

From "Recording" to "Generating": The Reconstruction of Film Lens Aesthetics by Generative AI

Fang Cao

Communication University of Nanjing, Nanjing, Jiangsu, 211172, China

ABSTRACT

Generative AI subverts the traditional film aesthetic paradigm based on photographic indexicality through algorithmic image synthesis, realizing a paradigmatic shift from "recording reality" to "generating reality" in film creation. Starting from the perspective of film ontology, this paper analyzes the technical logic of generative AI's image generation and its impact on the traditional "myth of total cinema". By exploring the changes in film lens language in composition, movement, light and shadow, and montage under the intervention of generative AI, it reveals the triple reconstruction logic of AI film aesthetics: the ontological shift from photographic ontology to algorithmic generation theory, the subject migration from single human creator to human-AI hybrid subject, and the formal fission of lens language driven by technological affordance. Combined with practical cases such as Sora and Runway Gen-2, this paper also reflects on the crises of hyperreality and emotional defocus in AI-generated images, and puts forward the development prospect of human-AI collaborative creation, providing a new theoretical perspective for film aesthetic research in the AI era.

KEYWORDS

Generative AI; Film lens aesthetics; Image ontology; Algorithmic generation; Human-AI collaboration

1 The Theoretical Origin of Lens Aesthetics and the Technical Intervention of Generative AI

1.1 The Aesthetic Foundation of "Recording": Indexicality and the Myth of Total Cinema

The core of traditional film aesthetics lies in the indexicality of photographic images, which is the ontological identity between the image and the photographed object proposed by André Bazin in *The Ontology of the Photographic Image*. Bazin emphasized that photography's aesthetic value is to "reveal the truth" and preserve the reality of time, and this view constitutes the core proposition of traditional film aesthetics: the essence of film is the recording and restoration of physical reality.

A deeper connotation of Bazin's theory is that the "myth of total cinema" is the original driving force of film development—human's eternal desire to reproduce a complete and realistic illusion of the external world, and technology is only a secondary realization means. Based on this ontology, traditional lens language forms fixed aesthetic characteristics: the contingency of composition from the interception of real fragments, the authenticity of light and shadow from physical light sources, and the temporal and spatial continuity of long takes from the respect for real time and space, which Siegfried Kracauer summarized as the "recovery of physical reality".

1.2 The Technical Logic of "Generating": From Probabilistic Prediction to the World Model

The image generation logic of generative AI is essentially different from photographic recording. Mainstream models such as Sora adopt Transformer-based diffusion models, which are probabilistic generative models that analyze statistical laws in massive training data to generate video frame by frame or in short segments independently. This process is an active construction based on data distribution, not a passive recording of the real world.

AI-generated images are not simple material collage. Taking the Liaozhai-style model in CIVITAI as an example, the model can generate new visual content that has never appeared in the original data through algorithmic prediction, which means the core of "generation" is to predict the "future environment" based on the causal chain of environment and action. Runway's "General World Models" proposal points out that the ultimate goal of AI-generated images is to become a world simulator, where the physical, imaginary and virtual worlds overlap and coexist. This realizes the aesthetic shift of film images from "the truth captured by the lens" to "the truth calculated by algorithms".

2 The Paradigm Reconstruction of Film Aesthetics Driven by Generative AI

2.1 The Disenchantment and Reenchantment of Image Ontology

The intervention of generative AI has fundamentally deconstructed the indexicality of traditional photographic images. When images are no longer imprinted on photosensitive media by light, but "predicted" by algorithms, the

ontological connection between images and reality is cut off. Bazin's photographic image ontology is loosened: images are no longer the "mummy" of time, but the "illusion" synthesized by algorithms, and the traditional aesthetic demand of "representing reality" falls into aesthetic dislocation in the AI generation world.

However, AI images do not completely negate the ontological connotation of film images, but realize reenchantment in a new dimension. The "Platonic Representation Hypothesis" proposed by the MIT team points out that different AI models are moving towards a unified representational reality, meaning AI-generated images are constructing a new "myth of the complete world". What AI images depict is no longer the "happened reality", but the "possible reality" or "surreality", which expands the aesthetic connotation of film images from the reproduction of reality to the construction of reality.

2.2 The Migration and Game of Creative Subjects

Traditional film creation is a director-centered single subject model, with the camera as an extension of the human creator, forming an "I-it" binary relationship with the photographed object. Generative AI has broken this model, and the creative subject of film has evolved into a multi-subject pattern including "limited writers", "intelligent subjects" and "director-authors". In the text-to-video mode, human creators who provide prompt words are only "customizers" and "finalizers", and the real author of AI films is the human-AI hybrid subject formed by the cooperation of humans and intelligent models.

AI, as a "technical other", forces human creators to transform into Meta-Creators. The core function of creators has shifted from the executor of lens language to the designer of creative rules—they need to master prompt engineering to guide AI generation, and undertake the dual responsibilities of content screening and humanistic value injection. Human-AI collaboration has formed a symbiotic mode of "mutual domestication": AI provides a "possibility space" of visual materials through data drive, and humans endow AI-generated content with a meaning framework based on narrative logic.

3 The Formal Fission of Film Lens Language Under Generative AI

3.1 Composition Logic: From "Art of Subtraction" to "Art of Addition"

Traditional film composition is an art of subtraction—photographers intercept limited fragments from the real world through the viewfinder, and the frame is a clear boundary for reality interception, with the contingency of composition coming from the capture of real scenes.

AI image composition is an art of addition that constructs the entire visual scene from scratch, bringing two core changes: first, it presents an extreme stylization tendency, which can accurately imitate and strengthen specific visual styles such as Wes Anderson's symmetrical composition and cyberpunk neon tones; second, it can generate extreme perspectives beyond the human eye's physiological limits, such as drone perspective and macro perspective, expanding the boundary of film visual expression. More importantly, the generativity of AI makes the frame boundary permeable and fluid—Sora's continuous generated video replaces the traditional frame logic with generative domain logic, and images become a variable field continuously generated by algorithms.

3.2 Movement Grammar: From Physical Movement to Imaginative Movement

Traditional film lens movement is based on physical laws, with each movement such as push, pull and pan corresponding to specific mechanical devices, and the audience experiences the "physical presence" through the camera's physical movement.

AI-generated lens movement shows the paradoxical unity of physical continuity and visual collage. AI long takes are dynamic images formed by the collage of semantics, pixels and viewpoints, and the audience experiences a "changing scenes with steps" wandering perspective, which is close to the scattered perspective of traditional Chinese scroll painting. In addition, AI movement grammar follows the associative laws of the imaginary world, creating "impossible scheduling" that cannot be realized in reality—such as Sora 2's "latte art vortex transforming into an aerial coastline", which breaks the causal laws of the physical world and realizes the visual expression of surreal movement.

3.3 Light and Shadow Texture: From Physical Lighting to Stylized Rendering

Traditional film light and shadow is based on the actual existence of physical light sources, and photographers shape space and depict characters by controlling the intensity, direction and texture of natural or artificial light, with the authenticity of light and shadow coming from its correspondence with the physical world.

AI image light and shadow has the characteristic of stylized rendering—creators can accurately control light effects such as golden hour rim light and volumetric light through semantic description of prompt words, without relying on the actual arrangement of physical light sources. AI can not only imitate the light and shadow styles of classic films such as film noir's high-contrast lighting and Hollywood Golden Age's soft light, but also present a stylistic purity beyond physical possibilities: the "golden hour" light that is difficult to maintain in reality can be infinitely extended in AI images, and multiple conflicting light source effects can be seamlessly integrated.

3.4 Montage Logic: From Editing Splicing to Continuous Generation

Traditional montage is based on the break and stitching between lenses, and edit points, transitions and rhythm constitute the core of film time aesthetics, whether it is Eisenstein's "conflict montage" or Bazin's long take aesthetics, they all presuppose the editability of lenses as basic units.

AI images have subverted the traditional montage logic in two aspects: first, AI models tend to generate long takes with temporal and spatial continuity because of their continuous prediction generation logic, and montages emphasizing lens discontinuity are relatively rare; second, AI adopts "morph transition" to realize lens conversion—one scene smoothly transforms into another, making the traditional edit points blurred or even disappear. The rhythm of film has thus shifted from the "rhythm of editing" created by discrete lens units to the "rhythm of generation" formed by the continuous flow of algorithmic generation.

4 Aesthetic Reflection: Crises and Opportunities of AI-Generated Film Images

4.1 The Crisis of AI Image Aesthetics: Hyperreality and Emotional Defocus

AI-generated images are a typical form of hyperreality proposed by Jean Baudrillard—they are more realistic than real shots in physical details, but lose the ontological connection with reality, forming a "paradox of truth". When the audience realizes that the "reality" in the image is only an algorithmic fiction, the depth and authenticity of emotional investment are questioned, resulting in emotional defocus.

In addition, the technical defects of AI-generated images further damage the viewing immersion: the accidental breakdown of physical laws, the minor disorder of detail logic, and especially the Character Drift problem (the inconsistency of character appearance in different lenses) destroy the character identity established by traditional films, which has become a major challenge for professional AI film production.

4.2 The Liberation of AI Image Aesthetics: Breaking Material Barriers and Visual Imagination

Generative AI has opened up an unprecedented possibility space for film aesthetics by breaking the material and technical barriers of traditional film creation. Expensive location shooting, complex special effects production and large-scale mass actor scheduling in traditional film can be realized at low cost through prompt words, enabling independent and experimental films to obtain the visual expressiveness that only major film studios could achieve in the past.

AI also realizes the embodiment of subconscious images—styles such as surrealism and stream of consciousness that rely on symbolic techniques in traditional film can be directly transformed into visual images by AI, and dreams, illusions and memories can be presented intuitively, showing a de-anthropocentric artistic imagination. In addition, the combination of AI real-time generation technology and extended reality (XR) technology makes film images overflow the formal boundary, and the audience changes from passive viewers to participants of image generation, expanding the connotation of "filmicity" itself.

4.3 The Development Prospect: Human-AI Collaborative Creation

The most likely development direction of future film is the hybrid form of recorded images and generated images. Professional AI film and television workflows such as Utopai Studios are exploring a "Planning-Rendering decoupling" architecture: the upper sequence planner acts as the "director's brain" responsible for narrative planning and consistency constraints, and the lower conditional renderer focuses on high-quality image generation. This architecture defines the core of human-AI collaboration: humans are responsible for creativity, planning and value judgment, and AI is responsible for execution, optimization and effect presentation.

Under the framework of human-AI collaboration, the subjectivity of human creators is not eliminated, but reconstructed and enhanced. The transformation from "creator" to "Meta-Creator" puts forward new requirements for the creator's ability structure—they need to integrate artistic intuition, narrative ability and AI technical logic control ability, and become a "technology-art complex" that is neither a vassal of technology nor an isolated island of humanism.

5 Conclusion

The intervention of generative AI has realized a fundamental paradigm shift in film lens aesthetics from "recording" to "generating", which is reflected in three interrelated levels: at the ontological level, the traditional photographic ontology centered on indexicality is replaced by the algorithmic generation theory, and the film image has shifted from the reproduction of "happened reality" to the construction of "possible reality"; at the subject level, the single human creator has evolved into a human-AI hybrid subject, and human creators have completed the transformation to Meta-Creators, forming a human-AI symbiotic mode of mutual domestication; at the formal level, the film lens language has undergone comprehensive fission, with composition, movement, light and shadow and montage forming new aesthetic rules driven by generative AI technology.

Generative AI brings both crises and opportunities to film aesthetics: the hyperreality of AI images may lead to emotional defocus of the audience, and technical defects such as Character Drift restrict the development of professional production; on the other hand, it breaks the material barriers of film creation, releases the visual imagination of surrealism, and expands the boundary of film art. The future of AI film aesthetics lies in the dynamic balance of human-AI collaboration—abandoning both the blind defense of traditional film aesthetics and the blind pursuit of technical possibilities, and rediscovering and confirming the potential of human creation in the technological revolution. Only by anchoring humanistic values in the algorithmic logic and integrating artistic innovation with technical development can we truly explore the unique aesthetic connotation of AI-generated film images and promote the innovative development of film art in the AI era.

About the Author

Fang Cao, female, Han ethnicity, Nanjing, Jiangsu Province, Master's degree. Associate Professor at the School of Cultural Management, Communication University of China, Nanjing. Research interests: Digital Media Arts.

References

- [1] Wang, Y. W. (2026). Collage in the long take: Traditional Chinese scroll painting in AIGC films. *Film Art*, (1), 110-118.
- [2] Xu, L. J., Zhao, Z. Y., & Li, M. (2025). Research on narrative strategies and cultural identity construction of AIGC films. *Film Literature*, (15), 46-50.
- [3] Chen, X. G., & Du, T. X. (2025). From the "myth of total cinema" to the "world vision" of AI-generated images: Image ontology of film industrial aesthetics in the artificial intelligence era. *Ethnic Art Studies*.
- [4] Xu, J. (2025). The subjectivity status and aesthetic criticism of artificial intelligence films. *China Literary and Art Criticism*, (7).
- [5] Kang, Q. (2025). The deconstruction and reconstruction of the creative subject of film and television photography by generative AI: A practical reflection based on Midjourney generating film storyboards. *Journal of News and Communication Science*, 13(5), 749-753.
- [6] Huang, K., et al. (2025). FilMaster: Bridging Cinematic Principles and Generative AI for Automated Film Generation. *arXiv preprint arXiv: 2506.18899*.
- [7] Bazin, A. (2005). *What is Cinema?* (Vol. 1). University of California Press.
- [8] Kracauer, S. (1960). *Theory of Film: The Redemption of Physical Reality*. Oxford University Press.