Academic Education Publishing House
Industry Science and Engineering Vol. 1 No. 5, 2024

Research on Mixed Reality Experience Models Driven by Generative Artificial Intelligence

Le Zheng, Xuan Li
Xi'an Eurasia University, Xi'an, Shaanxi, China

Abstract: In today's era of rapid technological advancement, Generative Artificial Intelligence (AIGC) and Mixed Reality (MR), as two significant technologies at the forefront of innovation, are transforming our ways of life and work at an astonishing pace. Research on the integration of AIGC and MR, focusing on new technologies and application methods, is increasingly becoming a new trend. This paper reviews the development of AIGC and MR technologies, explores the technical principles and application methods of AIGC in Mixed Reality (MR), and proposes specific design solutions for MR experience models driven by AIGC. These efforts aim to provide new perspectives for AIGC-driven MR solution models and to explore more possibilities for the development of digital media technologies.

Keywords: AIGC; MR; Mixed Reality; Artificial Intelligence

1. Introduction

1.1 Research Background

Generative Artificial Intelligence (AIGC), as a branch of artificial intelligence, primarily uses machine learning models to generate new content such as images, videos, text, and audio. AIGC technology is a way to simulate human thinking and perception capabilities to produce realistic images and animations. It leverages technologies like Deep Learning and Neural Networks to automatically analyze large amounts of data and images, generating high-quality visuals based on learned patterns and rules. The development of AIGC technology can be traced back to the mid-20th century, but the real breakthrough came in the early 21st century, especially after the introduction of Generative Adversarial Networks (GANs) in 2014. GANs significantly improved the quality of image and video generation by creating realistic data through the interplay of generators and discriminators. Other important generative models, such as Variational Autoencoders (VAEs) and autoregressive models, have also played critical roles in generating images, videos, text, and audio. VAEs generate new samples by learning latent representations of data, offering advantages in generating diverse samples and manipulating latent spaces. Autoregressive models, such as Transformer models, are widely used in natural language processing tasks. For example, the GPT series models can generate high-quality text and are also applicable to image generation tasks [1-3].

The application of AIGC in Mixed Reality (MR) has significantly improved the efficiency and quality of content generation. For example, text-to-image technology can quickly generate realistic virtual scenes based on user descriptions, and image-to-image technology can transform simple sketches into detailed images, greatly facilitating design and creation. AIGC also enhances the user interaction experience in MR; through natural language processing technology, users can interact with MR systems using voice or text commands. Emotion recognition technology can analyze users' expressions and emotions, allowing real-time adjustments to MR content to provide a more personalized experience. AIGC-driven content management systems enable automated content generation and updates, reducing the need for manual intervention and increasing production efficiency. Real-time generation and feedback mechanisms dynamically adjust and optimize MR content based on user behavior data and feedback, ensuring that users always have the best experience. As technology continues to advance, the integration of AIGC and MR will have a profound impact across more fields, bringing richer and more realistic digital experiences [4, 5].

1.2 Research Objectives and Significance

This study aims to explore how Generative Artificial Intelligence (AIGC) can enhance the quality of content generation and user interaction experience in Mixed Reality (MR), achieving more realistic and immersive MR experiences. The research focuses on the applications of AIGC in areas such as text-to-image, image-to-image, text-to-video, image-to-video, and text-to-3D model generation, optimizing MR content creation and management processes. Additionally, the study will investigate how AIGC technology can enhance the intelligence of MR systems, allowing users to interact with virtual environments more naturally. Through case studies and user testing, the study will evaluate the practical application effects of AIGC in MR projects, analyzing its strengths and weaknesses in different application scenarios, and providing data support and theoretical foundations for future research and practice [6].

Driving technological innovation is the core objective of this research. By integrating AIGC with MR technologies, this study aims to foster technological innovation in both areas, bringing new breakthroughs in digital content generation and user interaction. The goal is to achieve automated and efficient MR content generation, reduce manual intervention, increase production efficiency, and lower development costs, thereby delivering tangible economic benefits to related industries. Additionally, this research seeks to provide richer and more personalized user experiences, meeting the demand for high-quality, immersive MR content and enhancing user satisfaction and engagement. The study will also explore the applications of AIGC in fields such as education, healthcare, entertainment, and industrial design, expanding its application scenarios and promoting the adoption and use of MR technology across more industries [7].

1.3 Innovation and Application Value

This paper systematically explores the application of AIGC in MR, proposing innovative methods for integrating technologies such as text-to-image, image-to-image, text-to-video, image-to-video, and text-to-3D model generation into MR experiences. The combination of these technologies not only optimizes the MR content generation process but also enhances the user interaction experience. By exploring the integration of AIGC and MR technologies, this research provides new ideas and directions for technological innovation in both fields [8].

2. Overview and Applications of Generative Artificial Intelligence Technology

2.1 Major AIGC Technologies: GANs, VAEs, Autoregressive Models, etc.

Generative Adversarial Networks (GANs) consist of two components: a generator and a discriminator. These components engage in a competitive process to generate high-quality data. GANs have achieved remarkable success in areas such as image generation, image restoration, and super-resolution processing. Typical variants include Conditional GANs (cGANs) and Cycle-Consistent GANs (CycleGANs), which are optimized for specific tasks.

Variational Autoencoders (VAEs) generate new samples by learning the latent representations of data, making them suitable for generating diverse samples and manipulating latent spaces. Although the quality of their generated outputs may not match that of GANs, VAEs have the advantage of producing a wide variety of samples.

Autoregressive models generate data using a sequence-to-sequence approach and are widely used in natural language processing (NLP) tasks. Transformer models, such as the GPT series, are typical examples of these models. They not only generate high-quality text but are also applicable to tasks like image generation.

2.2 Current Applications of AIGC in Different Fields

AIGC excels in areas such as image generation, image restoration, and style transfer, and is widely used in fields like artistic creation, advertising design, and film production. AIGC is also extensively applied in text generation, machine translation, and dialogue systems, enhancing the efficiency and quality of automated writing and translation. Through text-to-video and image-to-video generation, AIGC shows great potential in video production, animation generation, and film special effects. Text-to-3D model generation technology is crucial in fields such as game development, virtual reality, and industrial design, significantly improving the efficiency of creating 3D content.

3. Overview and Applications of Mixed Reality (MR) Technology

3.1 Basic Principles and Development Trends of MR Technology

Mixed Reality (MR) technology is an advanced interactive technology that combines the features of Virtual Reality (VR) and Augmented Reality (AR). MR technology achieves the coexistence and interaction of physical and digital objects by integrating virtual content with the real environment. The core technologies of MR include high-precision sensors, computer vision, and graphics processing. MR technology is evolving towards more natural and immersive experiences. Hardware devices like Microsoft HoloLens and Magic Leap One are continually being upgraded, offering higher resolution, wider fields of view, and greater computing power. On the software side, MR applications are increasingly integrating AI technologies, enhancing the intelligence of content generation and user interaction.

3.2 Applications of MR in Education, Healthcare, and Entertainment

MR technology is widely applied in the field of education, enhancing teaching effectiveness and student engagement through virtual laboratories, historical reconstructions, and scientific simulations. For example, virtual anatomy experiments help medical students gain a direct understanding of human anatomy. In healthcare, MR plays a crucial role in surgical planning, intraoperative navigation, and rehabilitation training. By providing real-time visualization of a patient's internal structures, MR assists doctors in performing precise surgeries, while virtual rehabilitation training offers personalized recovery plans for patients. In the entertainment sector, MR technology is extensively used in games, movies, and theme parks. Virtual reality games provide immersive experiences, augmented reality films offer interactive experiences for viewers, and theme parks use MR devices to enhance the appeal of attractions.

4. Feasibility Study on the Application of AIGC in Mixed Reality (MR)

4.1 Methods of Applying AIGC in Mixed Reality (MR)

The integration of Generative Artificial Intelligence (AIGC) with Mixed Reality (MR) significantly enhances the quality and efficiency of content generation and user interaction. In terms of content generation, text-to-image technology allows for the rapid creation of realistic scenes and objects based on user descriptions, increasing the efficiency and flexibility of content creation. Image-to-image technology can convert simple sketches into detailed images, facilitating design and creation. Text-to-video technology can generate dynamic video content from textual descriptions, making it suitable for dynamic backgrounds or scenario simulations in MR. Image-to-video technology can transform static images into dynamic videos, enriching character movements and scene transitions. Text-to-3D model technology can automatically generate 3D models based on descriptions, accelerating content creation and meeting personalized needs.

4.2 Methods of Applying AIGC in Mixed Reality (MR)

In the field of education, text-to-image technology can generate images based on descriptions of historical events or natural phenomena, helping students understand and learn concepts more intuitively. In advertising design, this technology can automatically generate visual content that matches the ad copy, enhancing the efficiency of ad production and creative expression. Additionally, artists can use text descriptions to generate inspirational images, aiding the artistic creation process.

In terms of style transfer, image-to-image technology can transform photos into artistic style images, such as converting real-life photos into paintings in the style of Van Gogh. For image restoration, this technology can repair old or damaged images, restoring them to their original appearance. Additionally, in augmented reality applications, this technology can convert simple sketches into high-quality images, thereby enhancing the user experience.

In virtual tours, text-to-video technology can generate virtual tour videos based on textual descriptions, such as museum or city tours, thereby enhancing the user experience. In educational video production, this technology can automatically create instructional videos that showcase scientific experiments or historical events, supporting the teaching process. In marketing and promotion, product promotional videos generated from product descriptions can significantly enhance marketing effectiveness.

In animation production, image-to-video technology can convert static character images into animations, reducing the workload and cost of animation creation. In video special effects production, this technology can generate special effects videos, such as transforming static scenes into dynamic ones. In virtual reality applications, this technology can generate dynamic video content, enhancing the user's immersive experience.

In game development, text-to-3D model technology can automatically generate 3D models of characters and scenes based on game settings, thereby speeding up the game development process. In virtual reality applications, this technology can generate 3D content for virtual environments, such as virtual furniture and buildings, enriching the virtual experience. Additionally, in industrial design, this technology can create 3D models of products based on design requirements, aiding in design and presentation.

5. Designing AIGC-Driven Mixed Reality Experience Models

5.1 User Needs Analysis and Experience Design Principles

When designing AIGC-driven Mixed Reality (MR) experiences, it is crucial to understand the needs and expectations of the users. This involves a detailed analysis of the target user group, including factors such as age, interests, and usage scenarios. Methods such as user research, surveys, and interviews can be used to gather specific requirements and suggestions for MR experiences. Based on the analysis of user needs, experience design principles are formulated. The design should ensure a seamless transition between the virtual and real worlds, providing a highly immersive experience. Users should be able to interact with the MR environment through natural interaction methods such as voice and gestures. The system dynamically adjusts content based on user preferences and behaviors, offering a personalized experience. The interface should be simple and intuitive, allowing users to easily engage with the MR experience.

5.2 Presentation Methods of AIGC-Generated Content in MR

The virtual scenes generated by AIGC should be presented in high resolution with rich details to ensure realistic and visually appealing effects. For example, virtual cityscapes created using GANs technology can be displayed to users through high-precision display devices. The presentation of virtual objects should also be highly interactive, allowing users to interact with them through touch, voice, or gestures. Virtual characters generated using text-to-3D model technology can enable users to engage in conversations through voice commands or control their movements using gestures. Additionally, the dynamic effects generated by AIGC should be smooth and natural. For instance, animation scenes generated using text-to-video technology should play seamlessly in the MR environment, without delays, ensuring a fluid and immersive experience.

5.3 Design and Optimization of User Interaction Modes

In designing the interaction scheme, emphasis should be placed on natural user interaction modes, allowing users to interact with the MR environment through voice, gestures, eye movements, and more. By using natural language processing technology, users can control virtual objects or adjust scenes with voice commands. Emotion recognition technology can be employed to analyze users' expressions and emotions in real-time, dynamically adjusting MR content and interaction methods. For example, if a user shows signs of confusion, the system can provide appropriate assistance and guidance. Additionally, a timely and effective feedback mechanism should be designed, enabling users to receive immediate feedback and responses during their interactions, thus enhancing the fluidity and satisfaction of the interactive experience.

5.4 Collection and Utilization of User Feedback Data

Using sensors, cameras, microphones, and other devices, real-time collection and recording of user feedback such as eye movement, gestures, and voice commands can be achieved. Machine learning and data analysis techniques can then be used to analyze this feedback data, extracting valuable insights to guide system optimization and improvement. By analyzing user behavior patterns, common issues and needs can be identified, which helps in optimizing MR content and interaction design. Based on user feedback data, the MR experience models and content generation strategies can be continuously refined to ensure the system consistently enhances the user experience. Regular updates and adjustments to virtual scenes, along with the addition of new interactive objects and features, should be implemented during routine maintenance to meet the evolving needs of users.

5.5 User Experience Evaluation and Satisfaction Analysis

By conducting user surveys and collecting feedback, the effectiveness of AIGC applications in MR experiences can be analyzed. Results from questionnaires and interviews indicate that users generally express high satisfaction with content generated by AIGC. Detailed analysis includes user evaluations of content generation speed, the quality of images and videos, and the naturalness of interactions. Statistical data show that content generated using AIGC technology significantly enhances user immersion and engagement.

6. Conclusion and Outlook

6.1 Research Conclusions

This study systematically explored the application of Generative Artificial Intelligence (AIGC) in Mixed Reality (MR) and found that AIGC offers significant advantages in enhancing content generation efficiency and user interaction experience. AIGC technology can automatically generate high-quality images, videos, and 3D models, reducing manual intervention and production costs while increasing content diversity and personalization. Additionally, AIGC optimizes user interaction with MR systems through natural language processing and emotion recognition technologies, enhancing user immersion and satisfaction.

Data from user surveys and analyses of practical application cases show that content generated using AIGC surpasses traditional methods in terms of visual effects, interactivity, and personalization. Furthermore, users demonstrated higher satisfaction with AIGC-driven MR experiences, indicating that the potential of AIGC in MR applications has been practically validated. These findings provide important references for future research and practice.

6.2 Future Research Directions

Future research should continue to explore the deep integration of AIGC and MR technologies to drive technological innovation. Specific directions include optimizing AIGC algorithms to improve the real-time performance and quality of content generation, enhancing the application of natural language processing and emotion recognition technologies to increase the intelligence of user interactions, and utilizing multimodal learning and cross-domain transfer techniques to enrich the diversity and authenticity of AIGC-generated content. Additionally, there should be a focus on the upgrading of hardware devices, such as enhancing the computational power and sensing accuracy of MR equipment, to support more complex and refined AIGC-generated content.

6.3 Future Research Directions

Future research should continue to explore the deep integration of AIGC and MR technologies to drive technological innovation. Specific directions include optimizing AIGC algorithms to improve the real-time performance and quality of content generation, enhancing the application of natural language processing and emotion recognition technologies to increase the intelligence of user interactions, and utilizing multimodal learning and cross-domain transfer techniques to enrich the diversity and authenticity of AIGC-generated content. Additionally, there should be a focus on the upgrading of hardware devices, such as enhancing the computational power and sensing accuracy of

MR equipment, to support more complex and refined AIGC-generated content.

References

[1] Lu Siyi. Beyond Boundaries—Applications and Prospects of MR Technology in Contemporary Art Exhibitions. Screen Printing, 2024(06): 112-115. DOI:10.20084/j.cnki.1002-4867.2024.06.032.

[2] Liu Yunli, Liang Dong, Chang Lin, et al. Redefining the Boundaries of In-Depth Reporting in the AIGC Era. Media, 2024(12): 20-22.

[3] Qu Enxi. Research on AIGC Information Falsehood in the Era of Social Media. Dongbei University of Finance and Economics, 2023. DOI:10.27006/d.cnki.gdbcu.2023.000233.

[4] Tai Yuquan. Application of Generative AI Technology in News Creation. Television Technology, 2024, 48(07): 103-105. DOI:10.16280/j.videoe.2024.07.025.

[5] Xing Xiangying. Development and Current Status of MR Technology and Its Applications in Education. China Modern Educational Equipment, 2024(06): 31-33. DOI:10.13492/j.cnki.cmee.2024.06.011.

[6] Yuan Yishuai. Application of Mixed Reality Technology in the Broadcasting and Television Industry. Television Technology, 2024, 48(07): 109-111. DOI:10.16280/j.videoe.2024.07.027.

[7] Zhou Na. Key Technologies of Mixed Reality-Based Surgical Navigation. North China University of Technology, 2022. DOI:10.26926/d.cnki.gbfgu.2021.000047.

[8] Zhang Tongyue. Practical Research on AIGC Technology Empowering Digital Media Art Creation. Shanghai Conservatory of Music, 2024. DOI:10.27319/d.cnki.gsyyy.2024.000132.