Available on: https://jsshrs.ep-journals.org/ E-ISSN: 3051-3561, P-ISSN: 3051-3553

Journal of Social Science and Human Research Studies

Volume : 02 Issue : 10 October-2026

Revisiting the Uncanny Valley in the AI Era: A Review of Visual Artifacts and Aesthetic Perceptions in AI-Generated Characters

Muhan Li

No. 28293, Jingshi Road, Huaiyin District, Jinan City, Shandong Province, 250117, China

Published Date: October 31, 2026 **Article DOI: 10.65150/EP-jsshrs/V2E10/2026-18**

ABSTRACT: The rapid evolution of generative artificial intelligence (AI), particularly diffusion-based image synthesis models, has fundamentally transformed digital character design and visual storytelling. However, the frequent occurrence of "visual artifacts"—such as anatomical distortions, unnatural skin smoothness, rigid facial expressions, and latent space blending—presents significant aesthetic and psychological challenges. This paper provides a critical, multidisciplinary review that re-examines the classic "Uncanny Valley" hypothesis within the contemporary context of AI-generated media. Moving beyond the conventional engineering perspective that views these artifacts merely as technical deficiencies to be eliminated, this study investigates their potential reinterpretation through the lenses of post-digital media art, glitch aesthetics, and visual semiotics. We analyze how audience perception, cognitive friction, and emotional engagement are affected by these synthetic cues. Furthermore, we explore how human designers utilize advanced node-based control frameworks (e.g., ComfyUI, ControlNet) to negotiate, fix, or intentionally embrace these algorithmic anomalies. Finally, we establish a conceptual framework for human-AI co-creation and highlight future research trajectories for character narrative design in generative media.

KEYWORDS: Generative AI, Uncanny Valley, Visual Artifacts, Digital Character Design, Post-Digital Aesthetics, Visual Storytelling, Human-AI Collaboration.

INTRODUCTION

Digital character design has long functioned at the intersection of technological advancement and artistic expression. From early 3D computer graphics (CG) in cinematic visual effects to modern real-time rendering in interactive game engines, the pursuit of realistic and emotionally expressive characters has driven media technology forward (Catmull, 1978; Prince, 2011). In recent years, the emergence of deep generative models—specifically Generative Adversarial Networks (GANs) and Latent Diffusion Models (LDMs)—has catalyzed a paradigm shift in visual asset generation (Goodfellow et al., 2014; Rombach et al., 2022). Tools built upon diffusion architectures, such as Stable Diffusion, Midjourney, and node-based pipeline environments like ComfyUI, allow artists to generate complex character concepts rapidly through textual prompts and structural conditioning controls (Zhang, Rao, & Agrawala, 2023).

Despite these remarkable computational capabilities, AI-generated character outputs routinely exhibit distinct "visual artifacts." These range from structural inaccuracies—such as distorted hands, extra limbs, and inconsistent eye gaze—to subtle texture anomalies, including unnatural skin glossiness, lack of sub-surface scattering, and strange latent space blendings where clothing merges seamlessly with body anatomy (Borji, 2023). While engineering research focuses primarily on reducing these artifacts through model architecture scaling, higher-quality dataset curation, and fine-tuning methods like LoRA (Low-Rank Adaptation) (Hu et al., 2021), these visual flaws remain prevalent in current production workflows.

From a human-computer interaction (HCI) and visual arts perspective, these visual artifacts are not merely computational errors; they are distinctive perceptual triggers that fundamentally alter human emotional engagement and aesthetic interpretation. This paper addresses three primary research objectives:

1. To re-evaluate the Uncanny Valley hypothesis (Mori, 1970) in light of current generative AI media, identifying how "algorithmic uncanniness" differs from traditional 3D CGI or robotic non-realism.
2. To contextualize AI artifacts through Post-Digital Aesthetics (Cascone, 2000; Pepperell, 2003), examining how technical flaws can be re-appropriated as novel visual style markers and surrealist storytelling devices.
3. To formulate a conceptual framework for Human-AI Collaborative Design, illustrating how artists navigate these artifacts using controllable node-based interfaces (e.g., ComfyUI) to balance technological efficiency with creative autonomy.

THE UNCANNY VALLEY RE-EXAMINED IN THE AGE OF GENERATIVE AI

The "Uncanny Valley" hypothesis, originally formulated by Japanese roboticist Masahiro Mori (1970), posits that as a humanlike entity's affinity and realism increase, human emotional response becomes increasingly positive until a specific threshold is reached. At this critical juncture, minute imperfections in appearance or movement provoke profound feelings of eerie revulsion, cognitive dissonance, and existential discomfort among

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

observers (Mori, MacDorman, & Kageki, 2012). Historically, this phenomenon was analyzed within robotics and 3D computer-generated imagery (CGI), where physical movement mechanics (e.g., rigid micro-expressions) or rendering artifacts (e.g., dead eye phenomenon, incorrect skin subsurface scattering) caused the immersion failure (Seyama & Nagai, 2007; MacDorman & Ishiguro, 2006).

In the contemporary landscape of generative AI, particularly deep learning models operating within implicit high-dimensional space (Latent Space), the nature of this aesthetic friction undergoes a fundamental transformation. Unlike 3D CGI, which constructs digital assets through explicitly defined geometric topology, rigging hierarchies, and physical lighting calculations, diffusion-based AI synthesizes visual output through probabilistic pattern matching and noise reduction (Rombach et al., 2022). Consequently, the resulting "Algorithmic Uncanny" does not stem from mechanical stiffness or low polygon rendering, but rather from local semantic incoherence, probabilistic hallucination, and topological anatomical violations (Borji, 2023; Kätsyri et al., 2015).

Table 1. Comparative Analysis of Classical Uncanny Valley vs. Algorithmic Uncanny in Generative AI

Dimension	Classical Uncanny Valley (3D CGI/Robots)	Algorithmic Uncanny (Generative AI)
Primary Source Trigger	Rigid facial movements Artificial eye movements and blink patterns Surface skin without subsurface scattering	Local anatomical hallucinations Latent space texture bleeding/blending Structural impossibility (e.g., 6 fingers)
Underlying Technology	Geometric mesh modeling, manual rigging, keyframe animation, physics rendering.	Latent Diffusion Models (LDMs), GANs, probabilistic pixel de-noising.
Psychological Friction	Violation of physical movement expectation and biological realism.	Violation of spatial-semantic coherence and perceptual category boundaries.
Perception Characteristics	Global consistency with local mechanical flaw (looks real until it moves).	Local details broken within a globally hyper-realistic photo/art style.
Primary Mitigation Method	Advanced blendshapes, facial mocap, subsurface scattering (SSS) shaders.	Node-based conditioning (ComfyUI/ControlNet) Inpainting, LoRA fine-tuning, IP-Adapter.

The psychological discomfort induced by AI-generated character artifacts can be explained through two complementary theoretical frameworks: Perceptual Mismatch Theory (Kätsyri et al., 2015) and Category Ambiguity (Burleigh, Schoenherr, & Lacroix, 2013).

When viewing an AI-generated character produced by state-of-the-art diffusion models, the human visual system is presented with hyper-realistic surface renders—such as realistic pore structures, filmic lighting, and intricate hair details. However, this global realism often coexists with subtle, local spatial violations, such as an extra digit on a hand, asymmetry in pupil dilation, or a garment line dissolving seamlessly into the character's skin tissue. According to Perceptual Mismatch Theory, this sharp inconsistency between high-level photo-realism and low-level anatomical impossibility causes severe cognitive friction in human visual processing networks (Kätsyri et al., 2015). The human brain experiences a conflict between the automatic "human detection" module and the "error-sensing" visual cortex, resulting in a sudden drop in aesthetic affinity—a digital manifestation of Ernst Jentsch's classical concept of *Unheimlich*(the unhomely or eerie) (Jentsch, 1906; Freud, 1919).

Furthermore, in node-based workflows like ComfyUI, when artists blend multiple conditioning inputs (e.g., mixing pose estimations via ControlNet with face-swapping nodes via IP-Adapter), the output frequently oscillates across category boundaries—neither purely human nor clearly stylized artifact. This category ambiguity forces the audience into an active state of cognitive re-evaluation, disrupting immersion and heightening the perception of algorithmic presence (Ferrey et al., 2015).

DECONSTRUCTING AI VISUAL ARTIFACTS: FROM TECHNICAL FLAWS TO AESTHETIC MARKERS

In the domain of generative media, visual artifacts are commonly defined as unintended digital noise, spatial distortions, or semantic hallucinations introduced by algorithmic processing (Borji, 2023). Within diffusion-based character generation workflows (e.g., Stable Diffusion, Midjourney, or custom ComfyUI node setups), these artifacts exhibit specific visual patterns attributable to the underlying latent space operations.

Latent Diffusion Models (LDMs) compress image data into lower-dimensional latent representations, applying iterative denoising guided by text prompts and conditioning vectors (Rombach et al., 2022). When cross-attention mechanisms fail to align complex spatial semantics with high-dimensional prompt embeddings, structural anomalies manifest. These range from anatomical violations (e.g., polydactyly, distorted limb connections) to surface-level rendering anomalies (e.g., plastic skin gloss, unnatural subsurface light absorption, and latent texture bleeding where garments fuse with body tissues) (Borji, 2023; Zhang et al., 2023).

Table 2. Taxonomy of AI Visual Artifacts, Technical Roots, and Post-Digital Reinterpretation

Artifact Type	Technical / Latent Root Cause	Post-Digital Aesthetic & Narrative Reinterpretation
Anatomical Distortion (e.g., Extra digits, limb fusing)	Attention map misalignment; lack of 3D structural priors in 2D LDM training datasets.	Symbol of post-human body mutation, alienness, or psychological instability in cyberpunk/surrealist narratives.
Synthetic Plasticity (Unnatural skin texture & gloss)	Over-smoothing in VAE decoding; loss of high-frequency spatial skin detail during latent compression.	Metaphor for hyper-commercialization, artificiality, or virtual identity in dystopian/cybernetic character design.
Latent Texture Bleeding (Clothing & skin fusion)	Cross-attention leakage across token embeddings (e.g., "fabric" merging with "skin" latent vectors)	Glitch aesthetic; visual representation of digital entropy, memory corruption, or

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Corresponding Author: Muhan Li

		boundary dissolution between virtual entity and environment.
Asymmetric / Gaze Disconnect (Eye gaze misalignment)	Inconsistent ocular feature extraction during noise reduction; lack of gaze-vector constraints.	Expression of cognitive alienation, uncanny distress, or possession by algorithmic/artificial entities.

While conventional computer graphics engineering treats artifacts as errors to be removed via higher-resolution training datasets, LoRA fine-tuning, or ControlNet pose constraints (Hu et al., 2021; Zhang et al., 2023), the field of Post-Digital Media Art offers an alternative critical paradigm.

Coined by Kim Cascone (2000), "post-digital aesthetics" refers to an artistic stance that no longer fetishizes digital perfection, but rather highlights the failures, glitches, and material traces of computational media. In music and visual media, this manifested as *Glitch Art*—where digital noise and compression artifacts were elevated to stylistic signifiers of technological transparency (Cascone, 2000; Pepperell, 2003).

When applied to AI-generated character design, this framework re-contextualizes the Uncanny Valley. Rather than striving for absolute photorealism (which often results in banal or uncanny representations), avant-garde media artists embrace algorithmic artifacts as authentic stylistic markers of human-AI hybridity.

For instance, the "plastic smoothness" of AI faces or the surreal fusion of biological skin with metallic garments (Table 2) can be weaponized in character narrative design to evoke themes of post-humanism, algorithmic surveillance, and virtual alienation (Haraway, 1991; Hayles, 1999). In this context, the visual artifact shifts from a technical deficiency to an intentional semiotic tool that subverts traditional linear storytelling.

THE DOUBLE-EDGED SWORD IN CHARACTER NARRATIVE AND EMOTION

While post-digital media artists reframe AI artifacts as novel stylistic elements, these visual anomalies present substantial risks in commercial character animation and narrative filmmaking. According to Narrative Transportation Theory (Green & Brock, 2000), story world engagement relies on the audience maintaining an uninterrupted state of cognitive immersion. When an AI-generated character displays an unintentional visual artifact—such as a morphing finger during a dramatic gesture or pupil asymmetry during an emotional close-up—it triggers a perceptual disruption known as "break of immersion" (BOI) (McGloin, Farrar, & Krcmar, 2011).

Furthermore, the overuse of unrefined generative models introduces a critical phenomenon termed "Identity Dilution."Because Latent Diffusion Models sample from millions of web-scale visual pairs during noise reduction, generated characters often suffer from a homogenization of facial features, lighting, and pose conventions—often termed the "AI homogeneous look" or "prompt-driven genericism" (Borji, 2023). This uniformity diminishes character distinctiveness,impeding long-term audience empathy and character recognition (MacDorman & Chattopadhyay, 2016).

Conversely, when intentionally orchestrated by designers, algorithmic artifacts can function as powerful visual metaphors. In genre-specific contexts—such as science fiction, psychological horror, dystopian cyber-fiction, and surrealist animation—the inherent uncanny quality of AI artifacts aligns with narrative themes of memory loss, artificial consciousness, and post-human transformation (Haraway, 1991).

For instance, in character designs depicting virtual entities, Androids, or psychologically fractured protagonists, retaining subtle structural distortions or latent texture bleedings serves as a visual signifier of internal conflict or computational glitching (Cascone, 2000). By framing algorithmic flaws as thematic devices, creators transform the Uncanny Valley from an obstacle into an expressive visual language.

HUMAN-AI COLLABORATIVE STRATEGIES AND FUTURE TRENDS

To navigate the balance between generative efficiency and precise artistic control, contemporary character production pipelines are shifting toward node-based conditioning architectures. Environments such as ComfyUI allow designers to isolate specific stages of the latent diffusion process, applying localized constraints through tools like ControlNet (for pose and edge preservation), IP-Adapter (for character identity locking), and Region-Prompt Inpainting (Zhang et al., 2023; Hu et al., 2021).

By replacing single-text-prompt generation with modular node graphs, designers gain fine-grained agency over visual artifacts. Artists can explicitly choose which latent anomalies to eliminate through targeted image-to-image refinement and which to preserve for stylistic expression.

Table 3. Evolutionary Trajectory of Generative AI in Character Design Pipelines and Human-AI Agency

Epoch / Paradigm	Core Technology Architecture	Character Design Workflow & Human-AI Agency
Era 1: Early GANs (2014 - 2020)	Deep Convolutional GANs (DCGAN) , StyleGAN (Goodfellow et al., 2014)	-Low-resolution, black-box image generation. - Zero spatial control; passive selection from random latent space iterations.
Era 2: Prompt-Driven Diffusion (2021-2022)	Latent Diffusion Models (LDM), Midjourney, Stable Diffusion 1.5 (Rombach et al., 2022)	- High-resolution, text-to-image synthesis. - Text-prompt steering; frequent artifact generation due to lack of structural spatial control.
Era 3: Node-Based Controllable AI (2023 - Present)	LDM + ControlNet, IP-Adapter, Inpainting, LoRA adaptation (Zhang et al., 2023; Hu et al., 2021)	- Modular node-graph workflows (e.g., ComfyUI). - High human agency: precise control over pose, identity and selective preservation/repair of artifacts.
Era 4: Future 3D/4D Generative Pipelines (Emerging)	Text/Image-to-3D, Gaussian Splatting, NeRF-based motion dynamics (Poole et al., 2022)	- Direct generation of animatable 3D rigged meshes with embedded physics and neural subsurface scattering shaders to eliminate uncanny artifacts.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Corresponding Author: Muhan Li

As illustrated in Table 3, the field is transitioning toward 3D-aware generative frameworks and real-time interactive character systems. Key future research directions include:

1. Neural Subsurface Scattering and Skin Shading: Integrating physical rendering equations into latent diffusion decoders to eliminate plastic skin artifacts at the model level.
2. Cognitive Ergonomics in AI Interface Design: Studying how node-based AI interfaces (such as ComfyUI) affect designers' cognitive load, creative autonomy, and decision-making speed (HCI perspective).
3. Cross-Media Character Identity Consistency: Establishing standardized LoRA and IP-Adapter frameworks that guarantee 3D-consistent character identity across sequential storyboards and animation frames.

CONCLUSION

This paper has provided a systematic and critical re-examination of the Uncanny Valley hypothesis within the evolving landscape of AI-generated digital character design. Moving beyond traditional engineering paradigms that reduce generative visual artifacts to mere technical bugs awaiting elimination, this review has articulated a nuanced dual-perspective framework grounded in post-digital aesthetics, visual semiotics, and human-computer interaction (HCI). We have demonstrated that while stochastic algorithmic errors—such as topological anatomical distortions, VAE over-smoothing, and cross-attention texture bleeding—severely disrupt cognitive immersion and dilute character identity in mainstream commercial media, they simultaneously possess deep expressive potential. When strategically wielded by designers, these algorithmic uncanny cues function as poignant visual metaphors for post-human condition, algorithmic surveillance, digital memory degradation, and psychological alienation in avant-garde narrative filmmaking and surrealist media art.

The theoretical and practical insights synthesized in this study yield several critical implications for digital asset production. As character generation transitions from unconstrained prompt-based synthesis to node-based controllable pipelines (e.g., ComfyUI, ControlNet, IP-Adapter) and neural 3D assets (e.g., Gaussian Splatting, NeRFs), the role of the human designer is fundamentally elevated from a passive prompt creator to a high-agency aesthetic orchestrator. Controllable AI architectures allow artists to dynamically negotiate with algorithmic unpredictability—selectively suppressing artifacts to safeguard narrative immersion or intentionally amplifying them as unique style markers.

Nevertheless, this review acknowledges certain methodological limitations. Given the hyper-rapid velocity of generative AI development, our qualitative taxonomy primarily reflects the current capabilities and constraints of 2D latent diffusion models and early-stage 3D generative frameworks. Quantitative empirical studies measuring audience eye-tracking responses, galvanic skin resistance (GSR), and electroencephalography (EEG) during exposure to specific AI visual artifacts remain scarce and warrant deeper investigation.

Building upon these insights, future research should focus on three promising trajectories:

1. Neural Rendering Integration: Incorporating physical rendering equations and neural subsurface scattering directly into diffusion decoders to eradicate synthetic skin glossiness at the architectural level.
2. Cognitive Ergonomics in Human-AI Workflows: Investigating how node-based generative interfaces influence designers' mental workload, creative autonomy, and decision-making efficiency within high-pressure studio production environments.
3. Cross-Media Narrative Consistency: Developing standardized LoRA and identity-locking protocols that preserve 3D-consistent character identity and intentional artifact stylization across sequential animation frames and interactive media.

Ultimately, bridging the gap between computational automation and artistic intentionality ensures that generative AI operates not as a homogenizing constraint, but as an empowering catalyst for human creative autonomy in the age of synthetic media.

REFERENCES

1. Borji, A. (2023). A categorical archive of ChatGPT failures. *arXiv preprint arXiv:2302.03494*.
2. Burleigh, T. J., Schoenherr, J. R., & Lacroix, G. L. (2013). Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Computers in Human Behavior*, 29(3), 759-771.
3. Cascone, K. (2000). The aesthetics of failure: "Post-digital" tendencies in contemporary computer music. *Computer Music Journal*, 24(4), 12-18.
4. Catmull, E. (1978). A hidden-surface algorithm with anti-aliasing. *ACM SIGGRAPH Computer Graphics*, 12(3), 6-11.
5. Ferrey, A. E., Burleigh, T. J., & Fenske, M. J. (2015). Stimulus-category competition, inhibition, and affective devaluation: A novel account of the uncanny valley. *Frontiers in Psychology*, 6, 249.
6. Freud, S. (1919). Das Unheimliche [The Uncanny]. *Imago*, 5(5), 297-324.
7. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems (NeurIPS)*, 27, 2672-2680.
8. Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, 79(5), 701-721.
9. Haraway, D. (1991). A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. *Simians, Cyborgs, and Women: The Reinvention of Nature*, 149-181. Routledge.
10. Hayles, N. K. (1999). *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. University of Chicago Press.
11. Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., ... & Chen, W. (2021). Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*.
12. Jentsch, E. (1906). On the psychology of the uncanny. *Psychiatrisch-Neurologische Wochenschrift*, 8(22), 195-198.
13. Kätsyri, J., Förger, K., Mäkäräinen, M., & Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology*, 6, 390.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Corresponding Author: Muhan Li

14. MacDorman, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies*, 7(3), 297–337.
15. McGloin, R., Farrar, K. M., & Krcmar, M. (2011). The impact of controller naturalness on spatial presence, gamer enjoyment, and perceived realism in a tennis simulation video game. *Presence: Teleoperators and Virtual Environments*, 20(4), 309–324.
16. Mori, M. (1970). Bukimi no tani [The uncanny valley]. *Energy*, 7(4), 33-35.
17. Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2), 98-100.
18. Pepperell, R. (2003). *The Posthuman Condition: Consciousness Beyond the Brain*. Intellect Books.
19. Poole, B., Jain, A., Barron, J. T., & Mildenhall, B. (2022). Dreamfusion: Text-to-3d using 2d diffusion. *arXiv preprint arXiv:2209.14988*.
20. Prince, S. (2011). *Digital Visual Effects in Cinema: The Seduction of Reality*. Rutgers University Press.
21. MacDorman, K. F., & Chattopadhyay, D. (2016). Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition*, 146, 190–205.
22. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10684-10695.
23. Seyama, J., & Nagai, R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial faces. *Presence: Teleoperators and Virtual Environments*, 16(4), 337-351.
24. Zhang, L., Rao, A., & Agrawala, M. (2023). Adding conditional control to text-to-image diffusion models. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 3836-3847.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Corresponding Author: Muhan Li