Research on Multimodal Game Scenarios Based on XR Technology

Yixin Zhang [a,*], Weilun Qi [b]

[1]Lanzhou University, YixinZhang, Lanzhou, 730000, China
[2]Qinghai Normal University, Weilun Qi, Qinghai, 810000, China

[a]2897457622@qq.com, [b]1742628919@qq.com

Abstract. Virtual reality technology is a high-tech technology developed based on computer technology, which has high simulation characteristics and can bring people a real-world experience. This technology has significant application advantages in the field of games. This paper introduces the advantages of XR technology applied to game scenes from four aspects: multi-perceptual, immersive, interactive, and imaginative, and analyzes the technical difficulties and optimization measures in multi-modal game scenes from four aspects: 3D space modeling and rendering technology, data diversity and integration, computing resource demand and real-time interaction, enhancing user perception efficiency and alleviating cognitive load, so as to provide a reference for the research of multi-modal game scenarios under XR technology.

Keywords: AR, VR, multimodal human-computer interaction, game scenes

1 Introduction

Since the 90s of the 20th century, the game design industry has undergone a revolution centered on "experience", and the game experience of players has gradually been regarded as a key issue in game development. While the industry and academia generally agree that player experience is critical to game design, the concept of "experience" has been interpreted in a wide range of ways, and there is no consensus among different researchers on its core and research areas. In the past ten years of research, domestic and foreign scholars such as Robin Hunicke and Marc LeBlanc[1] have conducted in-depth analysis and discussion on player experience design and game experience design, and put forward unique insights on game experience problems in the real world. With the popularity of the concept of the metaverse, extended reality (XR) technology and intelligent tools have extended human gaming activities to the virtual space, bringing a new gaming experience, which not only provides a new research path for experience problems in game design, but also triggers new theoretical and practical challenges. This paper is dominated by design theory, combined with multidisciplinary knowledge such as human-computer interaction and interaction design, and attempts to

M. F. Sedon et al. (eds.), *Proceedings of the 4th International Conference on Culture, Design and Social Development (CDSD 2024)*, Advances in Social Science, Education and Humanities Research 917,
https://doi.org/10.2991/978-2-38476-380-1_25

further expand the research scope of multimodal game scene research in the context of extended reality, so as to provide reference for relevant researchers and designers.

2 Advantages of XR Technology in Gaming Scenarios

Extended Reality (XR) refers to the creation of a real, virtual combination of digital environment through modern high-tech means with computer as the core, as well as a new human-computer interaction mode, to bring the user a seamless transition between the virtual world and the real world immersion, it is virtual reality (VR), augmented reality (Augmented Reality, AR) and mixed reality (Mixed Reality, MR) and other technologies.[2]Mixed Reality (MR) represents the evolution of the convergence of virtual reality (VR) and augmented reality (AR) technologies, which enables the dynamic integration of virtual entities and physical environments through advanced interaction mechanisms. The core of MR technology is that it can provide real-time virtual and real interaction capabilities, so that virtual objects can seamlessly connect and interact with objects in the real world. In addition, MR, VR, and AR all rely on advanced visual interaction technologies, which together build a multi-dimensional perceptual framework that enables users to smoothly transition between virtual and reality, resulting in a deeply immersive gaming experience. XR has reshaped the field of practical activities and created a new form of experience: between the real world and the virtual world, with the increase of the concentration of virtual information, the degree of virtual reality of experience has increased in the direction of virtuality.[3] The application of XR technology to game scenarios has many advantages, which are analyzed from four aspects: multi-perception, immersion, interactivity, and conception.

2.1 Multi-Perceptuality

The core of XR technology lies in the construction of a multi-modal interaction framework, that is, the integration of multiple sensory inputs and outputs to achieve a richer and more intuitive user experience. "Non-visual perception, as a supplement to visual perception, can improve the efficiency of users in analyzing and understanding data, such as allowing users to perceive data information outside the field of view or obscured when analyzing large or high-density data visualizations.[4]For players who cannot obtain visual perception, XR technology converts visual information into non-visual information through its non-visual perception interaction ability, realizes the all-round multi-perception experience delivery to players, and brings players an immersive experience in game design. This perceptual substitution mechanism involves a complex information encoding and decoding mechanism, and efficient algorithms and models need to be developed to ensure the accuracy and real-time performance of information transformation. For example, kine-appendage technology is an advanced human-computer interaction solution that enables the innovation of visual kinesthetic feedback by integrating a virtual additional limb into the user's avatar hand. This technology makes use of motion capture and force feedback algorithms to enable the attached limb to exhibit precise rotation, deformation, and other complex kinematic

changes as it interacts with the surface of the virtual object, which are key components of visual kinesthetic feedback. These visual changes not only enhance the user's sense of immersion, but also provide the user with an intuitive, haptic-based interactive experience by simulating real physical contact. In virtual environments, traditional input devices often do not provide sufficient kinesthetic feedback, making it difficult for users to perceive the texture, shape, and weight of virtual objects. Kine-AppEndage technology makes up for this shortcoming by simulating tactile signals, enabling users to perceive and manipulate virtual objects more accurately through visual and kinesthetic feedback. The application of this technology can significantly improve the user's input speed and accuracy in free-hand typing scenarios, as it allows the user to perceive the position and force of keys through visual and kinesthetic feedback without a physical keyboard.

2.2 Immersion

Immersion refers to the realistic feeling that users seem to be completely immersed in the virtual environment, can perceive and manipulate various objects in the virtual world, and can actively participate in various events. The immersive characteristics of this virtual reality communication are mainly reflected in the perception system and behavior system of the user's body.[5]Oliver Grau argues that from dioramora, panoramas, cinema to virtual reality, "it shows the extraordinary efforts made by human beings in various periods of history to create the maximum illusionary space by various technological means".[6]The game scene built by XR technology incorporates an immersive sense of realism and multi-dimensional audio-visual stimulation, so that the player's mind can experience it in a vast, free and imaginative game world, which is incomparable to traditional games.

2.3 Interactivity

Through XR technology, the player's interaction with the game environment has never been more natural and intuitive. This interaction includes not only basic gesture control and voice commands, but also more advanced interactions such as eye tracking, face tracking, and gaze tracking, which provide a new way of interacting with input and allowing players to interact with VR content in a more comfortable and natural way. For example, the application of eye tracking technology in game scenes can be summarized as enhancing the immersion and interactivity of the game by accurately capturing the player's eye focus, enabling features such as foveated rendering, eye control of character or object interaction, and optimization of image quality. In addition, XR technology ensures that player movements are synchronized with the virtual world through a graphics transformation pipeline and a rendering pipeline, achieving a low-latency response that further enhances the sense of immersion and presence. The combination of these technologies not only improves the intuitiveness and convenience of game operation, but also promotes the innovation of game interaction, providing players with a new and multi-dimensional game experience.

2.4 Ideational

The background of the game draws on the essence of literary creation and weaves it into an engaging game screen; Draw on the subtleties of painting and architecture to create a lifelike play environment; Using the skills of sculpture art, carefully carved out the game characters with distinctive personalities; And use the power of music to create an immersive game atmosphere and a rich world of imagery. Take Assassin's Creed: Odyssey, for example, a game known for its ancient Greek setting and magnificent open world. The game's architecture and natural landscapes, such as the Parthenon in Athens and the majestic walls of Sparta, are imbued with classical beauty, creating an atmosphere of epic battles through carefully designed light and sound effects. The game's heroes, such as the protagonist Alexios or Kassandra, not only possess extraordinary combat skills, but also carry the idea of exploring history and pursuing freedom. The new characters, such as mythical creatures and historical figures, have been given unique looks and costumes while maintaining their historical background, demonstrating a deep understanding and pursuit of classical aesthetics. From the game's theme and plot, to the design of the scenes, characters, costumes, and props, the game's scenes show the blending of individual and collective imagery, creating a "second nature" imagery world that allows players to fully immerse themselves in this vibrant virtual world.

3 Construction of Multimodal Game Scenes in the Context of Artificial Intelligence

The application of artificial intelligence (AI) in the field of virtual reality is expanding day by day, pushing the boundaries of the virtual world. The integration of AI algorithms makes virtual characters more intelligent and interactive, and they are able to respond dynamically based on the player's actions and conversations. By fine-tuning keywords and parameters, AI can enrich the details of virtual environments, making them more vivid and realistic, resulting in a more immersive experience for users. In addition, AI plays a key role in the content generation and rendering process, which can significantly improve the user experience by optimizing resource allocation, improving rendering efficiency, and reducing latency. These applications of AI not only enhance the interactivity of the virtual world, but also provide the possibility to create more complex and diverse virtual scenes, making the VR experience more immersive.

4 Technical Difficulties and Optimization Measures in Multimodal Game Scenarios

Virtual reality technology can simulate the real environment and build a dynamic three-dimensional game scene, and players can enter the interactive interface to interact with the characters and environment in the game as long as they wear the device, so that players can have an immersive game experience.[7]There are still many problems in the

construction of multimodal game scenes, which limit the development of virtual reality games. The problems can be summarized as follows.

4.1 3D Space Modeling and Rendering Technology

XR technology uses high-precision 3D modeling and rendering technology, including polygon modeling, surface modeling, solid modeling, parametric modeling, procedural modeling, digital sculpting and skeletal animation modeling. Rendering techniques include ray tracing, rasterization, global illumination, ambient occlusion, path tracing, real-time rendering, volumetric rendering, and photorealistic rendering, all of which work together to take the entire model from creation to photorealistic rendering. These technologies are masterfully constructed to create a highly realistic game environment that provides players with a near-realistic, immersive experience. From the majestic natural landscape to the bustling scene of the city, to the exquisite architectural details and interior layout, the application of XR technology enables these complex scenes to be accurately reproduced, greatly enhancing the player's sense of immersion in the game and leading the player into a virtual dimension full of infinite possibilities.

4.2 Data Diversity and Integration

In the world of multimodal game design, dealing with data from different sources is a core technical challenge. The game world incorporates several types of data such as visual, auditory, haptic feedback, and textual information, each with its own specific characteristics and processing needs. For example, image data needs to be represented in high-dimensional space, while audio data needs to be analyzed for its time series. The differences in format, dimension, and statistical properties of these data types make them difficult to integrate and process directly. So, how to deal with this problem? First, the application of deep learning techniques, especially convolutional neural networks (CNNs) and recurrent neural networks (RNNs), is used to extract features from image and audio data, respectively, and convert them into unified data representations. Secondly, by constructing a multi-modal fusion network structure, such as a dual-stream network or a multi-branch network, the parallel processing and effective integration of different modal data are realized. In addition, the introduction of attention mechanism enhances the ability of the model to identify and integrate key information in different modal data, which is conducive to improving the accuracy and stability of the model and establishing an integration and collaborative processing mechanism.

4.3 Calculate Resource Requirements and Interact with Real-Time

In the field of multimodal game scene design, the high demand for computing resources is a key challenge, especially when games need to process and respond to player actions in real time. Multimodal data processing not only involves huge amounts of data and complex models, but also requires systems that can respond quickly to player interactions to keep the game fluid and immersive. First, by using model compression tech-

niques, such as parameter sharing, knowledge distillation, and network pruning, the number of parameters and computational complexity of the model are effectively reduced. These technologies allow models to be lightweighted while maintaining high performance, making them suitable for running on resource-constrained devices. Second, develop lightweight network architectures, such as MobileNet and ShuffleNet, that are designed for mobile devices and edge computing environments to reduce reliance on computing resources while maintaining high efficiency. The adoption of these network architectures enables multimodal gaming to deliver a gaming experience on mobile platforms such as smartphones and tablets that rivals high-end consoles or PCs. In addition, to further improve the inference speed of the model, parallel computing and hardware acceleration technologies such as graphics processing units (GPUs) and tensor processing units (TPUs) can be utilized. These techniques can significantly improve the computational efficiency of models, especially when dealing with graphics-intensive game scenarios such as real-time ray tracing and complex physics simulations. The implementation of these measures helps multimodal games maintain high performance, effectively reduce the consumption of computing resources, improve the real-time responsiveness of the system, and promote players to enjoy a lag-free, high-frame-rate gaming experience whether it is in intense combat scenarios or puzzle-solving sessions that require delicate operations. Through the application of these technologies, it is possible to achieve a richer and more realistic game world, while ensuring real-time interaction and dynamic feedback.

4.4 Enhance User Perception Efficiency and Alleviate Cognitive Load

When users are in a virtual environment for a long time, excessive acceptance of complex information and perception of virtual objects will have a certain fatigue impact on the user's perception ability, so that the information acceptance will be discounted, so it is necessary to consider the relevant virtual experience and interaction process from the perspective of the user. XR technology needs to comprehensively consider the relationship between the individual (subject), computer (media), and environment (real or virtual life world), and make certain design changes. For example, Half-Life: Alyx, developed by Valve Corporation, is a virtual reality (VR) game that was launched in 2020. As the latest installment in the Half-Life series, the game is tailored for the VR platform and is considered by the industry to be an iconic achievement in the history of VR gaming. In Half-Life: Alyx, players will delve into an immersive game world created by cutting-edge VR technology. The development team fully considered the perceived fatigue that players can cause when they were in a virtual environment for a long time, and designed a variety of interaction and rhythm control mechanisms to reduce the visual and cognitive burden, such as arranging short breaks between intense battles and puzzle solving sessions, and guiding the player's gaze and movement through well-designed environments. Half-Life: Alyx demonstrates the synergy of XR technology between the user, the computer, and the environment. Through VR headsets and controllers, the game provides an intuitive and natural interactive experience that allows players to interact with objects in the virtual world. In addition, the game's environment is designed to be both realistic and highly immersive,

making it feel as if the player is actually in the game world. During the design process of Half-Life: Alyx, Valve paid special attention to the player's experience and comfort. To reduce the risk of perceived fatigue and information overload, the game introduces a dynamic rhythm design to ensure that players don't feel overly tired during long gaming experiences. At the same time, the game also optimizes the proportions of objects in the virtual environment and the interactive interface to make it closer to the intuition of the real world, thereby reducing the cognitive burden on players.

5 Cross-Domain Technology Convergence of XR Technology

Take a deep dive into the use of XR technology in gaming, a field that combines the best of science, art, technology, and design. XR technology provides players with a new immersive experience by seamlessly combining virtual elements with the real world. With the continuous development of XR technology, although the current games developed by XR technology do not form a smooth interaction, an interdisciplinary research approach has become the key to the development of this field. As the law of scientific development shows, highly differentiated scientific fields form a unified whole in synthesis, and this multifaceted understanding is the basis for achieving interdisciplinary integration and innovation.[8]The mixed application of multimodal technology has gradually penetrated into various fields and has had a positive impact, which can help the in-depth connection of various industries, further promote the integration of disciplines, and present a new ecology, new modality, and new business format.

6 Conclusion

In the era of spatial computing, XR technology has brought players an interactive experience of multi-modal game scenes. As an extension and innovation of traditional games, by deepening the player's embodied experience in the multi-modal narrative language environment, this not only enhances the player's multi-perceptual interaction, but also promotes the deep integration of art and technology, thus providing players with a rich artistic experience. In virtual interactive games, if scene design can help the game resonate with the player's emotions, it will make the game have more added value in addition to entertainment value.[9]With the iterative upgrading of technology, the game scene is developing in the direction of intelligence, diversification and high immersion, heralding the arrival of a more realistic and interactive game experience.

Acknowledgments

In early November, I completed this paper. I didn't know what writing meant to me, but all I knew was that I couldn't stop writing. In every quiet moment of writing, conversing with myself allows me to glimpse my true self. Thank you to my best friend for helping

me. You told me how to organize my thoughts and accompanied me every day. Your appearance is like the best gift given to me by heaven; Thank you, Teacher Wang Jing, for your guidance. You have broadened my horizons, provided me with ideas, and enriched my perspective on the world. Only then have I been inspired, and this article has gradually been written from conception to writing. It is my great fortune to have met a good teacher in life.

References

1. Hunicke R , Leblanc M , Zubek A R .MDA: A formal approach to game design and game research[C]//Challenges in Games Ai Workshop.2004.
2. FAN Liya, MA Jieyuan, LIU Xitao, et al. 2023年extended reality (XR) hot spot[J].Science & Technology Review, 2024, 42(1):296-305.FAN L Y, MA J Y, LIU X T, et al. Review on Hot Spots of Extended Reality(XR)in 2023[J]. Science&Technology Review, 2024, 42(1):296-305.
3. JIANG Ke, ZHANG Jingyu, YAN Yupeng, et al.Virtual Experience Design in the Context of Extended Reality[J].Packaging Engineering,2024,45(20):43-48+57.DOI:10.19554/j.cnki.1001-3563.2024.20.004.
4. Tao Jianhua, Wu Yingcai, Yu Chun, et al. Chinese Journal of Image and Graphics,2022,27(06):1956-1987.)
5. Hang Yun, Su Baohua. Modern Communication(Journal of Communication University of China),2007,(06):21-24.)
6. Oliver Grau, Virtual Art: From Illusion to Immersion, translated by Chen Ling, Tsinghua University Press, 2007, p. 4.
7. Hu Zhizhong. Science and Technology Innovation and Application,2022,12(26):193-196.DOI:10.19981/j.CN23-1581/G3.2022.26.048.
8. Wang Daqian. From Immersion to Multidimensionality: A Study on Multimodal Digital Exhibition Hall Integrating VR, AR and AI Technologies[J].Art Education Research,2024,(13):83-85.)
9. Bai Yuwei, Ding Ni, Zhou Wen. Analysis on the scene design of virtual interactive experience game: A case study of VR game Top Floor[J].Modern Film Technology,2019,(06):45-49.)