Vol. 11 (2026): Proceedings of the 2nd International Conference on Social Sciences and Cultural Studies (ICSSCS 2026)
ISSN: 3105-0522 P-ISSN：3105-0514
DOI: https://doi.org/10.70267/icsscs.20260109

Human–AI Co-Creation in Film Production: Media Innovation and Aesthetic Governance in AIGC-Supported Filmmaking

Shiyu Huang*

School of Journalism and Communication, West China Normal University, Nanchong, Sichuan, 637002, China

**Corresponding author: Shiyu Huang*

Abstract

Generative artificial intelligence (AIGC) is rapidly transforming film production by introducing new modes of media innovation and aesthetic experimentation. Rather than functioning merely as a technical substitute, AIGC reconfigures creative workflows through diverse forms of human–AI co-creation. This paper examines how aesthetic decision-making, creative control, and cultural meaning are negotiated within AIGC-supported film production. Drawing on a comparative case study of the Chinese AI-assisted short film The Cipher in the Notes and the Singapore–Malaysia co-produced feature film Madame Ching, the study analyzes two contrasting production pathways: constraint-based generation for historical reconstruction and prompt-driven generation for stylized visual innovation. Using textual analysis and production documentation, the paper compares the cases across four dimensions: stages of technological intervention, human–AI division of labor, configuration of aesthetic control, and cultural–semantic calibration. The findings reveal that AIGC enables novel forms of visual experimentation and accelerates aesthetic iteration, while simultaneously shifting the director's role toward the governance of generative aesthetics. However, both cases expose shared challenges related to realism, cultural specificity, and ethical governance. By foregrounding media innovation and aesthetic processes, this study contributes to ongoing discussions on creative authorship, visual culture, and responsible AI adoption in contemporary media production.

Keywords

AIGC, media innovation, film aesthetics, human–AI co-creation, creative governance

1. Introduction

Recent advances in generative artificial intelligence have significantly expanded the expressive and productive capacities of contemporary media technologies. In film production, AIGC systems are increasingly employed not only for post-production enhancement but also for early-stage concept development, visual design, and narrative experimentation. The enhanced capabilities of generative AI in text, image, and video generation have enabled it to transition from a 'post-production supplement' to playing a more active role in the pre-production, production, and marketing phases of films. Rather than engaging in the grand debate about 'AI replacing creators,' a more actionable research approach focuses on practical implementation: how AIGC (AI-generated content) is integrated into workflows, the division of responsibilities between humans and systems, and the allocation and realignment of key decision-making authority. From the perspective of media

innovation, AIGC represents a shift in how images are conceived, generated, and evaluated within creative industries (Manovich, 2001, Jenkins, 2006).

This article does not re-examine the deterministic question of whether artificial intelligence will replace human creators. Instead, it explores from an aesthetic and process-oriented perspective how human creators and generative systems collaborate in practice, how aesthetic decisions are distributed in the combination of humans and artificial intelligence, and how creative authority is reconfigured. Therefore, the core research question of this article is: When artificial intelligence-generated content participates in film production, how are aesthetic innovation and creative control jointly produced and managed?

2. Research Design and Analytical Framework

This study employs a comparative case study methodology, focusing on two Chinese-language film projects situated in different production contexts. The Cipher in the Notes is a Chinese AI-assisted short film that emphasizes historical authenticity through carefully constrained generative processes. Madame Ching is a feature film co-produced in Singapore and Malaysia that foregrounds stylized visual spectacle through prompt-driven generative tools. Different from the existing researches which mostly stay in the technical review or value judgment, this paper takes the verifiable analysis dimension to disassemble the creative process, and tries to answer the main question of "how to implement the human-machine collaboration", and discusses its influence on the film aesthetics, industrial organization and copyright governance.

The study employs a combination of comparative case study and textual analysis. The materials include: the completed film, publicly available production interviews and reports, as well as accessible behind-the-scenes materials. The analytical framework consists of four dimensions: (1) stages of technological intervention across the production pipeline; (2) human–AI division of creative labor; (3) configuration of aesthetic and decision-making control; and (4) cultural–semantic calibration in generated imagery. This framework enables a systematic comparison of how media innovation unfolds through different modes of AIGC deployment.

3. Media Innovation through Generative Pathways

A persistent misunderstanding of AIGC equates generative output with the direct reproduction of training data. From a technical standpoint, diffusion models and generative adversarial networks generate images by learning statistical distributions rather than copying specific visual instances (Ho et al., 2020, Goodfellow et al., 2014). This generative logic underpins new forms of media innovation by enabling rapid aesthetic exploration.

The Cipher in the Notes exemplifies a constraint-based generative pathway. By curating authorized historical images, documentary records, and textual materials, the production team employs AIGC to iteratively explore visual reconstructions while maintaining stylistic and historical consistency. Here, AIGC functions as an aesthetic exploration engine constrained by human-defined cultural and historical parameters. The core data sources, authorization, and screening scope are clearly defined as follows: The character data is primarily sourced from the officially authorized portrait rights of the late performing artists Li Yalin and Liu Shilong, with AI learning and replication based on their representative works (e.g., "The Youngsters in Our Village" and "Heroic Offspring") and publicly available visual materials. The screening scope is focused on the images of the artist's golden period of art career, ensuring the authenticity and representativeness of the digital image. The historical scene data is based on the existing historical image archives of the 1943 song creation site in Tangshang Village and the 1989 Beijing street style for AI training and restoration. The historical scene data is based on the existing historical image archives of the 1943 song creation site in Tangshang Village and the 1989 Beijing street style for AI training and restoration. The narrative material of the literature originates from the oral history and public reports of Cao Hongwen, daughter of Cao Huoxing, which includes details such as 'bean selection' and 'leaf command,' all of which are documented in public records. The script is strictly based on verified historical events and character relationships to ensure the authenticity of the narrative framework. In this project, AIGC primarily supports the initial conceptualization and scenario planning phases. However, its outputs still require human team validation and redesign before proceeding to subsequent production stages.

In the exploration of visual style, The Queen of the Sea Zheng Yisao has practiced a path of "native style generation". The production team opted not to develop a dedicated model, instead leveraging existing general-purpose image generation tools like Midjourney and Stable Diffusion. Through the carefully designed prompt engineering, the rough lines of Qing Dynasty woodcut, the freehand blank of Chinese painting and the dynamic tension of martial arts films are fused and reshaped. This combination isn't just a simple addition, but rather a unique visual aesthetic that emerges from AI's generative logic. The historical theme and the visual impact of genre films are intertwined, forming a "strong typification and strong stylization" appearance that is both ancient and full of screen vitality. In terms of visual presentation, the AI intermediary plays a pivotal role. In this approach, two core tasks are the continuous iteration of prompts and the maintenance of consistency with the film's overall visual style. The team will enter specific instructions to guide the generative model in synthesis. The team must conduct numerous iterations during this process, continuously refining the prompt to better align with the desired aesthetic vision. Meanwhile, maintaining this highly customized style throughout the 70-minute film to prevent stylistic drift or quality fluctuations has become the core challenge and critical process in production. In this project, AIGC technology plays a pivotal role in translating abstract artistic concepts into visual drafts. It effectively connects the creative conception of the director and the art director with the final image realization, making the highly complex aesthetic intention of "making the Qing Dynasty woodcut come alive and staging a jianghu epic" technically possible for large-scale production and fine-tuning.

Both examples demonstrate AIGC's edge in solution generation and iteration speed. Its value lies not in producing a' final shot' upfront, but in dramatically shortening the cycle from conceptualization to visual proposal through low-cost trial-and-error, thereby expanding the visual solution space. This technology enables small teams to explore and make decisions more frequently and at lower costs during the conceptual design and initial visualization phases of traditional processes.

Meanwhile, the two cases also revealed different types of bottlenecks: The first category involves stability issues in continuous motion interacting with complex physical systems (e.g., the continuity between human movements and ocean waves). The other is the long-term maintenance cost of cross-shot style consistency and detail logic. This leads to two comparable paradigms: The first is "enhancing controllability through data and rule constraints". The second approach is "style-driven production through a system of universal tools and prompt words". The two have different requirements on personnel structure, cost structure and quality control mechanism.

4. Aesthetic Governance and Creative Roles

In both cases, AIGC has accelerated visual iteration and expanded the range of aesthetic possibilities available to filmmakers. However, this expansion also necessitates the establishment of new forms of aesthetic governance. The director's role increasingly shifts from direct orchestration toward the evaluation, selection, and regulation of AI-generated outputs, reflecting broader transformations in information-intensive creative work (Floridi, 2014).Therefore, aesthetic governance itself becomes a creative meta-practice, and the creative core of the director becomes the shaping of the boundary and the giving of meaning to the algorithmic possibilities, and its artistic authority is re-established through this new aesthetic decision-making and frame control.

The two parts jointly present a cyclic collaborative structure of "creativity - generation - screening - optimization", in which roles such as prompt engineer/AI art designer undertake the pivotal function of connecting aesthetic intentions with model outputs. It is important to note that collaboration does not mean "the automatic transfer of subjectivity to AI": at the levels of responsibility and attribution, the final decision-making, compliance review, and risk-bearing entities should still be clearly defined. To avoid conceptual generalization, this article describes AI as "the generation system in the production process/non-human actor", and limits "subjectivity" to the allocation of decision-making power that can be held accountable, attributed, and governed, rather than an abstract anthropomorphic expression.

Emerging roles such as prompt designers and AI artists mediate between human creativity and machine generation, forming iterative collaboration loops of ideation, generation, selection, and refinement. Importantly, creative agency does not transfer to AI systems themselves; instead, agency remains anchored in accountable human decision-making structures (Barthes, 1977, Jenkins, 2012).

5. Aesthetics, Realism, and Cultural Semantics

The Code in the Notes is designed to create an immersive experience through chronological coherence and dense details. For instance, in close-up shots recreating historical figures, AI analyzes archival footage of deceased artists to generate lifelike skin textures, clothing folds, and era-appropriate eye lighting, creating an illusion of 'confronting historical figures in person.' The film "Queen of the Sea: Zheng Yisao" creates its genre appeal through high-concept visual spectacles. For instance, in the climactic scene where Zheng Yisao stands at the ship's bow with her sleeves and waves swirling in surreal tones, though not physically real, the stylized composition generates intense dramatic tension. These digital images are not only a visual reproduction, but also an activation and re-encoding of the deep structure of cultural memory, which gives new semantic levels to historical symbols in the contemporary context. This technology-driven creative mode essentially constructs a "digital realism" aesthetic, which challenges the boundary between traditional documentary and fiction, and opens up a new dimension of expression between the sense of reality and artistic expression.

From an aesthetic perspective, these two films present a stark contrast in their portrayal of realism versus visual spectacle. The Cipher in the Notes seeks immersive realism through historical detail density, while Madame Ching constructs a heightened, stylized visual world that prioritizes affective impact over physical plausibility. AIGC-generated images achieve credibility through stylistic coherence rather than photographic indexicality, complicating classical theories of cinematic realism (Bazin, 1967, Baudrillard, 1994).

Stylized generation also amplifies the risk of cultural homogenization and symbolic misalignment. For example, when the coastal figures of the Qing Dynasty are generated, there may be a mixture of the Tang Dynasty's curtain hat style, the Ming Dynasty's cloud and thunder patterns, and the pan-East Asian architectural components. or in the depiction of specific weapons, the distinctions between different periods of design, such as the ring-headed sword and the willow-leaf sword, are blurred. The case shows that this problem is often alleviated by the refinement of prompt words, the constraint of reference images and the post-correction, but it also means that the accuracy of cultural semantics is transformed into new labor and new costs. Without careful calibration, generative systems may blend visual elements from different historical periods or cultural contexts. Addressing such issues requires sustained human intervention through prompt refinement, reference constraints, and post-generation correction, effectively transforming cultural accuracy into an additional layer of creative labor (Zhao, 2023, Fan, 2025).

6. Ethics and Responsible Media Innovation

The expansion of AIGC in film production raises pressing ethical and legal concerns, particularly regarding training data provenance, authorship attribution, and cultural representation. Ongoing copyright litigation surrounding generative image systems illustrates the evolving and uncertain regulatory landscape. In this context, responsible media innovation depends on project-level governance mechanisms, including transparent data sourcing, authorization procedures for personal likenesses, traceable generative workflows, and clear disclosure practices (Zhou et al., 2024, Yao, 2025, Xiao, 2024).

The current copyright litigation related to generative images (such as Andersen v. Stability AI and other cases) indicates that the main points of contention involve issues such as the source of training data, the definition of reproduction and derivative works, and the substantial similarity between the output and the original work (Stability AI Midjourney & DeviantArt Litigation, 2023). Currently, there is no unified judicial judgment standard worldwide, and the judgment results show a dynamic evolution feature. This legal environment's uncertainty constitutes a persistent systemic risk in this field, and industry practitioners need to closely monitor the development of case precedents and legal interpretations in different jurisdictions and use them as an important basis for risk assessment.

To address these risks, the viable approach is to establish a comprehensive compliance framework that spans the entire project lifecycle.This includes clear registration of the source of training data and definition of the scope of authorization, and establishing a process for obtaining explicit prior consent for the use of specific personal elements such as real human portraits and voices. At the technical level, implement version retention and traceability for generated content to ensure that key steps from initial prompts to final outputs are traceable and auditable. During the release phase, establish and enforce a unified policy for prompts and

credits, such as clearly labeling 'AI-generated' information. Meanwhile, relevant industry organizations have discussed such issues and issued preliminary initiatives, such as the ethical governance seminar minutes publicly released by institutions like the China Artificial Intelligence Industry Development Alliance, which provide valuable reference frameworks for the industry.

7. Conclusion

By comparing The Cipher in the Notes and Madame Ching, this paper identifies two representative modes of AIGC-supported media innovation in film production: constraint-based generative reconstruction and prompt-driven aesthetic stylization. The first approach enhances controllability through data and rule constraints, with AIGC primarily handling initial scheme expansion and style restoration. The second approach employs general-purpose generation tools and a cue word system to drive stylized production, with the cue word team serving as the key intermediary for aesthetic realization. Both modes highlight the growing importance of aesthetic governance and human oversight in AI-assisted creativity. The study argues that AIGC should be understood not merely as a production tool but as a media innovation infrastructure that reshapes aesthetic practices, creative labor, and cultural meaning-making in contemporary filmmaking.

References

Barthes, R., (1977). *Image-Music-Text*, (Vol. 6135). New York, NY: Macmillan.

Baudrillard, J., (1994). *Simulacra and simulation*. Ann Arbor, MI: University of Michigan Press.

Bazin, A., (1967). *What Is Cinema?* Oakland, CA: University of California Press.

Fan, Y. S., (2025). AI-driven restructuring of China's film industry ecosystem: A study on the driving forces of transformation and the mechanism of factor flow. *Academic Research***,** no. 6, pp. 47-52.

Floridi, L., (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford: Oxford University Press.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., (2014). Published. Generative adversarial nets. *Advances in Neural Information Processing Systems*, Red Hook, NY,. Curran Associates, Inc., pp. 2672–2680.

Ho, J., Jain, A. and Abbeel, P., (2020). Published. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, Red Hook, NY. Curran Associates, Inc., pp. 6840-6851.

Jenkins, H., (2006). *Convergence culture: Where old and new media collid*. New York University Press.

Jenkins, H., (2012). *Textual poachers: Television fans and participatory culture*. London: Routledge.

Manovich, L., (2001). *The Language of New Media*. Cambridge, MA: MIT Press.

Stability AI Midjourney & DeviantArt Litigation, (2023). *Class action complaint* [Online]. Available: https://stablediffusionlitigation.com/ [Accessed 6 February 2026].

Xiao, X., (2024). The boundaries and aesthetic characteristics of AI art enabling film creation from the perspective of artistic truth. *Movie Literature***,** no. 8, pp. 40-46.

Yao, R., (2025). AI playwriting from the perspective of film industry aesthetics: Theoretical stimulation, practical coupling, and problem generation. *Ethnic Art Studies,* vol. 38**,** no. 2, pp. 88-95.

Zhao, Y., (2023). Human-computer Co-creation, data fusion and multimode model: Film art and cultural industry criticism of generative AI. *Contemporary Cinema***,** no. 8, pp. 15-21.

Zhou, W., Liu, W., Guo, F. and Chen, H., (2024). AI, film, AI film, and discussions about the future. *Contemporary Cinema***,** no. 5, pp. 4-13.

Funding
This research received no external funding.

Conflicts of Interest
The authors declare no conflict of interest.

Acknowledgment
This paper is an output of the science project.