Research on the Integration of Stable Diffusion and Digital Human Applications

— Taking the Inheritance and Promotion of Horse-Shovel Face Culture as an Example

Xiaoshan Xu*

Xi'an University of Posts and Telecommunications, Xi'an, China

*993783686@qq.com

Abstract. This paper examines the use of AIGC and virtual digital humans in preserving Shehuo Ma Shao face painting, a traditional cultural art form. It assesses the current state, significance, obstacles, and potential of this integration, with the goal of offering fresh perspectives and strategies for cultural heritage. Technological progress, particularly in AIGC's algorithms and data handling, and the interactive, lifelike nature of virtual humans, are seen as opportunities to enrich and evolve traditional culture. The study aims to understand how this technological blend can revitalize cultural practices for the modern era, serving as both a safeguard for heritage and a catalyst for cultural innovation and national pride in the digital age.

Keywords: AIGC; Virtual Digital Human; Horse Ladle Face Painting; Digital Human Integration Applications;

1 Introduction

On July 10, 2023, the National Internet Information Office and six other departments issued the "Interim Measures for the Management of Generative Artificial Intelligence Services," promoting the innovative use of generative AI in various sectors to create positive content and build an application ecosystem. The rise of AIGC technology has revitalized the cultural field, using digital technology to transform and develop China's traditional culture. AIGC's algorithms and data processing enrich cultural content, while virtual digital humans provide immersive experiences. The integration of AIGC and virtual digital humans presents new opportunities for traditional culture's inheritance and development. Research on AIGC virtual digital humans aims to inject new vitality into traditional culture, ensuring its continuity and prosperity in the new era. This represents both the preservation and innovation of traditional culture, as well as an initiative to enhance national cultural confidence in the digital age.

© The Author(s) 2025
A. Y. M. A. Islam et al. (eds.), *Proceedings of the 2025 4th International Conference on Science Education and Art Appreciation (SEAA 2025)*, Advances in Social Science, Education and Humanities Research 952,
https://doi.org/10.2991/978-2-38476-452-5_59

2 Introduction- Research Foundation of AIGC Integration with Virtual Digital Humans

2.1 Concepts of AIGC and Virtual Digital Humans

Artificial Intelligence Generated Content (AIGC) refers to the creation of various types of content, including text, images, audio, and video, using artificial intelligence technologies. AIGC technology relies on advanced fields such as deep learning, natural language processing, and computer vision. Through the learning and training of vast amounts of data, it can discern users' needs and intentions and create content that possesses a certain degree of innovation and uniqueness.

Virtual digital humans are virtual avatars that possess human-like characteristics, with the core being the simulation of human features and digital representation[1]. In terms of appearance, virtual digital humans have human-like looks, voices, forms, and movements; from a technical perspective, they are virtual characters constructed in the non-physical world using comprehensive technological means such as deep learning, computer graphics, and speech synthesis.

2.2 The Current State of Mashao Face Painting Inheritance

Horse-dipper face masks, a handicraft from Shaanxi, use wooden horse-dippers as a base and feature vibrant, colorful patterns with cultural significance. These masks, inspired by mythological and historical themes, reflect local aspirations and are a unique folk art form. They serve as aesthetic objects and valuable resources for studying folk culture[2]. With time, these masks have evolved to include contemporary elements, such as characters from popular novels, and have been transformed into cultural products and digital collectibles. Innovators like Xue Yabing and Zhang Xing have integrated these masks with new technologies and platforms, expanding their market presence and cultural impact through digital innovation.

2.3 The Opportunity for Integration of AIGC with Virtual Digital Humans in Traditional Culture Applications

In the digital age, traditional culture faces new challenges and growth opportunities. AIGC technology and virtual humans are expanding the ways cultural heritage is shared and preserved. AIGC creates high-quality digital assets that make traditional culture more engaging. Virtual humans, with their lifelike qualities, interactively connect with audiences across time and space. These technologies can be used to create digital ambassadors for culture, offering immersive experiences that deepen public understanding of its history and art. Virtual figures like "Xiao Zheng" and "Xijiajia" are increasing interest in science and marketing, respectively, while "Huaxizi" and "Xiaoyang" are enhancing program engagement and brand promotion[3]. "Liuyexi" has gained popularity on social media through creative videos. Virtual humans are providing innovative applications for marketing, education, and entertainment, offering richer, personalized consumer experiences.

The success of "Xiao Zheng," a virtual digital human by Xinhua News Agency and Tencent, highlights the potential for digital avatars focused on traditional culture. Creating such an avatar could immerse audiences in traditional worlds, spreading cultural knowledge. It could also follow Baidu's "Xijiajia" model, engaging in interactive content like live streaming traditional crafts or promoting traditional products. These innovative methods could enhance the visibility and preservation of traditional culture.

3 Integration Application Models and Cases of AIGC with Virtual Digital Humans

3.1 AIGC Assisted Creation and Innovation of Traditional Culture

The rapid advancement of AIGC technology offers new opportunities for the growth of traditional culture. NLP is essential for organizing and mining knowledge from traditional texts, enabling in-depth analysis and understanding of cultural materials. For example, it can parse traditional opera scripts to extract elements for new opera creation. Image recognition and generation technology has revolutionized the visual representation of traditional culture by digitally acquiring and analyzing cultural artifacts with high precision. This technology can identify features of folk paper-cutting art and generate new works that blend traditional and modern aesthetics, meeting current market demands.

Speech synthesis and recognition technologies have shown significant advantages in the dissemination of traditional culture. For cultural forms that rely on oral tradition, such as folk tales, ballads, and musical performances, speech synthesis technology can transform written content into vivid speech, giving traditional culture new vitality[4]. At the same time, speech recognition technology can efficiently record the live performances and narratives of folk artists, ensuring that these precious sound materials are preserved and passed down. For example, with Mongolian long songs, through speech technology, ancient long songs have been digitally preserved and, with the aid of speech synthesis technology, turned into audiobooks, allowing more people to appreciate their unique charm.

Virtual digital human technology can showcase traditional culture with lifelike images and interact in real-time with the audience. AIGC technology provides powerful technical support for the performance of virtual digital humans. Virtual digital humans can use AIGC technology to write engaging explanatory texts, answer audience questions in real-time, and perform or display content based on audience needs. For example, the "Traditional Wisdom Inheritance" project uses AIGC technology to analyze and organize embroidery techniques, then has virtual digital humans demonstrate and explain to the audience, enhancing the effectiveness of spreading embroidery skills; the "Folk Music Revival Plan" uses speech synthesis and recognition technology to digitally preserve nearly extinct folk music and promotes it through virtual digital humans on online platforms, attracting the attention and participation of young people.

However, in the process of using AIGC technology to assist in the creation and innovation of traditional culture, there are also some challenges. The quality and accuracy

of data are particularly critical; inaccurate or incomplete data may lead to incorrect analysis and creative outcomes. Additionally, how to maintain the unique value of traditional culture during the innovation process is a question that requires careful consideration.

3.2 The Role of Virtual Digital Humans in the Display and Inheritance of Traditional Culture

Virtual digital humans, as an emerging technological form, are gradually demonstrating their unique value in the display and inheritance of traditional culture. This article will delve into the application of virtual digital humans in traditional culture, analyzing based on their classification characteristics.

Virtual digital humans can be designed with various personality traits for different applications. For instance, IP-type virtual humans can serve as ambassadors for traditional culture, like a virtual opera character with a unique story and personality, engaging audiences through social media and live streaming. Additionally, second avatar-type virtual humans can offer traditional skill inheritors a new platform to showcase their crafts in virtual spaces, such as live-streaming paper-cutting and allowing viewers to interact and participate.

In terms of graphical dimension, 2D virtual digital humans offer advantages in production cost and expression, particularly suited for traditional cultures with simple lines and unique styles. They can effectively showcase characteristics of folk painting and traditional embroidery through animated short films and online tutorials. On the other hand, 3D virtual digital humans are superior in realism and visual impact, ideal for complex traditional crafts like wood carving and ceramic making. They allow audiences to view the intricate details and processes from all angles, providing a deeper insight into the subtleties of these crafts.

Service-type virtual digital humans offer convenient support for traditional culture by answering questions and providing information. Identity-type virtual humans act as ambassadors for traditional culture, like virtual idols incorporating traditional music into performances to increase its appeal among youth[2]. Non-interactive virtual humans serve as introductory displays of traditional culture, while interactive ones engage with the audience, offering guidance and feedback on traditional crafts.

Intelligent-driven virtual digital humans, leveraging their powerful intelligent algorithms and learning capabilities, continuously optimize the ways of displaying and explaining traditional culture. Based on audience feedback and needs, they adjust content and form to provide more personalized and accurate services.

Human-driven virtual digital humans can fully utilize human emotions and expressiveness, conveying more authentic and profound emotions when displaying traditional culture. For instance, when telling the stories behind traditions and the experiences of inheritors, human-driven digital humans can evoke audience empathy in a more moving way.

In summary, the various classification characteristics of virtual digital humans provide a rich variety of means and approaches for the display and inheritance of traditional culture. By reasonably utilizing different types of virtual digital humans, we can break

through the limitations of time and space, allowing more people to understand and appreciate traditional culture, and injecting new vitality into the inheritance and development of traditional culture.

3.3 Analysis of Successful Cases of Virtual Digital Human Applications

Against the backdrop of the digital age, the emergence of virtual digital humans has opened new prospects for information dissemination, cultural heritage, and the entertainment industry. Through in-depth studies of successful historical cases of virtual digital humans, we have summarized a series of key success factors and lessons that can be learned. Taking the virtual digital human "Tianyu" as an example, the key factors of its success mainly include the following:

Firstly, a unique and eye-catching image design. The appearance of the virtual digital human "Tianyu" not only has high aesthetic value but also integrates elements of traditional Chinese Dunhuang flying apsaras, making it stand out among many virtual digital humans[5]. Its delicate facial features, elegant hairstyle, and splendid attire are all carefully designed to be in line with public aesthetics while containing unique cultural connotations. This indicates that a successful virtual digital human image needs to strike a balance between aesthetic appeal and cultural depth, attracting the audience's attention while conveying specific cultural values.

Secondly, compelling stories and plotlines. Content creation is one of the core elements of a virtual digital human's success. Through a series of engaging stories and plotlines, the virtual digital human "Tianyu" showcases its rich personality and emotions. Whether narrating ancient myths or demonstrating traditional crafts, she presents them in a lively, interesting, and easily understandable way, sparking the audience's keen interest and emotional resonance. High-quality content should not only possess knowledge and informational value but also touch the audience's emotions, prompting them to be willing to invest time and energy to pay attention and participate.

Thirdly, precise market positioning. The virtual digital human "Tianyu" clearly targets young people interested in traditional culture and seeking fresh experiences as its audience. In response to the characteristics and needs of this group, "Tianyu" has innovated in its communication forms and channels, using emerging media such as social media and short video platforms to disseminate in a fragmented and interactive manner, quickly accumulating a large number of fans and attention. This shows that virtual digital humans must clearly understand their audience and choose appropriate communication channels and strategies based on the characteristics of the audience.

During the launch of "Tianyu," we observed the advantages of virtual digital human production in the current mainstream market. Firstly, virtual digital human production led by humans has a clear advantage in detail control. For example, in terms of image design, professional designers, with their rich experience and keen artistic perception, meticulously design the appearance and attire of virtual digital humans to ensure that every detail is consistent with the overall artistic style and theme. Secondly, in terms of motion and expression design, human design pays more attention to the expression and transmission of emotions, making the performance of virtual digital humans more impactful and persuasive.

4 Challenges of Integrating AIGC with Virtual Digital Humans

4.1 Technological Bottlenecks of Virtual Digital Humans

Virtual digital human technology is proving valuable across various sectors, including new media. In entertainment, virtual idols are popular for their idealized images and talents, serving as effective brand ambassadors. In education, virtual teachers offer flexible, personalized instruction, catering to student needs. In e-commerce, virtual hosts provide continuous live streaming with engaging product promotions.

Virtual digital human technology has several shortcomings. Technically, character generation lacks expressiveness and precision, voice synthesis struggles to mimic natural human voices, and facial animation is limited by sensor granularity, motion capture imperfections, and software constraints, resulting in poor audiovisual consistency. Natural language processing limitations also hinder virtual humans from fully understanding complex emotions and semantics, which affects communication quality. Emotional expression is another bottleneck, as these digital humans cannot accurately interpret or express emotions and feelings. Compared to traditional hosts, virtual hosts need improvement in expression and analytical judgment, limiting their use in high-emotional interaction and complex judgment scenarios.

Martin, author of "Game of Thrones," and the "New York Times" have criticized OpenAI for potential misuse of virtual digital humans, including the creation of misleading deepfake videos[4]. These technologies can distort information and harm societal integrity. To address these issues, increased investment in research and development is needed, focusing on areas like the metaverse, and solving key technical challenges such as sensors, AI chips, and display equipment. Advanced modeling and emotional AI algorithms must be developed, along with a comprehensive emotional dictionary and improved production technologies for virtual characters. Additionally, management systems and laws should be established to ensure the ethical application of virtual digital humans, protect intellectual property, and prevent consumer deception, allowing for their healthy integration into society and media ecosystems.

4.2 Design Principles for Virtual Digital Human in the Integration Process

Designing a virtual ladle character requires historical accuracy and cultural heritage representation. Use technology to blend traditional and modern aesthetics, incorporating creative colors and shapes for visual appeal. Create interactive scenes for user engagement and optimize performance for a smooth, high-quality experience. Balance cultural respect, artistic value, interactivity, practicality, and technical performance to craft a character with depth and charm. The 3D character design process starts with AI-generated content, inputting ladle features to produce concept images. Stable Diffusion analyzes ladle color data for a modern, soft color scheme and reinterprets traditional patterns innovatively for the virtual character design.

RODIN software converts 2D images into 3D models, allowing adjustments to symmetry, edges, and patterns. Users can customize model complexity and materials, then download and refine in 3D software[1]. For motion capture, AIGC technology offers solutions for facial and full-body dynamics. Plask.AI captures full-body movements from live video, creating a skeletal action framework for 3D models, which can be fine-tuned for realism. HeyGen uses Surreal Engine to generate AI characters with synchronized lip movements, offering various tones and languages for audio-visual consistency.

In terms of interactive scene design, through the design of interactive scenes, it can be integrated with the ladle virtual digital character image to present a holistic effect that is rich in cultural connotations and artistic beauty. For example, by inputting descriptions and specific themes related to ladle culture into Stable Diffusion, traditional folk event scenes or ancient workshops can be generated. After completing the scene concept sketches, detailed requirements such as color preferences, lighting conditions, and atmosphere settings can be inputted, and based on the characteristics of ladle face culture, related props and decorations, such as traditional brocade, can be generated to give viewers a more immersive viewing experience.

5 Conclusion-Development Strategies and Future Prospects

AIGC technology and virtual digital humans can revitalize Ma Shao Face Painting culture by creating works that blend tradition with modern appeal. Virtual digital humans can effectively demonstrate the history and process of this art form, enhancing audience engagement. They also enhance virtual exhibitions with VR and AR, offering personalized experiences. AIGC can design related products and promote them online, balancing cultural preservation with commercial success. Nonetheless, accuracy, respect for the culture, and ethical considerations must be prioritized. Overall, this fusion will inject new life into Ma Shao Face Painting culture.

Acknowledgments

Research Results of the Innovation Fund Project of Xi'an University of Posts and Telecommunications on the Application of Stable Diffusion and Digital Human Integration - Taking the Inheritance and Promotion of Ma Shao Culture as an Example

Project number: CXJJZW2023009

References

1. Wang Xinyu, Wang Shenghua. Research on the Media Image Shaping and Cultural Representation of National Style Virtual Digital Humans - Taking "Tianyu" as an Example [J]. Audio-Visual, 2024, (10): 149-152. DOI:10.19395/j.cnki.1674-246x.2024.10.016.

2. Lin Ruyun, Li Sa. The Application and Enlightenment of Domestic Virtual Digital Humans in the Dissemination of Intangible Cultural Heritage [J]. Hundred Flowers, 2024, (08): 97-99.
3. Chang Hong. The Application of Virtual Digital Humans in the Inheritance and Development of Intangible Cultural Heritage [J]. People's Forum, 2024, (02): 103-105.
4. Yin Shuo. Dynamic Protection of Intangible Cultural Heritage Based on Virtual Digital Humans [J]. Art Market, 2023, (04): 92-95.
5. Tu Bingqian, Zeng Xin. Research on Interaction Design Strategies of DigitalHumans Based on Generative Artificial Intelligence [J]. Hunan Packaging, 2025, 40(02): 132-135.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.