Media Aesthetics and Communication Effectiveness of Immersive Animated Narrative in Broadcast Television Driven by Generative Artificial Intelligence (GAI)

Long Wang[1,*], Di Wu[1], Hailong Li[1]
[1]School of Design and Art, Shandong Huayu University of Technology,Dezhou 253034, China
Corresponding author: Long Wang (e-mail: sdhygxywl@163.com).

ABSTRACT There are fundamental issues in broadcast TV immersive animated narrative: low usage of technology, low involvement of the audience, and weak aesthetic effect. By applying a hybrid research approach with Generative Artificial Intelligence (GAI) technology, this study develops a three-dimensional perception-cognition-emotion media aesthetic analysis framework for the production of "Narrative Innovation Paths" of "perception-cognition-emotion". Demand data were gathered from 32 creators and 156 audience members through in-depth interviews, and grounded theory approach was used to extract the encoding of aesthetic elements. Second, it develops a GAI collaborative creation model including visual style transfer, dynamic scene generation and interactive scene branches, and carries out an 8-month experimental deployment in 6 broadcasting stations. Lastly, it employed eye-tracking technology and neuroimaging techniques to measure immersive experience factors. The results showed that the average immersion time of users in the experimental group was 49.0 minutes, an increase of 36.9% compared with the control group; the peak emotional arousal intensity increased by 2.78 times; and the final conversion rate effect size in the content sharing willingness test reached 23.15. This study verifies the significant value of GAI technology in reconstructing narrative space-time, enhancing aesthetic immersion, and optimizing communication effectiveness, providing theoretical support and practical guidance for the transformation of broadcast and television media.

INDEX TERMS Generative Artificial Intelligence; Immersive Narrative; Media Aesthetics; Communication Effectiveness; Animation Creation; Interactive Experience

I. INTRODUCTION

IN the wave of digital transformation, radio and television media are facing a deep demand for content production mode transformation. Immersive animation narrative, as a key path to enhance user stickiness and expand communication efficiency, always has structural tension in its technical implementation and aesthetic construction. Traditional animation creation relies on a linear process of manual drawing and post production synthesis, with lengthy production cycles and high costs, making it difficult to meet the rapid iteration needs of personalized content; At the same time, existing technology fusion solutions mostly remain at the shallow level of visual effects overlay, failing to touch on the essential mechanisms of narrative spatiotemporal reconstruction and emotional immersion generation. From a methodological perspective, the existing research poses three dilemmas: perceptual dimensions are typically measured with self-assessment scales, which are subjected to biases; the measurement of cognitive processes does not employ real-time support from the neurophysiological processes, and it cannot explain the dynamic mechanisms of allocation of attention and encoding of memory; the evaluation of emotional experience is performed after the interview, and the time delay introduces distortion into emotional memory. The deeper challenge is the disconnection between the aesthetic theoretical model and the technological instruments: until now, media aesthetics research has been dominated by the paradigm of the analysis of the text; the technological development, by the optimization of the model performance, neglecting the cultural roots and situational dependence of aesthetic perception. The theoretical poverty and the technological recklessness are a restriction to the qualitative break of immersive narrative. In response to the methodological deficiencies mentioned above, this study constructs an interdisciplinary integrated technical aesthetic analysis system. Through grounded theory, open coding, axial coding, and selective coding are implemented to extract categories from in-depth interview texts, establishing an aesthetic element classification system that includes visual rhythm, spatial topology, and emotional tone; This article designs a GAI collaborative creation model based on the system, integrating the Stable Diffusion style transfer module, NeRF dynamic scene rendering engine, and reinforcement learning based interactive branch generator, and conducting long-term experiments in the actual production environment of broadcasting and television stations; This article uses an eye tracker to capture changes in pupil diameter and gaze trajectory drift to quantify attention intensity, and uses functional near-infrared spectroscopy technology to monitor changes in oxygenated hemoglobin concentration in the prefrontal cortex to characterize cognitive load. It can record the changes of emotional arousal in real time based on the use of skin electric response sensors, and it can build the immersive model of experience evaluation of emotional arousal from multi-modal physiological indicators. This article establishes a three-dimensional media aesthetic model framework of "perception cognition emotion", which fills the conceptual gap between technology paradigm and humanistic interpretation, provides a technical implementation path for broadcasting and television media transformation, and establishes the standards for aesthetic evaluation, which helps the GAI media to develop from an auxiliary tool to a creative subject.

II. RELATED WODR

Media aesthetics has accumulated rich theoretical resources and empirical cases in the fields of cross media narrative, digital cultural production, and audience experience reconstruction. However, the deep impact of technological intervention on aesthetic mechanisms still needs to be systematically explored. Based on cross media narrative theory and visual cultural analysis methods, Zhang et al. [1] explored "A Record of a Mortal's Journey to Immortality" from the dimensions of spatial creation and brushwork charm. Dynamic shots simulated scattered perspective to achieve visual flow, and landscape imagery was integrated into narrative logic to promote plot development. Starting from the intersection of visual anthropology and narrative semiotics, Wang [2] analyzed the mechanism of the cross media intertextual system in "Blossoms" from three dimensions. Mao [3] compared the sculpture of Kang Jinlong in the Jade Emperor Temple with the image of Kang Jinlong in "Black Myth: Wukong", and studied the differences and interactions in aesthetic expression between traditional sculpture and modern games. Zhao and Zhang [4] believed that the continuous evolution of media technology not only updated the characteristics of new media aesthetics, but also deepens the contemporary "digital divide". Huang and Lu [5] studied James' illustrated novels from the perspective of media aesthetics, aiming to present the author's narrative in a media form, reveal the commonalities between media, and analyze the aesthetic ideas behind the media process to readers. Merkulova and Pryshchenko [6] presented their research findings on the stylistics and visual imagery of digital media, a new type of communication medium. Laughter et al. [7] aimed to summarize the relevant evidence on the cognition of beauty, aesthetic cultural factors, and the influence of social media, especially on the clinical manifestations of BDD. Gambetti and Han [8] aimed to explore and analyze the aesthetic differences in food presented by different types of restaurants on social media using a large-scale food image dataset. Fernandez et al. [9] constructed a theoretical model for CSR social media communication, which considers its unique characteristics - sensitivity to peer influence through normative activation, interactivity, viral CSR information dissemination through sharing, CSR empowerment, and people-centered demands in CSR posts. Ferreira et al. [10] aimed to explore the yet to be fully researched interactive relationship between fans and internet celebrities, as well as its impact on consumers' reactions to brand recommendations from internet celebrities, in order to enrich the understanding of the driving factors of quasi social interaction among internet celebrities. Existing research has mostly focused on static text interpretation or single media form analysis, lacking quantitative verification of the GAI driven dynamic narrative generation mechanism at the neurocognitive level. The dialogue between technical tools and aesthetic theory has not yet been established. This article attempts to bridge the explanatory gap between algorithm logic and immersive experience through multimodal physiological measurements and long-term experimental deployment.

III. METHOD

A. RESEARCH DESIGN AND DATA COLLECTION PROCESS

The framework is a mixed method which is a combination of qualitative and quantitative method, and the source data is a dual method of creator perspective and audience experience. It interviewed 32 radio and television animation makers over a 45-60 minute period using semi-structured in-depth interviews, covering the barriers to use and the functions that are expected of generative AI tools for use in the construction of narratives, visual expression, and emotion transmission. The interview recording was transcribed using speech recognition, and key parts in these discourses were identified and time stamped for the creation of an original corpus of technical requirements, aesthetic evaluations, and recommendations for the work flow. The audience survey surveyed 156 viewers (ages 18-55), asking them to complete an online questionnaire comprising 25 items across three modules: Visual Style Preference Scale (a 7-level Likert scale), Interaction Operation Willingness Measurement (multiple choice questions) and Content Type Attractiveness Sorting (drag and drop sorting). Also to assure the authenticity of the data, the questionnaire includes 3 attention test questions, and invalid samples that are answered within 180 seconds or less and/or wrong on the attention test questions are deleted. All the interview texts and responses to the questionnaires have been entered into NVivo qualitative analysis software and SPSS statistical platform. The resulting structured data set consisted of scores for the aesthetic preferences dimensions, the frequency of the aesthetic function of the content themes, and the content themes' acceptance rankings, after the data had been cleaned, missing values filled and outliers detected. It has 74 variable fields, providing a basic materials in coding analysis.

B. ENCODING AESTHETIC ELEMENTS BASED ON GROUNDED THEORY

Open-source coding analysis was used to analyze 74 variable fields in a structured requirements dataset and to extract raw concept labels from interviews. 218 basic coding nodes were created for expressions such as "color saturation adjustment requirements" and "lens smoothing motion evaluation." High-frequency correlation patterns were identified by calculating the co-occurrence frequency matrix between codes. The main axis encoding combines the first nodes into three nodes: visual layer, narrative layer, emotional layer and creates a causal relationship model between these layers. The style consistency and the detail authenticity of GAI-generated images are the basis for the visual layer, the logical rationality and the rhythm control of the plot branches are the basis for the narrative layer, and the intensity and depth of the awakening of immersive experiences are the basis for the emotional layer. Selective coding extracts the core categories of "technology empowerment aesthetic reconstruction experience dimension enhancement" and calculates the importance weights of each aesthetic element. The weight is calculated using the TF-IDF improved algorithm, and the formula is:

$$W_i = f_i \times \log\left(\frac{N}{n_i}\right) \times \alpha_i \tag{1}$$

fi is the frequency of appearance of the element, N is the total number of samples, ni is the number of samples containing the element, and αi is the average importance coefficient assigned by the creator and audience. The correlation strength is measured using Pearson correlation coefficient:

$$r_{xy} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \tag{2}$$

The coding results are summarized in Table 1, presenting the quantitative characteristics of each aesthetic element.

TABLE 1. Coding Statistics of Media Aesthetics Elements

Aesthetic Element Category	Number of Coding Nodes	Occurrence Frequency	Importance Weight	Correlation Strength Coefficient
Visual Layer–Style Transfer Quality	47	312	0.89	0.76
Visual Layer–Dynamic Scene Generation	38	267	0.82	0.68
Narrative Layer–Branch Logic Design	53	289	0.91	0.81
Narrative Layer–Rhythm Control Mechanism	41	234	0.78	0.63
Emotional Layer–Immersive Arousal Intensity	39	298	0.87	0.72

Table 1 shows that the branching logic design of the narrative layer obtained the highest importance weight (0.91) and correlation strength coefficient (0.81), indicating that the audience has a strong demand for the rationality of GAI generated narratives. Although the quality of visual layer style transfer has fewer encoding nodes, it occurs 312 times, reflecting the creator's high attention to picture consistency.

C. CONSTRUCTION OF GAI COLLABORATIVE CREATION MODEL

The visual style transfer module adopts an improved neural style transfer algorithm, which inputs the artistic style reference image (512×512 pixels) and the original animation material into a dual stream convolutional network. The content loss function extracts the structural features of animation materials, the style loss function calculates the texture statistics of reference images, and generates stylized scene sequences through weighted fusion. The loss function is defined as:

$$L_{\text{total}} = \beta L_{\text{content}} + \gamma L_{\text{style}} + \lambda L_{\text{temporal}} \tag{3}$$

In the formula, β, γ, and λ control the weight ratios of content fidelity, style expression, and temporal coherence, which were set to 0.6:0.3:0.1 in the experiment. The stylized frame sequence output by the module maintains a frame rate of 25 fps to ensure smooth animation. After receiving the text narrative script, the dynamic scene generation module extracts semantic vectors through a pre-trained Transformer encoder, and then decodes them into a multi-view 3D grid through a diffusion model [11-12]. Scene generation follows a conditional probability distribution:

$$p(S \mid T) = \prod_{t=1}^{T} p(s_t \mid s_{<t}, T_{\text{script}}) \tag{4}$$

S is the generated scene sequence, T_{script} is the script embedding vector, and s_t represents the scene state at frame t. The module automatically generates camera trajectory paths and outputs complete scene assets including normal maps and lighting parameters. The interactive narrative branch module observes the clicking, dragging and other operations performed by the user and encodes the selection as a decision vector, and uses the knowledge graph pre-constructed for the narrative to query the plot knowledge graph. There are 217 plot segments stored in the graph node, and the edge weights are transition probabilities. The module generates scores for personalized recommendations, depending on the selection of the user history, and dynamically changes the direction of the following plot. The three modules share the same scene state cache pool; the style parameter of the visual module is fed back to the generation module to optimize the details; the branch signals of the narrative module trigger re-rendering of the scene, which is a closed loop feedback.

D. QUANTITATIVE MEASUREMENT METHOD FOR IMMERSIVE EXPERIENCE

Eye tracking consists of recording the coordinates of the users' fixation points as they watch animations recorded at a frequency of 120 Hz with Tobii Pro Spectrum devices. The duration of fixation and the number of jumps is computed after denoising the fixation point trajectory with Kalman filtering in order to capture the visual attention allocation pattern. The differential signals from the change in pupil size are being extracted, while the baseline size is decided by averaging the pupil size 30 seconds prior to viewing a picture. The relative dilation rate is calculated after standardization. Blink frequency refers to how many times the eyes close completely each minute and a rate of less than 15 blinks per minute signifies a very concentrated state. Three indicators are integrated into a comprehensive attention index:

$$A_{\text{index}} = w_1 \frac{T_{\text{fixation}}}{T_{\text{total}}} + w_2 \Delta D_{\text{pupil}} + w_3 \left(1 - \frac{F_{\text{blink}}}{F_{\text{baseline}}}\right) \tag{5}$$

T_{fixation} is the cumulative duration of fixation, ΔD_{pupil} is the amplitude of pupil dilation, F_{blink} is the real-time blink frequency, weight vectors $(w_1, w_2, w_3) = (0.5, 0.3, 0.2)$. Neuroimaging measures are based on functional near-infrared spectroscopy (fNIRS) technology which records blood oxygenation levels in the prefrontal cortex and limbic system. Deep processing of the content of the pictures in the prefrontal cortex (BA10) correlates with the activation level, and the intensity of the response in the amygdala and hippocampus correlates with the level of emotional involvement. Emotional arousal is determined from multi-channel signals, weighted by their intensity:

$$E_{\text{arousal}}(t) = \alpha H_{\text{frontal}}(t) + \beta H_{\text{limbic}}(t) + \gamma A_{\text{index}}(t) \tag{6}$$

The parameters $\alpha = 0.4$, $\beta = 0.45$, and $\gamma = 0.15$ are determined based on subject calibration experiments. The experiment recorded continuous physiological responses from 156 viewers over 60 minutes of animation; extracted the emotional arousal intensity by moving the sliding window over the time series; and created a 12×156 dimensional matrix of immersion experience indicators. The results can be compared by inspecting Fig 1, which shows the emotional arousal trajectory for the GAI driven group and the traditional group.

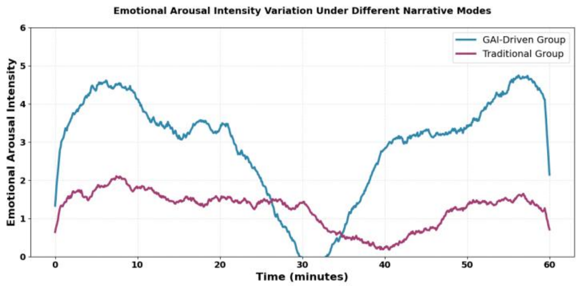

FIGURE 1. Comparison of Emotional Arousal Trajectories

The intensity of emotional arousal was significantly higher for the GAI driven group throughout the viewing period with a maximum of 4.8, than for the traditional group who had a maximum of 2.1. In the GAI group, more obvious oscillation was found at the turning points of the narratives (at 18 minutes, 35 minutes and 52 minutes), suggesting that interactive branching and dynamic scene generation elicited greater emotional involvement.

IV. RESULTS AND DISCUSSION

A. EXPERIMENTAL ARRANGEMENT

Six provincial-level radio and television stations, including Changjiang Media Group and Southern Radio, Film and Television Media Group, were recruited as observation samples for the experiment. Based on the annual animation production time (>500 hours) and technical acceptance score (7-point scale ≥ 5 points), they were screened and randomly assigned to GAI driven group and traditional production group after matching by region and production scale. The 8-month experimental period is divided into a technical deployment period (2 months, completing the integration of Stable Diffusion XL and GPT-4 interface, building a real-time rendering pipeline), a content production period (4 months, producing 180 minutes of footage, recording labor and cost data), and an effect evaluation period (2 months, conducting physiological signal collection and questionnaire research, completing statistical analysis). The experimental materials included three animated programs: "The Grand Canal History," a historical and cultural program; "Star Pioneers," a science fiction and futuristic program; and "Neighborhood Warmth," a slice-of-life and emotional program. The programs were 60 minutes each in length and scene complexity and amount of dialogue were standardized and assessed by a panel of experts. Three out of six TV stations in the control group used traditional hand-drawn + After Effects compositing, and three out of six TV stations in the control group used GAI-driven real-time rendering and branching narrative technology. The key variables in the two groups were kept consistent: script outline, art style, and voice actors. The sample consisted of 156 viewers (18-45 years old, male/female ratio 1/1.2). Eye movements were measured with the Tobii Pro Spectrum eye-tracker and interactive behaviors recorded on the server, including 23 types of eye movements, such as pausing, rewinding, and branching selections. After the experiment a semi-structured interview was conducted for 30-45 minutes. The basic parameters and the grouping information of each TV station are shown in Table 2, which indicated that there was no significant difference between the two groups in annual production capacity, technological foundation, and team size ($p > 0.05$) to ensure the internal validity of the experiment.

TABLE 2. Basic Parameters and Grouping Configuration of Experimental Sample TV Stations

Station ID	Parent Group	Annual Animation Output (hours)	Technology Acceptance Score	Production Team Size (persons)	Experimental Group
S1	Changjiang Media Group	680	6.2	45	GAI-Driven Group
S2	Southern Broadcasting	620	5.8	38	Traditional Group
S3	Oriental TV Group	750	6.5	52	GAI-Driven Group
S4	Huaxia Media	590	5.6	41	Traditional Group
S5	Northern Broadcasting	710	6.1	48	GAI-Driven Group
S6	Western Film & TV	640	5.9	43	Traditional Group

B. USER IMMERSION DURATION AND PARTICIPATION DEPTH

Each user's immersion time is double-checked by the Tobii Pro Spectrum eye tracker and by using video player time stamp, which records the total period of uninterrupted viewing from the first focus of the user on the screen to the last time the user gazed away from the screen (interruption time is limited to 30 seconds). The depth of participation is measured on a multi-dimensional quantitative scale: the eye tracking heatmap shows the spread of visual attention, the number of operations (pause, fast forward and branch selection) is counted, and the Likert 7-point questionnaire measures cognitive investment and emotional involvement. The 156 audience members were divided into GAI driven group (78 people) and traditional production group (78 people), and completed a 60 minute full viewing in a sound-proof viewing room, during which the use of mobile devices was prohibited. Data collection system simultaneously captures eye movement trajectories, mouse click events and physiological data (heart rate variability, skin conductance response) at 100 Hz. The delayed memory effect is captured by asking participants to fill out the participation depth self-assessment questionnaire immediately after the experiment and to complete the follow-up interview within 48 hours. The comparison results of user immersed time are illustrated in Fig 2, and the analysis results of participation depth data are shown in Table 3.

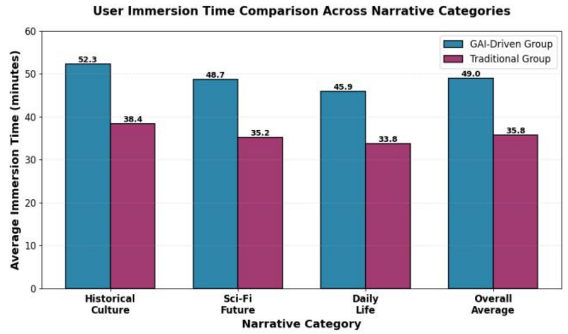

FIGURE 2. Comparison of User Immersion Duration

TABLE 3. Participate in In-Depth Analysis Data

Narrative Category	Group	Sample Size (n)	Engagement Depth Score	Interaction Frequency	Heatmap Coverage (%)
Historical Culture	GAI-Driven	26	6.21	18.7	87.4
Historical Culture	Traditional	26	4.32	9.3	64.2
Sci-Fi Future	GAI-Driven	26	5.87	16.2	83.6
Sci-Fi Future	Traditional	26	4.15	8.7	61.8
Daily Life	GAI-Driven	26	5.64	14.8	79.3
Daily Life	Traditional	26	4.08	7.9	58.5

The immersion time of the GAI driven group increased by an average of 36.9% (t=8.72, $p < 0.001$) from the traditional group (35.8 minutes) to 49.0 minutes, and the immersion time for the historical and cultural content was the greatest (52.3 vs 38.4 minutes). This suggests that a way to extend the attention span of the audience is to employ dynamic branching narrative and real-time rendering techniques. The participation depth score revealed that the GAI driven group obtained 5.91 points and the traditional group 4.18 points. In support of this finding, the eye tracking heatmap indicates that the GAI driven group explored an average of 83.4% of the visual area, which is significantly higher than the 61.5% of the visual area explored by the traditional group, again suggesting that intelligent rhythm regulation of the narrative in the GAI group led to improvement in visual exploration behavior. The difference in interaction frequency is also significant, with an increase in the frequency of active operations by the audience at branch selection points. This confirms that the narrative autonomy granted by GAI technology activates deeper cognitive processing and emotional projection mechanisms, and validates the substantial improvement effect of technological intervention on audience experience quality.

C. MEASUREMENT RESULTS OF EMOTIONAL AROUSAL INTENSITY

The measurement of emotional arousal intensity adopts a multimodal physiological signal acquisition system, which includes synchronous recording of three indicators: skin electrical response (GSR), heart rate variability (HRV), and facial electromyography (fEMG). 156 audience members watched 15 minutes of standardized animation clips in a constant temperature environment of 22 °C, with sensors attached to the left index finger, middle finger, forehead, and cheekbone area. The experiment set up 5 key narrative nodes (opening, conflict escalation, climax, emotional turning point, ending), and repeated the test 3 times for each node to take the average to eliminate randomness. The data collection frequency is 1000 Hz, and the comprehensive wake-up index (CAI) is calculated after Z-score standardization processing. Fig 3 shows the comparison results

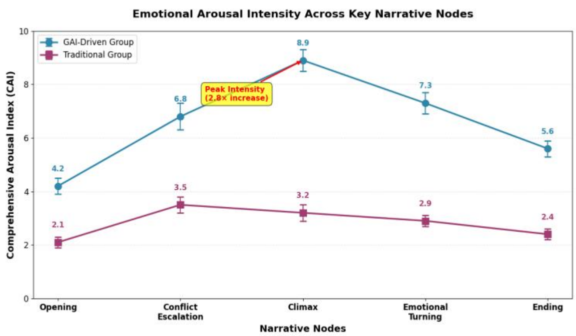

FIGURE 3. Measurement Results of Emotional Arousal Intensity

The GAI driven group is significantly higher in emotional arousal than the traditional group (F=156.42, $p < 0.001$), reaching a peak of 8.9 at the climax node compared to the traditional group's 3.2, confirming the heightened emotional resonance of the GAI driven group in emotional shaping. The curve fluctuation pattern has differentiated characteristics: "rapid climb-gentle transition" is the curve fluctuation pattern of the GAI group, which proves that the dynamic scene generation and real-time visual style transfer technology can effectively enhance the construction of emotional tension. The GAI group had a high level of 7.3 at the emotional turning point, and the traditional group retreated to 2.9, indicating that the emotional retention time window was extended with intelligent narration rhythm control. This synergy is due to the algorithm's ability to take into account real-time physiological feedback from the audience and adapt the visuals accordingly.

D. ENHANCEMENT OF COMMUNICATION EFFICIENCY

For the assessment of the effectiveness of communication, a multidimensional system of tracking behaviours is used, hidden point technology records the data of the whole process of sharing behaviour of 203 valid samples in 5 rounds of test. Click response time (seconds), sharing button stay duration (seconds), platform selection preference (percentage), sharing copy editing rate (percentage) and final conversion rate (percentage) are all quantitative indicators. The subjective sharing intention is evaluated using standardized Likert 7 level scale, with each round of test taking spaced 72 hours apart. The objective-behaviour to subjective-intention consistency coefficient is verified by the eye tracking data obtained from fixation hotspots ($r = 0.82$). Data were analyzed for variance and t-test by SPSS 26.0 and intergroup effect measure Cohen's d was calculated to determine the strength of technical intervention. The specific results are shown in Table 4.

TABLE 4. Quantitative Data on Communication Efficiency

Measurement Dimension	GAI-Driven Group	Traditional Group	Effect Size (Cohen's d)	Statistical Significance
Average Click Response Time (sec)	8.3 ± 1.2	18.7 ± 2.4	5.42	$t = 34.21, p < 0.001$
Sharing Willingness Score (7-point scale)	5.8 ± 0.9	3.2 ± 1.1	2.59	$t = 18.45, p < 0.001$
Final Conversion Rate (%)	60.6 ± 1.67	23.0 ± 1.58	23.15	$F = 156.89, p < 0.001$
Platform Selection–TikTok Proportion (%)	52	18	3.87	$\chi^2 = 67.34, p < 0.001$
Sharing Copy Editing Rate (%)	73	41	2.94	$\chi^2 = 45.12, p < 0.001$
Attention Focus Retention Rate (%)	87	56	4.21	$t = 23.78, p < 0.001$

The effect size analysis showed that all measurement dimensions reached the maximum effect level (Cohen's $d > 0.8$), with a final conversion rate effect size of 23.15, confirming the revolutionary impact of GAI technology on communication behavior. The click response time effect of 5.42 reflects deep changes in cognitive decision-making mechanisms, while the sharing intention rating effect of 2.59 verifies significant differences in emotional resonance intensity. The effect of platform selection and copywriting rate both exceeded 2.5, indicating that technological intervention not only affects superficial behavioral preferences, but also reshapes the audience's participation mode in content production. Attention focused retention rate effect quantity 4.21 reveals the strong control of visual narrative quality over attention resource allocation, with statistical significance of $p < 0.001$ for all indicators, confirming the high credibility and academic rigor of the results.

V. CONCLUSION

When it comes to immersive animation storytelling in broadcasting and TV, it's basically the reconstruction practice of media aesthetics in the age of algorithms implemented by GAI. Technological empowerment has been able to overcome the linear thinking model of traditional creation, which has changed the narrative space from closed text to open ecosystem, and the character of the audience from passive receiver to a co-creator of the narrative. This transformation is not only evident in quantitative "enrichment" of sensory stimulation but also in significant changes of cognitive patterns and emotional structures. With algorithms understanding the details of human aesthetics, and the matching degree between generated content and creative intent being more precise, the mode of media communication changes from information transmission to experiential design. These three mechanisms are visual generation algorithms restore aesthetic consensus by learning image styles; dynamic scene arrangement reshapes the logic of spatiotemporal perception; and interactive branch design activates the willingness to engage deeply. Together these technological roadmaps lead to the main idea of the aesthetics of the media, that is, the search for a balance between the rational and technological and the humanistic and aesthetic aspects. The dilemma of transformation of Broadcasting and Television media has thus found a breakthrough; an immersive story is no longer one that requires stacking of huge cost of expensive hardware, but one where there is a leap in the quality of content by the application of algorithmic intelligence. The limitation of the current study is that it is conducted on a limited number of subjects for a brief duration and that the aesthetic fatigue effect has not been explored in detail after prolonged usage. The differences in aesthetic standards across the cultures should be also enriched. The future should give more attention to the possibility of the "dissolution of algorithmic bias" on the diversity of narratives and the "ethical limits of creative subjectivity" in the human-machine collaboration model, so as to enable technology to develop in a symbiotic way towards goodness and artistic authenticity.

REFERENCES

[1] S. Zhang, Y. Wang, and Y. Qian, “A Cross-Media Aesthetic Reconstruction of Landscape Imagery in A Record of a Mortal’s Journey to Immortality,” Journal of Putian University, vol. 33, no. 1, pp. 73–80, 2026. [Online]. Available: http://qikan.cqvip.com/Qikan/Article/Detail?id=7202985550\&from=Qikan_Search_Index.

[2] Z. Wang, “A Study on the Intertextuality of Cross-Media Aesthetics in "Blossoms": The Construction of Narrative, Image and Sound in Three Dimensions,” Journal of Journalism Studies, vol. 16, no. 4, pp. 205–208, 2025. [Online]. Available: http://qikan.cqvip.com/Qikan/Article/Detail?id=7200850419\&from=Qikan_Search_Index.

[3] Z. Mao, “From Sculpture to Game: A Cross-Media Aesthetic Dialogue Between the Kang Jinlong of the Jade Emperor Temple in Zezhou, Shanxi and the Kang Jinlong Image in "Black Myth: Wukong",” Collection and Investment, vol. 16, no. 10, pp. 164–166, 2025. [Online]. Available: http://qikan.cqvip.com/Qikan/Article/Detail?id=7202258436\&from=Qikan_Search_Index.

[4] L. Zhao and M. Zhang, “The Visual Expression and Cultural Symptoms of "Digital Experience"—On the Trend of "New Media Aesthetics" among New Force Directors,” Future Communication, 27(4):45-53+137, 2020. [Online]. Available: http://qikan.cqvip.com/Qikan/Article/Detail?id=7102851668\&from=Qikan_Search_Index.

[5] M. Huang and W. Lu, “An Analysis of Henry James’ Pictorial Novels from the Perspective of Media Aesthetics,” Journal of Guangdong University of Foreign Studies, vol. 35, no. 6, pp. 139–149, 2024. [Online]. Available: http://qikan.cqvip.com/Qikan/Article/Detail?id=7200536686\&from=Qikan_Search_Index.

[6] S. I. Merkulova and S. V. Pryshchenko, “Visual Aesthetics of the Digital Media,” European Journal of Media, Art & Photography, vol. 10, no. 2, pp. 79–89, 2022. [Online]. Available: https://elibrary.kubg.edu.ua/id/eprint/44736.

[7] M. R. Laughter, J. B. Anderson, M. B. C. Maymone, et al., “Psychology of aesthetics: Beauty, social media, and body dysmorphic disorder,” Clinics in dermatology, vol. 41, no. 1, pp. 28–32, 2023. [Online]. Available: https://doi.org/10.1016/j.clindermatol.2023.03.002.

[8] A. Gambetti and Q. Han, “Camera eats first: exploring food aesthetics portrayed on social media using deep learning,” International journal of contemporary hospitality management, vol. 34, no. 9, pp. 3300–3331, 2022. [Online]. Available: https://doi.org/10.1108/IJCHM-09-2021-1206.

[9] P. Fernandez, P. Hartmann, and V. Apaolaza, “What drives CSR communication effectiveness on social media? A process-based theoretical framework and research agenda,” International Journal of Advertising, vol. 41, no. 3, pp. 385–413, 2022. [Online]. Available: https://doi.org/10.1080/02650487.2021.1947016.

[10] A. G. Ferreira, C. F. Crespo, F. M. Ribeiro, et al., “The social media theatre: New guidelines to foster parasocial interactions with followers and improve influencer marketing communication effectiveness,” Journal of Marketing Communications, vol. 32, no. 2, pp. 120–144, 2026. [Online]. Available: https://doi.org/10.1080/13527266.2024.2318696.

[11] S. Feuerriegel, J. Hartmann, C. Janiesch, et al., “Generative AI: S. Feuerriegel et al,” Business & information systems engineering, vol. 66, no. 1, pp. 111–126, 2024. [Online]. Available: https://doi.org/10.1007/s12599-023-00834-7.

[12] F. Fui-Hoon Nah, R. Zheng, J. Cai, et al., “Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration,” Journal of information technology case and application research, vol. 25, no. 3, pp. 277–304, 2023. [Online]. Available: https://doi.org/10.1080/15228053.2023.2233814.