example of the local location context. The white ball is the user's head and the user is now sitting in IPLAB. The green line is the field of view of the device.

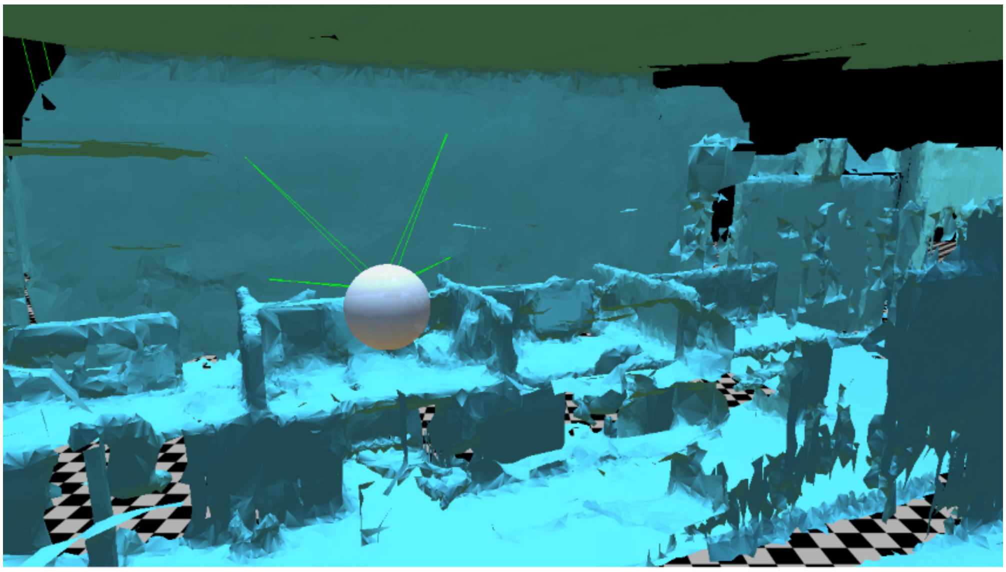


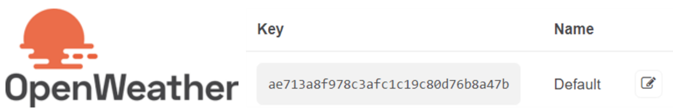
Fig. 6.5 Local location information from HoloLens 2

Weather

Weather information is a non-negligible factor that affects the schedule. The difference in weather may cause changes in the schedule, and may also cause cancellation or change of the schedule.

The weather conditions in the current time and space and the weather forecasts in the future are all worthy of consideration. To obtain this information, we use the weather service from OpenWeather.

The first is to register an account and generate your own api key.



(a) OpenWeather (b) API key

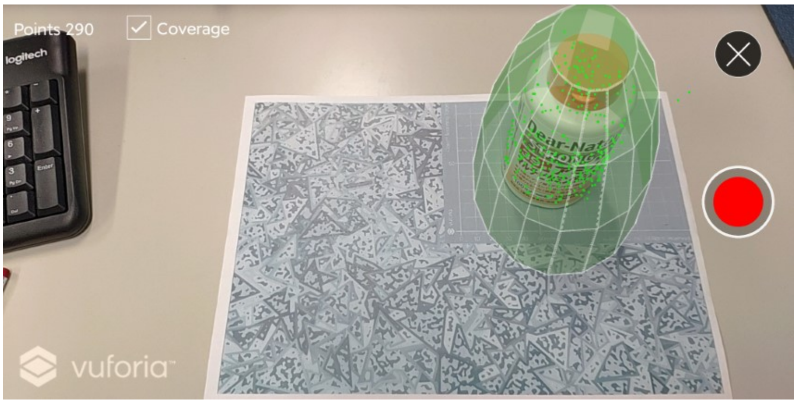
Fig. 6.6 Get API key from OpenWeather

By using our own API key, we can query current weather information and future forecast data. We can use the UnityWebRequest to query the message from official web. The service returns data in JSON format. We use Newtonsoft.JSON to deserialize the data and analyze the data.

3D Object

Objects in the environment are also contextual information of the system. On the one hand, it can affect the schedule; on the other hand, when the virtual assistant is conducting task guidance, it can combine the identified objects to give targeted guidance.

We use Vuforia to realize our object recognition function. First, we need to scan the object in 3D to get a file in .od format. We upload the file to get the target ID, and then download and apply the recognition database to our application to realize the object recognition function. The figure 6.7 shows part of the leading steps for 3D object recognition using Vuforia: scan the object and then get Object Target Databases generated by Vuforia. The figure 6.8 shows the recognition of medicine bottle and cup in real environment.



(a) Scan the object

Type: Object

Status: Active

Target ID: 6e8745e92b2744bba87ccd08a5b28e60

Added: May 14, 2021 08:28

Modified: May 14, 2021 08:29

(b) Object info

Fig. 6.7 3D object recognition using Vuforia

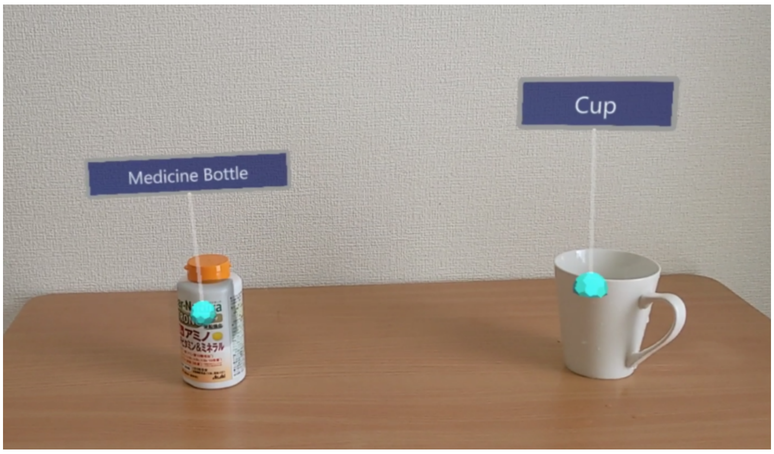


Fig. 6.8 Recognition of medicine bottle and cup

6.4.2 Context Matching

The user sets or determines his own goals and tasks in advance, and decides under what circumstances to let the system remind and guide himself to start the tasks.

What we are matching here is the current context environment and the context preset by the user. When the user's current context meets certain specifications, the user will be arranged with a suitable schedule, and then the virtual assistant will be "awakened" to provide services.

The preset task information can be recorded in Microsoft Azure Table Storage. It can provide online NoSQL database service.

6.5 Interactions Between User and Avatar

In order to make the avatar more anthropomorphic and provide a better sense of realism, we provide voice interaction and non-verbal interaction (the body movements of the avatar) between the user and the avatar.

6.5.1 Voice Interactions

Let the Avatar Speak

Letting the avatar speak will provide a more natural interactive experience. We can use Azure speech to convert between speech and text. As shown in the figure 6.9, we will write the SubscriptionKey and ServiceRegion parameters applied from the official website, and then we can set the text information we want to replace. Other similar text-to-speech services are also available. Next, we can assign the voice to the avatar, then he can "speech".

```
// Creates an instance of a speech config with specified subscription key and service region.
// Replace with your own subscription key and service region (e.g., "westus").
// The default language is "en-us".
var config = SpeechConfig.FromSubscription("eb7cdd3189ea4be28653df7a67d47d19", "eastus");

// Creates a speech synthesizer using the default speaker as audio output.
using (var synthesizer = new SpeechSynthesizer(config))
{
    // Receive a text from console input and synthesize it to speaker.
    Console.WriteLine("Ok, I will show the Guidance.");
    Console.Write("> ");
    string text = Console.ReadLine();

    using (var result = await synthesizer.SpeakTextAsync(text))
```

Fig. 6.9 Convert text to Speech

Speech Recognition

In order for the avatar to "understand" and respond to the user's words, the system needs to recognize the user's speech. We can turn on the Speech Input option in MRTK, enter keywords or key sentences, and turn on the system's microphone permissions. In this way, when the user speaks and triggers the set keywords, the avatar and system response can be triggered.

6.5.2 Avatar Motion Control

We added actions to the avatar in the Rigging Avatar and Making Animation section, such as drinking water. When the user interacts with the avatar, the system can recognize the user's voice commands or actions, so that the avatar's actions can be switched to achieve more natural interaction and feedback.

We can use Unity's Animator Controller component to organize and switch avatar actions. When the input meets certain conditions, the state can be switched between different actions. The figure 6.10 shows a controller interface. Different states can be connected by arrows, indicating that an action can be transformed into an arrow-pointing action under certain conditions.

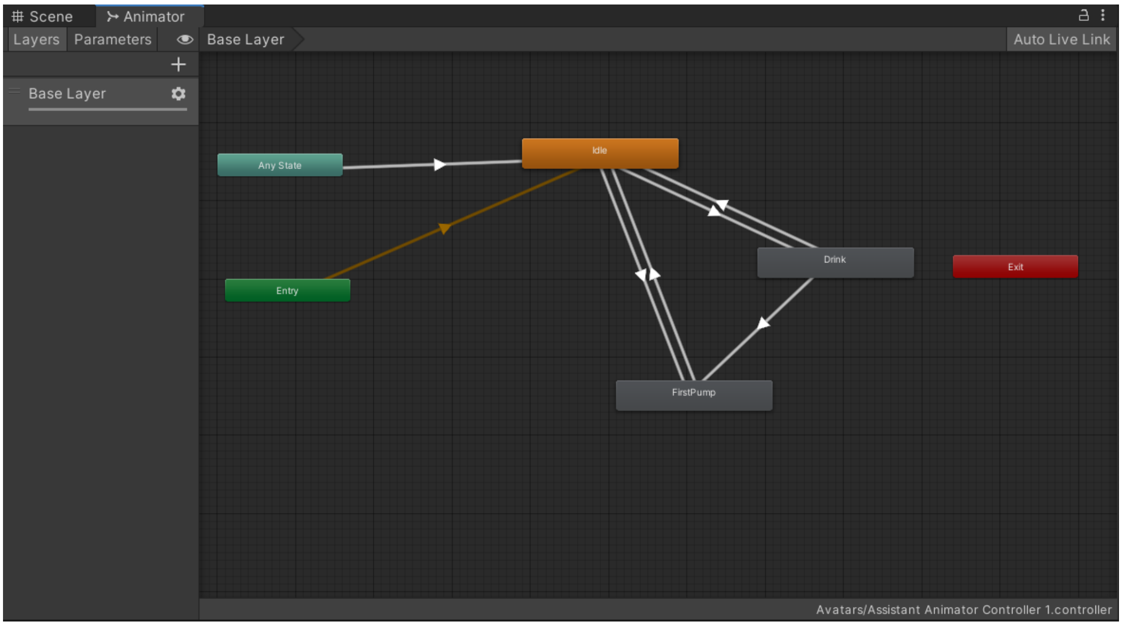


Fig. 6.10 Animator controller

6.5.3 User-Object Touch Detection

We need to detect the user's actions to determine the completion of the user's task. In order to meet the requirements that are simple and fast, only HoloLens can be used.

We design and apply the method to detect the user's touch action on the items. When the user implements the action of touching items in a certain order at a certain time, we can determine the corresponding task schedule completed by the user.

In the previous article, we mentioned the use of Vuforia to realize object recognition. In order to detect the touch action, we automatically add a transparent container with a collision body to the recognized object.

We also add the NearInteractionTouchable component from MRTK to the user's hand. When the user touches the recognized object with his hand, the two collide, and the system judges that the touch has occurred. The figure 6.11 gives an example on touch the medicine bottle action, when user touches the bottle, the system recognizes this event, and said the the object is touched.

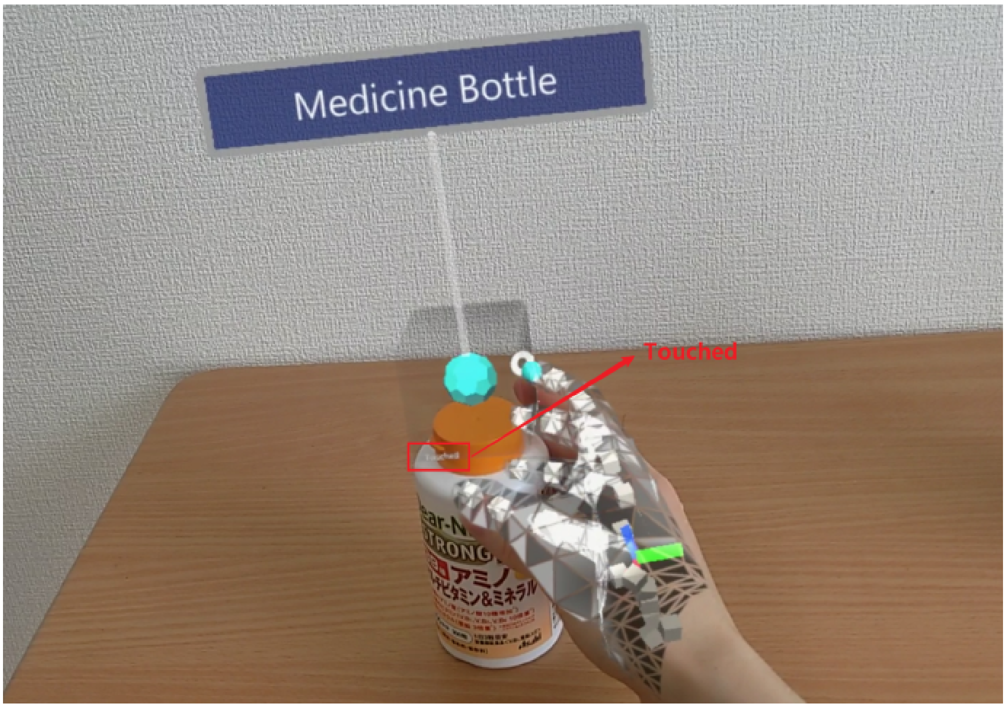


Fig. 6.11 Touch the medicine bottle (the red annotation is not displayed in HoloLens)

6.6 Assistant's Indoor Navigation

When users do not know how to reach their destination, we can provide indoor navigation service. The navigation service consists of two parts: setting the route and navigation. The user directly uses the navigation part.

We mainly use two data structures to store information related to the spatial anchor.

- Spatial anchor mainly contains anchor-related data information, including name, ID, and anchor position.
- The Edge contains the information of the edges in the graph composed of spatial anchors, and its edge length represents the spatial distance between two adjacent nodes.

Spatial anchor and Edge information are stored in azure table storage in the form of tables, we can use API key to query and change these information through the network.

6.6.1 Set the Anchors

The main operation of route setting is to place a spatial anchor in the space, and store the anchor and edge information in the Azure Table Storage database.

First of all, the system queries the database whether there is a data table, if not, create an empty table, then you can directly place the anchor, if there is, select an anchor as the starting point to maintain the connectivity of the route.

After placing these anchors in the space, we register them in Azure Spatial Anchor to obtain a durable and shareable experience. After successful registration, we can get a unique SpatialAnchorID for each anchor. The figure 6.12 shows the data format.

At the same time, since we set that every anchor must have at least one adjacent anchor connected to it, we can get the edge information of the graph composed of all anchors created. The figure 6.13 shows the data format.

We store the Spatial anchor data and edge data obtained in the Azure Table Storage cloud storage service, so that users can query the anchor data in other machines.

PartitionKey	RowKey	AnchorPosition	Name	SpatialAnchorId
main	f2.0	(-0.2, -1.4, 2.6)	f2.0	3cfb3a02-9878-49a1-9594-641caa315306
main	f2.1	(0.1, -1.4, 5.3)	f2.1	9e93f6e2-24dc-43a6-89d2-a3fff95db247
main	f2.2	(-1.8, -1.4, 6.8)	f2.2	fc47e2f4-e8a9-457e-aa5a-af4750ecafb0
main	f2.3f	(-4.0, -1.4, 6.9)	f2.3f	92138cd3-740e-48c8-9cb8-c4d1783cc961
main	f2.4	(-6.3, -1.5, 7.0)	f2.4	269e8db7-4c0f-4a57-a30e-a16b27f392b3
main	f2.5	(-8.9, -1.5, 7.5)	f2.5	f5f8a46f-cd9f-4188-9007-ff0abc7eb7fc
main	f2.6	(-11.5, -1.5, 7.9)	f2.6	0ec871ce-a424-4226-a6f8-643465b56d16
main	n258.door2n	(-13.5, -0.7, 7.0)	n258.door2n	0a4e283e-611f-4e8c-8109-3e7166c5293f

Fig. 6.12 Spatial anchors data

RowKey	Name	SpatialAnchorId	ConnectedName	ConnectedSpatialAnchorId	Distance
f2.0,f2.1	f2.0	3cfb3a02-9878-49a1-9594-641caa315306	f2.1	9e93f6e2-24dc-43a6-89d2-a3fff95db247	2.75558161735535
f2.1,f2.2	f2.1	9e93f6e2-24dc-43a6-89d2-a3fff95db247	f2.2	fc47e2f4-e8a9-457e-aa5a-af4750ecafb0	2.46606802940369
f2.2,f2.3f	f2.2	fc47e2f4-e8a9-457e-aa5a-af4750ecafb0	f2.3f	92138cd3-740e-48c8-9cb8-c4d1783cc961	2.21369409561157
f2.3f,f2.4	f2.3f	92138cd3-740e-48c8-9cb8-c4d1783cc961	f2.4	269e8db7-4c0f-4a57-a30e-a16b27f392b3	2.31574511528015
f2.4,f2.5	f2.4	269e8db7-4c0f-4a57-a30e-a16b27f392b3	f2.5	f5f8a46f-cd9f-4188-9007-ff0abc7eb7fc	2.58192181587219
f2.5,f2.6	f2.5	f5f8a46f-cd9f-4188-9007-ff0abc7eb7fc	f2.6	0ec871ce-a424-4226-a6f8-643465b56d16	2.60301399230957
f2.6,n258.door2n	f2.6	0ec871ce-a424-4226-a6f8-643465b56d16	n258.door2n	0a4e283e-611f-4e8c-8109-3e7166c5293f	2.37327885627747

Fig. 6.13 Edge data of spatial anchors

6.6.2 Wayfinding

Wayfinding, or navigation, is a function directly used by users. When users arrive at an unfamiliar indoor scene, they need to determine the route to the destination, and our system provides assistant navigation.

Users can communicate with virtual assistants and seek navigation assistance. We use MRTK's speech command to identify the user's words. After the virtual avatar receives the word from the user, it recognizes it, triggers the keyword and opens the following event.

The first stage is the preparation stage. At this stage, the system will query the spatial anchor table from Azure Table Storage. After the system receives the anchor list returned by the Azure cloud service, the user selects the starting point and destination. According to the information of the two endpoints and the anchor map, we calculate the path according to the shortest path algorithm, and get the id of the anchor on the path. Next, we query these ids in the Azure Spatial Anchor service and load these spatial anchors into the real scene. The figure 6.14 shows the workflow.

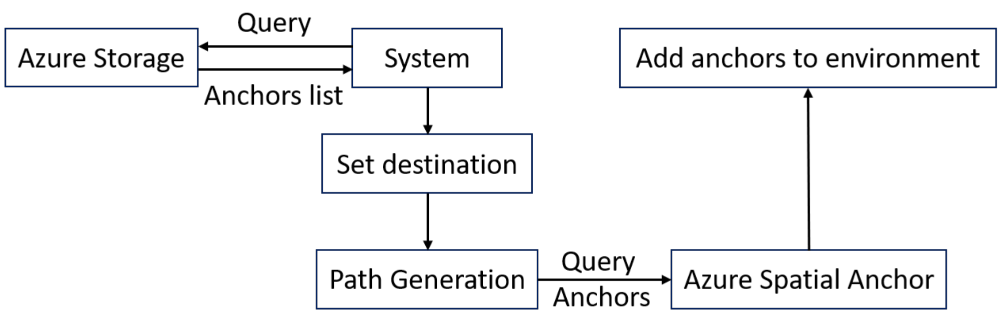


Fig. 6.14 Preparation stage

When the anchors have been loaded into the scene, the system starts to navigate. These anchors also play a role in judging the user's current location in the environment. We add collision bodies to these anchors. When the user contacts the collision body, the system will determine that the user has passed the position and lead the user to the next anchor. In the process of the user's foreground, there will be arrows indicating the direction, and at the same time there will be a virtual avatar standing at the anchor point in front, indicating "come here".

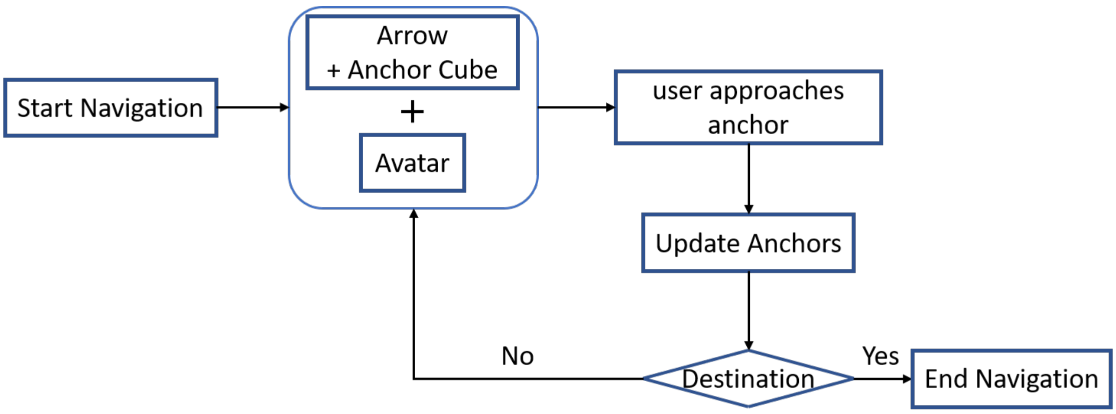


Fig. 6.15 Navigation stage

The last stage is to reach the end and end the navigation. After the user arrives at the destination, the system will pop up fireworks special effects, and the virtual avatar will also give feedback: applauding while saying "Congratulations, you have arrived at the destination".

Chapter 7

Conclusion and Future Work

7.1 Conclusion

In this thesis, we design and implement an AR scheduling system, which could be context-aware and use personalized avatar as virtual assistant.

In general,our system provided users with intuitive and efficient guidance and instructions based on the user's context. We use augmented reality technology to display personalized human avatars as virtual assistants in a realistic environment and interact with the user.

The personalized avatars can have the same body shape and appearance as the full body photo provided by the user. To make the avatars interact more naturally with the user, we add movement and voice to them. The avatars act as virtual assistants in the system, providing better reminders, guidance and motivation to the user.

Our system recognizes the user's current context, such as time, location, weather, objects and their status information. Depending on the preset tasks and the current scenario, the system can generate different schedules. In the scenario recognition stage, we need to use the device sensors and recognition methods to identify and judge the current situation. After arranging the schedule according to the scenario, the system will call out the corresponding incarnation as an assistant. The assistant can give guidance to the user through voice and action and give feedback based on the user's behavior. The system can also provide spatial anchor-based indoor navigation in situations where the user does not know the indoor route.

We also apply the proposed context-aware schedule assistant system to several typical daily scenarios. One is to remind and guide user to take medicine at home. We also consider a scene that remind user to participant in a meeting at school when he is at home, in which the reminder time and content may vary depending on the weather. And the system also provide user with indoor navigation when user arrive at school, where he is not familiar with.

7.2 Future Work

Although we have proposed a system that can act as an intuitive scheduling assistant, there are still some things worth to research.

We can consider taking more contextual information into account, such as connecting to the IoT devices to get more data. We can also consider more application scenarios to accommodate the many different needs of people's schedules.

References

[1] Max V Birk, Cheralyn Atkins, Jason T Bowey, and Regan L Mandryk. Fostering intrinsic motivation through avatar identification in digital games. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 2982–2995, 2016.

[2] 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.

[3] S Nate, P Navele, V Mote, and Laxman S Naik. Smart reminder application with gps system. *International Research Journal of Engineering and Technology (IRJET)*, 3(4):131–134, 2016.

[4] Nur Rokhman and Lubab Saifuddin. Location and time based reminder system on android mobile device. In *2016 2nd International Conference on Science in Information Technology (ICSITech)*, pages 147–151. IEEE, 2016.

[5] Mei-Ying Wang, PH Tsai, Jane W-S Liu, and John K Zao. Wedjat: a mobile phone based medicine in-take reminder and monitor. In *2009 Ninth IEEE International Conference on Bioinformatics and BioEngineering*, pages 423–430. IEEE, 2009.

[6] 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.

[7] Ronald T Azuma. A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, 6(4):355–385, 1997.

[8] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, volume 2351, pages 282–292. International Society for Optics and Photonics, 1995.

[9] Xuan Wang, Chunmeng Lu, Soh Masuko, and Jiro Tanaka. Interactive online shopping with personalized robot agent. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*, pages 1–2, 2018.

[10] 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.

[11] 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.

[12] Nadia Magnenat-Thalmann and Daniel Thalmann. Virtual humans: thirty years of research, what next? *The Visual Computer*, 21(12):997–1015, 2005.

[13] Thomas Waltemate, Dominik Gall, Daniel Roth, Mario Botsch, and Marc Erich Latoschik. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE transactions on visualization and computer graphics*, 24(4):1643–1652, 2018.

[14] Alexandru Eugen Ichim, Sofien Bouaziz, and Mark Pauly. Dynamic 3d avatar creation from hand-held video input. *ACM Transactions on Graphics (ToG)*, 34(4):1–14, 2015.

[15] 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.

[16] Gediminas Adomavicius and Alexander Tuzhilin. Context-aware recommender systems. In *Recommender systems handbook*, pages 217–253. Springer, 2011.

[17] Frada Burstein, Arkady Zaslavsky, and Nitin Arora. Context-aware mobile agents for decision-making support in healthcare emergency applications. In *International Workshop on Context Modeling and Decision Support: 05/07/2005-05/07/2005*. CEUR Workshop Proceedings, 2005.

[18] Bin Guo, Ryota Fujimura, Daqing Zhang, and Michita Imai. Design-in-play: improving the variability of indoor pervasive games. *Multimedia Tools and Applications*, 59(1):259–277, 2012.

[19] Jens Grubert, Tobias Langlotz, Stefanie Zollmann, and Holger Regenbrecht. Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE transactions on visualization and computer graphics*, 23(6):1706–1724, 2016.

[20] Enes Yigitbas, Ivan Jovanovikj, Stefan Sauer, and Gregor Engels. On the development of context-aware augmented reality applications. In *IFIP Conference on Human-Computer Interaction*, pages 107–120. Springer, 2019.

[21] J Zhu, Soh-Khim Ong, and Andrew YC Nee. A context-aware augmented reality assisted maintenance system. *International Journal of Computer Integrated Manufacturing*, 28(2):213–225, 2015.

[22] Shahriar Nirjon, Jie Liu, Gerald DeJean, Bodhi Priyantha, Yuzhe Jin, and Ted Hart. Coin-gps: Indoor localization from direct gps receiving. In *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*, pages 301–314, 2014.

[23] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer. Indoor human navigation systems: A survey. *Interacting with Computers*, 25(1):21–33, 2013.

[24] Chouchang Yang and Huai-Rong Shao. Wifi-based indoor positioning. *IEEE Communications Magazine*, 53(3):150–157, 2015.

[25] Mahmut Aykaç, Ergun Erçelebi, and Noor Baha Aldin. Zigbee-based indoor localization system with the personal dynamic positioning method and modified particle filter estimation. *Analog Integrated Circuits and Signal Processing*, 92(2):263–279, 2017.

[26] Adam Satan and Zsolt Toth. Development of bluetooth based indoor positioning application. In *2018 IEEE international conference on future IoT technologies (Future IoT)*, pages 1–6. IEEE, 2018.

[27] Alwin Poulose and Dong Seog Han. Hybrid indoor localization using imu sensors and smartphone camera. *Sensors*, 19(23):5084, 2019.

[28] Jesús Gimeno, Cristina Portalés, Inmaculada Coma, Marcos Fernandez, and Bibiana Martínez. Combining traditional and indirect augmented reality for indoor crowded environments. a case study on the casa batlló museum. *Computers & graphics*, 69:92–103, 2017.

[29] Patrick Hübner, Martin Weinmann, and Sven Wursthorn. Marker-based localization of the microsoft hololens in building models. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42(1), 2018.

[30] Umair Rehman and Shi Cao. Augmented-reality-based indoor navigation: A comparative analysis of handheld devices versus google glass. *IEEE Transactions on Human-Machine Systems*, 47(1):140–151, 2016.

[31] Xiaochen Zhang, Xiaoyu Yao, Yi Zhu, and Fei Hu. An arcore based user centric assistive navigation system for visually impaired people. *Applied Sciences*, 9(5):989, 2019.

[32] Patrick Herbers and Markus König. Indoor localization for augmented reality devices using bim, point clouds, and template matching. *Applied Sciences*, 9(20):4260, 2019.

[33] Patrick Hübner, Kate Clintworth, Qingyi Liu, Martin Weinmann, and Sven Wursthorn. Evaluation of hololens tracking and depth sensing for indoor mapping applications. *Sensors*, 20(4):1021, 2020.

[34] Microsoft Azure. Spatial anchors, 2021. https://azure.microsoft.com/en-us/services/spatial-anchors/, Last accessed on 2021-06-25.

[35] Adobe. Mixamo, 2021. https://www.mixamo.com/, Last accessed on 2021-06-25.