

Cross-cultural Interpretation of Visual Symbols and Aesthetic Translation in K-pop MV from the Perspective of Multimodal AI

Chenyuan Zhang

Higher School of Cultural Policy and Management in the Humanitarian Sphere, Lomonosov Moscow State University, Moscow, Russian Federation, 119991, Russia

Abstract: The continuous evolution of digital intelligent technologies has profoundly transformed the creative logic and production models of film and television imagery, driving the entire visual culture industry toward digital transformation. As a core vehicle for South Korea's cultural export, K-pop music videos leverage a mature industrial production system and a youth-oriented visual style to amass vast audiences in overseas markets such as Europe, America, and Southeast Asia, thus possessing inherent advantages in cross-cultural communication. Drawing on multimodal discourse analysis and cross-cultural reception aesthetics, this paper examines the underlying mechanisms through which multimodal AI enables Korean visual aesthetics to overcome geographical dissemination barriers and successfully enter international markets, by analyzing recent popular K-pop MVs and their current global distribution. Multimodal AI can deconstruct, reassemble, and fine-tune visual elements within music videos, precisely aligning with aesthetic differences among regional audiences, effectively mitigating common aesthetic disparities and cognitive biases in cross-cultural communication, and facilitating the adaptation of traditional Korean aesthetics to the digital dissemination ecosystem. Compared to conventional, monolithic models of cultural export, AI-enhanced communication pathways address the limitations of poor adaptability and content homogenization in traditional approaches, offering viable theoretical support and practical insights for the global promotion of East Asian pop culture and innovation in digital visual aesthetics.

Keywords: Multimodal AI; K-pop MV; Visual sign; Aesthetic Translation; trans-culture communication.

1. Introduction

The widespread adoption of digital networks has driven the evolution of communication forms. Visual communication, with its intuitive and powerful advantages, has gradually replaced traditional text-based communication and has become the main medium for cross-border cultural exchanges among young people. Text-based communication has limitations such as high cognitive barriers and fixed presentation forms, which are difficult to adapt to the fragmented and fast-paced nature of new media dissemination. Visual image content is lively and highly adaptable, and it perfectly aligns with the current laws of online communication, thus becoming the core medium for popular culture to go global. Relying on a mature industrial production system and unique Korean regional aesthetics, K-pop MVs have long been deeply rooted in the global youth entertainment market, maintaining stable overseas dissemination popularity, and serving as an important channel for South Korea to export its local culture and enhance its international cultural influence [1].

The existing research on K-pop MVs in China still has significant limitations. Most studies focus on the interpretation of visual symbols, cultural connotations, and the review of communication effects, following traditional directions. The research perspectives are relatively fixed. Few studies combine the development background of intelligent technology to explore the internal connection between multimodal AI and MV aesthetics innovation and overseas breakout dissemination. There is no systematic framework for intelligent technology-enabled cross-cultural communication research, and there are obvious research gaps.

In response to this research situation, this paper selects the hit MVs of leading idol groups such as LE SSERAFIM and BLACKPINK as research samples. It uses multimodal discourse analysis and cross-cultural reception aesthetics as theoretical support to explore the evolution characteristics of Korean visual symbols in the digital context. The research breaks away from the traditional static symbol analysis approach and focuses on the complete logic of visual reconstruction and cultural translation in K-pop MVs after the intervention of AI technology. It also combines industry creation chaos to identify dissemination pain points and proposes targeted optimization strategies to fill the gaps in existing research [2].

Clear definition of core concepts is a prerequisite for ensuring the rigor of academic research and avoiding research deviations. Based on the research theme and the industry practice of AI-enabled image creation, this paper clearly defines the three core concepts of multimodal AI, K-pop MV visual symbols, and aesthetic translation.

Unlike ordinary image editing software with a single function, multimodal AI is a comprehensive intelligent system that is suitable for the entire process of image creation. This technology can recognize, decompose, and reorganize the color, light, composition, and scene information in MVs, achieving refined polishing of the images. At the same time, it can adjust the visual presentation effect according to the aesthetic characteristics of different regions, serving as the core technical support for this research. K-pop MV visual symbols specifically refer to visual elements that carry Korean local cultural heritage and highlight regional characteristics, including Korean traditional architecture, folk patterns, ethnic costumes, and local natural landscapes, and are the core identifiers of Korean popular culture that

distinguishes it from European and Japanese pop images [3].

The aesthetic translation referred to in this paper focuses on the two-way adaptation and transformation process in cross-cultural communication scenarios. By adjusting the visual form and content expression of images through intelligent technology, it can not only meet the differentiated aesthetic habits of audiences in different countries and regions but also flexibly convey the core of Korean local culture, reducing the cognitive barriers of cross-cultural communication and promoting the global recognition of the Korean aesthetic system.

2. Multimodal AI Enables Digital Reconstruction of Visual Symbols in K-pop Music Videos

2.1. The AI modernization transformation of traditional Korean symbols

The traditional Korean visual symbols rooted in local folk culture are the core resources for K-pop MVs to establish a differentiated aesthetic system and break through in the overseas market. Classic elements such as Korean classical architecture, folk patterns, and traditional clothing have accumulated the folk customs and cultural traits of various eras in Korea, constructing a highly recognizable visual narrative system, and forming the core differentiated advantage of K-pop MVs in the global visual market.

Before the popularization of intelligent technology, the creation of K-pop MVs relied entirely on on-site shooting. Objective factors such as site conditions, production costs, and the difficulty of restoring ancient architectural scenes greatly restricted the visual presentation effect of traditional folk symbols [4]. The produced images had a stale texture and a monotonous form. Even if they retained the original flavor of traditional culture, they were difficult to meet the aesthetic demands of contemporary young people for novelty, refinement, and individuality, and could not adapt to the fragmented global communication rhythm of the new media era, directly resulting in the narrow overseas dissemination scope and limited international influence of early Korean traditional aesthetics.

The application of multimodal AI effectively solved this communication dilemma, providing a new technical path for the modernization activation of Korean traditional cultural elements and their international dissemination. The creative team can utilize functions such as AI image quality restoration, color optimization, and virtual modeling to upgrade the image texture while retaining the core of Korean elegant aesthetics, transforming the profound and obscure folk elements into digital visual content suitable for modern dissemination [5]. Popular works such as LE SSERAFIM's "ANTIFRAGILE" and BLACKPINK's "How You Like That" have reconstructed Korean ancient buildings and folk patterns through AI modeling, allowing niche traditional cultural elements to enter the international spotlight in a youthful and trendy form. This intelligent creative mode breaks through many limitations of traditional shooting and realizes the modern transformation of Korean traditional aesthetics, promoting the long-term adaptation of local cultural symbols to global entertainment dissemination scenarios.

2.2. Construction of the multimodal system of digital virtual symbols

The digital transformation of traditional Korean visual symbols by multimodal AI has revitalized the dormant local cultural resources. On this basis, the establishment and improvement of the digital virtual symbol system have further enriched the aesthetic expression dimensions of K-pop MVs, becoming an important incremental advantage for their long-term dominance in global popular visual culture. Intelligent technology covers the entire production process of MVs, from creativity to shooting and post-production, completely breaking the objective constraints of traditional live-action shooting, and opening up creative space for the modernization and innovation of Korean aesthetics [6].

The traditional live-action shooting mode has significant creative shortcomings: scene construction, picture presentation, and narrative expression are all constrained by real conditions, with low creative freedom, and are prone to problems such as monotonous pictures and homogenized content, making it difficult to match the fast-paced and high-emotional tension artistic characteristics of popular music. Advanced technologies such as AI virtual scene construction, digital modeling, and intelligent light and shadow rendering effectively break through the limitations of traditional creation, can create diversified immersive virtual-real fusion scenes, break free from the constraints of venues, funds, and real-site restoration, and significantly enrich the picture layers and narrative dimensions of MVs. At the same time, AI can integrate massive dissemination data from overseas social platforms, accurately capture the audio-visual preferences of audiences in different regions, and adjust the picture composition, color ratio, and light and shadow style in a targeted manner, allowing the works to retain the Korean aesthetic background while adapting to the diverse aesthetic systems of the world [7].

The current industry has formed a mainstream creative model of "traditional folkloric base + digital virtual innovation". The combination of virtual and real creates a creative mode that takes into account local cultural heritage, visual innovation highlights, and market communication advantages. By analyzing a large number of industry cases, it can be found that intelligent-empowered MV works have a full visual layer, novel and unique styles, and are in line with the individualized aesthetic needs of contemporary youth, achieving a balance between artistic quality and market popularity, which is also the core reason why K-pop MVs have long occupied the mainstream position in the global youth entertainment market.

3. The Aesthetic Translation Mechanism of K-pop Music Videos from the Perspective of Multimodal AI

3.1. Surface Aesthetics: Standardized Translation of Visual Modalities

Intelligent technology has significantly enhanced the visual quality and expressiveness of K-pop music videos. However, a sophisticated visual presentation does not equate to the success of cross-cultural communication. The differentiated aesthetic system nurtured by regional cultures has always been the core obstacle for Korean aesthetics to enter the international market. The aesthetic perception of the audience

stems from long-term cultural and life accumulation. The segmented aesthetic preferences formed by different regional groups often lead to problems such as poor market acceptance and unsatisfactory communication effects of high-quality Korean-style video works in overseas markets.

The aesthetic preferences of global audiences have distinct regional characteristics, and the visual preferences vary significantly in different regions. European and American audiences prefer simple compositions and visually impactful images, favoring high-contrast, neat and grand visual styles; Southeast Asian audiences are more receptive to images with high saturation and lively and vivid atmospheres; East Asian audiences, who have been deeply immersed in Eastern aesthetics, tend to recognize visual expressions with warm tones, detailed elements, and implicit [8]. This natural aesthetic gap can easily cause overseas audiences to feel alienated from the original ecological Korean traditional images and have difficulty forming emotional resonance, directly weakening the actual effectiveness of cross-cultural communication.

Multimodal AI can effectively solve these aesthetic adaptation problems and has strong cross-regional aesthetic translation and adaptation capabilities. This technology uses deep learning of a large amount of high-quality visual materials to extract universal aesthetic laws and makes fine adjustments to the composition of MV images, color matching, and narrative rhythm. Unlike the one-sided optimization of artificial creation that deliberately caters to overseas aesthetics and discards local characteristics, AI intelligent adaptation is more gentle and moderate. It can narrow the regional aesthetic gap while retaining the recognition and cultural core of Korean aesthetics and quickly gain aesthetic recognition from global audiences, laying a solid foundation for the subsequent cross-domain penetration of deep cultural connotations.

3.2. Deep Aesthetics: Flexible Transposition of Cultural Connotations

Simple superficial visual optimization can only enhance the value of the work and improve the shallow viewing experience of the audience, but it is difficult to achieve the deep transmission and long-term penetration of cultural value. The core of Korean aesthetics' aim to establish a foothold in overseas markets and continuously enhance the international discourse power of Korean culture lies in completing the flexible cross-border transmission of deep cultural connotations. This is also the core value of applying multimodal AI to aesthetic translation.

Traditional official cultural output often adopts a one-way indoctrination model, with rigid forms and insufficient entertainment value, resulting in low audience acceptance and difficulty in achieving deep implementation and long-term dissemination of culture. K-pop MV uses audio-visual images as a flexible communication carrier, which can subtly convey the core of local culture. The intervention of AI further amplifies this communication advantage [9]. The creative team uses AI to break down and reorganize visual symbols, naturally integrating the enterprising, inclusive, and diverse spiritual traits of contemporary Korean youth culture into scene performances, picture details, and stage narratives, allowing cultural output to deeply align with visual art expression and completely avoiding the drawbacks of blunt preaching.

Some Korean traditional visual symbols carry heavy

regional historical connotations. Overseas audiences generally lack corresponding cultural cognitive reserves. Directly presenting them in their original form is likely to cause misunderstandings and cultural misinterpretations in understanding. To address this communication pain point, AI can strip away the obscure niche historical attributes of the symbols, extracting universal emotional cores such as growth, freedom, and healing, significantly reducing the cognitive threshold for cross-cultural communication [10]. Based on overseas audience research data, this gentle cultural transformation model can enable audiences from different cultural backgrounds to break through geographical cognitive limitations, intuitively perceive the artistic value and spiritual core of the work, and achieve seamless, high-quality, and deep cross-cultural communication without barriers.

4. The Current Issues and Optimization Paths of Cross-cultural Dissemination of K-pop MV Aesthetics

4.1. The current challenges in cross-cultural communication

Multimodal AI has provided strong impetus for the aesthetic innovation and overseas dissemination of K-pop MVs, effectively promoting the global popularization of Korean pop culture. However, in the process of rapid technological advancement and the comprehensive intelligence of the industry, industry such as the commercialization of creativity and the templateization of works have emerged, seriously restricting the high-quality overseas dissemination of Korean pop culture.

After a long period of industrialization development, the Korean entertainment industry has become relatively mature [11]. The flow thinking has long dominated the creative orientation of the industry, and the standardized output on the assembly line has become the mainstream model. Most creative teams overly rely on AI visual effects, taking short-term visual impact and online popularity as the core creative goals, while neglecting the in-depth exploration and innovative expression of local cultural symbols. This reversed creative concept directly leads to serious homogenization of industry works, the continuous weakening of the unique regional cultural characteristics of Korean pop culture, and the gradual reduction of the recognition of Korean culture among overseas audiences. Some teams, in order to conform to the universal aesthetic standards of the world, deliberately delete or weaken the core Korean cultural elements, and ultimately produce MVs with exquisite visuals and rich special effects, but with empty content and lacking local cultural heritage and spiritual core [12].

The core cause of these industry lies in the one-sided interpretation and excessive reliance of the industry on the value of AI empowerment. Intelligent technology is always an auxiliary tool for cultural innovation and cannot replace the core value of culture itself. Blindly stacking special effects and chasing short-term traffic not only lowers the overall artistic level of K-pop MVs, exhausts the market reputation of Korean pop culture, but also hinders the deep overseas penetration of Korean culture and is not conducive to the long-term cultivation and steady improvement of Korea's international cultural influence [13].

4.2. Optimization Strategies for Enhancing Aesthetic Communication Enabled by AI

To break through the industry predicament of homogenization and cultural hollowing in K-pop MV creation, we do not need to abandon intelligent technology as a modern creative tool. The key lies in clarifying the primary and secondary relationship between technology and culture. The industry should adhere to the creative principle of "culture as the foundation and technology as the application", allowing intelligent technology to serve the inheritance, innovation and overseas dissemination of local culture, rather than dominating the overall creative direction.

AI itself has the ability to optimize the current situation of industry creation and solve development problems. The final communication effect presented by the technology depends entirely on the creative concepts and application methods of the creative team [14]. By relying on AI differentiated modeling technology, the creative team can combine the brand positioning, artistic style and audience characteristics of different idol groups to create an exclusive and personalized visual aesthetic system, completely breaking the standardized and repetitive creation mode, and fundamentally alleviating the problem of industry homogenization. With the help of AI intelligent screening and semantic analysis technology, high-quality cultural symbols that have both Korean local recognition and are suitable for global audiences can be accurately identified, promoting the integration of traditional aesthetics into digital image creation in a young and modern form, and achieving a balanced bidirectional relationship between cultural inheritance and artistic innovation [15].

At the same time, AI big data can precisely outline the global audience profile and clearly capture the aesthetic preferences of different regions and different circles of audiences. The creative team can rely on the data analysis results to flexibly adjust the visual presentation form, picture details and narrative rhythm of the MV, precisely matching the diversified aesthetic needs of global audiences, improving the accuracy, adaptability and appeal of the overseas dissemination of Korean aesthetics, and building a virtuous development pattern of technology empowerment and cultural deepening in the industry.

5. Conclusion

The comprehensive study of the entire text indicates that multimodal AI has already transcended the attributes of a single imaging production tool and has become the core technological driving force that promotes the aesthetic iteration of K-pop MVs and supports the long-term overseas dissemination of Korean culture. Intelligent technology empowers the creation of K-pop MVs throughout the entire process, mainly through optimizing visual symbols, innovating aesthetic paradigms, and translating cultural connotations. These three paths are interconnected and deeply coordinated to build a complete intelligent communication system. It not only achieves global aesthetic adaptation at the visual level, enhancing the international perception of the works, but also realizes the flexible cross-cultural transmission of the cultural core, effectively eliminating aesthetic barriers and cognitive barriers in cross-cultural communication, and comprehensively improving the quality and coverage of Korean visual culture's overseas dissemination.

From the perspective of the long-term development of the industry, if Korean popular culture wants to achieve high-quality and sustainable global development, it must abandon the utilitarian creative thinking that prioritizes popularity and reasonably define the application boundaries of AI technology, while adhering to the core foundation of local culture. Only by taking into account the picture texture, cultural depth, and global aesthetic adaptability, and creating high-quality images that are mutually empowered by technology and culture, can the international competitiveness of Korean popular culture be continuously consolidated, providing practical reference for the cross-regional dissemination of global popular culture and the innovation of digital image aesthetics.

This research has certain limitations. The study only focuses on the aesthetic empowerment role of multimodal AI and does not introduce Japanese and European popular imaging works for a horizontal comparison. At the same time, it lacks quantitative research data from the field of audiences to support, and cannot accurately verify the actual communication effect of aesthetic translation. The research dimensions still have room for expansion. Future research can broaden the analytical perspective, add cross-regional work comparisons and quantitative empirical research on audiences, and further improve the academic research system on the empowerment of intelligent technology for the cross-cultural dissemination of popular culture.

References

- [1] Huebscher, N. (2025). Standing on the shoulders of giants: Commemorative mimesis and vocal assertion in Red Velvet's borrowing of J. S. Bach's air from orchestral suite no. 3. *Music Theory Spectrum*.
- [2] Wulandari, S. (2023). Hybridization K-pop aesthetics and Western pop culture in BTS' music video 'Butter' through circuit of culture. *Journal of Language, Identity & Culture*.
- [3] Flinchum, E., Roy, E., & Arif, R. (2024). Displays of parasocial interaction in K-Pop: A content analysis of YouTube comments on BTS's music videos. *Korean Journal of Communication*.
- [4] Putra, I. G. A., & Sari, N. K. (2025). Reclaiming nudity: Feminism, commodification, and the male gaze in (G)I-DLE's Nxde. *Humanities and Social Sciences*.
- [5] Lie, J. (2012). What is the K in K-pop? South Korean popular music, the culture industry, and national identity. *Korea Observer*, 43(3), 339–363.
- [6] Oyebamiji, G. A., & Olutayo, G. O. (2025). A multimodal discourse analysis of the music video 'IBA'. *Multimodal Communication*, 14(2), 137–160.
- [7] Aljohani, A. (2025). Multimodal insights: Enhancing cultural promotion through analysis of Saudi Arabian audiovisual productions. *Humanities and Social Sciences Communications*.
- [8] Bae, S., Jung, T., Cho, J., et al. (2025). Effects of meta-human characteristics on user acceptance: From the perspective of uncanny valley theory. *Behaviour and Information Technology*, 44(4), 731–748.
- [9] Andrade López, J., & Flexer, A. (2025). On computational approaches to pop music culture. *Journal of New Music Research*.
- [10] Davies, T., et al. (2025). From sound to sight: Towards AI-authored music videos. *Journal of Audio Research*.
- [11] Mayer, R. (2019). Visual aesthetics and computer vision in popular music album artwork. *Journal of Popular Music Studies*, 31(3), 1–26.

- [12] Dorochoewicz, G., & Kostek, B. (2019). Automatic analysis of visual features in music-related images. *Journal of the Audio Engineering Society*.
- [13] Yoon, K., Min, W., & Jin, D. Y. (2020). Consuming the contra-flow of K-pop in Spain. *Journal of Intercultural Studies*, 41(2), 132–147.
- [14] Lee, H. (2022). “Yea I’m a f* tomboy”: “Girl crush,” postfeminism, and the reimagining of K-pop femininity. *Feminist Media Studies*, 23(8), 2142–2159.
- [15] Kroo, J. (2020). K-pop idols in Japan: The translation of Korean masculinities in music videos. *East Asian Journal of Popular Culture*, 6(1), 27–43.