

The Transformation of Micro-Drama Creation Models Driven by AI Video Generation Technology

Xu Zhang*

Huanghe S & T University, Zhengzhou, 450000, China

*Corresponding author: m15837300919@163.com

Abstract: *The iterative breakthroughs in AI video generation technology are driving a systematic transformation in the creation models of micro-dramas from the underlying logic. This study focuses on the evolutionary trajectories across three dimensions: technological logic, content production, and creative subjects. At the level of technological logic, generative models reshape the creative process through the reconstruction of narrative grammar, the establishment of human-machine collaboration mechanisms, and the role transition from auxiliary tools to generative subjects. At the level of content production, algorithmic regulation of plot tension, infinite generation of spatiotemporal scenes, and algorithmic synthesis of character images drive the production model toward intelligent transformation, with textual features exhibiting tendencies toward modularity, programmability, and surreal splicing. At the level of creative subjects, the creators' functions shift toward aesthetic arbiters, technical literacy is redefined, and generative technologies give rise to aesthetic characteristics such as “algorithmic smoothness,” establishing a bidirectional symbiotic relationship between technological iteration and the creative ecosystem. The study reveals that AI video generation technology has evolved into a core driving force within the creative process, and the deep integration of algorithmic logic with creative logic is redefining the formal boundaries and developmental pathways of micro-dramas.*

Keywords: *AI video generation; micro-dramas; creation models; narrative grammar; human-machine collaboration; technological leap; intelligent content production*

Introduction

As an emerging form of digital audiovisual content, micro-dramas are distinguished by their high-frequency narration and strong emotional density, and they have risen rapidly in the field of content consumption. With the expansion of the market scale, micro-drama creation faces an inherent tension between production capacity and content quality, as the traditional linear production model struggles to balance efficiency with innovation. The breakthrough in AI video generation technology offers a solution to this problem: generation systems based on diffusion models, Transformer architectures, and multimodal alignment algorithms have already achieved the direct synthesis of coherent narrative segments from text, and the integration of this technology is moving from the level of tool assistance to the core aspects of the creative process. This transformation not only improves production efficiency but also concerns the reconstruction of narrative logic, the reorganization of visual presentation, and the repositioning of creators' functions. Systematically examining the transformation paths of micro-drama creation models driven by AI video generation and revealing the coupling mechanism between technological logic and creative logic hold significant academic value for understanding the technological shift in digital content production. Starting from the intervention mechanism of technological logic, this study sequentially analyzes the intelligent shift in content production models and the generative patterns of textual features, and then explores the transformation of creative subjectivity and the evolution of industrial forms, in an effort to construct a systematic analytical framework encompassing the three dimensions of technology, content, and subjects.

1. The Technological Logic and Coupling Mechanism of AI Video Generation in Micro-Drama Creation

1.1 The Reconstruction of Micro-Drama Narrative Grammar by Generative Models

Through deep learning of massive narrative data, generative models are reshaping the narrative

grammar of micro-dramas from the underlying coding level. Video generation technologies based on Transformer architectures and diffusion algorithms can automatically identify and extract narrative units with high communication efficiency, transforming the traditionally linear dramatic structure into quantifiable and calculable plot modules. This algorithm-driven narrative generation mechanism enables the plot progression of micro-dramas to no longer rely entirely on the creator's subjective conception; instead, it forms an automated path of “hotspot identification-plot synthesis” based on data training, thereby rendering the density of dramatic conflict and the arrangement of reversal points programmable^[1].

The reconstruction of narrative grammar is further manifested in the compression of the temporal dimension and the reinforcement of the emotional dimension. Through algorithmic modeling of audience attention, generative models can achieve the distribution of high-frequency emotional stimulation points within an extremely short narrative duration, and the progressive logic of introduction, development, climax, and denouement in traditional narratives is replaced by a modular superimposition mechanism of pleasure points. Consequently, the narrative structure of micro-dramas exhibits distinctly algorithm-optimized characteristics: the logical coherence of plot development gives way to real-time adjustments based on data feedback, and the combination of narrative units shifts from linear sequencing to reconfigurable parallel arrangements.

1.2 The Human-Machine Collaboration Mechanism in the Dimension of Visual Presentation

At the level of visual presentation, AI video generation technology constructs a bidirectional collaboration mechanism between the creator's intent and the algorithm's generative capabilities. Through prompt engineering, creators exert targeted control over visual styles, scene compositions, and lighting relationships, while generative models perform feature mapping and pixel synthesis within the latent space, transforming abstract linguistic descriptions into concrete visual sequences. This process forms a cyclic feedback loop of prompt optimization and image iteration: the creator's aesthetic judgments are technically realized through algorithmic parameters, and the stochastic generation results of the algorithm, in turn, inspire the extension and adjustment of visual imagination.

The establishment of the human-machine collaboration mechanism further alters the generation path of visual elements. In terms of scene construction, the application of conditional control technologies such as ControlNet enables character movements and background generation to maintain consistency in spatial logic, while the real-time rendering of virtual scenes and the flexible invocation of digital assets break through the temporal and spatial constraints of physical shooting. The visual presentation of micro-dramas is no longer a passive recording of the real world; instead, it becomes a composite product of algorithm-based generation and human intervention. The visual texture shifts toward being calculable and controllable, and the negative correlation between the degree of spectacle in scenes and production costs is thereby restructured.

1.3 The Technological Leap from Auxiliary Tool to Generative Subject

The role of AI video generation technology within the micro-drama creation process exhibits a clear trajectory of transition from peripheral assistance to core generation. In the early stages of technological integration, generative models primarily performed auxiliary functions such as image quality enhancement, shot restoration, or simple transition synthesis, with their role limited to optimizing the efficiency of existing creative workflows. As video generation models achieved breakthroughs in key technical dimensions such as temporal coherence, multimodal alignment, and physical law simulation, AI began to acquire the capability to independently generate complete narrative segments, enabling both the coherent generation of character movements and the automatic synthesis of shot sequences to be accomplished at the algorithmic level^[2].

This technological leap has triggered a corresponding adjustment in the structure of the creative subject. The traditional linear division of labor model, centered on screenwriters, directors, and post-production editors, is gradually shifting toward a parallel collaborative relationship between creators and algorithmic systems, with AI transformed from a tool that is invoked into a collaborative subject possessing content-generation capabilities. The creation of micro-dramas no longer depends entirely on physical shooting and manual production; instead, it has entered an algorithm-driven stage based on model training, parameter tuning, and generated content selection, with the generative capacity of the technological system and the aesthetic control of creators jointly constituting the dual fulcrums of the new creative workflow.

2. The Intelligent Shift in Content Production Models and Textual Characteristics

2.1 Algorithmic Optimization of Script Generation and Plot Tension

The integration of AI video generation technology is propelling micro-drama script creation into an algorithm-optimized phase. Generation systems based on large language models and reinforcement learning frameworks can perform feature extraction and pattern recognition on vast amounts of script data, automatically constructing a narrative parameter system that includes conflict density, reversal frequency, and emotional curves. Script generation is no longer a creative process that relies solely on the inspiration of the creator; instead, it has formed an automated path that can be quantitatively regulated based on data training, where the introduction, development, climax, and denouement of the plot are transformed into structural units that algorithms can recognize and calculate. This technological intervention provides a precise quantitative basis for the creation of dramatic tension, allowing the arrangement of conflict nodes and the occurrence of emotional peaks to be directionally configured according to audience preference data.

This algorithmic optimization mechanism is further manifested in the dynamic regulation of plot tension. Through real-time analysis of audience feedback data, generative models can iteratively refine the release rhythm of suspense elements and the distribution density of emotional resonance points within the script, thereby endowing the script content with a data-driven capacity for self-evolution. The plot structure of micro-dramas no longer maintains a fixed form; instead, it can be adaptively adjusted at the algorithmic level to match the acceptance thresholds of different audience groups. The generative logic of dramatic tension shifts from traditional linear progression toward modular combination and programmable regulation, significantly strengthening the technical attributes of script creation^[3].

2.2 Virtual Production and the Infinite Generation of Spatiotemporal Scenes

The maturation of virtual production technology provides an infinite generation path for the spatiotemporal construction of micro-dramas. Scene generation algorithms based on neural radiance fields and diffusion models can automatically synthesize high-fidelity three-dimensional spatial environments according to textual descriptions, effectively eliminating the physical constraints of shooting locations and the costs of set construction through algorithmic generation capabilities. The combination of modular storage of scene elements with real-time rendering technology enables the reuse and recombination of the same digital assets across different episodes, significantly enhancing the efficiency and flexibility of spatiotemporal construction. The scene presentation of micro-dramas is no longer constrained by the physical existence of the real world; instead, it has entered a stage of virtual construction based on algorithmic synthesis.

The infinite generation of spatiotemporal scenes further transforms the visual textual characteristics of micro-dramas. Algorithms can construct historical scenes, fantasy spaces, or micro-worlds in real time according to the needs of the plot, greatly expanding the breadth and freedom of visual presentation. Scene transitions no longer rely on physical cuts between shots; instead, they can achieve seamless transitions and stylistic gradations through algorithms, ensuring technical support for the continuity of spatiotemporal flow. Consequently, the visual space of micro-dramas exhibits distinct surreal characteristics, with the degree of spectacle in scenes and the intensification of narrative emotion forming a deep technical-level coupling, and the programmability of spatiotemporal construction becomes an internal driving force in the evolution of the text's visual style.

2.3 Algorithmic Synthesis of Character Images and Emotional Projection

The algorithmic synthesis of character images serves as an important manifestation of AI video generation technology's integration into micro-drama creation. Digital human synthesis systems based on generative adversarial networks and three-dimensional reconstruction technologies can achieve fully automated generation across the entire process, from facial features and micro-expression dynamics to sequences of limb movements. The construction of character images no longer relies on the live performances of human actors; instead, it enables parametric control of image characteristics at the algorithmic level, allowing features such as age, temperament, and emotional expression to be directionally synthesized according to the requirements of the script. This technological approach significantly enhances the controllability and reproducibility of character images, enabling the same digital image to maintain visual consistency across different episodes.

The relationship of emotional projection between character images and the audience undergoes a reconstruction in the context of algorithmic synthesis. The micro-expressions and eye movements of synthetic characters are products of feature extraction and model training based on massive amounts of real performance data, with their modes of emotional expression technically pursuing a high degree of approximation to human performance. The audience's emotional identification with synthetic characters no longer rests on the improvisation and irreproducibility of live performances; instead, it shifts toward a technical recognition of algorithmically generated emotions. Consequently, the character images in micro-dramas become composite products of data training, where the authenticity of their emotional expression stems from the algorithm's precise modeling of performance patterns rather than the individualized interpretation of specific actors, and the logic of characterization shifts from the recording of performance to algorithmic synthesis^[4].

3. The Transformation of Creative Subjectivity and the Evolution of Industrial Forms

3.1 The Transformation of Creators' Functions and the Reconstruction of Technical Literacy

The deep integration of AI video generation technology triggers a structural transformation in the positioning of creators' functions. The functional boundaries of roles such as screenwriters, directors, and post-production editors in the traditional creative process gradually blur under the infiltration of algorithmic generative capabilities, shifting the core work of creators from execution-level tasks such as camera shooting, scene scheduling, and editing synthesis toward the selection, refinement, and aesthetic control of AI-generated content. The initial drafting of scripts, the visual presentation of storyboards, and even complete narrative segments can be automatically completed by generative models, thereby shifting the role of creators significantly toward that of “prompt engineers” and “aesthetic arbiters,” with their focus centered on the parameter tuning of generative models, the precise design of prompt syntax, and the evaluation of visual consistency in generated results, rather than the manual production of specific frames or physical shooting. This transformation implies that the technical execution aspects of the creative process are effectively taken over by algorithmic systems, allowing creators to concentrate their efforts on higher-level tasks such as conceptual ideation and aesthetic direction.

The reconstruction of creators' technical literacy has become an inherent requirement and foundational support for this round of functional transformation. Traditional narrative skills and visual aesthetic capabilities need to form a new composite structure with algorithmic literacy, requiring creators to possess a fundamental understanding of the underlying operational logic of generative models, master the technical methods of regulating visual style, character forms, and scene composition through prompt engineering, and be capable of making professional judgments and precise corrections regarding the performance of algorithmically generated content in terms of narrative coherence, accuracy of emotional expression, and consistency of visual style. Technical literacy is no longer regarded as a supplementary skill or optional capability for the post-production stage; instead, it constitutes a core competency element running through the entire process from conceptual ideation and content generation to final product optimization, and the efficiency of collaboration and the quality of output between creators and algorithmic systems directly depend on the degree of development and the level of application of this new form of literacy. The pace of technological iteration requires creators to maintain a posture of continuous learning, and their competency structures need to be dynamically adjusted to synchronize with algorithmic evolution^[5].

3.2 Aesthetic Characteristics of Micro-Dramas Driven by Generative Technology

The integration of AI video generation technology has given rise to a unique system of aesthetic characteristics that distinguishes micro-dramas from traditional film and television. In the visual dimension, algorithmically generated content exhibits a distinct tendency toward “surreal splicing,” where different spatiotemporal scenes, stylistic elements, and visual symbols can achieve seamless fusion and instantaneous transformation within the same narrative space, shifting the logical connection of scene transitions from physical realism to visual fluidity and emotional coherence, while the degree of spectacle in images is systematically enhanced at the technological level. This visual characteristic stems from the generative model's capacity for reorganizing intrinsic features and transferring styles from massive image datasets, rather than from passive recording of the real world through shooting, thereby shifting the logic of scene construction from representation to generation and from imitation to creation. Consequently, the visual space of micro-dramas possesses a high degree of plasticity and

programmability, where the combination of visual elements breaks free from the constraints of physical laws, forming a new visual language based on algorithmic logic^[6].

In the dimension of emotional expression, micro-dramas have developed an aesthetic style characterized by “algorithmic smoothness.” The trajectory of emotional transitions in synthetic characters, the dynamic presentation of facial micro-expressions, and the emotional progression of plot fluctuations are all based on expressive patterns and statistical features formed through data training. The abruptness of emotional expression and the traces of individualized performance are effectively dissolved through algorithmic averaging, and the emotional curve exhibits smoothly transitioning characteristics that have been quantitatively regulated. The emotional transmission in micro-dramas presents highly controllable, predictable, and repeatable technical features, where the intensity level of dramatic conflict and the rhythmic nodes of emotional evocation can be precisely quantified and directionally configured at the algorithmic level. This technical aesthetic is neither akin to the documentary reproduction of real emotions in traditional film and television nor to the external superimposition of visual spectacle through digital effects; rather, it represents a new aesthetic form shaped by the internal operational logic of generative models and the principles of data training, with its uniqueness stemming from the profound reshaping of emotional expression by the technological system.

3.3 The Symbiotic Evolution of Technological Iteration and the Micro-Drama Creation Ecosystem

The continuous iteration of AI video generation technology and the structural evolution of the micro-drama creation ecosystem exhibit an evident symbiotic relationship and bidirectional shaping characteristics. Upgrades and advancements in the technological system are directly reflected across various stages of the creative process and the evolutionary trajectory of content forms. Breakthroughs in key technical dimensions of video generation models, such as generation resolution, temporal coherence, multimodal alignment accuracy, and physical law simulation, continuously expand the boundaries of content expression and the realm of visual presentation possibilities for micro-dramas. Concurrently, the specific demands of the creation ecosystem regarding technological outputs guide the direction of algorithm research and the focus of technological optimization. The content characteristics of micro-dramas, namely their high frequency, short production cycles, intense emotions, and high density, serve as an important reference framework for the training and optimization of video generation models, as well as directional requirements for technological iteration. A dynamic mechanism of mutual shaping and coordinated evolution is thus formed between technology and creation.

The internal structure of the creation ecosystem undergoes continuous reorganization and functional differentiation during this process of symbiotic evolution. The content supply relationship shifts from independent production by professional institutions toward collaborative production between institutions and algorithmic systems, the creative cycle is significantly compressed, and the characteristics of content iterability, reusability, and modular combination become important technical features for the efficient operation of the ecosystem. The enhancement of generative capabilities brought by technological iteration and the ecosystem's demand for content volume and update frequency form a positive feedback loop that reinforces each other, with the morphological and technical characteristics of micro-dramas continuously adjusting and being redefined alongside algorithmic evolution. Technical bottlenecks, such as long-term coherence in narrative logic, nuanced layering in emotional expression, and the maintenance of character consistency across scenes, consistently impose constraining factors on the creation ecosystem, yet such technical constraints, in turn, become directional research and development requirements for the next round of algorithmic breakthroughs and intrinsic driving forces for technological advancement, thereby ensuring the dynamic balance and coordinated evolution between technology and creation are continuously maintained and progressively advanced to higher levels.

Conclusion

The transformation of micro-drama creation models driven by AI video generation technology is, in essence, a deep penetration of algorithmic logic into the core aspects of content production. At the level of technological logic, generative models reshape the underlying operational mode of the creative process through the reconstruction of narrative grammar, the establishment of collaborative mechanisms in visual presentation, and the transition from auxiliary tools to generative subjects. At the

level of content production, the algorithmic optimization of plot tension in scripts, the infinite generation of spatiotemporal scenes, and the algorithmic synthesis of character images drive a systematic evolution of textual features toward modularity and programmability. At the levels of creative subjects and industrial forms, the transformation of creators' functions toward aesthetic arbiters catalyzes the reconstruction of technical literacy, micro-dramas develop an aesthetic system centered on “surreal splicing” and “algorithmic smoothness,” and a bidirectional symbiotic mechanism is established between technological iteration and the creation ecosystem. Looking ahead, AI video generation technology will continue to evolve in the directions of narrative logic control and emotional expression simulation, with creation models correspondingly advancing toward deeper levels of human-machine collaboration. The tension between the controllability of content generation and the unpredictability of creativity will become a core issue continuously confronted by technological evolution and creative practice, and the transformation of creation models will be embodied in the deep integration of technological logic and creative logic as they collectively explore new content forms.

References

- [1] Duan, Mei. “Intelligent Media Revolution and Cultural Transcoding: The Technological Construction and Cultural Paradigm Reconstruction of Short Video Creation in the Era of Generative AI.” *New Media Research*, vol. 12, no. 1, 2026, pp. 102-105.
- [2] Han, Xiaoyi, and Dong Yuzhi. “Research on the Application of AI Generation Technology in Short Video Content Creation.” *Media Forum*, no. 19, 2025, pp. 9-11.
- [3] Lu, Kuan. “A Preliminary Study on the Technical Principles of AI-Generated Video.” *Proceedings of the 2025 Annual Conference of Tianjin Electronic Industry Association*, edited by Tianjin Haihe Media Center, 2025, pp. 124-129.
- [4] Chen, Xin. “An Exploratory Study on the Innovative Application of AI Video Generation Technology in Short Video Production.” *News Tide*, no. 4, 2025, pp. 3-5+10.
- [5] Xie, Bo, and Li Yujing. “The Generation Logic and Collaborative Governance of Security Risks in AI-Generated Video Technology.” *Journal of Beijing Police College*, no. 2, 2025, pp. 1-9.
- [6] Ni, Junjie. “Let Creativity Continuously Emerge: The Application of AIGC Video Generation Technology.” *China Information Technology Education*, no. 7, 2024, pp. 69-73.