

Identification of Issues and Innovation Strategies of Chinese AI Cultural Tourism Short Videos

Yun Li^{1*} Yanshuang Zhao²

School of Journalism Communication, Anhui University, 111 Jiu Long Road, Hefei 230011, China

* E-mail of the corresponding author: 40425396@qq.com

Abstract

In the era of video content, short videos have emerged as a new channel for cultural tourism marketing. The integration of short videos and cultural tourism has elevated marketing strategies to new heights, offering fresh perspectives for expanding the dissemination and marketing models of the cultural industry. AI technology empowers the cultural tourism sector, driving cultural heritage preservation and innovation. This paper examines the 2024 representative "AI XXX" series of Chinese cultural tourism short videos (such as "AI China", "AI Guangxi", "AI Qingdao", "AI You. Nanjing", "AI Ningbo, Love Ningbo!", and "AI Hainan") as case studies. It explores the visual artistry and existing challenges of AI-driven cultural tourism short videos and proposes future innovation pathways for reference.

Keywords: AI, cultural tourism short videos, innovation

DOI: 10.7176/NMMC/110-01

Publication date: November 30th 2025

1. Development Background of AI Cultural Tourism Short Videos

With the widespread application of advanced AI models such as ChatGPT and Midjourney, and driven by policies promoting the integration of culture and tourism, local cultural and tourism bureaus, mainstream media, and other entities have actively utilized AIGC (AI-generated content) technology to create imaginative and visually impactful AI cultural tourism short videos. These videos combine AI technology and visual art, offering audiences a dreamlike visual experience through the lens of AI. By leveraging tools like Midjourney, Stable Diffusion, Animatediff, and Runway, these videos vividly showcase the profound depth of Chinese culture and the unique charm of regional characteristics.

The diverse application scenarios constructed by short videos provide a new media context for reshaping cultural tourism resources, fostering deeper integration of culture and tourism. With the development of mobile internet and 5G technology, users increasingly prefer obtaining information and entertainment through short video platforms, which has laid the technical and user foundation for the popularity of cultural tourism short videos. Data shows that the scale of China's digital cultural tourism market grew from 787.05 billion yuan in 2017 to 969.81 billion yuan in 2022, an increase of 57.89%. By the end of 2023, the market size reached 1.15 trillion yuan, a year-on-year increase of 18.58%, accounting for over 30% of the entire cultural tourism industry. In 2024, China's digital cultural tourism market is expected to exceed 1.2 trillion yuan, setting a new historical record (2025 Analysis of Market Development Status, Supply and Demand Pattern and Future Trend of Digital Cultural Tourism, 2025).

"AI + Cultural Tourism" not only helps disseminate rich historical and cultural content but also provides users with virtual reality experiences, such as virtual tourism and historical scene restoration, enhancing user engagement and immersion. AI cultural tourism short videos contribute to improving urban cultural tourism services by exploring the unique cultural and tourism connotations, transforming "traffic" into "retention." The application of AI in tourism marketing offers innovative content marketing solutions, revolutionizing the way destinations are promoted.

2. Visual Art of AI Cultural Tourism Short Videos

AI cultural tourism short videos integrate the latest AIGC (AI-generated content) technology, leveraging AI's imaginative capabilities to tell compelling Chinese stories. Through fantastical illusions, dynamic time-travel effects, and striking visual impacts, these videos construct a powerful visual experience that transcends the

boundaries between reality and imagination. They open a door to infinite visual possibilities, playing a significant role in promoting cultural heritage and innovation.

2.1 Algorithm-Based: Computational Aesthetics

Computational image aesthetics refers to the process of enabling computers to simulate human visual and aesthetic thinking for making aesthetic decisions about images, thereby building a bridge between computers and visual art works(Wang Weining,2012).

AI cultural tourism short videos utilize the principles of computational aesthetics to create or enhance visual content, making it more aligned with human aesthetic preferences. By analyzing and learning from vast amounts of visual data, AI can identify visual elements and combinations that are more likely to evoke a sense of beauty, applying them to the generation and editing of video content to improve its appeal and artistic quality. For example, "AI China" unfolds with seven panoramic shots of the country's landscapes, presenting iconic natural scenery from a grand perspective that highlights the vastness and magnificence of China's mountains and rivers. By analyzing and integrating extensive literature, historical data, social media content, and more, AI delves into the folklore, historical stories, local cuisine, and artistic culture of each province, selecting the most representative and scenes with visual impact to comprehensively showcase the unique features of 34 provinces, municipalities, and autonomous regions. This provides audiences with a fresh cultural experience. In the creation of "AI China", over 200 landmark images were generated, and deep learning models were employed to produce city landscape pictures that are both realistic and artistically expressive. The team designed SD artistic fonts for the names of more than 40 cities, presenting the natural landscapes and cultural characteristics of each region. This resulted in a fantastical and aesthetically pleasing visual effect that captivates viewers.

2.2 Supported by Technology: Innovative Expression of Thoughts and Feelings

The application of AIGC technology in the field of short video of cultural tourism is not only a technological innovation, changing the way of content creation and dissemination, but also a profound change in the way of expression of thoughts and emotions.Wen Fengming and Wen Yuqun believe that Vincent Video has subverted the traditional short video production paradigm from four aspects: interactive planning, intelligent production, low-threshold post-production and virtual presentation(Wen Fengming,2024). AI cultural tourism short videos are supported by technology, constantly enriching content production concepts and expressions, breaking creative barriers and thinking boundaries, not only broadening the boundaries of imagination, but also allowing ideas and emotions to reach the hearts of the audience.

AI cultural tourism short videos allow AI to give free play to its creativity. While deeply exploring the cultural characteristics and advantages of various regions, AI cultural tourism short videos endow the region with unique perspectives and emotions, resulting in unique charm and ideological value of high conception, deep emotion, multi-level and multi-dimensional. AI enables cultural and tourism short videos to realize the blending of ancient and modern times, endow the audience with visual enjoyment through time and space, and trigger cultural perception and experience.By dynamically adjusting the scene," AI My China" realizes the dialogue between ancient and modern times in the same scene and the new integration of reality and cyber. The new visual experience is more in line with the preferences of young users. It integrates diverse excellent cultures of China from a macro perspective, which is no longer a traditional and single list of cultural elements.Through AI analysis of the uniqueness of local cultures, such as customs of different nationalities, local traditional skills, these elements are combined in a personalized way to show an inclusive, unique, vibrant and dynamic image of China.

"AI You. Nanjing" interweaves the actual situation and goes deep into the historical and cultural details of Nanjing, from the historical relics of the ancient capital of the Six Dynasties to the scientific and technological innovation of modern Nanjing, through the way of comparing ancient and modern, it shows the changes of Nanjing, vividly captures the ancient and modern style of Nanjing, and excavates the tenacity and vitality of Nanjing in the historical changes through AI. It conveys the heavy sense of history of Nanjing and the vigorous vitality of modern development. "AI Qingdao" focuses on the blending of Qingdao's marine culture and modern urban culture, and shows Qingdao's ancient and modern urban charm from a unique perspective. The unique perspective of thought and emotional expression makes the audience deeply feel the charm of regional characteristics, and generate curiosity and yearning for these places.

2.3 Driven by creative social hot spots: Interactive Integration

AI cultural tourism short video combines the current popular cultural phenomena or hot topics, driven by creative performance, and continuously provides magical and interesting content, which will make the image presentation full of new ideas. The personalized video generated by AIGC has a better insight into the aesthetic needs of young people and can create a highly interactive experience. Guided by the youth plus thinking, the production team of “AI MY China” has skillfully used the elements that young people think are interesting and topical, collected a lot of data before video production, and creatively integrated the hot content of social media platforms with the power of AI. For example, the scene of giant pandas gathering around a table to play mahjong, or the terracotta warriors and horses holding electronic menus to act as waiters, enhances the entertainment and interaction of AI cultural tourism short videos. “AI My China” interprets Shanghai with Xiaolongbao + Modern City + Flowers and Heilongjiang with the typical symbols of XXL frozen pears and skiing. It successfully combines the hot topics of Shanghai's Flowers Fever, Heilongjiang's cultural tourism explosion and frozen pears out of the circle. The presentation of strange fantasies catches the audience's attention and produces a strong interaction. The short videos of “AI Guangxi”, “AI Hainan” and “AI Qingdao” highlight the famous tourist punch cards, and redraw the scenery of these tourist punch cards through AI technology, so that the audience can feel the different styles of familiar and unfamiliar tourist punch cards, which increases the communication creativity and interactive effect of short videos of cultural tourism. The short video “AI Guangxi” shows the rich national cultural elements of Guangxi, such as folk song culture, national costumes, etc which shows the unique charm of national culture and conforms to the hot trend of national style.

3. Problems in AI Cultural Tourism Short Videos

AI cultural tourism (CT) short videos effectively integrate diverse CT resources and creatively present the unique cultural elements of different regions. These videos play a positive role in destination promotion, tourism marketing, improving CT services, and optimizing tourist experiences. However, as the application of AI in the CT industry remains in its early stages, several prominent issues persist, including homogenized creative expression, limited cultural depth, insufficient authenticity in emotional expressions, and a lack of immersive experiences.

3.1 Homogenized Creative Expression and Lack of Content Innovation

The creation of AI CT short videos often suffers from homogenized creative expression, leading to a lack of content innovation. When a particular AI CT short video achieves high exposure and traffic, production agencies and creators tend to imitate and replicate similar trending videos to garner attention, rather than focusing on differentiated and unique creativity. For instance, videos such as “AI Ningbo, Love Ningbo!” and “AI You, Nanjing” exhibit similar content, style, and presentation features as “AI My China”.

In the AI XXX series of short videos, natural landscapes, cultural history, and modern urban elements collectively form the shared visual components of “AI My China” and other regional CT videos. Taking “AI My China” as an example, this 2-minute-57-second video captures frames at the second level and eliminates highly repetitive images, resulting in 119 unique frames. Among these, virtual creative elements account for 45%, cultural history for 30%, natural landscapes for 16%, and modern urban elements for 9%. However, AI regional CT videos rarely incorporate or present virtual creative elements to reflect the cultural characteristics and distinctiveness of local regions.

In the AI XXX series, the application of AI technology remains confined to the “technical effects” level, lacking layered narratives and originality. For example, videos like “AI My China” and “AI You, Nanjing” integrate SD (Stable Diffusion) 3D artistic text of regional landmarks to display city names—a seemingly innovative approach. Yet, as this feature becomes a standard template in subsequent videos, it diminishes the novelty experienced by audiences. This overuse inevitably reduces viewers' sense of freshness and engagement.

3.2 Challenges in Representing Complex Cultural Elements and Limitations in Cultural Depth

AI-generated text-to-video and image-to-video tools struggle to capture the deep connotations of cultural elements due to their reliance on algorithms and existing datasets. The “output” of AI-generated videos is heavily influenced by “input,” such as textual prompts and preset materials. As a result, their outcomes are confined to the limited information within their source material and lack a profound understanding of cultural contexts.

The AI XXX series of cultural tourism short videos, while capable of presenting fantastical and surreal visuals with strong visual impact, face significant limitations in interpreting complex cultural elements and conveying the humanistic spirit inherent in regional cultural landscapes. For instance, “AI You, Nanjing” attempts to showcase the cultural heritage of the ancient Six Dynasties capital. However, the visuals presented—such as the Nanjing City Wall and Fuzi Temple—are simplified, symbolic representations. These visuals are delivered through static images with minor dynamic effects, failing to delve into the unique cultural connotations and characteristics created by the blending of Nanjing's history, culture, and modern technology.

In terms of cultural storytelling, AI CT short videos tend to directly display an accumulation of visually rich scenes rather than using specific narratives or characters to convey emotions or ideas. In the AI XXX series, each individual scene lasts no more than three seconds. While the rapid transitions and superficial content provide a powerful visual impact, they fail to highlight the uniqueness and cultural depth of the featured cities.

3.3 Insufficient Authenticity in Emotional Expression and Limited Immersion

AI cultural tourism short videos, rooted in technological frameworks, inherently struggle with authentic emotional expression. While the virtual visuals generated by AI may deliver a strong visual impact, they often lack emotional realism, making it difficult for audiences to develop a sense of immersion. For example, in “AI Ningbo, Love Ningbo!”, the depiction of the city's fishing port uses animated simulations to present dynamic effects like waves crashing and boats rocking. However, the visuals are overly idealized and deviate significantly from actual scenes. Local audiences in Ningbo may find it hard to relate to such depictions, as the nuanced emotions and human warmth that define these settings are absent, stripping the scene of its inherent significance.

The virtual nature of AI-generated content naturally lacks the authenticity associated with real-world experiences. Without real people and scenes involved, the narratives of CT short videos struggle to resonate with viewers on an emotional level. For instance, while “AI My China” and similar videos employ AI technology to showcase China's magnificent landscapes and rich cultural heritage, the scenes are entirely computer-generated “data.” This approach fails to capture the vitality and humanistic warmth of real cultural environments, such as the life-like atmosphere and the interpersonal connections inherent in those settings. Consequently, these videos lack the “human touch,” often appearing rigid and artificial, which can evoke resistance or discomfort among audiences, thereby limiting their immersive experience.

AI cultural tourism short videos should prioritize maintaining a natural and authentic feel, avoiding excessive use of special effects and filters. Over-reliance on such techniques risks alienating viewers by making the content seem overly artificial or detached from reality.

3.4 Lack of Immersive Experiences and Limited Interactivity

Although AI-generated technology introduces a new visual dimension to CT short video production, there remains significant room for improvement in terms of immersive experiences and interactivity. For instance, What is Nanjing includes 29 iconic cultural and scenic elements of Nanjing, such as Linggu Temple, Jinling, and the Yangtze River section in Nanjing, within a brief 1-minute-52-second runtime. However, its presentation relies on one-way delivery, where audiences passively “view” the content without truly immersing themselves or engaging interactively to experience the cultural essence.

AI technology has lowered the barriers to video creation, bringing innovative approaches to CT short video production and enabling broader participation. By leveraging AI to generate unique video effects and virtual characters, the expressive forms of short videos have been enriched. For example, AI cultural tourism short videos could utilize AI-generated interactive buttons, allowing audiences to “travel” to a historical period of their choice. Based on their selections, the video could display the natural landscapes and cultural features of the corresponding era, offering an interactive cultural exploration experience. Additionally, audiences could be encouraged to share on social media why they chose a particular historical period and what they would most like to do if they were to travel back in time.

By incorporating these interactive features, AI cultural tourism short videos can foster deeper audience engagement, enhance immersion, and transform passive viewing into an active, participatory experience.

4. Innovative Paths for AI Cultural Tourism Short Videos

With the rapid advancement of artificial intelligence (AI) technology, AI-driven CT short videos are

spearheading a revolution in visual and content creation through their imaginative breakthroughs. Although the application of AI in CT videos demonstrates immense potential, its current development stage still faces challenges in improving the natural fluency of video content, the authenticity of emotional expression, and collaboration with human creators. How to ensure content innovation and audience engagement while avoiding homogenization requires continuous exploration. How to integrate AI technologies with human creativity needs exploration as well.

4.1 Integrating Technology and Creativity for New Visual Experiences

Media researcher Joshua Meyrowitz argued that changes in media forms create new systems of information and corresponding contexts. Similarly, Robert Scoble emphasized the importance of virtual environments, based on the "five forces of context"—mobile devices, social media, big data, sensors, and location-based systems. AI development has blurred the boundaries between virtual and real worlds. From real-world scenes to virtual settings, augmented reality (AR), virtual reality (VR), and AI-generated hybrid environments, scene construction is breaking traditional physical spatial limits and entering a new dimension of virtual-real integration.

Currently, AI CT short videos lack thematic diversity, often resulting in monotonous viewing experiences for audiences. Future developments should focus on deeply exploring thematic content. China's rich traditional culture provides a vast reservoir for creative expression. By delving into the profound meanings of traditional culture and innovating perspectives to achieve stylized expression, AI cultural tourism short videos can unlock new developmental pathways. Traditional culture encompasses material culture, institutional culture, and spiritual culture, with diverse forms such as classical literature, poetry, music, and opera. By exploring regional histories, folklore, and stories of famous figures, content can be enriched. For instance, videos could explore the cultural heritage and stories of ancient villages or highlight the unique traditional crafts of a specific region and their modern evolution.

Through cross-temporal dialogues between history and modernity, AI cultural tourism short videos can address contemporary societal issues. By using AI technology to guide topic selection, perspectives, and narrative styles, the boundaries of cultural tourism content can be expanded. While AI cannot fully replace human creativity and emotional resonance, the development of AI CT videos must focus on storytelling with depth and warmth, presenting content that resonates both intellectually and emotionally.

4.2 Personalized Demand Fulfillment: Integrated Temporal-Spatial Experiences

Space and environment are traditionally defined as physical concepts; however, in AI-generated scenarios, time and space are redefined with newfound flexibility, emphasizing an integrated, real-time temporal-spatial experience. This approach focuses on constructing "in-the-moment" scenarios for audiences. The innovative development of AI cultural tourism short videos requires enhanced audience interaction. By leveraging AI to analyze dissemination data and user feedback, creators can better understand market demands and audience preferences, offering more personalized, interactive viewing experiences and tailored services. For example, enabling users to create their customized cultural tourism short videos or integrating their images into the content enhances both interactivity and engagement. Such features can further connect with audiences through online and offline events, such as virtual tourism check-ins or on-site cultural experiences.

Through algorithmic recommendations and data analysis, AI can align audience needs, behaviors, and scenarios to create highly personalized temporal-spatial experiences. Using AI-driven algorithmic recommendations, the system can predict and suggest content that matches an audience's desires for specific times and spaces. For instance, if an audience member wishes to engage in a particular type of tourism activity at a given moment, AI can recommend relevant videos or services. In the future, based on historical behavior and environmental data, AI could even generate unique virtual scenes tailored to the user's preferences, enabling "anytime, anywhere" personalized interaction.

This advanced temporal-spatial adaptability not only improves the precision of scenario dissemination but also imbues scenario construction with unprecedented flexibility and immediacy. Such innovations have the potential to revolutionize audience engagement, making cultural tourism experiences more immersive, relevant, and emotionally resonant.

4.3 "Human-Led, AI-Assisted": Deep Cultural Exploration in Cultural Tourism Short Videos

In the human-AI collaboration framework for AI cultural tourism short videos, humans occupy the central role, while artificial intelligence serves as a supporting tool. Humans remain at the core of content creation for cultural tourism videos, where creativity and emotional expression are the essence of the work. The application of AI technology neither challenges the centrality of humans in artistic creation nor overturns the existing cultural and artistic production systems(Ni Wan,2024).Instead, it acts as an extension of existing creative thought processes and production relations. AI can assist in processing video materials and providing data analysis, but the ultimate creativity and artistic expression must be driven by human agency. It is the synergy between AI's data-driven rationality and human experience-driven sensibility that enables the creation of cultural tourism short videos with greater depth and cultural significance.

AI-generated scenarios not only bring technological breakthroughs but also demonstrate innovative applications in human-AI collaboration. Traditionally, humans have been the sole creators of scenarios, with the designs and constructions tailored to human audiences, largely reliant on human design and imagination. With AI's integration, this process becomes increasingly intelligent and automated. AI can adjust scenario designs in real-time based on user behaviors and feedback, even autonomously generating scene content that aligns with user needs. Contemporary initiatives, such as the City Imagination: AI Cultural Tourism Short Video Creative Competition, highlight how AI serves as a catalyst for creative generation. The separation of technology and creativity into distinct yet complementary tracks signifies a future where scenario dissemination evolves beyond pre-designed frameworks, advancing through human-machine collaboration to produce emotionally rich and data-informed scene reproductions.

By leveraging technologies like natural language processing, AI can identify user preferences, while human creators incorporate real-world experiences to foster meaningful dialogues with audiences, environments, and society. Together, humans and AI dynamically adjust and generate scenes, achieving "contextualized" dissemination. Looking ahead, collaborations between humans and intelligent systems will enable the co-design of virtual spaces, augmented reality (AR) experiences, and even real-time generation of interactive content. These diverse forms of human-AI collaboration transcend the linear understanding of scenario construction, injecting greater possibilities into scenario dissemination and redefining the potential for cultural tourism content.

4.4 Adaptation and Application of the "Five Forces of Context": Technology-Driven Future Development

At the technical level, the future development of AI cultural tourism short videos can achieve higher quality and greater impact through the application and integration of the "five forces of context" with AI generation software. Based on big data technology, the development of AI cultural tourism short videos can proceed in two key directions. First, by collecting and analyzing extensive data, audience interests and viewing habits can be precisely identified, enabling the use of tools like ChatGPT, Wenxin Yiyan, and Qianyi Tongwen for generating virtual text and assisting with video script creation(Yang Yujie,2024). Second, big data optimization enhances video content and recommendation mechanisms, providing localized and targeted features to improve the audience viewing experience.

Mobile devices and sensor technologies enhance the real-time and interactive capabilities of videos. For example, augmented reality (AR) technology can be employed for virtual tours and interactive experiences. In the future, AI video tools such as Sora will empower ordinary users by providing technical tools to create content. Through user-generated content (UGC) and its dissemination on social media platforms, the reach of AI-generated videos can be expanded, reducing barriers to content production and fostering a positive dissemination effect.

Furthermore, by incorporating positioning systems and integrating geographic information with tourism resources, AI can deliver personalized travel recommendations and customized services, enriching the audience's tourism experiences. These applications of the "five forces of context"—big data, mobile devices, sensors, social media, and position—represent a technology-driven map for the evolution of AI CT short videos, offering new opportunities for innovation and audience engagement.

References

- 2025 Analysis of Market Development Status, Supply and Demand Pattern and Future
Trend of Digital Cultural Tourism Industry[OL]https://www.chinairn.com/hyzz/20241203/16474712.shtml
Wang Weining, Yi Jingjian, He Qianhua. Research Progress of Computable Image Aesthetics[J].Chinese Journal

of Image and Graphics,2012,17(08):893-90

Wen Fengming, Wen Yuqun.Research on the Path of Wensheng Video AI Enabling Short Video Content Production[J].Audio-visual world,2024,(03):31-34.

Ni Wan, Guo Jialiang. Exploring the Nature and Value of AI Artistic Creation from a Media Perspective [J]. Modern Communication, 2024, 46(08): 95-103.

Yang Yujie, Fan Gang. From Video Models to Application Scenarios: How AIGC Influences the Audiovisual Industry [J]. Film and Television Production, 2024, 30(03): 15-23.