Journal of American Science

Websites:
http://www.jofamericanscience.org
http://www.sciencepub.net

Emails:
editor@sciencepub.net
sciencepub@gmail.com

MARSLAND PRESS
Multidisciplinary Academic Journal Publisher

Research on Interaction Mode of Mixed Reality System Based on Embodied Cognition Theory

Ziyu Hu [1], Na Li [2]

1. School of Architecture, Harbin Institute of Technology, Heilongjiang Harbin, China;
2. School of Humanities, Social Sciences and Law, Harbin Institute of Technology, Heilongjiang Harbin, China)

Abstract: Embodied cognition theory reveals the important role that the physical structure of human body plays in the cognitive process, and inspires people to study people's cognitive behavior from the perspective of physiological experience and mental state. Mixed reality technology combines the advantages and characteristics of virtual reality and augmented reality, combines 3D virtual picture with digital reality, and builds an interactive information feedback channel among users, the real world and the virtual world. Under the effect of mixed reality technology, people can interact with the digital information in the virtual space in reality. From the perspective of embodied cognition theory, based on people's physical experience, this paper studies the interaction design model of mixed reality space. By analyzing the embodied mechanism and cognitive logic of mixed reality system in the interaction process, it summarizes the user experience model of mixed reality system, which is more conducive to clarifying the interaction target of mixed reality system. Create mixed reality systems that better meet people's needs.

[**Ziyu Hu, Na Li; Research on Interaction Mode of Mixed Reality System Based on Embodied Cognition Theory.** *J Am Sci* 2022;18(12):30-35]. ISSN 1545-1003 (print); ISSN 2375-7264 (online). http://www.jofamericanscience.org. 04. doi:10.7537/marsjas181222.04.

Key words: Embodied cognition, mixed reality, user experience model, Human machine interaction

Introduction:

In recent years, with the development of computer technology and interactive ways based on mixed reality technology, MR Has gradually become a research hotspot in various industries, widely used in medical imaging, clinical simulation, aerospace, mechanical systems, education and teaching and other fields. The interactive mode based on mixed reality, through the digital information presentation, can not only effectively improve the system performance, but also expand the cognitive ability of users. People can make better decisions by simulating different situations, thus helping to improve people's ability to solve problems. In mixed reality systems, massive data, real-time synchronization and dynamic feedback are finally presented to the user in the form of an interactive interface. The performance of mixed reality system, user experience feedback and user data reception all depend on the presentation of visual information, which has become a common focus of research in computer science and other fields. In order to enrich the application scenarios of mixed reality system and improve user experience, in addition to the breakthrough of technical difficulties, the use of pain points is also an important link that cannot be ignored. How to make a more scientific layout of mixed reality interface, achieve a truly effective combination of digital information and the real world, and improve the efficiency of user positioning information and information collection? Especially in the medical, military, national defense and other important fields, therefore, it is very important to study the information interface presentation mode and scientific and efficient interaction mode of mixed reality system, and put forward the information presentation mode and interaction design principle that can better meet people's needs.

I. The theory of embodied cognition is proposed

1.1 The development of Merleau-Ponty's phenomenology of perception

Phenomenology, first proposed by Husserl, is also called "descriptive phenomenology". It holds that people should abandon existing knowledge, experience and memory when understanding the world, and only focus on directly experienced phenomena and describe them. After Husserl, Merleau Ponty attempted to reveal the deeper nature of phenomenology from the perspective of perception. He criticized Descartes' "mind-body dualism" and held that the separation of body and consciousness would hinder people's exploration of the correct relationship between body and mind. Body and consciousness are inseparable in the cognitive process, and people perceive the world with the help of the body, so the body is the basis for people to understand the world. The interaction between the body and the outside world shapes people's behavior,

and cognition, body and the world are mutually unified[1]. Merleau Ponty's "phenomenology of perception" is a further development of Husserl's phenomenology and the direct ideological source of embodied cognition theory.

Different psychologists have different definitions of embodied cognition, but they all focus on the relationship between the body and the environment, affirming the important role the body plays in the formation of cognition. In general, embodied cognition theory summarizes the characteristics of cognition as embodied cognition, situational cognition and dynamic cognition[2]. The embodiment of cognition means that the intuitive feeling of the body at the physical level will affect the formation of cognition. For example, people will feel relaxed when they touch a soft medium, while they will feel nervous when they touch a hard medium. In the process of cognition formation, the sensory experience brought by the touch of human physiological structure will map the first impression of something in the human brain. Situational cognition means that people's cognition of things is affected by the surrounding environment. The relationship between cognition and situation has been studied in behaviorism and constructivism. The dynamic nature of cognition means that people's cognitive behavior will change over time. For example, when people contact a new sport for the first time, they will feel strange and stiff, and most of their attention will be on their own body. However, with the integration of body, scene and time, people will not only have a more relaxed and clearer cognition. And you can see more of what's going on around you while you're doing this exercise. People's cognitive state and cognitive experience are not invariable, and will change with the changes of various elements.

1.2 Cognitive logic in mixed reality system interaction

The concept of interaction design was proposed by Bill Moggridge, one of the founders of IDEO. The initial purpose of interaction design was to combine computer software with user interface design. It was not until the 1990s that the research focus of interaction design gradually developed to "human-oriented" human-computer interaction. Its theory is widely used in digital information interaction, intelligent somatosensory interaction, multi-modal interaction and other technical fields. The concept of interaction design has become an independent discipline. Compared with ergonomics theory, interaction design lays more emphasis on human behavior, cognition, psychology and so on, adhering to the "human-oriented" design concept. With the rapid development of virtual reality technology, in addition to breakthroughs in key and difficult technologies, interactive experience has also become the focus of research. Because there are many factors in virtual reality system, such as immature hardware terminal equipment, large demand for computing resources, network environment restrictions and other factors, unnatural interactive experience is also an important reason why it cannot be truly popularized and serve the masses in a short time. Compared with virtual reality, mixed reality technology uses 3D modeling, real-time tracking and registration, intelligent sensor interaction, 3D projection and other technologies to generate a reality-like environment at the cross boundary between the real world and the virtual world, enabling users to gain a sense of immersion while enhancing the authenticity of experience. Users can live closer to the experience of the real world in the mixed reality system, and portable terminal devices also make the interaction process more natural. To study the interactive cognition process of users in the mixed reality system, the concept of "sense of presence" is indispensable.

The cognitive process of users in mixed reality system mainly includes the following stages: receiving signals, gaining presence, input signals, and gaining immersion. The sense of presence is an important indicator to measure user experience, which refers to the natural experience generated when users perceive the existence in the virtual environment through sensor equipment. Only on the basis of the sense of presence can users generate emotional experience in the process of interacting with the virtual system. In the field of mixed reality interaction, the main research directions focus on gesture recognition, remote assistance, multi-user interaction, virtual-real interaction, etc. [3]

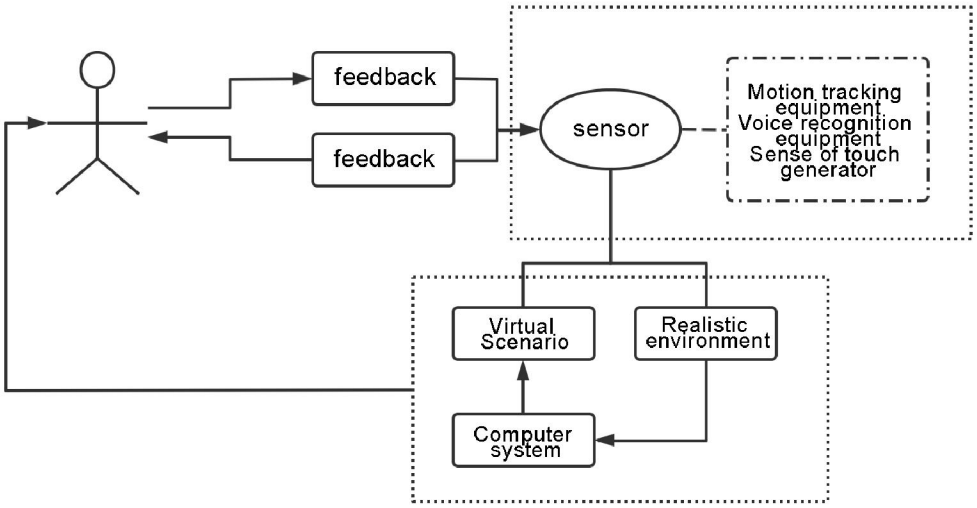

Figure 1 Mixed reality system cognitive model

2. User experience model of mixed reality system

The concept of user experience was first created in the 1940s as an important index to measure the usability of products in the field of human-computer interaction. With the continuous development of technology, the theory of user experience has also been gradually improved, among which, the widely recognized definition refers to the physical and psychological feelings accumulated by users through interaction with a product in the process of contact and use. It includes sensory impression, cognitive memory, physical feeling and mental experience. Bernd H. Schmitt studied user experience from the perspective of psychosociology and proposed a user experience evaluation system based on sensory, emotional, thinking, behavior and correlation [4]. When users use products to gain experience, it is also the cognitive process of users. The user experience model of mixed reality system can be divided into three parts: sensory experience, interactive experience and behavioral experience.

2.1Sensory experience

By wearing sensing devices, users interact with computer-generated 3D models in real scenes to generate visual, tactile, auditory and other sensory impressions. The comfort, immersion and authenticity of sensing equipment are important factors that affect user experience. Comfort refers to the control factors, such as thermal management and ergonomics, that affect the wearing experience of display devices. Immersion and authenticity refer to the ability of the system and equipment to process and interpret the user's environment in real time. Field of vision, field depth and pixel density are all important indicators to measure the immersion brought by the system to users. Microsoft Hololens is one of the most complete, widely used and technologically advanced headsets (mixed reality head-mounted displays) in the field of mixed reality at present. It has been widely used in manufacturing, engineering construction, healthcare, education and other industries. Hololens can track the user's line of sight and real-time position, generate holographic projection of 3D virtual objects and project it into the user's eyes. Therefore, image clarity, color effect and other factors are positively correlated with user's immersion. If the virtual object has obvious color difference and low clarity, users will quickly distinguish the virtual image from the real environment and cannot get immersion from the system. Resulting in a poor experience. Nreal Light is a mixed reality glasses released by Nreal in 2019. Compared with Hololens, NREAL uses Birthbath optical solution to present a clearer virtual image without rainbow interference, enabling users to receive light signals closer to the real world.

2.2Interactive experience

The sensor of the mixed reality system can automatically locate the position of other objects in and around the room, and project 3D virtual images on the surface or inside of objects by using holographic projection technology. Users often interact with virtual objects through gestures, gaze or voice commands. Common gestures include left swipe, right swipe, zoom in, zoom out, click, etc. In the interaction process, the timeliness of system feedback, the fluency of virtual object interaction and the overall interactivity of the system are important factors that affect user experience. RoboRaid is a shooting game based on Hololens mixed reality platform. After registering and logging in, players need to first look around and ask the device to complete the environment scanning. Due to the plot and game mechanism setting, the game can only start by scanning

more than three walls (or wall-like occlusion). The Hololens projects 3D virtual alien creatures onto the walls. You can either shoot them to kill them or shoot them at the walls. The device projects deep 3D virtual destruction onto the walls, and you can hear the realistic sound of the walls cracking. Hololens will recognize the player's hand movements, track the position of the hand and the player's line of sight, and detect and project the aiming center according to the real-time position. When the player's thumb and index finger touch, the firing function will be triggered along the center direction, and the hit object will respond immediately (explosion disappears, etc.), providing the player with timely feedback and smooth interactive experience.

Figure 2, Figure 3 RoboRaid game interface

2.3Behavioral experience

The operability, learnability and usability of the system are also important factors affecting user experience. The operation mechanism of the mixed reality system should follow the people-oriented design concept and fully consider the physiological structure, living habits and behavioral habits of the human body. User behavior of mixed reality system is characterized by the combination of virtual and real. Users are in the real world and virtual world at the same time, and can receive signals and messages from two different attribute worlds at the same time. Therefore, the system is required to correctly handle the relationship between virtual and real things, including the physical position relationship between objects, user action prediction and timely interactive feedback. The physical position relationship between virtual and real objects includes distance, perspective, occlusion, etc., which depends on the environmental scanning in the early stage of the system and the three-dimensional position tracking in the late stage. While scanning the real space, the existing objects are marked, and the virtual objects are projected on the surface or floating in front of the objects, so that the players can easily find and successfully interact with them. The wrong positional relationship can affect the speed and experience of the interaction. When setting the projection position of virtual image, the system should reasonably predict the user's subsequent actions. Different action paths and actions correspond to different interaction triggering ways.

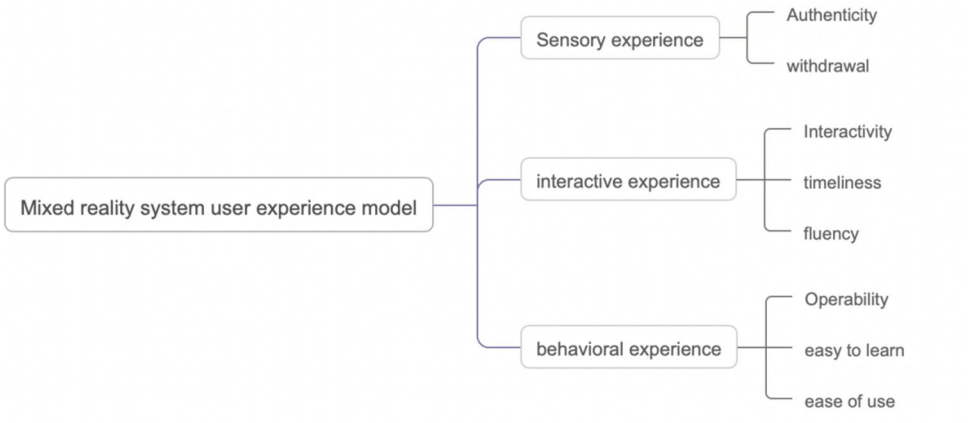

Figure 4 User experience model of mixed reality system

3. Interactive target of MR System under the guidance of Embodied cognition theory

3.1people-oriented, environmental integration, low load information presentation.

The interaction model of the mixed reality system focuses on the connection between the physical world and the digital world. The realization of simple instinctive interaction is an important concept throughout the interaction design of the mixed reality system. The system should try its best to ensure that users can focus on the actual task rather than the redundant information brought by the environment. Adhere to the user experience-centered design, according to people's life and cognitive habits, the system internal operation behavior presets, mixed reality interface information in terms of data sources have multiple sources; In terms of information characteristics, it is real-time and dynamic. In the way of information presentation, it has the characteristics of overlap, multi-dimension and space[5].

3.2 Higher frame rate, faster speed and more realistic hologram.

Holograms in virtual reality system are objects that appear around users and are composed of light and sound. Hologram technology is a technology that uses interference and diffraction principles to record and reproduce the real three-dimensional images of objects[6]. Holograms respond to the user's gaze, gestures and voice commands and interact superficially with real objects around them. HoloLens provides a detailed representation of the actual surfaces in the system's surrounding environment through spatial mapping that helps anchor objects in the user's world and makes full use of real-world depth cues to enhance the authenticity of the hologram and the confidence of the system. The main factors affecting spatial mapping are as follows: positioning, occlusion and closure, physical simulation, navigation, and visualization effect.

Generally, holograms can be located in two ways, "body lock" and "display lock". "Body locking" refers to binding the hologram to the user's body or staring state, and the image will move according to the user's Angle of view. "Display lock" binds the hologram to the display system in a way that doesn't work well with virtual reality, where users can easily lose sight of the target and get bored. When the user is in the mixed reality system, the attention of the external elements will still be diverted, causing the focus of interest to shift. The use of "body lock" can ensure that the user does not lose the hologram in the process of action. Spatially mapped surfaces can block holograms and have a significant impact on their perceptual authenticity, helping to create an instinctive sense of truly being in the same physical space as the user. Physical simulation can enhance the presence of holograms in the user's physical space and provide applications with opportunities for natural and common interactions of physics. Navigation and visualization should minimize visual clutter through visual surfaces that allow the application to share an understanding of the environment with the user.

3.3Convenient operation, reasonable feedback, multi-mode interaction.

Due to its limited information flux, it is difficult for a single mode signal to fully represent the user's design intention, which can be effectively overcome by means of multi-mode fusion[7]. With the continuous progress of technology, human-computer interaction mode has developed from the initial command line interface and graphical user interface to the natural human-computer interaction stage represented by three-dimensional user interface and multi-channel user interface. Multi-modal refers to the integration of multiple senses, that is, human-computer interaction through text, speech, vision, action, environment and other ways. The current multi-mode interaction modes used in mixed reality systems mainly include hand operation system, eye movement operation system and voice control system:

3.3.1 Hand direct operation system combined with virtual and real

The hand and motion controller mode is one of the most widely used modes in mixed reality systems. The system can recognize the user's hand and interpret it as the skeleton model of the left and right hands. The spherical touch body is bound at the position of the fingertip. The user can interact with the object at a great distance. The use of hardware devices such as motion controllers can extend the user's physical capabilities, improve the accuracy of the user's interaction with the hologram, and provide real-time tactile feedback to the user.

3.3.2Eye-based gaze and tracking

Eye tracking and voice input are collectively referred to as manual free mode. Eye tracking technology enables users to interact with the system smoothly when it is not convenient to use their hands. The device can bind the interactive interface to human eyes through head eye tracking, so as to realize dynamic interface tracking, intelligent page turning, intelligent scrolling and other functions, as if users are not deliberately interacting with the system. It's a combination of the two, a synchrony. The way of gaze input has the advantages of high speed directivity, effortless, implicit and predictive. The eye muscle is the fastest muscle in the human body. The system can infer the object that the user is paying attention to through the user's eye movements, and determine the target that the user will interact with.

3.3.3 Voice input and control

Language is the most natural way for us to express our intentions. The use of voice input in mixed reality system can save interaction time and has a powerful

emotional effect on users' experience and perception. Voice control is also one of the most common interaction modes in the field of AI. Changes in size, orientation, scene interface switching, function selection, etc. Voice dictation allows users to free their hands, enter text into the system more efficiently, and quickly locate targets. Voice input not only improves user operation efficiency, but also has many drawbacks. First, users need to change their language habits and try to fit the built-in voice commands of the system to improve the success rate of voice control. Secondly, in order to avoid receiving wrong instructions, users should first wake up the voice function when they input and control the voice. For example, before using the voice function of iPhone, users should first "hey siri", and only after the correct password can they wake up the intelligent voice system: In addition, timely feedback and scientific guidance are also more important. Repeated input of voice is often more annoying to users than words or images. Hololens adopts the voice model of "See it, speak it out". Users can speak voice commands that can be recognized by the system through the content displayed on the label, which improves the success rate of command triggering. Therefore, voice commands in the mixed reality system should follow the requirements of concise meaning, simple vocabulary, non-destructive, consistency, etc. Only in this way can voice control be truly used to improve user experience[8].

Conclusion

The ultimate goal of mixed reality system interaction is to achieve natural and instinctive interaction, so that users can rely on intuition rather than learning guidance and easily use the system to achieve their goals. With the innovation of interactive equipment and interaction technology, the traditional interaction model has been unable to meet the needs of mixed reality system. Multi-modal interaction will be the focus of research and application in the field of mixed reality interaction. Hat the balance between artistry and authenticity of virtual scenes should be further studied in the future, the timeliness and innovation of news story content should be further improved.

References

[1] ZHU Xiangrong. Research on the Body Phenomenology of Merleau-Ponty [D]. Beijing: Minzu University of China, 2018.

[2] Li Qiwei. Cognitive Revolution and the Second Generation of Cognitive Science. Acta Psychologica Sinica, 2008 (12) : 1315-1320.

[3] XIE Guojun. Research and implementation of real-time Virtual-real interaction technology in Mixed reality [D]. Xian university of electronic science and technology, 2021. DOI: 10.27389 /, dc nki. Gxadu. 2021.000463.

[4] Research on Profit model of Augmented Reality Publishing [J]. Yang Xiaoxin, WU Haozhen. Science and Technology and Publishing. 2020(11)

[5] Yan Ketong, Shao Jian, Liu Ke. Research on Man-Machine Interface Design for Mixed Reality [J]. Development and Innovation of Electromechanical Products,2021,34(01):26-28.

[6] HUANG Wenjun. The application of Hologram technology in Wisdom Education [J]. Science and Technology Communication, 201,13(22):166-168.

[7] Niu Hongwei, Hao Jia, Cao Bening, Long Hui, Zhang Feifan, Wang Guoxin. Multi-modal Intelligent Interaction Framework for Product Conceptual Design and Its Implementation [J]. Computer Integrated Manufacturing Systems :1-22.

[8] Yuehong Chen, Xinjing Gao, Enze Zhang. From Sympathize to Share & Spread: Virtual Reality News Design Based on Text-to-scence Conversion Technology and Communication Effect Evaluation Strategy [J]. Researcher, 2021;13(5) :104-111.

12/22/2022