Magic Tour: An AR System Combined With Tangible Maps for Sharing Navigation

FU QIDONG

44191623-5

Master of Engineering

Supervisor: Prof. Jiro Tanaka

Graduate School of Information, Production and Systems

Waseda University

September 2021

Abstract

In recent years, pedestrian navigation has gradually become one of the most important services in people's lives. And with the popularity of smart phones, there are more and more digital map AR navigation based on smart phone platforms. However, studies have shown that only using digital maps and AR view for navigation has gradually exposed many problems. The tangible physical map plays and continues to play an important role in contemporary cartography. The tangible physical map based on the cartographic tradition and enhanced through new digital interactive forms can supplement and advance the digital cartography presented on the computer screen.

Therefore, we design and developed a mobile AR system combined with tangible maps for sharing navigation. On the basis of retaining tangible maps, we use augmented reality technology to make up for the shortcomings of paper maps that cannot update information in real time, transform the key information contained on the map from a 2D plane to a 3D model, and introduce the GPS signal of the smartphone as Positioning basis, real-time display of the user's relative position on the tangible map.

In addition, the system allows the user to freely choose the destination, and can display the navigation path on the screen of the smartphone in the form of augmented reality. Taking into account the user's usage scenario, we designed the system to support multi-person sharing, and users in the same group can view each other's location and travel route. We hope to help users have a better interactive map navigation experience through this system.

Keywords: Augmented Reality; tangible map interaction; positioning; sharing navigation

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor, Prof. Jiro Tanaka, who provided me much guidance and valuable advice on my research. It's a great honor to have him led me into the field of HCI. After two years of study, I deepened my understanding of human-computer interaction. It is his patience and kindness that helped me keep making progress on my study. Without him, my master research would not have gone that smoothly during these days.

Secondly, I want to thank all IPLAB members. I am very happy to meet you and have spent precious time together in these two years. You not only provided me with so much help in my studies, but also gave me great happiness in my life away from home. Hope we can all have a bright future. Last but not least, I want to thank my parents. Thanks for your financial support and encouragement throughout my school life. I am grateful to have your always warm company.

Contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Introduction

In recent years, pedestrian navigation has become one of the most important services in people's urban life. Now, more and more people use mobile phones to get maps and directions on the go. Current navigation applications usually use Global Navigation Satellite System (GNSS) positioning technology to safely guide users to their destinations in an outdoor environment [1].

Smart phones are the most widely used portable devices, and current pedestrian navigation applications are almost all developed for smart phone platforms. The user can check the digital map of the area at any time through the smart phone and get his own position. Recently, mobile phones and handheld devices have become powerful enough to provide augmented reality (AR) experiences [2]. The system can be used to display virtual representations of points of interest (POI) superimposed on the real world, and applications such as Layar, Wikitude, and Junaio have become very popular. There are many potential applications, such as for travel guides [3], as a restaurant finder, providing directions to the subway, and even art installations. In all these cases, the mobile AR system can guide the user to a specific destination, and the navigation can be visualized in the wearable display by fusion with the real scene. AR-based navigation can greatly improve convenience and reality.

However, studies have shown that there are major problems with only using digital maps and AR views for navigation. The tangible physical map plays and continues to play an important role in contemporary cartography [4]. The tangible physical map based on the cartographic tradition and enhanced through new digital interactive forms supplements and advances the digital cartography presented on the computer screen. However, tangible maps can also be valuable assets for learning geography in specific ways. Overall, tangible interfaces seem to offer broad prospects for geographic learning because they reduce the patterns on the interface and promote sensory participation. This is a natural way for people to learn, they promote spatial tasks, and they provide coupled physical object control. Opportunities and manipulation of its digital representation, and encourage learning by providing an inherent social interface [5].

Being able to interact more naturally with digitally rich spaces can enhance our spatial thinking and encourage creativity, analytical exploration and learning. The interaction in the tangible interface is mainly done through touch. However, touch is not only a computing input device, but also a human feeling, which is over-underestimated as a learning tool. One fundamental difference between AR and maps for navigation is that maps provide a top down exocentric view of the environment, while AR adds virtual location information to an egocentric view of the real world.

In this research, we developed an augmented reality navigation system based on real tangible maps and smartphones. Users can enhance the display of area information on the map through touch interaction with the tangible map, and view their current geographic location. In addition, a more natural way to select a destination is designed, based on the destination to quickly calculate and display the navigation path, and supports multi-person sharing. The system aims to provide users with a more convenient and tangible navigation interaction experience.

1.2 Organization of the thesis

The rest of the thesis is organized as follows: In chapter 2, we will introduce the background of the thesis. In chapter 3, we will introduce some related works. In chapter 4, we will describe the goal and approach of our research and show the novelty of our research as well. In chapter 5, we will introduce the concept of the system design. In chapter 6, we will show the implementation details for the system development. In chapter 7, we will make a conclusion and talk about future possibilities.

Chapter 2

Background

2.1 Augmented Reality Navigation

Augmented reality (AR) offers firms new possibilities in delivering content to consumers [6] by superimposing virtual 3D objects on the actual visible world that register and interact with virtual images in real time [7]. In the last decade, AR has seen a boom in commercial applications since smartphone and mobile devices have become ubiquitous as a means of searching for information [8].

AR navigation is the combination of AR technology and navigation functions, and the panoramic route and destination location can be seen through the phone camera. The user's route is displayed in real-time on the phone camera, and there is no need to repeatedly check the map. Even the "road idiot" who has no sense of direction can find the way home.

However, compared with ordinary navigation products, users may have become accustomed to watching while walking, picking them up when they need to confirm the direction, and most of the time they hang their hands down instead of picking them up and checking all the time. Especially after the online voice broadcast, ordinary navigation can almost complete the function of correctly guiding the user to the destination in the pocket.

However, AR navigation requires the user to hold the mobile phone in front of the line of sight at all times, and because of the visual positioning technology, the user needs to align the scanning point very accurately to successfully locate it. The difference in the position of the line of sight will also cause the main interface of the entire App to be greatly different from ordinary navigation products [9].

At present, the usage scenarios and viewpoint processing of AR navigation can be divided into two directions for discussion: outdoor navigation and indoor navigation [10]. In outdoor navigation scenarios, GNNS is still used to locate the user's position, and virtual indicators are superimposed on the real scene. In the indoor navigation scene, the displayed virtual content is similar. The main difference lies in the method of positioning. The current mainstream method is to use the mobile phone WiFi signal to perform positioning. Recently, there are also methods to enhance positioning in complex indoor environments through inertial sensors and RGB cameras [11]. However, when using AR navigation, the user's field of view is greatly restricted, because the route needs to be checked by superimposing the smartphone screen on the real street view [12].

2.2 Tangible Map Interaction

Tangible interaction includes the user interface and interaction methods, emphasizing the tangibility and importance of the interface, the physical manifestation of data, the whole body interaction, the embedding of the interface, and the interaction of the user in the real space and context. We are born to perceive our limbs, joints and muscles with the help of our five senses. We feel that should be used as the input method of mixed reality. In this way, we don't need to rely on external hardware at all, and only rely on our own body to interact. This method allows people to interact more naturally with the body, and we foresee that it has considerable potential.

Paper cartographic maps provide highly detailed information visualization with unparalleled fidelity and information density. In addition, the physical properties of paper provide simple interactions for browsing maps or focusing on personal details, managing concurrent access by multiple users, and general scalability. However, printed maps are static displays. Although computer-based map displays can support dynamic information, they lack the good characteristics of the above-mentioned real maps [13].

Through the use of map-based desktop applications, compare users' preferences for tangible objects and touch buttons [14]. Although people generally expect finger touch as a control mechanism for the touch table, tangible object interaction is another possibility that is rarely realized. However, related research has studied the role of tangibility in problem-solving tasks by using multi-touch or tangible interface to observe logistics apprentices [15]. The results show that tangibility helps them complete tasks better and obtain higher learning benefits. In addition, groups that use tangible interfaces collaborate better, explore more alternative designs, and think that solving problems is more interesting.

For some special scenes and special people, the interaction with tangible maps is also impossible to complete with digital maps. The interactive method that uses tangible maps is closer to people's interactive experience and expectations, and uses specific learning availability to achieve pleasant learning [16]. And it is an effective and interesting static tactile guidance that can help them learn more effectively.For people with visual impairments, providing mobile augmented reality that can stimulate tactile and audio sensations is an innovative and low-cost solution [17]. Facts have proved that tangible interaction is an effective way to help visually impaired users manipulate spatial representation [18].

The tangible user interface aims to make the interaction between humans and computers more intuitive by giving digital data a physical form and a digital dimension of the physical form. By coupling physics and numbers through a feedback loop, it can directly and very naturally interact with the body using highly developed kinesthetic intelligence. Many geographically tangible user interfaces with continuous shape displays [19] have been proposed that are of particular interest.

2.3 Shared Display

People often choose to travel with companions. In scenes that require navigation, there are often not only one user, but even multiple people in different locations. If personal information can be disclosed and public information can be personalized, it will be beneficial to the cooperation of team members [20].

For example, winter sports such as skiing and snowboarding are usually group activities. Traditionally, groups of skiers and snowboarders use paper maps or large-scale maps mounted on boards near the ski lifts to help make decisions: which slope to go next, where to have lunch, or where What dangers should be avoided when leaving the piste. Allow skiers and snowboard enthusiasts to share such content on printed panoramic resort maps, enrich these static maps with personal content (such as pictures, previously taken routes, or encountered dangers), making decision-making more intuitive and convenient [21].

When smart phones have become popular, users have long been accustomed to using handheld computers or mobile phones as the selection technology of paper maps for interactive devices. Affected by this, when participants are asked to imagine how to express a series of queries, they tend to interact with one click [22]. Participants can also accept other interaction methods, including tracking routes, circled areas, using paper menus to restrict queries, and selecting multiple non-adjacent map icons, but these are more or less affected by the operating mode of the smartphone. Therefore, as a carrier of content display, smart phones are a very easy-to-use choice.

Chapter 3

Related Work

3.1 Related work of Handheld AR Navigation

Recently, mobile-phone based outdoor augmented reality (AR) systems have become readily available. One of the most popular applications are AR browsers that show virtual points of interest (POIs) overlaid on top of the phone's camera view. These virtual cues can be used to guide people to the POIs. However, the usefulness of AR systems for guiding users to POI has not yet been evaluated, especially when compared to map interfaces.

Andreas Dünser et al.[2] present results of a user study comparing navigation with information typically provided by currently available handheld AR browsers, to navigation with a digital map, and a combined map and AR condition. They found no overall difference in task completion time, but found evidence that AR browsers are less useful for navigation in some environment conditions. They also found that navigation performance differed significantly with gender for the Map and AR+Map interfaces, but is very similar across gender for the AR interface. Users preferred the combined AR+Map condition, and felt that there were significant problems with using the AR view alone for navigation.

People often search for local information (e.g., a restaurant, store, gas station, or attraction) from their mobile device. Jaime Teevan et al.[23], via a survey of 929 mobile searchers at a large software company, that local searches tend to be highly contextual, influenced by geographic features, temporal aspects, and the searcher's social context. While location was reported to be very important, respondents looked for information about places close to their current location only 40% of the time. Instead, they were often in transit (68% of our searchers) and wanted information related to their destination (27% of searchers), en route to their destination (12%), or near their destination (12%). Additionally, 63% of participants' mobile local searches took place within a social context and were discussed with someone else. They discussed these findings to present a picture of how location, time, and social context impact mobile local searches.

3.2 Related work of Tangible Map Interaction

The Tangible Media Group at the MIT Media Laboratory moved from GUIs to tangible user interfaces (TUIs) in the mid-1990s [24]. TUIs represented a new way to embody Mark Weiser's (former chief scientist at Xerox PARC) vision of ubiquitous computing by weaving digital technology into the fabric of the physical environment, rendering the technology invisible [25].

Rather than make pixels melt into an interface, TUIs use physical forms that fit seamlessly into a user's physical environment. TUIs aim to take advantage of these haptic-interaction skills, an approach significantly different from GUIs. The key TUI idea remains: give physical form to digital information [26], letting serve as the representation and controls for its digital counterparts. TUIs make digital information directly manipulatable with our hands and perceptible through our peripheral senses through its physical embodiment.

Tangible physical maps couple physical landscape model with digital information and can become an invaluable asset for learning geography in an embodied way. The work of George et al. [4] is to create and evaluate an easily constructible 3D tangible map for elementary students. The main differentiation of our approach are two: a) they suggest a new interaction style on the map for learning geography purposes, the use of trips with fingers b) the proposed setting can be recreated and reprogrammed even by students in primary education. 58 4th grade students participated in a pilot study. The participants played with the FingerTrips environment in 24 sessions and in groups of 2 or 3. Students supported that FingerTrips are very pleasant and easy to use, promote efficient and effective learning and help them to learn faster and to learn more things than with the traditional means. Such an approach differs drastically from traditional means of learning, is closer to students' interactive experiences and expectations, gamifies learning and exploits embodied learning opportunities.

David H. Uttal [27] considered the relation between the development of spatial cognition and children's use of maps and models. A new theoretical perspective is presented that takes into account the influences of maps on the development of spatial cognition. Maps provide a perspective on spatial information that differs in important ways from the perspective gained from direct experience navigating in the world.

Using and thinking about maps may help children to acquire abstract concepts of space and the ability to think systematically about spatial relations that they have not experienced directly. In addition, exposure to maps may help children to think about multiple spatial relations among multiple locations. The results of previous studies that have demonstrated developmental differences in children's cognition of large-scale environments are examined from this theoretical perspective. This review suggests that the development of spatial cognition consists partly of the acquisition of models of large-scale space, and that maps influence the development of the modern Western model.

3.3 Related work of Display Shared Content on Map

In natural face-to-face collaboration, people use speech, gesture, gaze, and nonverbal cues to attempt to communicate. In many cases, the surrounding physical world and objects also play an important role, particularly in design and spatial collaboration tasks [28].

Real objects support collaboration through their appearance, physical affordances, such as size and weight, use as semantic representations, and ability to create reference frames for communication. In contrast, most computer interfaces for co-located collaboration create an artificial separation between the real world and the shared digital task space.

People looking at a projection screen or crowded around a desktop monitor are often less able to refer to real objects or use natural communication behaviors. Observations of the use of large shared displays have found that simultaneous interaction rarely occurs due to the lack of software support and input devices for co-present collaboration [29].

AR technology promises to enhance such face-to-face communication. AR interfaces blend the physical and virtual worlds so real objects can interact with 3D digital content and improve users' shared understanding. Tangible interaction methods [26] can be combined with AR display techniques to develop interfaces in which physical objects and interactions are as important as the virtual imagery.

Such interfaces naturally support face-to-face collaboration; for example, in a collaborative urban design application users might sit at a real table and see virtual building models appear in the middle of the table. If these buildings were then attached to physical objects, the users could pick up and place the buildings on a real street map.

Groups of skiers and snowboarders traditionally use folded paper maps or board-mounted larger-scale maps near ski lifts to aid decision making: which slope to take next, where to have lunch, or what hazards to avoid when going off-piste. To enrich those static maps with personal content (e.g., pictures, prior routes taken, or hazards encountered), Anton Fedosov et al. [30] developed SkiAR, a wearable augmented reality system that allows groups of skiers and snowboarders to share such content in-situ on a printed resort map while on the slope. They thought that the SkiAR system could improve group decision making, interaction, communication, and, eventually, provide a safer environment for skiers and snowboarders.

Chapter 4

Research Goal and Approach

4.1 Research Goal

This research aims to provide multiple users with an interactive and shared navigation experience combined with tangible maps.

Although the existing paper tourist maps provide overall information about the scenic spot, they lack real-time display of the location of tourists. Tourists may get lost, and it is very difficult for partners in different locations to meet in a crowded place.

In order to enable users to know the location of themselves and their companions at any time and assist in decision-making, we propose a navigation system that combines tangible maps to enhance regional information and provide real-time location. The system aims to provide users with an interactive and shared navigation experience.

- Improve the display of information on the maps;
- Provide real-time positioning and display corresponding location on the physical maps;
- Realize the display of shared AR content among multi-users;

4.2 Research Approach

4.2.1 Approach

Paper cartographic maps provide highly detailed information visualization with unparalleled fidelity and information density. In addition, the physical properties of paper provide simple interactions for browsing maps or focusing on personal details, managing concurrent access by multiple users, and general scalability.

As the most popular display device at this stage, smart phones have the advantages of being easy to use, easy to operate, and easy to carry. With the development of hardware and software, smart phones already support AR functions and can achieve very good visual effects.

Therefore, on the basis of retaining tangible maps, we use augmented reality technology to make up for the shortcomings of paper maps that cannot update information in real time, transform the key information contained on the map from a 2D plane to a 3D model, and introduce the GPS signal of the smartphone as Positioning basis, real-time display of the user's relative position on the tangible map.

In addition, the system allows the user to freely choose the destination, and can display the navigation path on the screen of the smartphone in the form of augmented reality. Taking into account the user's usage scenario, we designed the system to support multi-person sharing, and users in the same group can view each other's location and travel route. We hope to help users have a better interactive map navigation experience through this system.

The AR navigation system is based on:

Fig. 4.1 System components

- Observe AR content and control location sharing through mobile phones;
- Touch tangible maps to trigger the display of AR content;
- Store and transmit data via cloud server;

In order to achieve the research goal, we designed the system. The detailed system can be divided into three parts: map display, real-time location tracking and sharing AR navigation between different users.

Map display: Using foldable paper with a border as the base of the map, users can carry it with them and interact with tangible maps at any time.

(1) Using Vuforia-based image tracking technology, users can freely set up the park map they need to display;

(2) Papers with different borders display different regional maps, including global maps and local maps;

(3) Through the virtual button technology, the user can set interactive options at different positions on the tangible map to enhance the display of the information contained in the current map.

Real-time location tracking: The user's location is determined by calling the GPS signal of the user's mobile phone. By comparing the longitude and latitude coordinates of the user and the map display area, the user's relative position on the tangible map is calculated and displayed.

Sharing AR navigation between different users: Use Photon PUN2 to build a cloud server to store and transmit the location information and interaction trigger information of multiple users.

(1) Users can set up shared rooms by themselves, and verify their privacy by password;

(2) Users in the group can share each other's location and display them in real time;

(3) User interaction on the map will be synchronized to the server to realize remote sharing navigation.

4.2.2 Novelty

The novelty of this research mainly reflects in these aspects:

1. We design and implemented a mobile AR system combined with tangible maps to realize real-time positioning and navigation;

2. We emphasize the tangible interaction between the user and the real map to enhance the tour experience;

3. We realize remote sharing navigation by displaying shared AR content on multiple tangible maps.

Chapter 5

System Design

5.1 System Overview

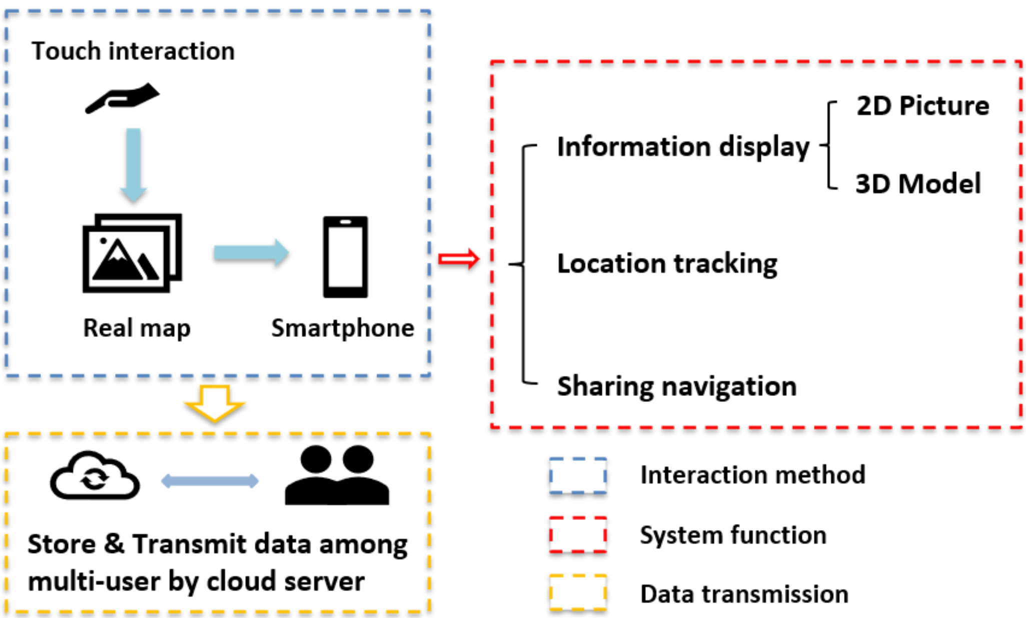


Fig. 5.1 System Design Overview

In this section, we will introduce the system design of our augmented reality navigation system from three aspects: interaction method, system function and data transmission. Fig 5.1 shows the overall structure of our system.

The interaction method between the user and the system is mainly realized through touch interaction. The system allows users to trigger different preset functions by touching different locations on the real map, and display virtual content through the user's smart phone. The interaction between the user and the real map is the main body of the system, and the smartphone is used as a tool for playing virtual content.

As an augmented reality navigation system that combines real maps, it has three main functions.

First, the system can realize the information expansion of existing maps. This means that when a user uses a real map, the information that can be obtained will not be limited to the map itself. Through augmented reality technology, the matching environment information can be superimposed on the real map, including the name and model of each building. The display of this information includes 2D pictures and 3D models. On this basis, users can interact with virtual information to enhance their understanding of their surrounding environment.

Second, the system can provide users with real-time location tracking. The system uses the user's smart phone to receive GPS signals to determine the user's location. According to the calculation of the relative position, the system can display the user's position on the real map in real time. Users can observe their location from the perspective of God, enhancing the user's sense of immersion when using real maps.

Third, the system supports multi-person sharing mode. Multiple users can create their own rooms and set a password to ensure the privacy of location sharing. Users in the same room can share their location with each other, and can use their mobile phones to take photos of their current location and share them to discuss the next destination. The system will calculate the shortest path based on the target point selected by the user and display it to the user to guide the user to the destination.

5.2 Interaction Method

5.2.1 Interaction Process

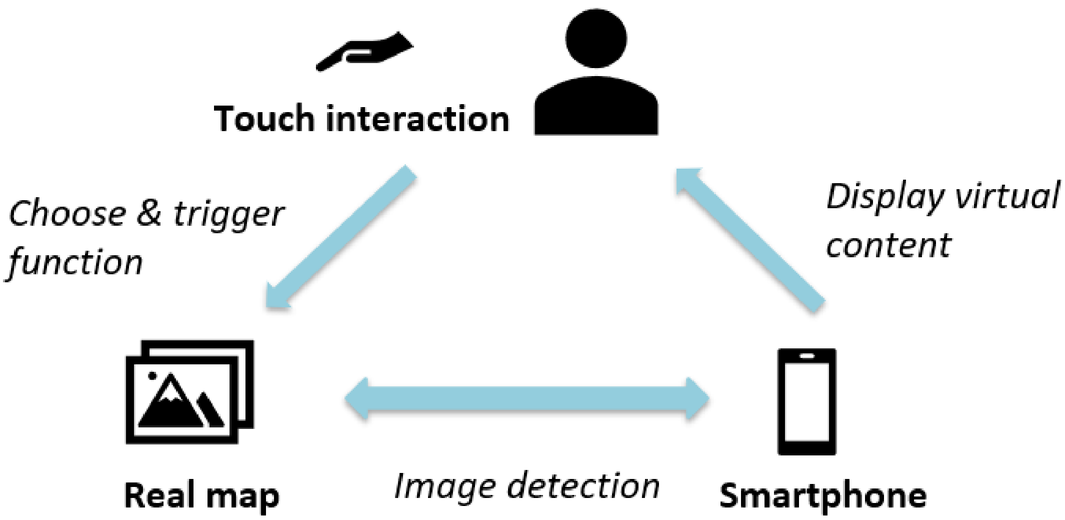


Fig. 5.2 Interaction design concept

The interaction of the system consists of three parts. The user's touch is used as the triggering method of interaction; the real map is used as the object of interaction; the smart phone is used as the device for image detection and display. The interaction design concept is shown in Fig. 5.2, which shows the interaction process of the augmented reality navigation system.

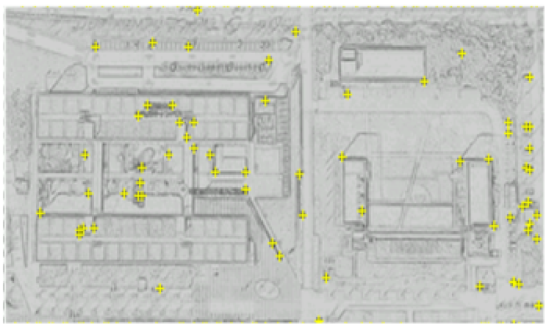
The interaction method of the system is mainly touch interaction, emphasizing the interaction between users and physical objects. We hope that users can get a more convenient and intuitive interactive experience when using real maps, that is, when users want to obtain information about a certain place, they only need to touch or indicate with their hands. The user selects and triggers various functions of the system through touch interaction. The user's smart phone is used as an image detection tool to determine which function the user has selected, and is used as a playback device to display augmented reality virtual content.

5.2.2 Touch Trigger Design

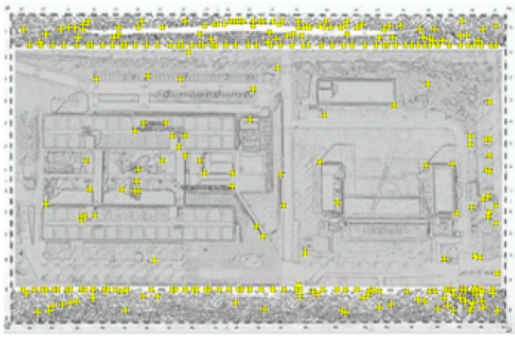
According to user surveys, when using real maps, most users will subconsciously point out the location of interest with their fingers. Based on this result, the system uses a touch-triggered design to allow users to interact more naturally.

On the other hand, the real map is the object of user interaction, and we have also made some adjustments to it.

For computer vision, if the system wants to recognize different maps, it needs to contain enough feature points. However, the number of feature points contained in the general real map is not enough. As shown in Fig 5.3.(a), it is difficult for the computer to recognize this type of map, so we have modified it: in keeping the original real map Based on the information, a border with rich feature points is added to it, as shown in Fig 5.3.(b). The purpose is to allow the computer to recognize the target more quickly.



(a) Real map only

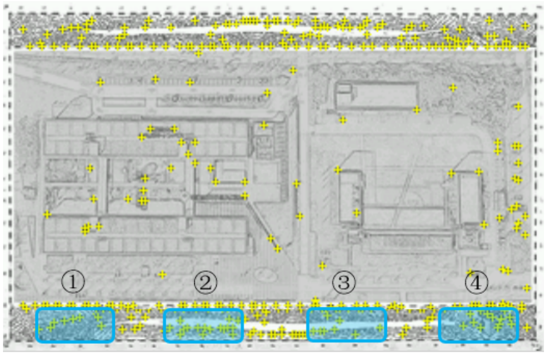


(b) Real map with frame

Fig. 5.3 Feature points of maps

For the original real map, due to the small number of feature points, the function triggers that can be set are limited accordingly, and feature points that are closer together may cause interference and affect the stability of the system. After adding an additional frame, the system can use additional feature points to set more interactive options, enrich system functions, avoid false triggers of different functions, and improve system stability.

The system uses additional borders to set up multiple virtual buttons, and the user can touch to realize the occlusion of feature points to trigger the corresponding functions, as shown in Fig 5.4.(a). We added 3D text to the corresponding button to prompt the user to select different functions. The actual running effect is shown in Fig 5.4.(b).



(a) Concept design

(b) Running effect

Fig. 5.4 Function option

5.2.3 Display Device Choice

Considering the usage scenarios of the system, the equipment must not only be indoors, but also must work normally outdoors. Therefore, it is required that the equipment can display virtual content normally in a strong light environment. In addition, the user may need to walk around during use, some head-mounted devices may be inconvenient, and the user may not need to use the display device all the time. Since the system needs to implement navigation functions, the device must be able to obtain the user's current geographic location. In the sharing mode, users need the network for data transmission.

Based on the above factors, we choose smart phones as the display device and use it as an image detection device to reduce the use of additional unnecessary equipment and reduce the user's operational burden.

5.3 System Function

As an augmented reality navigation system that supports multi-person use combined with real maps, the main function of the system is to enhance the information display of the real map and realize multi-person positioning and navigation on the real map. As shown in Fig 5.4, we have set four functional options for the system:

- Locate: Show user's location on the real map.
- Display: Expand the information on the real map.
- Sharing: Support multiplayer mode, know the location and route of partner.
- Navigate: Provide dynamic shortest path navigation.

The above four function options can be summarized into three parts of the system function: information display, location tracking and sharing navigation.

5.3.1 Information Display

Generally, ordinary paper maps can only contain limited environmental information, and are limited to picture information printed on paper. When a user makes a travel plan, he needs to have an in-depth understanding of one of the locations, and the user can only resort to the Internet or actual visits. This requires a lot of time and effort for users.On the other hand, although digital maps have become popular, users need to zoom in and out to switch to view the global and local information of the map. There is still room for improvement in this interactive way of information display.

The system retains the real paper map and has improved the information display of the map. Through augmented reality technology, the system can add all the information of the corresponding area of the map, such as detailed information and three-dimensional models of each building, the location of roads available for passage, and so on.

The user can activate this function by touching the display button. After it is turned on, the system will display the names and 3D models of all the buildings in the corresponding area on the map. Users can view the information of each location, and can also operate on the 3D model. As shown in Fig 5.5, the system prototype was developed for the Waseda campus. When the user selects the display function, the system will display the names and 3D models of the two landmark buildings on the Waseda campus. In addition, users can also view real-life photos of various locations to provide a reference for travel planning.

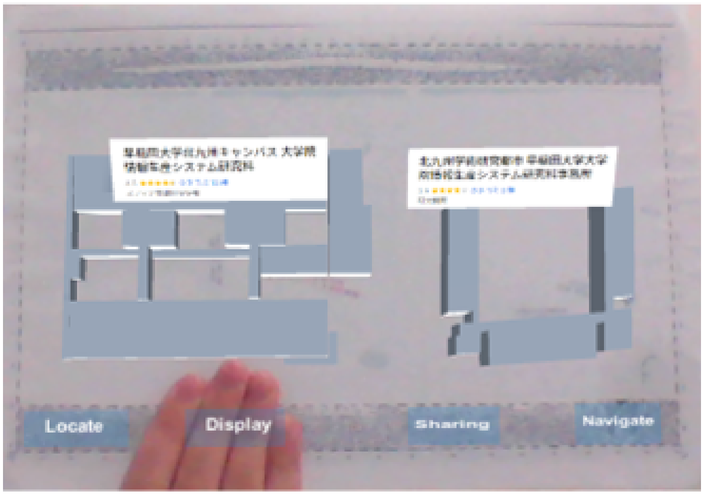


Fig. 5.5 Information display

5.3.2 Location Tracking

Taking into account the convenience of the user, the system uses as few external devices as possible. On the basis of using a smart phone as a display device, use its GPS positioning function to retrieve the GPS signal of the smart phone, and calculate the user's relative position on the real map through coordinate conversion.

In order to display the user's actual location on the real map, we need to obtain the actual longitude and latitude coordinates of the map corresponding area in advance, and establish a coordinate system in the map corresponding area. According to the GPS signal transmitted from the smart phone in real time, the user's real-time latitude and longitude information is determined, and on this basis, the user's relative coordinates in the corresponding area of the map are calculated, and finally displayed by the system.

Fig 5.6 shows the realization of the system. The system obtains the real latitude and longitude coordinates of the upper left and lower right corners of the map in advance, and then calculates the user's relative position based on the user's latitude and longitude coordinates.

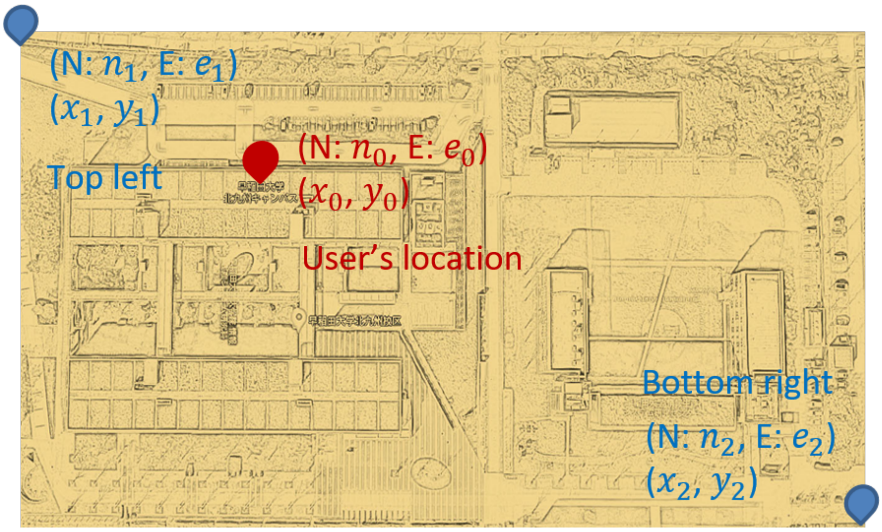


Fig. 5.6 Coordinate transformation

Calculation formula:

$$x_0 = x_1 + (n_0 - n_1)\frac{x_2 - x_1}{n_2 - n_1}$$

$$y_0 = y_1 + (e_0 - e_1)\frac{y_2 - y_1}{e_2 - e_1}$$

(N: n_i, E: e_i): longitude and latitude coordinates, (x_i, y_i): Unity world space.

When the user touches the locate button, the system will run the above process and display a red portrait and the user's name on the smart phone to mark the user's actual location. The actual effect of system operation is shown in Fig 5.7.

Fig. 5.7 Location effect

5.3.3 Sharing Navigation

The system supports multi-user mode. In order to ensure the privacy of users, users need to create their own rooms and set a password for joining. We have made a login interface for the system, as shown in Fig 5.8.

Fig. 5.8 Login interface

Users can customize their own name and room password, and enter the same password to enter the same room. The login process is shown in Fig 5.9. All users in the same room can share their location with each other, and can set a common target point, and display the travel route of all users.

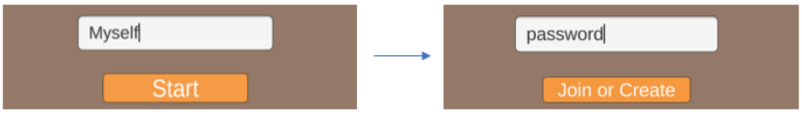
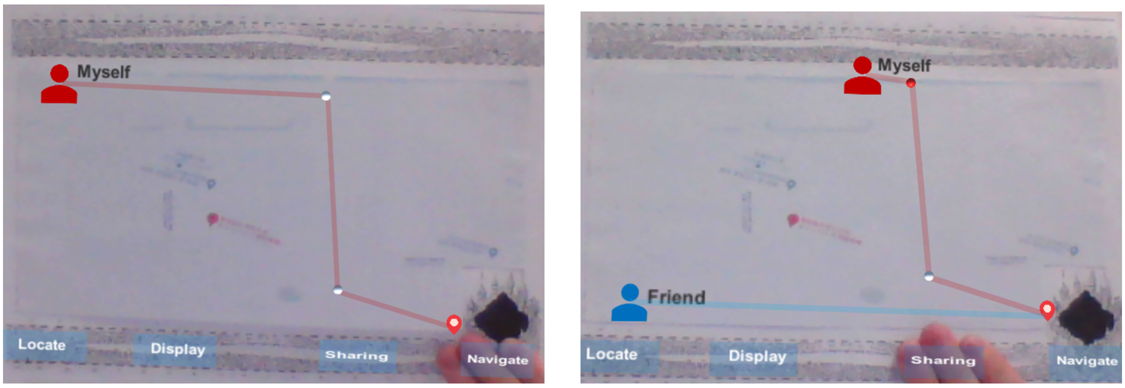


Fig. 5.9 Login process

When the user touches the navigate button, the system turns on the navigation function. The user can select the destination by touching the real map. After selection, the system will automatically calculate and give the shortest path from the current user to the destination. The actual running effect is shown in Fig 5.10. In a multi-user scenario, the system defaults that the user himself is a red portrait, and other users are blue portraits. Of course, the path display of different users will also use different colors to distinguish. In addition, the system will also display the nickname that each user has made.



(a) Individual (b) Multiplayer

Fig. 5.10 Route display

When the user's location changes, the system will synchronize the user's location changes and dynamically update the navigation path. For each path, the system will mark their turning point. When the user approaches the turning point of the route, the system will automatically mark it in red to remind the user, as shown in Fig 5.11.

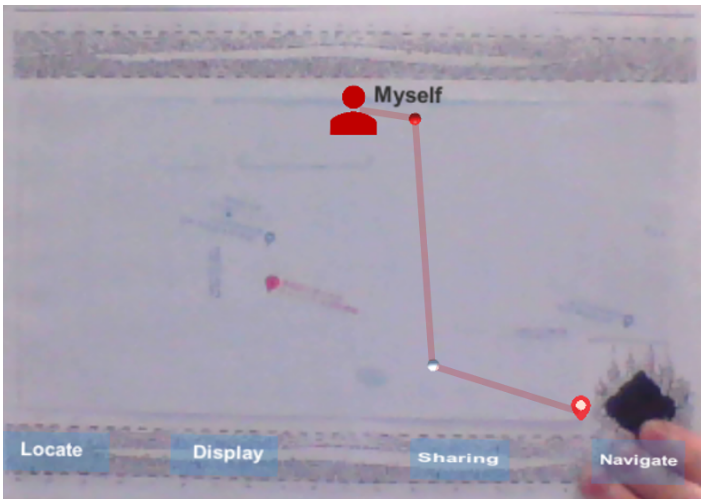


Fig. 5.11 Turning tip

5.4 Data Transmission

The data transmission of this system is realized through cloud server. The current system supports up to 20 users online at the same time. The location information of multiple users and the interaction information with the real map will be transmitted to the cloud for storage, and sent from the cloud to each client for synchronization.

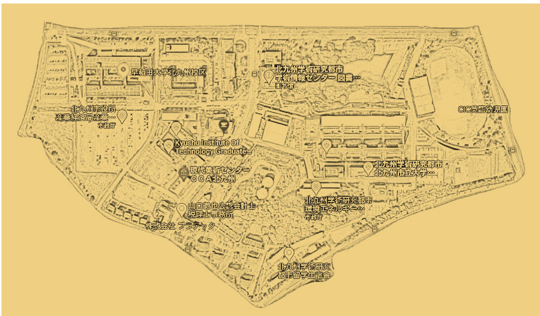
For location tracking, the system will automatically upload the GPS signal of the smartphone to the cloud in real time to synchronize the location information of each user; for interactive trigger signals, the system will only automatically upload the location of the target point set by the user. Then call local resources on each client to perform route calculations to reduce the amount of data transmission and reduce the pressure and delay of the cloud.

5.5 Usage Scenario

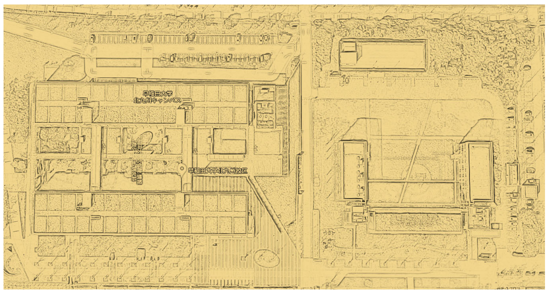
In order to facilitate users to understand this augmented reality navigation system, a description of the usage scenario is given next.

The usage scenario for the system is that multiple tourists in different locations in the same area visit at the same time and discuss the next destination.We selected Kitakyushu Science and Research Park as the place for users to use the map to visit, and designed a detailed interaction for the Waseda campus. The following description will be based on the background of tourists visiting the Waseda campus.

First, we need to make a sightseeing map of Kitakyushu Science and Research Park. We first use Google Maps to obtain a rough map of the map, and then use Photoshop software to extract the map of the target area from it to make the real map used by the system prototype.



(a) The first map of Kitakyushu Science and Research Park



(b) The first map of Waseda campus

(c) The current map of Waseda campus

Fig. 5.12 Map design

In the initial version of the map design, in order to create the parchment effect of the ancient map, a series of adjustments were made to the map. The overall map uses a yellow background and the lines are very clear to facilitate the system to distinguish, as shown in Fig 5.12.(a)(b). However, for For users, especially those who are not familiar with the area drawn on the map, it is very difficult to find their location, which weakens the original function of the map, so further adjustment is required. The real map at this stage uses part of the digital map familiar to most users for functional debugging, as shown in Fig 5.12.(c).

The system needs to be used in combination with real maps and applications installed on smartphones. All user interaction processes involved in the system are shown in Fig 5.13.

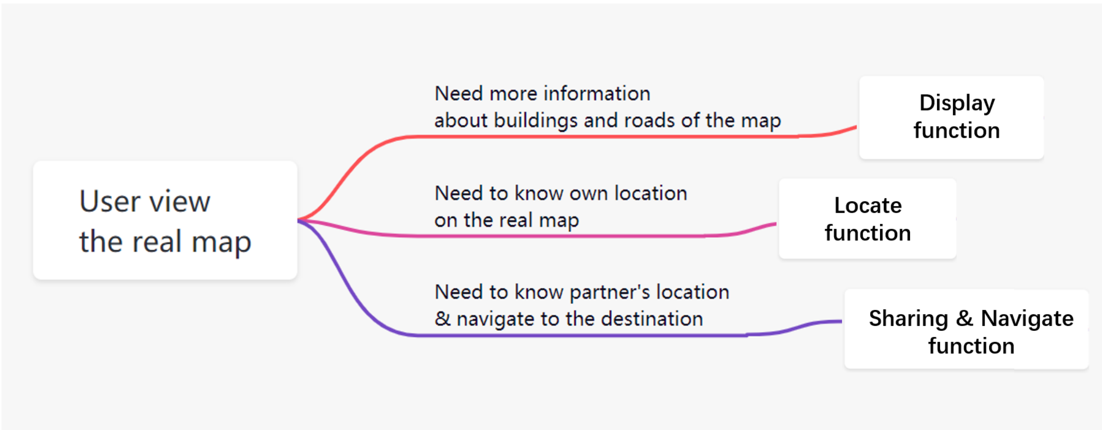


Fig. 5.13 Usage process

When the user arrives at the area drawn on the map, the user can first use the real map to observe the rough information of the area. If the user wants to know where he is, he can use the app to observe where he is on the map. The system will obtain the user's GPS signal through the smart phone, and calculate the user's relative position on the real map, and finally use augmented reality technology to superimpose the user's position on the real map. The location function process is shown in Fig 5.14.

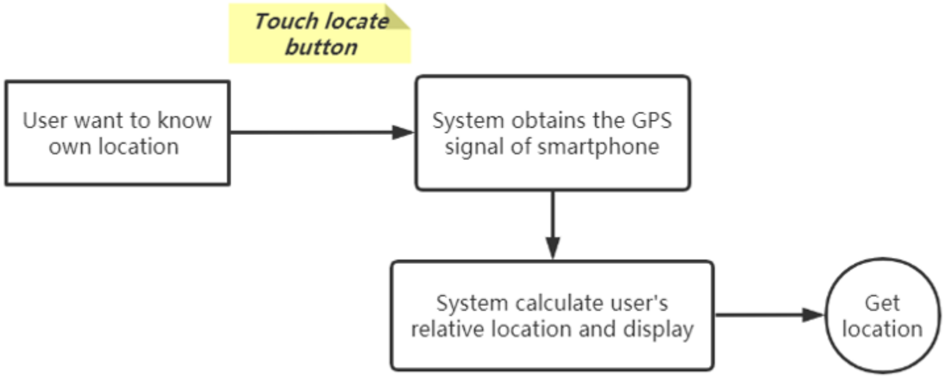


Fig. 5.14 Location function process

If the user is very interested in certain places in the area and wants to obtain more information, the user can trigger the display of additional information by interacting with the real map. The system will use augmented reality technology to display the building information and 3D model in the area through a smart phone. Users can also observe the building model further to deepen their understanding of their surrounding environment.The display function process is shown in Fig 5.15.

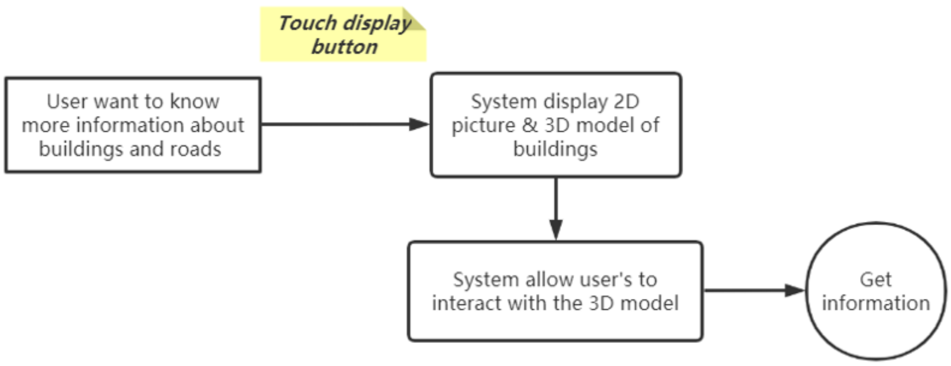
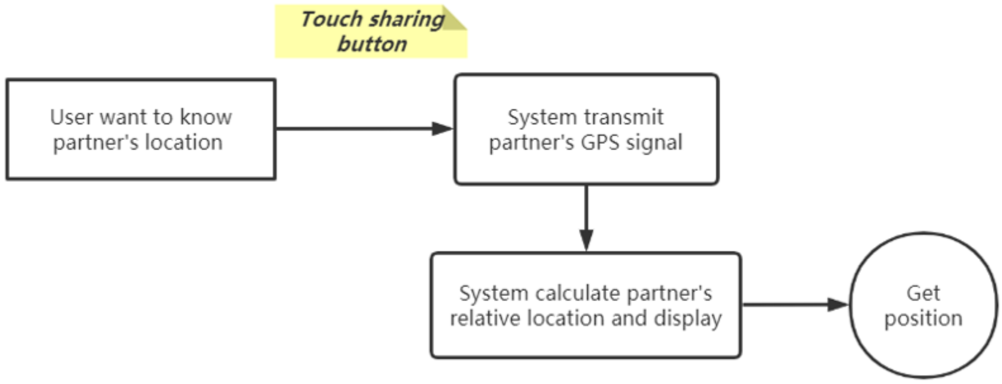
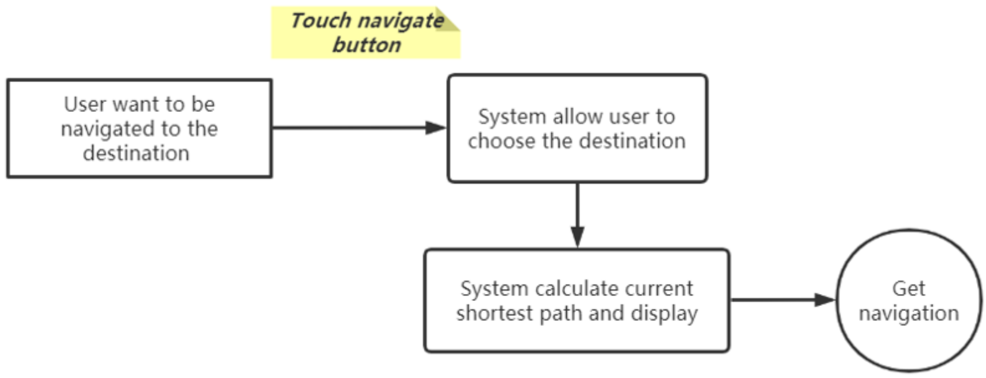


Fig. 5.15 Display function process

For user groups located in different locations, the location of each other can be observed. The system allows friends to create their own private rooms, and only users in the same room can share their location with each other. Users in the same room can set the same target point as the next destination for sightseeing. The system will superimpose the navigation route on the real map through the use of augmented reality technology through the smartphone, and give direction guidance at the turning point of the route.The sharing and navigation function process is shown in Fig 5.16.



(a) sharing function process



(b) navigate function process

Fig. 5.16 sharing and navigation function process

Chapter 6

System Implementation

6.1 System Hardware

The system uses Android smartphones as the identification and display device. The built-in GPS module of the smartphone provides the function of receiving and sending the user's geographic location. The camera of the smartphone is used to capture images for the system to recognize. The screen of the smartphone is used to display superimposed virtual content and as an interface for user interaction. The device used is shown in Fig 6.1.

Fig. 6.1 Android smartphones

Considering that the system needs to support multiple users, the upper limit of theoretical users that the system prototype can accommodate is 20. For the preview of the system function, using 2 smartphones is enough, so we set it as 2 users to use the system at the same time. One of the smartphones used is Xiaomi 11 and the other is Xiaomi 8 SE. Two smart phones need to transmit and receive data through the network. The specific device parameters are shown in Table 6.1.

Smartphone	Xiaomi 11	Xiaomi 8 SE
Operation System	MIUI 12.0.2.0	MIUI 12.0.2.0
SoC	Qualcomm(R) Snapdragon 888	Snapdragon 710
RAM	8.00GB	6.00GB

Table 6.1 The Android smartphones parameters

The core code of the system is written on a laptop PC, and it can be used as a monitoring station during the debugging process. The laptop PC uses the Windows operating system to ensure the compatibility of other software and interfaces. The relevant parameters of the laptop PC are shown in Table 6.2.

Operation System	Windows10.0.18362
CPU	Intel(R) Core(TM) i5-8300H CPU @ 2.30GHz
Graphics Card	GeForce GTX 1060 with Max-Q Design
RAM	16.00GB

Table 6.2 The PC information

6.2 Develop Environment

The Software Environment support is:

1. Unity 2019.4.20f1, it provides the development platform for applications based on Android which can be run on Android smartphones. Unity is also implemented with an XR SDK which helps our system to handle input and output events.
2. Vuforia Engine AR 8.5.9, it uses computer vision technology to recognize and track flat images and simple 3D images in real time, and allows the creation of virtual content associated with them, aiming to help create the interaction between real-world items and virtual items.
3. Photon Cloud service, it is a fully managed software as a service (SaaS) solution. The hosting, server operation and scaling are all taken care of by Exit Games. Among them, Photon Unity Networking (PUN) is a Unity 3D client plug-in, an API compatible with Unity Networking, connected to Photon Realtime for data synchronization.
4. Blender 2.92.0. We use Blender to extract the building and road information in the satellite photos of the area, create the corresponding 3D model and import it into Unity.
5. Google Map API. We use this API to get information about buildings in the map for displaying in our system.

Besides, we used Visual Studio 2019 with C# for scripting, debugging and simulation.

6.3 Cloud Server

In order to realize the various functions of the system, the position and interaction information between users need to be synchronized in real time. Considering that there may be usage scenarios with more than 2 users, it is not enough to rely solely on data transmission between smart phones, so we use Photon's cloud service [31] to provide the system with a cloud server to upload and download data.

The server engine encapsulates a series of network communication methods. The server engine usually includes load balancing technology to automatically complete the performance optimization and exception handling in network communication. The Photon server is built on the Windows operating system platform and supports reliable UDP, TCP, HTTP, and Web Sockets protocols.

The user's location information and interaction information will be monitored in real time and uploaded to the cloud server. Afterwards, according to the user's choice, the data can be downloaded to each client and displayed to achieve data synchronization. The communication between client and server is shown in Fig 6.2.

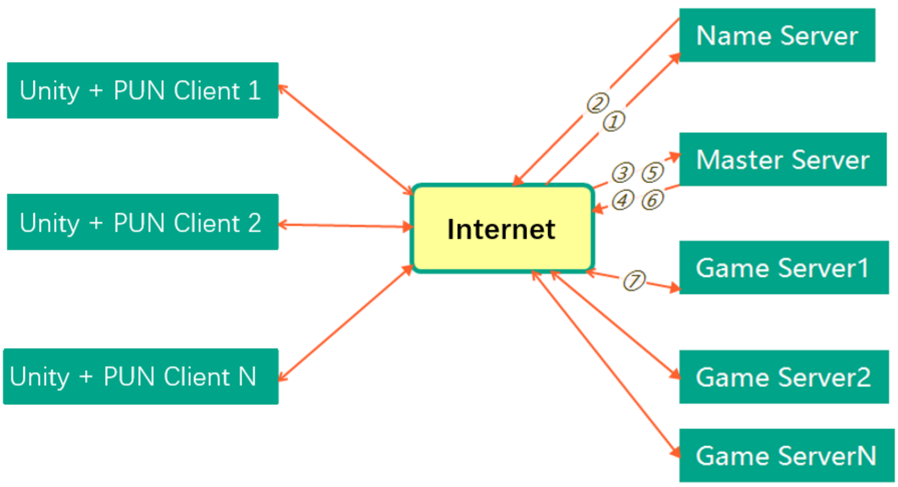


Fig. 6.2 Client-Server

1. The client sends the App id and version identifier of the lucky application to the Name Server server.

2. Name Serve will manage the Master Serve IP address of this version of the game application and send it to the client.
3. The Unity client connects to the master server to request the game lobby list of the master server.
4. The master server sends the game lobby list to the client.
5. The Unity client sends a request to the master server to create or join a game room.
6. The master server sends the address of the game server that manages the game room to the client.
7. The Unity client transfers data to other clients through the game server, and realizes the transfer of information between the unity clients.

The PUN network connection status is shown in Fig 6.3:

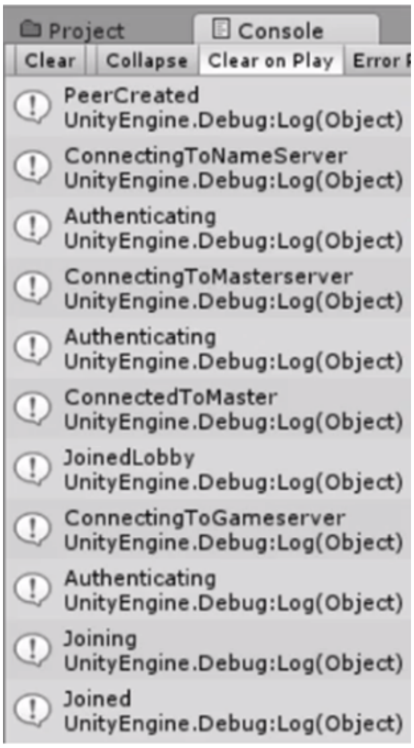


Fig. 6.3 Network connection status

1. PeerCreated: Unity client is not connected to Photon server.

2. ConnectingToNameServer: The Unity client is connecting to the Name Server of the Photon server.

3. Authenticating: The Photon server is authenticating the Unity client's request.

4. ConnectingToMasterserver: The Unity client obtains the Master Server address of this application from the Name Server and starts to connect to the Master Server.

5. Authenticating: The Photon server is authenticating the Unity client's request.

6. ConnectedToMaster: Unity client has connected to Master Server.

6.4 Virtual Button

In order to realize the interaction between the user and the real map, we use the virtual button mechanism provided by the Vuforia engine to create the touch interaction between the user and the real map and trigger the various functions of the system.

Virtual buttons [32] provide a useful mechanism for making image-based targets interactive. Handle the events with *OnButtonPressed* and *OnButtonReleased* when the button is visually obstructed from the camera. When creating a Virtual Button, the size and placement must be considered carefully with respect to the user experience. There are several factors that will affect the responsiveness and usability of Virtual buttons.

- The length and width of the button.

- The area of the target that it covers.

- The placement of the button in relation to both the border of the image, and other buttons on the target.

- The underlying area of the button has a high contrast and detail so that events are easily activated.

Attributes of an ideal virtual button are listed in the following Table 6.3.

Attribute	Suggestions
Size	Choose areas in the images that have dimensions of approximately 10 percent of the image target's size.
Shape	Make buttons easily identifiable to stand out from rest of image. Highlight active buttons in the augmentation layer to hint at active regions on the target.
Texture or contrast	Avoid defining buttons on low contrast areas of the targets. The underlying target area needs to have sufficient features to be evaluated.Choose a button design that is different in texture from the object that causes the occlusion.
Arrangement on the target	Arrange buttons around the target's borders with enough space between to avoid losing tracking when the end user presses a button.

Table 6.3 Virtual Button Attributes

In order to facilitate users to use the real map, we set virtual buttons corresponding to different functions in different positions of the lower border of the map to form an interactive interface that combines physical objects and virtual content. Regarding the trigger logic of the virtual button, we set it to trigger the corresponding function when the user touches the area corresponding to a certain virtual button and occludes the characteristic points contained in the corresponding area. And when the user removes the hand to expose the feature point again, it is still under this function until the user touches another button. The virtual button trigger logic used by the system is shown in Fig 6.4.

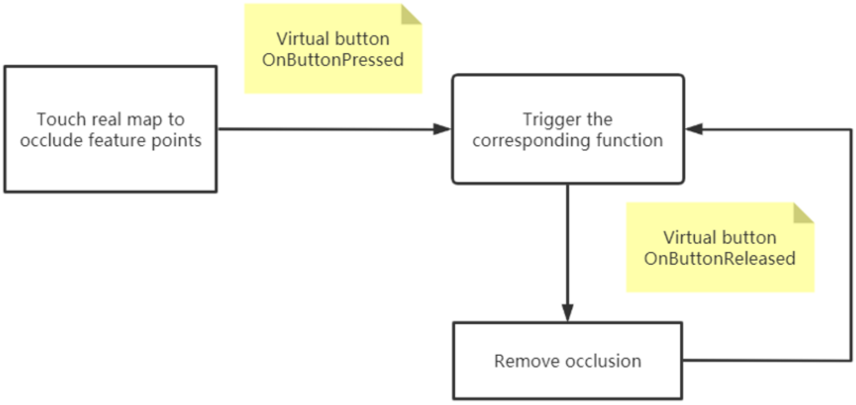


Fig. 6.4 Trigger logic

6.5 Positioning and Display

6.5.1 Obtain geographic coordinates

There are three ways to use a smart phone to achieve positioning.

- GNSS (Global Navigation Satellite System): refers to all satellite navigation systems, characterized by global coverage, all-weather, and high positioning accuracy (within 10 meters). The most famous and widely used GNSS is the American GPS, in addition to the Chinese Beidou, the Russian GLONASS, and the European Galileo.
- Base station location: Determine the location based on the location of the base station to which the mobile phone is currently connected. The mobile phone obtains location information from the base station through information such as country code, network code, and cell ID. The characteristic is that the positioning speed is fast, but the accuracy is the worst of the three methods, about tens of meters to several kilometers.
- WiFi location: Each WiFi hotspot has a unique MAC address. After turning on WiFi, the mobile phone automatically scans nearby hotspots and uploads WiFi hotspot location information, thus establishing a WiFi hotspot location information database. When the mobile phone is connected to a hotspot, the location information of all nearby

hotspots can be called from the database, and the approximate geographic location of the mobile phone can be calculated according to the signal strength of each hotspot. WiFi positioning accuracy is worse than GPS, about ten meters.

The system mainly uses GNSS and WiFi positioning, which are used in outdoor scenes and indoor scenes respectively.

Outdoor

The satellite positioning system used by this system is GPS.

GPS (Global Positioning System) is a satellite navigation and positioning system established by the United States. With this system, users can achieve all-weather, continuous, real-time three-dimensional navigation positioning and speed measurement on a global scale; in addition, users can also Carry out high-precision time transfer and high-precision precise positioning.

GPS system includes three major parts: space part-GPS satellite constellation; ground control part-ground monitoring part; user equipment part-GPS signal receiver.

The basic principle of the GPS navigation system is to measure the distance between a satellite with a known location and the user's receiver, and then synthesize the data of multiple satellites to know the specific location of the receiver. To achieve this goal, the position of the satellite can be found in the satellite ephemeris based on the time recorded by the on-board clock. The distance between the user and the satellite is obtained by recording the time it takes for the satellite signal to propagate to the user, and then multiplying it by the speed of light (due to the interference of the atmospheric ionosphere, this distance is not the true distance between the user and the satellite, but Pseudorange).

When the GPS satellite is working normally, it will continuously transmit navigation messages with a pseudo-random code composed of 1 and 0 binary symbols (pseudo code for short). Navigation messages include satellite ephemeris, working conditions, clock corrections, ionospheric delay corrections, atmospheric refraction corrections and other information. The role of the satellite part of the GPS navigation system is to continuously transmit navigation messages. However, since the clock used by the user receiver cannot always be synchronized with the satellite onboard clock, in addition to the user's three-dimensional coordinates x, y, z, a variable t, namely the time difference between the satellite and the receiver, must be introduced as an unknown number. , And then use 4 equations to solve these 4 unknowns. So if you want to know where the receiver is, you must be able to receive signals from at least 4 satellites. As Fig 6.5 shown below:

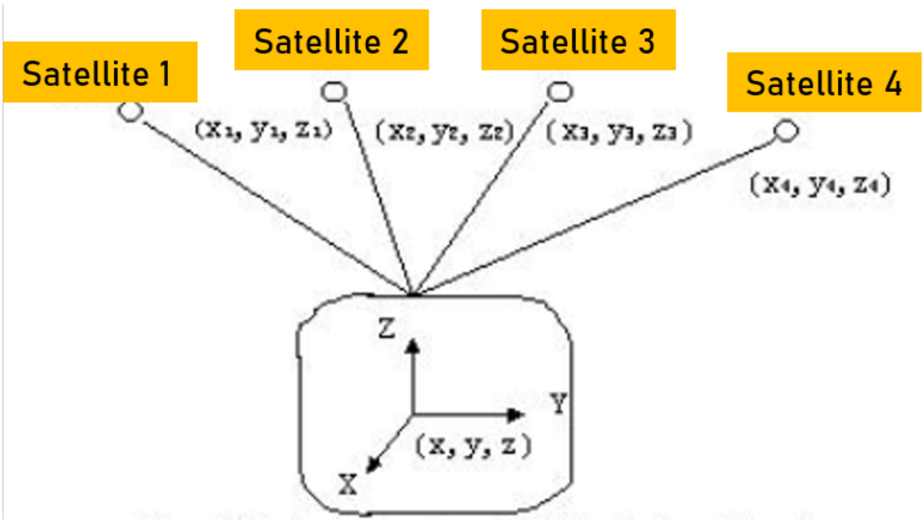


Fig. 6.5 GPS position calculation method

Calculation formula:

$$[(x_1-x)^2+(y_1-y)^2+(z_1-z)^2]\frac{1}{2}+c(v_{t_1}-v_{t_0})=d_1$$

$$[(x_2-x)^2+(y_2-y)^2+(z_2-z)^2]\frac{1}{2}+c(v_{t_2}-v_{t_0})=d_2$$

$$[(x_3-x)^2+(y_3-y)^2+(z_3-z)^2]\frac{1}{2}+c(v_{t_3}-v_{t_0})=d_3$$

$$[(x_4-x)^2+(y_4-y)^2+(z_4-z)^2]\frac{1}{2}+c(v_{t_4}-v_{t_0})=d_4$$

Solving x, y, z and t from the above four equations can be timed and positioned.

GPS positioning method, no sim card, no need to connect to the network, as long as you are outdoors, you can basically locate accurately anytime and anywhere.

Indoor

The indoor positioning of the system mainly relies on WiFi positioning.

The principle of WiFi positioning is shown in Fig 6.6:

Fig. 6.6 WiFi positioning

- Each wireless AP (router) has a globally unique MAC address, and in general, the wireless AP will not move for a period of time;
- When the device turns on WiFi, the wireless router will broadcast SSID by default (unless the user manually configures to turn off this function), and the broadcast frame contains the router's MAC address;
- The collection device can obtain the MAC information and signal strength information of the surrounding AP by receiving the broadcast information sent by the surrounding AP, upload the information to the server, and save it as a "MAC-Longitude" mapping after calculation by the server. When the collected information is enough At that time, a huge WiFi information network was established on the server;

- When a device is in such a network, the collected data that can mark AP can be sent to the location server. The server retrieves the geographic location of each AP, and combines the strength of each signal to calculate the device The calculation method of the location and return to the user equipment is similar to the calculation method of the base station positioning position, and it also uses three-point positioning or multi-point positioning technology, as Fig 6.7 shown;
- Location service providers must constantly update and supplement their own databases to ensure the accuracy of the data. When some WiFi information is not in the database, the location information of unknown WiFi can be inferred based on the location information of other nearby WiFi and uploaded to the server.

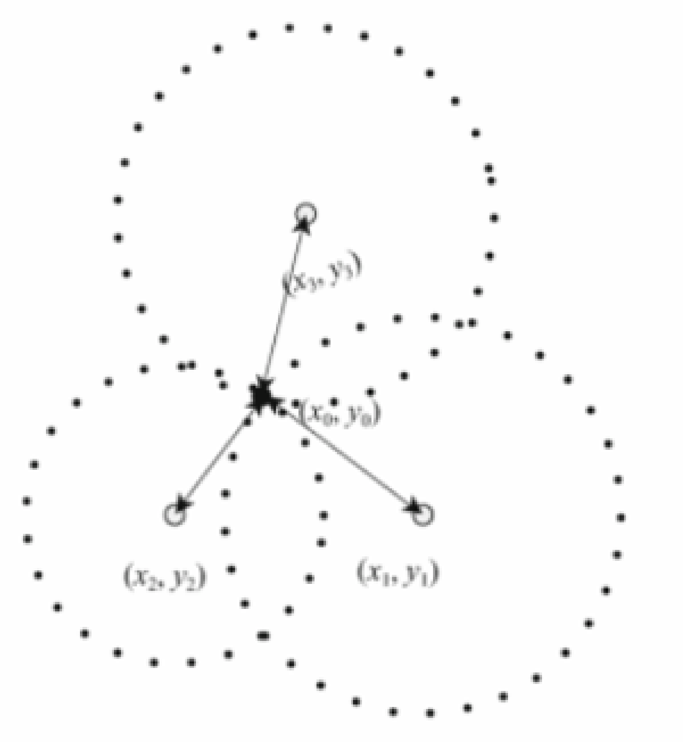


Fig. 6.7 Three-point positioning principle

Calculation formula:

$$(x_1 - x_0)^2 + (y_1 - y_0)^2 = (d_1)^2$$

$$(x_2 - x_0)^2 + (y_2 - y_0)^2 = (d_2)^2$$

$$(x_3 - x_0)^2 + (y_3 - y_0)^2 = (d_3)^2$$

Three-point positioning is to use the coordinate position information of three points to calculate the current position information. That is, the coordinates of the three points (x_1, y_1), (x_2, y_2), (x_3, y_3) and the distances d_1, d_2, d_3 from the desired position point (x_0, y_0) to the three points are known. Then make three circles with radii d_1, d_2, and d_3. According to the Pythagorean theorem, the formula for calculating the position of the unknown point is obtained. The problem of positioning is transformed into the problem of finding the coordinates of the intersection point of the circle, and finally the coordinates of the position point x_0 and y_0 are calculated.

AP location mapping data collection methods can be roughly divided into active collection and user submission.

Android mobile phone users will prompt whether to allow the use of Google's location service when turning on "Use wireless network location". If allowed, the user's location information will be collected by Google. iPhone will automatically collect WiFi MAC address, GPS location information, operator base station code, etc., and send it to Apple's server.

6.5.2 Calculate relative position

In this system, the user's location needs to be displayed on the corresponding location on the real map. This means that only obtaining the user's geographic location information is not enough. The system needs to perform a series of coordinate transformations to calculate the user's relative position on the real map.

The conversion method has been introduced in Fig 5.6, but the positioning error still needs to be reduced. The method that be used here is Delaunay triangulation.

Delaunay triangulation gives a definition of a "good" triangle mesh. Its outstanding characteristics are the empty circle characteristic and the maximum minimum angle characteristic. These two characteristics avoid the generation of long and narrow triangles. The feature of the empty circle is that for two triangles that share the same side, the circumcircle of any triangle cannot contain the vertices of the other triangle. This form of division produces the largest minimum angle. The result of Delaunay triangulation is shown in Fig 6.8.

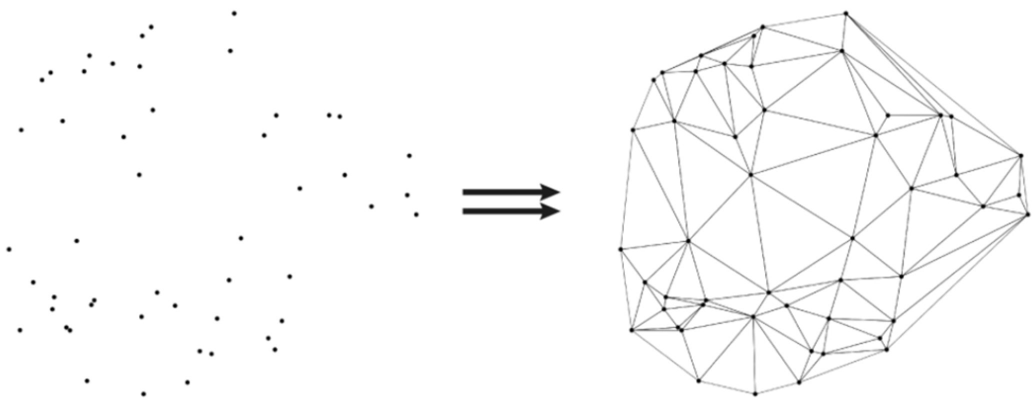


Fig. 6.8 Delaunay triangulation

The system prototype triangulates the corresponding area according to the real map. After receiving the user's geographic location and calculating the user's relative position, the system will determine which small area the user is in based on the triangulation result of the map. If there is a large deviation between the user's position in the real world and the user's relative position on the physical map after calculation, the user's position will be automatically corrected to the current triangle center position.

6.6 Generate 3D Model

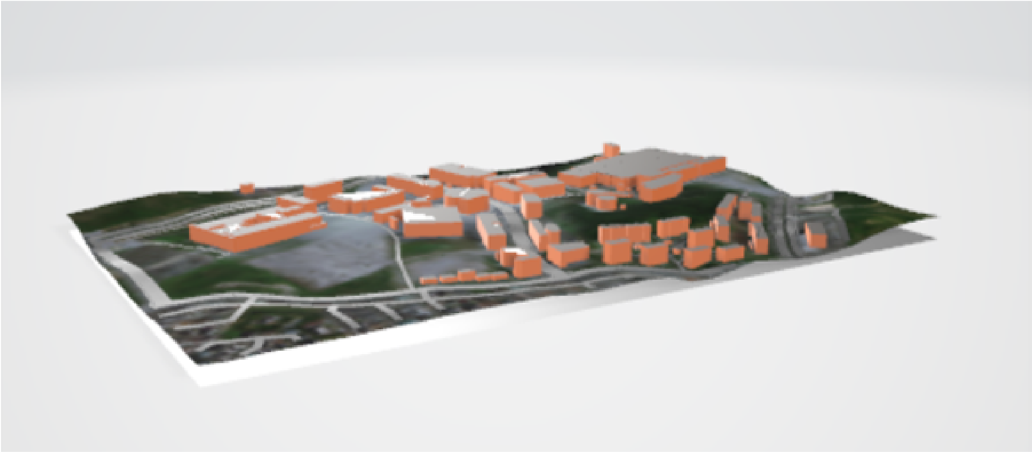
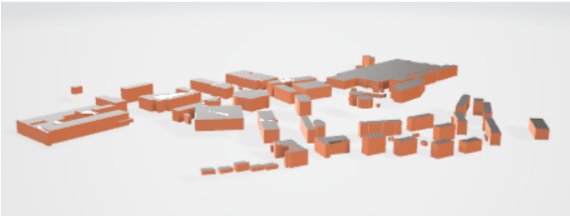
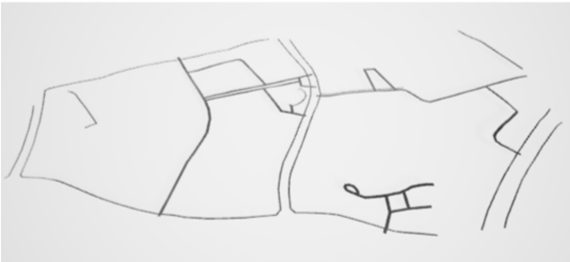


Fig. 6.9 3D terrain model

The system uses Blender [33] software to construct a three-dimensional terrain model based on the satellite photos of the area, which includes all the main buildings and roads in the area, as shown in Fig 6.9. After the 3D terrain model is rendered, the buildings and road models are extracted separately and imported into Unity to display as virtual content. Users can freely observe the model and perform some interactive operations. Fig 6.10 shows the main building model and road model in the corresponding area of the system prototype.



(a) Building model



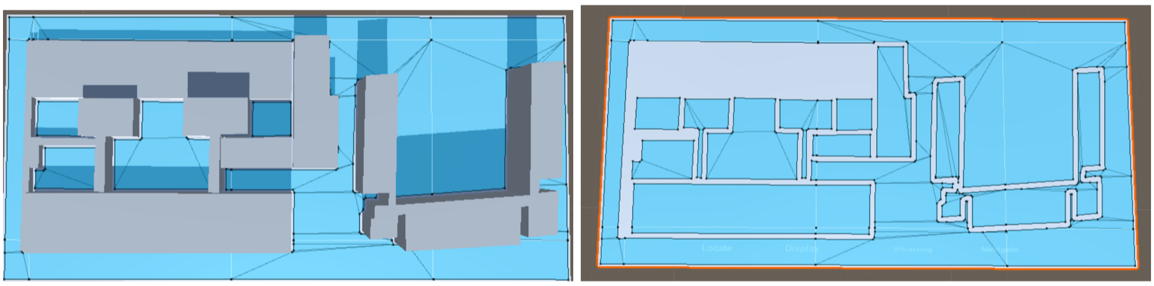
(b) Road model

Fig. 6.10 Extract part of the model

6.7 Navigation and Tips

In order to achieve the navigation function, we need to obtain the ground information of the corresponding area of the map, and roughly need to distinguish which areas are walkable areas and which areas are non-walkable areas. Then, based on this information, divide the entire map, and then use the mature A* algorithm to calculate the shortest path based on the positions of the initial point and the target point and finally display it to the user.

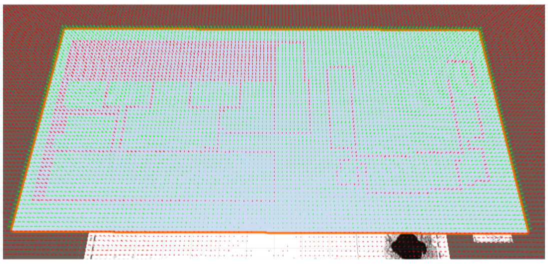
Create a topographic map of the corresponding area of the map, and divide the walkable area and the non-walkable area according to the location of the building. In the system prototype, the location of the building is simply specified as a non-walkable area, and the remaining areas are all walkable areas. Then the navigation grid NavMesh is generated according to the area division, as shown in Fig 6.11.



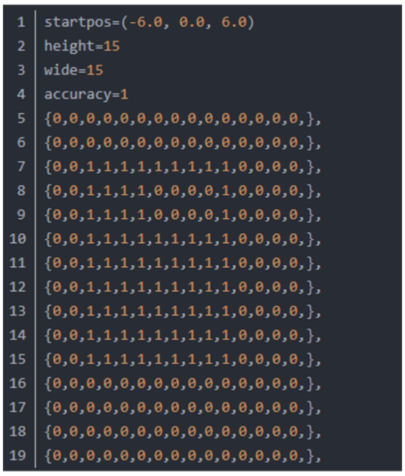
(a) With building (b) Hide building

Fig. 6.11 Generate NavMesh

And the area information is output and saved in the form of 0 and 1, where 0 means no walkable, 1 means walkable, and the output digital matrix and corresponding area are shown in Fig 6.12.



(a) Model effect



```
startpos=(-6.0, 0.0, 6.0)
height=15
wide=15
accuracy=1
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,},
{0,0,1,1,1,1,1,1,1,1,1,0,0,0,0,},
{0,0,1,1,1,1,0,0,0,0,1,0,0,0,0,},
{0,0,1,1,1,1,0,0,0,0,1,0,0,0,0,},
{0,0,1,1,1,1,1,1,1,1,1,0,0,0,0,},
{0,0,1,1,1,1,1,1,1,1,1,0,0,0,0,},
{0,0,1,1,1,1,1,1,1,1,1,0,0,0,0,},
{0,0,1,1,1,1,1,1,1,1,1,0,0,0,0,},
{0,0,1,1,1,1,1,1,1,1,1,0,0,0,0,},
{0,0,1,1,1,1,1,1,1,1,1,0,0,0,0,},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,},
```

(b) Digital matrix

Fig. 6.12 Extract area information

The system can calculate the shortest path by inputting the location of the starting point and the target point according to the area information. In addition, the system will also mark the turning point of the route. When the user approaches the turning point of the route, the turning point will change its color to remind the user. The actual operation effect is shown in Fig 5.11.The shortest path is calculated by the A* algorithm. A* algorithm is a very commonly used path finding and graph traversal algorithm.

It uses the following function to calculate the priority of each node.

$$f(n) = g(n) + h(n)$$

among them:

- $f(n)$ is the comprehensive priority of node n. When we choose the next node to traverse, we always choose the node with the highest overall priority (the smallest value).
- $g(n)$ is the cost of node n from the starting point.
- $h(n)$ is the estimated cost of node n from the end point, which is the heuristic function of the A* algorithm.

During the operation of the A* algorithm, each time the node with the smallest $f(n)$ value (the highest priority) is selected from the priority queue as the next node to be traversed.In addition, the A* algorithm uses two sets to represent the nodes to be traversed, and the nodes that have been traversed, which are usually called open_set and close_set.

The complete A* algorithm is described as follows:

- Initialize open_set and close_set;
- Add the starting point to open_set and set the priority to 0 (the highest priority);
- If open_set is not empty, select the highest priority node n from open_set:

    If node n is the end point, then:

        - Gradually track the parent node from the end point, until it reaches the starting point;

        - Return the found result path, the algorithm ends;

    If node n is not the end point, then:

        - Delete node n from open_set and add it to close_set;

        - Traverse all neighboring nodes of node n:

            If the neighboring node m is in close_set, then:

                - Skip, select the next neighboring node

            If the neighboring node m is not in the open_set, then:

                - Set the parent of node m to node n

                - Compute the priority of node m

                - Add node m to open_set

By adjusting the heuristic function, we can control the speed and accuracy of the algorithm. Because in some cases, we may not necessarily need the shortest path, but hope to find a path as soon as possible. Since movement in any direction is allowed on the map, Euclidean distance can be used.

Euclidean distance refers to the straight-line distance between two nodes:

$$\sqrt{(p_2.x - p_1.x)^2 + (p_2.y - p_1.y)^2}$$

Chapter 7

Conclusion and Future Work

7.1 Conclusion

In this paper, we introduced Magic tour, a mobile AR system that realize sharing navigation through interaction with tangible maps.

The focus of the system is on the interaction between the user and the tangible map, and provides navigation shared by multiple people. We believe that traditional paper maps have advantages that digital maps cannot replace. Users can better understand their environment through interaction with tangible maps.

Therefore, on the basis of the paper tangible map, we use augmented reality technology to enrich and supplement the detailed information of the map corresponding area that cannot be presented on ordinary paper. Users can view the building and road information in the area by touching the tangible map; users can also get their own location information on the tangible map and realize real-time tracking; in addition, the system also supports the navigation function, which will display the navigation route on the tangible map , And provide steering tips; Finally, the system supports multiplayer mode, different users can share location information and route information with each other, so that users can discuss the next travel decision.

7.2 Future Work

This research also has some limitations that could be addressed in the future studies.

Some limitations are due to existing augmented reality devices. In the current stage, we use smart phones as the display device. Although it is portable and easy to use outdoors, users still need to hold it with their hands, which has a certain degree of influence on the interaction with the physical map. If there is a wearable device that can still display better and is easy to carry in an outdoor strong light environment, such as the Google Glass under development, the interaction of the system will be more convenient.

Another aspect that the system can improve is the display effect of models and textures. In the current stage, system design and implementation are more focused on the way users interact with the system. The architectural model or texture displayed by the system prototype is relatively simple. If a better display effect is required, the fineness of the model and texture can be enhanced.

References

[1] Qinghua Zeng, Jingxian Wang, Qian Meng, Xiaoxue Zhang, and Shijie Zeng. Seamless pedestrian navigation methodology optimized for indoor/outdoor detection. *IEEE Sensors Journal*, 18(1):363–374, 2017.

[2] Andreas Dünser, Mark Billinghurst, James Wen, Ville Lehtinen, and Antti Nurminen. Exploring the use of handheld ar for outdoor navigation. *Computers & Graphics*, 36(8):1084–1095, 2012.

[3] Maha El Choubassi, Oscar Nestares, Yi Wu, Igor Kozintsev, and Horst Haussecker. An augmented reality tourist guide on your mobile devices. In *International Conference on Multimedia Modeling*, pages 588–602. Springer, 2010.

[4] George Palaigeorgiou, Anastasios Karakostas, and Keli Skenderidou. Fingertrips: learning geography through tangible finger trips into 3d augmented maps. In *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*, pages 170–172. IEEE, 2017.

[5] Magda Mpiladeri, George Palaigeorgiou, and Charalampos Lemonidis. Fractangi: A tangible learning environment for learning about fractions with an interactive number line. *International Association for Development of the Information Society*, 2016.

[6] Graeme McLean and Alan Wilson. Shopping in the digital world: Examining customer engagement through augmented reality mobile applications. *Computers in Human Behavior*, 101:210–224, 2019.

[7] Ronald Azuma, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon Julier, and Blair MacIntyre. Recent advances in augmented reality. *IEEE computer graphics and applications*, 21(6):34–47, 2001.

[8] Sandra Maria Correia Loureiro, João Guerreiro, Sara Eloy, Daniela Langaro, and Padma Panchapakesan. Understanding the use of virtual reality in marketing: A text mining-based review. *Journal of Business Research*, 100:514–530, 2019.

[9] Timothy Hyungsoo Jung, Sujin Bae, Natasha Moorhouse, and Ohbyung Kwon. The impact of user perceptions of ar on purchase intention of location-based ar navigation systems. *Journal of Retailing and Consumer Services*, 61:102575, 2021.

[10] Morten Fjeld, Nicholas Ironmonger, Sissel Guttormsen Schär, and Helmut Krueger. Design and evaluation of four ar navigation tools using scene and viewpoint handling. In *INTERACT*, pages 214–223, 2001.

[11] Georg Gerstweiler, Emanuel Vonach, and Hannes Kaufmann. Hymotrack: A mobile ar navigation system for complex indoor environments. *Sensors*, 16(1):17, 2016.

[12] Asami Tanabe and Yohsuke Yoshioka. Gazing pattern while using ar route-navigation on smartphone. In *International Conference on Applied Human Factors and Ergonomics*, pages 325–331. Springer, 2019.

[13] Gerhard Reitmayr, Ethan Eade, and Tom Drummond. Localisation and interaction for augmented maps. In *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)*, pages 120–129. IEEE, 2005.

[14] Mark Brown, Winyu Chinthammit, and Paddy Nixon. A comparison of user preferences for tangible objects vs touch buttons with a map-based tabletop application. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design*, pages 212–215, 2014.

[15] Bertrand Schneider, Patrick Jermann, Guillaume Zufferey, and Pierre Dillenbourg. Benefits of a tangible interface for collaborative learning and interaction. *IEEE Transactions on Learning Technologies*, 4(3):222–232, 2010.

[16] George Palaigeorgiou, Anastasios Karakostas, and Kyriaki Skenteridou. Touching and traveling on 3d augmented tangible maps for learning geography: The fingertrips approach. *Interactive Technology and Smart Education*, 2018.

[17] Ahmad Hoirul Basori. Hapar: handy intelligent multimodal haptic and audio-based mobile ar navigation for the visually impaired. In *Technological Trends in Improved Mobility of the Visually Impaired*, pages 319–334. Springer, 2020.

[18] Julie Ducasse, Marc Macé, and Christophe Jouffrais. From open geographical data to tangible maps: improving the accessibility of maps for visually impaired people. *ISPRS Geospatial Week*, 10(1):517–523, 2015.

[19] Anna Petrasova, Brendan Harmon, Vaclav Petras, Payam Tabrizian, and Helena Mitasova. *Tangible modeling with open source GIS*. Springer, 2018.

[20] Saul Greenberg, Michael Boyle, and Jason LaBerge. Pdas and shared public displays: Making personal information public, and public information personal. *Personal Technologies*, 3(1-2):54–64, 1999.

[21] Anton Fedosov, Ivan Elhart, Evangelos Niforatos, Alexander North, and Marc Langheinrich. Skiar: Wearable augmented reality system for sharing personalized content on ski resort maps. In *Proceedings of the 7th Augmented Human International Conference 2016*, pages 1–2, 2016.

[22] Derek Reilly, Michael Welsman-Dinelle, Colin Bate, and Kori Inkpen. Just point and click? using handhelds to interact with paper maps. In *Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*, pages 239–242, 2005.

[23] Jaime Teevan, Amy Karlson, Shahriyar Amini, AJ Bernheim Brush, and John Krumm. Understanding the importance of location, time, and people in mobile local search behavior. In *Proceedings of the 13th international conference on human computer interaction with mobile devices and services*, pages 77–80, 2011.

[24] Hiroshi Ishii. The tangible user interface and its evolution. *Communications of the ACM*, 51(6):32–36, 2008.

[25] Mark Weiser. The computer for the 21st century. *ACM SIGMOBILE mobile computing and communications review*, 3(3):3–11, 1999.

[26] Hiroshi Ishii and Brygg Ullmer. Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, pages 234–241, 1997.

[27] David H Uttal. Seeing the big picture: Map use and the development of spatial cognition. *Developmental Science*, 3(3):247–264, 2000.

[28] Mark Billinghurst and Hirokazu Kato. Collaborative augmented reality. *Communications of the ACM*, 45(7):64–70, 2002.

[29] Elin Rφnby Pedersen, Kim McCall, Thomas P Moran, and Frank G Halasz. Tivoli: An electronic whiteboard for informal workgroup meetings. In *Readings in Human–Computer Interaction*, pages 509–516. Elsevier, 1995.

[30] Anton Fedosov, Evangelos Niforatos, Ivan Elhart, Teseo Schneider, Dmitry Anisimov, and Marc Langheinrich. Design and evaluation of a wearable ar system for sharing personalized content on ski resort maps. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia*, pages 141–152, 2016.

[31] Photon Inc. Photon pun. https://www.photonengine.com/zh-CN/PUN, 2021.

[32] Vuforia Inc. Virtual button in vuforia. https://library.vuforia.com/articles/Solution/How-To-Implement-Virtual-Buttons.html, 2021.

[33] Blender Inc. Modules. https://wiki.blender.org/wiki/Modules, 2021.