Design of Smart Home Control System Based on XR

Shitian Zeng

School of Software & Internet of Things Engineer, Jiangxi University of Finance and Economics, Nanchang, 330000, China
Sakina_ove@outlook.jp

Abstract. In the 21st century, with the rapid development of science and technology, the whole world is accelerating the exploration of extended reality (XR) technology. The application of XR technology can be seen in education, medical treatment and other fields. Up to now, XR technology has become a hot point in scientific and technological research. In the same era, the Internet of things (IoT) technology is also developing rapidly. The integration of Internet of things technology and traditional home has promoted the great change of home life, which is a great leap from traditional home to smart home. This paper aims to combine the XR technology and the smart home system supported by the Internet of things technology, and puts forward the design of smart home control system in XR equipment (AR glasses in this paper as an example). Combined with gesture recognition, eye tracking and voice control, the virtual menu and control method of the smart home control system are designed in detail, and the virtual instructions are transformed into the control effect of the real home appliances through the communication protocol, and the users are given timely feedback. Through the design, the convenience of people's life in the era of widespread XR technology in the future will be improved.

Keywords: XR, Smart home control system, Gesture recognition, Virtual menu.

1 Introduction

XR technology, or extended reality technology, is a general term for all "blending of the virtual and real" technologies. It includes augmented reality (AR), virtual reality (VR), and mixed reality (MR), which is a collection of them, as well as all related technologies in the future. In modern society, XR technology is developing rapidly. It has not only become a hot topic of research, but also integrated into every aspect of people's lives. In the education field, VR technology creates enough scenes to simulate the real world for students, and it can change abstract concepts into corporeal things, stimulating students' enthusiasm and curiosity for knowledge in a variety of ways. By superimposing information in the real world, AR technology can provide students with additional information of what they are learing. It also achieves the gamification of education to cultivate students' teamwork ability and sets up more personalized learning plans [1]. In the architecture field, the application of VR technology to realize immersive 3D design provides a new way of design and display, greatly improves the

G. E. Atalay and A. Wang (eds.), *Proceedings of the 2026 3rd International Conference on Public Relations and Media Communication (PRMC 2026)*, Advances in Social Science, Education and Humanities Research 1032,
https://doi.org/10.2991/978-2-38476-609-3_68

communication efficiency among designers, customers and construction parties, making the complex design intuitive [2]. In the film and television field, XR technology and virtual shooting technology enable creative teams to create freely in virtual scenes and use far more than traditional shooting methods to writing their imagination [3]. In the field of power systems, MR technology is used to overlay grid resource data, facilitating the development of three-dimensional visualizations of grid resources and improving the efficiency of equipment maintenance and repairs [4].

In recent years, people's demand for living environment is no longer limited to living, but develops towards comfort, safety, convenience and other more aspects. Internet of things technology connects independent devices at home to achieve unified intelligent deployment [5]. It not only provides convenience for people's home life, but also improves comfort and happiness.

However, the application of XR technology mentioned above is more inclined to the scenes with practical significance, but ignores the more everyday aspects of home life. Therefore, according to the rapid development trend of XR, this paper believes that it is necessary to combine XR technology with smart home. Based on this, this paper proposes a design of smart home control system based on XR, and designs the system's virtual menu interface and specific control methods using gesture recognition, eye tracking, and speech recognition technologies. It aims to expand the application carrier of smart home control system and improve the convenience of people's life.

2 Overview of Related Technologies

2.1 Gesture Recognition Technology

Instead of using devices such as gloves to transmit gestures, this paper chooses a more convenient vision-based gesture recognition technology carrying a depth camera for visual positioning. The vision-based gesture recognition process includes depth perception, hand detection and gesture recognition. In particular, hand detection is divided into hand segmentation and hand tracking two sub-steps [6]. In this system, the user will use specific gestures to operate, so the steps of hand tracking will be strengthened to make corresponding judgments. Touchscreen operation has already been a part of daily life, which means finger gesture interaction has a higher level of social acceptance [7].

2.2 Eye Tracking Technology

Eye-tracking technology consists of eye-tracking data acquisition and gaze-tracking algorithms two components. The gaze-tracking algorithms can be divided into two methods based on corneal reflection and visible light. After obtaining the position of the user's eye gaze, the corresponding feedback can be designed [6]. For example, when the user looks at an option for more than three seconds, the option will increase the brightness or volume, indicating it is selected by the user. The main advantages of eye

tracking technology are rapidity, ease of use and intuitive. When the body cannot move in a large range or the hands are occupied, its utility will be maximized [7].

2.3 Speech Recognition Technology

The functions designed in this system are clear, as a result, users can use the system conveniently. Speech recognition technology generally consists of speech acquisition, voice preprocessing, feature extraction and algorithm recognition four steps. The microphone collects the user's voice, removes the noise, extracts the keywords in the voice information, and finally restores the user's instructions and needs as close as possible through the algorithm [6]. Voice interaction only is considered the most labor-saving way of interaction, because speaking in natural language requires almost no physical strength [7].

3 Design Description

Traditional AR glasses tend to rely on one single control method to recognize user inputs. In order to improve the accuracy of recognition and adapt to the preferences and needs of different kind of people, the system designed in this paper combines gesture recognition, eye tracking, and speech recognition technologies. Users can set their preferred primary identification method based on their individual preferences, and the remaining two methods will serve as secondary options.

In order to realize the above three technologies successfully, the system required AR glasses must have relevant field of view (FOV), sensors, central processing unit (CPU) with computing unit and other indispensable parts. Sensors mainly include depth sensors, cameras, microphones, etc. The computing unit is mainly responsible for the calculation of hand tracking and eye tracking.

3.1 System Usage Process

The process for the entire system is illustrated in the flowchart (Figure 1).

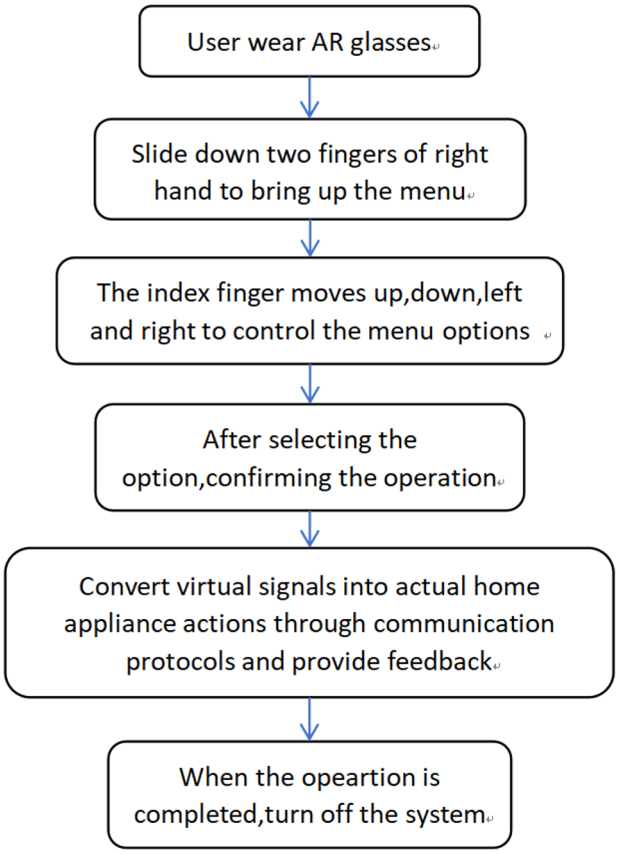

Fig. 1. Flowchart of smart home control system (Picture credit: Original).

3.2 Starting the System

After wearing compliant AR glasses that meet the requirements, the user can call out the main menu through the gesture of slide down with right index and middle fingers together. The virtual main menu interface will be displayed at the corresponding position (as shown in Figure 2) according to the recognized position of the user's initial finger tip. As a demonstration, the second option in the main menu is the smart home control system designed in this paper, and the last option is setting, which contains the functions of setting operation preferences and exiting the system.

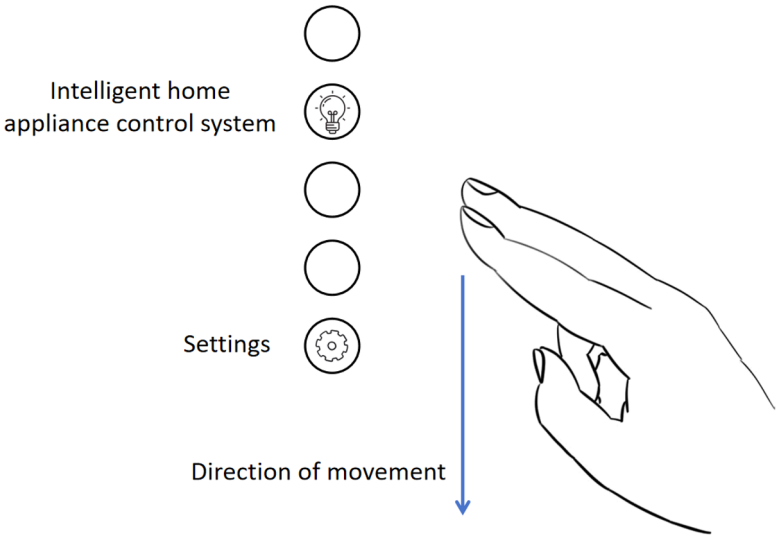

Fig. 2. Slide down two fingers of right hand to bring up the main menu (Picture credit: Original).

3.3 Control Options

After entering the system, use the gesture of only extending the right index finger as the default selection gesture. Ar glass monitors the direction and movement of the user's right index finger through the camera, and calculates a straight line that can indicate the direction with the tip of the index finger as the starting point (as shown in Figure 3).

If the predicted pointing line intersects with the virtual menu, highlight and enlarfe the intersecting option, indicating the option is preselected. In the process of preselection, the user can adjust the position by bending finger and moving hand up and down, so as to switch the options in preselection.

The menus designed for this system has polymorphism. The camera and algorithm of AR glasses will obtain the user's line of sight. When the user's line of sight looks at the option above the preselected option and stays for more than 3 seconds, the content displayed on the menu will be switched accordingly to switch the preselected option. The microphone on AR glasses can obtain the user's voice input. When the user want to use voice input, needs to say "enable voice mode" first. At this time, the speech recognition process starts to collect the user's voice. Since the speech recognition technology used by this system restores the user requests based on captured keywords, users can issue commands using everyday conversational phrases. For example, after enabling the voice mode, the user says "reduce the air conditioning temperature by 2 degrees", and the system will automatically call up the air conditioning control panel and reduce the set temperature by 2 degrees after collecting the user's voice command. Voice can also be used to control household appliances, and can also be used as an auxiliary choice under the premise of gesture control.

When the desired result is selected as a preselected option, move the index finger to the right for a certain distance (when it is greater than a certain threshold, the system will interpret the user's action as a selection gesture) to select it. If there is a Subdirectory, it will enter the next round of selection.

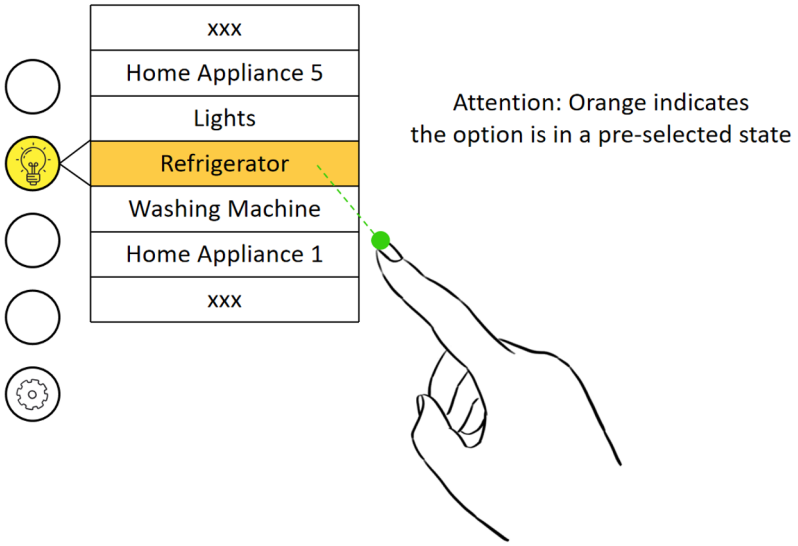

Fig. 3. Calculate the direction by fitting a straight line based on the fingertip direction (Picture credit: Original).

Because the smart home control system includes a variety of controllable devices, a circular roller was chosen as the way to display all options.

The assumption is to display five options. The default option is the third option (the middle item) in the preselected state, and the remaining options are omitted above or below the roller. The user can switch by moving fingers up and down. After making a selection, the user can still select by swiping to the right (as shown in Figure 4).

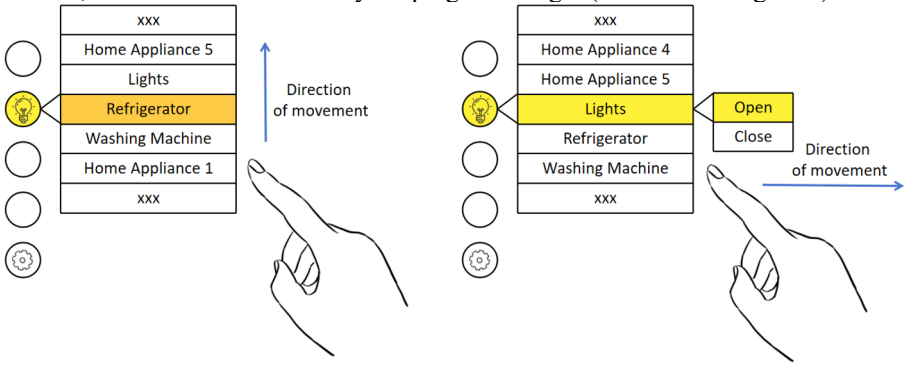

Fig. 4. Mobile selection and translation confirmation (Picture credit: Original).

3.4 Example Analysis

Taking the switching of the light as an example, this paper illustrates how to control household appliances. If the light option is selected in the system, and the indoor electric light is turned on, the default "on" is selected in the subdirectory of the light (as shown in Figure 4). To turn off the light, the user only needs to move finger down, and the system will switch "off" to the preselected state when it detects the gesture. Move finger to the right to confirm the selection. At this time, the new window confirmation operation "switch the light to 'off' mode" will pop up. Hang finger on the confirmation option for a few seconds to complete the confirmation, and the new window will automatically close.

In Laghari's paper, it is mentioned that the Internet of things can generally be divided into sensor and actuator devices, Internet gateway and data acquisition system, pre-processing edge IT and cloud and data center these four stages [8]. Zhou and Zhang's paper said that the working process of the smart home IoT remote control system is that the sensor transmits the collected data to the main control chip, and then transmits the data to the user through the network transmission technology and application layer protocol, and then the main control chip transmits the instructions to the actual home appliance to complete the remote control [9]. Based on these, the system designed in this paper will also realize remote control through a similar process.

After the user confirms the virtual instruction, it will be converted into an instruction of message queuing telemetry transport (MQTT) protocol and transmitted to the cloud. Then, the instructions are delivered to the home appliance for execution in the communication protocol with Wi-Fi as the medium, and the user feedback is given synchronously on the virtual menu.

3.5 Reselection and Shutdown System

After completing the control of one home appliance, the user may need to continue to control other home appliances. Because the properties of the previously selected list will be locked on the same straight line, the user can translate the index finger back from the end along the straight line and hover for a few seconds when it reaches the directory want to unlock. When the menu is unlocked, the subordinate directory will disappear. The remaining operations are the same as before.

When the user wants to shut down the system after completing all operations, there are two ways to exit. One is to enter the setting options at the end of the main menu, including the exit button. Another method is to use a specific gesture with right hand horizontally to the right with the palm facing outward. By using this gesture, the system will close immediately.

4 Application and Expansion

The switching of light has been described above as an example, but other household appliances are also mentioned in Figure 3. They can still be included in the system designed in this paper, such as air conditioners, washing machines, refrigerators and so

on. Therefore, in theory, the system can be used to achieve the unified control of household appliances in the whole house. The specific operation only needs to add all household appliances in the user's residence in the setting.

While expanding to a wider range of applications, the system can also be adapted to the centralized control of factory workshops and the overall allocation of materials after a certain degree of change.

For the applicable population of this system, at this stage, it has a more intuitive and significant effect on the patients and the elderly who have difficulty in moving. People with difficulty in moving do not need to go to a specific switch in person. They can easily control household appliances by gesture. With the advance of the times and the development of technology, when XR devices become as popular as today's mobile phones, it will become a general trend to combine the needs of daily life with them.

The system designed in this paper still has many shortcomings. For example, the ancient home appliances do not support the connection of smart home control system through the Internet of things, and different brands of home appliances support different communication protocols, resulting in incompatibility between home appliances and the system. During the operation of the XR device, the camera needs to be turned on for a long time to continuously record the situation at home. If the device system leaks secret securities, it will cause the user's privacy to be leaked, and may even be taken advantage of by other people with intentions to infringe on the user's privacy [10]. At the same time, the interface design of the system is too similar to the existing virtual screen interface, leading to a lack of gamification design.

For the third disadvantage mentioned above, this paper proposes an additional design which is to add the relevant content of "target task". The additional design can help patients with procrastination finish housework in time through a gamified process. Take washing clothes as an example. The user can set the task of washing clothes every x days. When the user starts the washing machine by operating the smart home control system, if there is a specified number of days between the last time the washing machine was started, the reward system in the system will jump out of the corresponding feedback to encourage the user to finish the housework on time. With this design, the gameplay of the system is increased.

5 Conclusion

In this paper, through the existing gesture recognition, eye tracking, speech recognition technology, the intelligent home appliance control system based on AR glasses is proposed, and the virtual menu interface and control method are designed detailly, and the process is demonstrated completely. The system proposed in this paper is mainly conducive to the operation of home appliances for the sick and elderly who are not convenient to move, and it is hoped that XR technology will be closely connected with people's daily life in the future. There are still some problems in this paper, such as the lack of adaptability between home appliances and the system, the protection of user privacy, and the monotonous interface design. The future research should investigate the adaptability of household appliances in the market and the system designed in this

paper, increase the logic of user privacy protection, and improve the interface design to make it more gamified. The above measures can be improved and optimized in the subsequent research and design.

References

1. Schwaiger, M., Krajnčan, M., Vukovič, M., Jenko, M., Doz, D.: Educators' opinions about VR/AR/XR: An exploratory study. Education and Information Technologies. Volume 29,24861–24880 (2024).
2. Zhou, M.: Application of VR Technology in Immersive Three-dimensional Architectural Design. Intelligent Building & City Information S1,40 – 42 (2025).
3. Xie, X., Xu, G., Li, X., Zhang, Y.: Research on the integration and application of XR and AIGC technology in film and television creation. West China Broadcasting TV 46(17),93 – 96 (2025).
4. Liu, J.: AR, VR and MR technology Application in power system. Wireless Internet Science and Technology 18(22),78 – 79 (2021).
5. Huang, D.: Application of Internet of things technology in smart home control system. In: Proceedings of the 9th Symposium on digital transformation and new productivity development in engineering field in 2025, pp.131 - 133. Preprint: Proceedings of the 9th Symposium on digital transformation and new productivity development in engineering field in 2025, Nanning (2025).
6. Chen, X.: Research on Multimodal Menu Based on One-handed Scenarios in Augmented Reality. Yanshan University (2024).
7. Wang, Z., Rao, M., Ye, S., Song, W., Lu, F.: Towards spatial computing: recent advances in multimodal natural interaction for Extended Reality headsets. Front. Comput.Sci.19(12),1912708 (2025).
8. Laghari, A.A., Li, H., Khan, A.A., Yin, S., Karim, S., Khani, M.A.K.: Internet of Things (IoT) applications security trends and challenges. Discover Internet of Things 4,36 (2024).
9. Zhou, F., Zhang, Y.: Design of smart home system based on IoT remote control. Practical Electronics 02,17 – 19 (2025).
10. Xie, P.: Analysis of privacy risk in VR news application. Media Observer 06,8 – 10 (2017).