

A Case Study of Intercultural Expression of Digital Exhibitions at the CGCM from a Multimodal Perspective

Chengkai Wang*

Ningbo University, Ningbo, Zhejiang, China

**Corresponding author: Chengkai Wang.*

Abstract

The Grand Canal of China, as a World Cultural Heritage, serves as a vital carrier for fostering international cultural communication through its digital exhibitions. While the scale of digital exhibitions in Chinese museums continues to expand, prominent issues still exist, including insufficient multilingual resources for international audiences, inadequate adaptation to cultural contexts, and underutilization of multimodal functions. Furthermore, case studies from a multimodal perspective remain scarce. Taking the digital exhibitions at the China Grand Canal Museum as the research subject, this study adopts methods of multimodal discourse analysis, case study and literature review. Six core exhibitions are selected to construct a Chinese-English parallel research corpus of over 21,000 Chinese characters and their corresponding English translations. The characteristics of intercultural expression are examined from three types of modalities. The findings reveal that the museum reaches a relatively high level of digitalization in China, yet deficiencies remain during the adaptation of English interactions. This paper analyzes the underlying causes and proposes optimization strategies, which not only expand the application scenarios of multimodal theory but also provide practical references for the international communication of Grand Canal culture.

Keywords

digital exhibition, multimodality, Grand Canal, international communication

1. Introduction

On May 31, 2021, it was emphasized at the 30th collective study session of the Political Bureau of the 19th CPC Central Committee that effectively communicating China's stories, spreading China's voice, and presenting a true, multi-dimensional, and panoramic view of China are core tasks for strengthening China's international communication capacity. The meeting also stressed the importance of promoting Chinese culture to go global and striving to build a credible, lovable, and respectable image of China [1]. In 2014, the Grand Canal of China was recorded on the World Heritage List. The international communication of its cultural value constitutes an integral part of the national strategy of bringing Chinese culture to the world.

Driven by China's national cultural digitalization strategy, the international communication of Grand Canal culture has undergone accelerated digital transformation. To a certain extent, the intercultural expression capacity of digital exhibitions shapes the effectiveness of cultural communication.

Currently, the scale of digital exhibitions in Chinese museums continues to grow. However, common problems still exist, such as insufficient multilingual resources for international audiences, inadequate cultural adaptation, and failure to make full use of multimodal advantages. Academic circles also lack systematic case studies from a multimodal perspective. Against this background, this paper takes the digital exhibitions of the China Grand Canal Museum (CGCM) as the research object. Based on the analytical framework of multimodal discourse, it focuses on two core research questions. Firstly, what multimodal characteristics do the digital exhibitions of the CGCM present, and what are their strengths and weaknesses in intercultural expression? Secondly, what are the causes of the current state of multimodal expression in digital exhibitions, and what optimization strategies can be proposed?

2. Literature Review

Modality refers to the channels and media of communication, including semiotic systems such as language, technology, images, colors, and music [2]. Multimodality describes a discourse that integrates multiple communication modalities (e.g., sound, text, and images) to convey information [3]. With the iteration of digital communication technologies and the expansion of semiotic forms, multimodal discourse has become the core carrier of contemporary cultural communication and meaning construction. Multimodal research has evolved from a marginal topic in semiotics to an interdisciplinary research direction spanning linguistics, communication studies, art studies, and museology.

In 2001, Kress & van Leeuwen first distinguished between modality and medium [4]. They proposed that modality refers to the semiotic resources that realize discourse and communication, which can be implemented through various production media, while medium refers to the material tools and carriers for semiotic production. Jewitt expanded the connotative boundary of modality, arguing that it covers all ideographic and communicative methods with organizational regularity, including both linguistic and non-linguistic signs [5].

Chinese scholars have further elaborated on the basis of Western theories. Starting from the essence of semiotics, Hu Zhuanglin proposed that modality is a semiotic system that can form contrastive and oppositional relations, while medium is the material means that carries semiotic traces [6]. The two constitute the internal connection between discourse and technology. Wei Qinrong further clarified the differences among modality, medium, and modality value. She specified that modality value measures the authenticity of semiotic signification and should be distinguished from modality in systematic analysis [7]. Zhang Delu argued that multimodal discourse is a discourse form that mobilizes multi-sensory channels and completes communication through the collaboration of diverse semiotic resources [8].

Overall, Western studies focus more on defining the open boundaries of modality from the perspective of semiotics itself, while Chinese studies emphasize the distinction and clarification of core concepts, highlighting the communicative function of modality and its adaptation to local contexts. Both have jointly clarified the conceptual foundation of multimodal research. However, no unified and targeted modality classification standard has been formed for specific scenarios such as museum digital exhibitions.

2.1 Multimodal Theoretical Framework

Halliday elaborated on the three metafunctions in systemic functional linguistics [9]. These three metafunctions (ideational, interpersonal, and textual) have been extended to all semiotic modalities, forming the core theoretical framework for multimodal discourse analysis.

Foreign multimodal research started earlier. In *Reading Images: The Grammar of Visual Design*, Kress & van Leeuwen constructed a complete grammatical framework for visual images, corresponding representational meaning, interactive meaning, and compositional meaning to the three metafunctions of language. This work marked the formal establishment of multimodal discourse analysis as a discipline. The two scholars further expanded the research scope later, putting forward four concepts (discourse, design, production, and distribution) as four levels of multimodal analysis.

Chinese multimodal research began in the early 21st century. Li Zhanzi systematically introduced multimodal discourse analysis theory into China [10]. Hu Zhuanglin systematically sorted out the process of multimodalization from the perspective of social semiotics, analyzed the differences of core concepts, and proposed the cultivation of multimodal literacy, which promoted the rapid development of Chinese research.

Zhu Yongsheng clarified the theoretical basis and research methods of multimodal discourse analysis, providing methodological guidance for local research. Zhang Delu's research is the most representative in Chinese multimodal discourse studies. In 2009, he constructed a comprehensive theoretical framework for multimodal discourse analysis, dividing multimodal discourse into cultural level, contextual level, content level, and expression level. In 2018, he further proposed a new model for multimodal discourse improvement, expanding the practical application dimension of research [11].

Compared with Western theoretical systems that focus on the formal grammatical construction of single modalities, Chinese studies emphasize multi-level and integrated analytical frameworks, strengthening the adaptation of culture and context. Existing theories provide a core paradigm for digital exhibition research. However, for digital spaces with stronger interactivity, their analytical dimensions still need to be further expanded and revised.

2.2 Multimodal Translation Research

Wang Honglin clarified three major research paths of multimodal translation from an academic history perspective, namely multimodal discourse research, social semiotics, and cultural semiotics [12]. He proposed development directions of interdisciplinary dialogue, cross-border method integration, and meta-theoretical integration. Wang Zhiguo revealed the internal mechanism of multimodal inter-semiotic transformation from a cognitive pragmatic perspective, and put forward two reasoning principles: optimal relevance and interpretive similarity [13]. He clarified the collaborative signification of multimodal symbols, providing a theoretical basis for the translation of combinations of multiple symbols such as digital exhibitions.

Among multimodal translation studies, the most fruitful achievements focus on film and television carriers. Xin Hongjuan and Chen Kexin took the external publicity documentary *Seasons of China* as an example, constructing an analytical model of contextualization, synchronization, systematization, and integration, and confirmed that the collaboration of multiple modalities such as subtitles, images, animations, and soundtracks can effectively break cultural barriers and improve the intercultural acceptance of external publicity content [14]. Lyu Jian and Wu Wenzhi, based on the comprehensive framework of multimodal discourse, analyzed the subtitle strategies of *The Flowers of War* from four dimensions: culture, context, content, and expression [15]. They verified the regulatory and supplementary effects of non-linguistic modalities such as images and sound effects on translation choices. Niu Qian took the historical film *Chang'an* as the research object and expounded the transmission of cultural connotation and dynamic adaptation of context, providing references for the translation practice of traditional cultural multi-media works [16].

In the field of stage performance, Chen Jirong focused on the multimodal characteristics of drama translation, proposing rhythm processing, character behavior design, and dramatic situation integration [17]. She emphasized that translation should balance the literariness of text and the performativity of stage, realizing the collaborative translation of written language and stage semiotic systems. This idea has reference value for museum digital exhibitions. Wu Yun and Mou Yiwu also incorporated multimodal translation into the national translation strategy framework [18]. They proposed that the overseas communication of Chinese stories should transcend the limitation of single-modal text narration, and realize innovative narration of cultural content through images, audio-visual media, and hypertext, so as to reduce cultural discount in intercultural communication. This provides methodological guidance for the intercultural expression of museum digital exhibitions.

Overall, as Chen Xi, Pan Hanting, and Pan Li pointed out in their review of Chinese and foreign research, the multimodal turn in translation studies has broken through the traditional language-centered paradigm [19]. The research scope has expanded from single interlingual conversion to inter-semiotic translation among modalities and media. However, Chinese research faces limitations such as fewer achievements, concentrated application scenarios in film and television, and weak basic theoretical construction. There is a research gap in multimodal translation studies of museum digital exhibitions.

2.3 Multimodal Intercultural Communication Research

Yuan Xiaolu and Nai Ruihua extended the multimodal analytical framework to international communication scenarios. Taking the BBC documentary *Chinese New Year* as an example, they constructed a multimodal analysis system of cultural China, and demonstrated the role of multimodal collaboration in

shaping national cultural image and improving overseas recognition from four levels including culture, context, content, and expression. Zou Li and Yu Hongliang, from the perspective of discourse system construction, pointed out that the overseas communication of Chinese culture needs to rely on multimodal interaction to realize intercultural interpretation [20]. Xin Hongjuan and Yang He took the international communication of *Chang'an* as a sample, extracting the narrative logic of consensus and empathy, the aesthetic logic of exchange and integration, and the expressive logic of technology empowerment [21]. They confirmed that starting from shared values, bridging with oriental aesthetics, and strengthening digital audio-visual effects can effectively break cultural barriers.

Xu Hejian et al. focused on the international communication scenario of the Grand Canal Cultural Belt, and proposed a communication framework of blending scene and emotion and mutual learning among civilizations [22]. They advocated strengthening appeal through emotional narration, highlighting recognition through landscape symbols, and expanding coverage through diversified presentation. Liao Xiang and Chen Jingwei, combined with local international communication practices, proposed that cultural international communication needs to build a collaborative pattern with multiple subjects, implement differentiated localization strategies, precisely target young audiences, and innovate media discourse expression [23]. These provide references for the design of digital exhibition communication systems. Liu Jia, targeting the overseas communication of intangible cultural heritage, proposed a construction scheme of a digital communication platform combining multimodal corpus and VR immersive experience [24]. She confirmed that technologies such as 3D modeling and virtual interaction can transcend temporal and spatial constraints, and enhance the intercultural communication and experience of cultural content. This provides technical references for the implementation of intercultural functions of museum digital exhibitions.

The above studies cover both macro communication strategies and micro case analysis. They jointly clarify the significance and value of multimodal expression in intercultural communication, and also provide references for the intercultural expression mechanism and optimization path of Grand Canal culture in digital exhibitions. Existing studies have made progress in the definition of core multimodal concepts, the construction of theoretical systems, and the expansion of application in different fields. Relatively mature research paradigms have been formed in directions such as multimodal translation and intercultural communication. However, the application fields of existing research still mainly focus on film and television translation, intangible cultural heritage communication, etc. There is still a lack of systematic research on museum digital exhibitions, a complicated cultural communication scenario with characteristics of narration in certain space, cultural relic symbols, and digital interaction. Meanwhile, most existing achievements on Grand Canal cultural communication focus on the discussion of culture itself. Few studies analyze the intercultural expression methods in digital exhibition scenarios from a multimodal perspective, leaving certain academic gaps. In view of this, this study can further expand the application field and boundary of multimodal discourse theory on the one hand, and on the other hand, provide theoretical basis and practical references for the international communication practice and intercultural translation of Grand Canal cultural heritage.

3. Research Methods

3.1 Research Object and Case Selection

Located in Yangzhou City, Jiangsu Province, the CGCM is a national first-class museum that officially opened to the public in 2021 as the landmark venue built for the Grand Canal National Cultural Park. Centered on Grand Canal culture, the institution integrates diverse core functions ranging from collection, exhibitions of cultural relics, academic research and public education to tourism service, leisure experience and cross-cultural exchange. To satisfy the demand of showcasing and spreading Grand Canal heritage culture, the museum has gradually built an integrated digital exhibition system covering offline physical exhibition spaces, online virtual scenes and cloud-based display platforms. This research targets Chinese-English bilingual materials within the museum's online digital exhibitions to explore multimodal intercultural expression strategies and corresponding international communication outcomes embedded in digital display designs.

Three criteria guide the selection of research cases throughout this work. First, the criterion of typical representativeness makes the CGCM an ideal research subject, as its digital exhibition infrastructure ranks among the top tier of Chinese museum practices. It has established a full-spectrum digital system spanning

physical exhibition halls, virtual cloud platforms and mobile terminals, and stands as one of very few canal-themed museums nationwide equipped with systematic bilingual digital display resources. Accordingly, its practices reflect the most advanced models for digitally communicating Grand Canal culture to global audiences and serve as an important benchmark for the industry as well as high academic research value. Second, the criterion of data accessibility guarantees the feasibility of multimodal discourse analysis, given that the museum's official website releases all cloud exhibition resources openly to the public. These open resources include bilingual introductions for permanent, temporary and thematic exhibitions, 360-degree VR panoramic virtual halls and supplementary digital footage, all of which form an open, transparent and complete dataset for research use. Third, the criterion of content coverage requires selected cases to incorporate multiple display categories, such as historical narration, cultural imagery interpretation, science and technology showcases and interactive experiences. Such diverse samples can fully demonstrate varied multimodal expression forms within digital exhibitions and typical scenarios of cross-cultural communication, which in turn ensure the generalizability and comprehensiveness of subsequent research conclusions.

Drawing on the three selection criteria outlined above, this study selects a combination of permanent exhibitions, thematic exhibitions and core interactive display programs hosted on the museum's official cloud exhibition platform as primary analytical samples. The selected materials encompass three permanent exhibitions titled *The Grand Canal: China's World Cultural Heritage*, *Boats on the Canal* and *Intangible Cultural Heritage of the Grand Canal*, alongside two thematic exhibitions named *Emperor Yang of Sui and the Grand Canal* and *The Forbidden City and the Grand Canal*, plus one interactive exhibit *The Mystery of the Canal: Supervisor of Canal in the Ming Dynasty*. The scope of analysis further extends to supporting VR panoramic exhibition halls and digital video materials attached to each listed exhibition. Throughout data extraction, the research singles out Chinese-English parallel texts carrying distinct cultural connotations and matches these linguistic materials with corresponding visual and digital multimodal assets, ultimately constructing a multidimensional and comprehensive research sample library.

3.2 Core Theoretical Framework

This study takes Zhang Delu's integrated theoretical framework for multimodal discourse analysis as its theoretical foundation and adapts this framework to fit the semiotic characteristics and communication contexts unique to museum digital exhibitions. It further incorporates Zhang's technology affordance theory to structure a systematic analytical framework organized around three interconnected layers: textual modality, visual modality and digital interactive modality.

The textual layer centers on discourse transformation and meaning adaptation within cross-cultural contexts, with analysis focusing on translation strategies adopted for bilingual exhibition texts. This section evaluates how culture-loaded terms is rendered in English translations and assesses whether supplementary explanations of cultural backgrounds effectively bridge comprehension gaps for foreign readers. Moving to the visual layer, the analysis explores collaborative expressive dynamics between linguistic and visual symbols, unpacking functional divisions and coordination patterns between images and written text while evaluating their complementary roles and value for intercultural communication. The third layer addresses digital interaction through the lens of technology affordance, examining immersive and interactive properties of digital formats including VR panorama galleries, hands-on exhibition installations and 5G live streaming, and measuring how such digital tools strengthen cross-cultural communication capacity.

This analytical framework retains the core logical underpinnings of multimodal discourse analysis while aligning closely with the practical traits of museum digital exhibitions, enabling structured, hierarchical evaluation of intercultural expression patterns featured in displays at the CGCM.

3.3 Specific Research Methods

Multimodal discourse analysis serves as the primary methodological tool of this research. Leveraging the analytical framework introduced above, the study dissects diverse multimodal symbols present in the museum's digital exhibitions, including written text, static imagery, audio clips and interactive digital interfaces, to examine the independent ideographic function of each modality, interactive relationships across different symbolic modes, and the effectiveness and internal mechanisms of intercultural expression achieved through multimodal coordination.

Case study functions as a complementary methodology alongside multimodal discourse analysis. By closely examining representative exhibitions such as *Love of the River*, *Boats on the Canal* and *Emperor Yang of Sui and the Grand Canal*, the research identifies actionable implementation paths for multimodal cross-cultural expression and pinpoints prevalent practical limitations in existing display designs.

Additionally, literature review constitutes another indispensable research technique. Through systematic collation of prior scholarship covering multimodal discourse analysis, multimodal translation and intercultural communication studies, this research clarifies definitional boundaries for core theoretical concepts and traces the evolution of relevant theories, laying a theoretical foundation for framework construction and subsequent analytical discussion. Beyond academic publications, the work also sorts through official policy documents and industry reports concerning Grand Canal cultural conservation, inheritance and museum digital transformation. This review of institutional materials contextualizes the research within real-world cultural development initiatives, ensuring all derived conclusions deliver tangible practical guidance and application value for museum practitioners.

3.4 Data Collection and Analysis

All primary data analyzed within this research are sourced from publicly accessible digital exhibition resources published on the CGCM's official website, with data collection concluding in February 2026. Three distinct categories of core data materials are collected and processed for subsequent analysis, the first of which consists of textual datasets. This textual corpus draws from bilingual introductory content for every permanent and thematic exhibition hosted in the website's cloud exhibition module, covering exhibition hall navigation copy, artifact descriptions, in-depth cultural interpretation and interactive operational prompts. During collection, the researcher intentionally filtered passages featuring cultural proper nouns, symbolic cultural imagery and historical allusions to guarantee the corpus contains sufficiently rich material for cross-cultural comparative analysis.

Second, visual datasets are compiled from visual semiotic assets embedded in digital displays, including photographs of cultural relics, digitally reconstructed historical scenes, original illustrative artwork and captured footage from circular-screen video installations. Each visual asset is paired and cataloged with the exact textual content displayed alongside it to create related image-text analytical units that supply raw visual evidence for the evaluation. Third, digital interactive datasets document structural design, user operation workflows and presentation formats of interactive digital exhibits such as panoramic VR halls, interactive installations of puzzles and 5G live broadcast modules. The researcher manually operated each digital feature online to record functional specifications and interpret core communicative values delivered through digital modalities.

To maintain analytical precision, all initially collected samples underwent secondary screening to prioritize exhibition units featuring complete Chinese-English parallel text, coordinated multimodal elements and prominent cross-cultural communicative significance. After filtering, the finalized research holds a Chinese-English parallel corpus of over 21,000 Chinese characters and their corresponding English translations, 122 matched image-text sample groups and six core digital interactive exhibition installations. Four sequential analytical stages govern processing of the compiled multimodal dataset. The research first categorizes and organizes all collected multimodal materials, annotating bilingual textual content sentence by sentence and flagging critical markers including cultural vocabulary and historical narrative references. In the second stage, single modality analysis is conducted across all samples following the established analytical framework. Third, the research shifts focus to multimodal coordination and cross-cultural communication outcomes. Building upon single modality evaluation results, researchers further investigate interactive links between separate modalities and apply Zhang Delu's modal relationship theory to classify inter-modal dynamics. The analysis distinguishes reinforced or non-reinforced complementary relationships from non-complementary overlapping or redundant modal interactions, which allows systematic diagnosis of structural weaknesses within the museum's existing digital displays. In the final analytical stage, the research unpacks root causes behind observed multimodal expression deficiencies and proposes targeted optimization strategies for cross-cultural multimodal exhibition design, thereby fully addressing the two central research questions raised at the outset of this paper.

4. Research Results

4.1 Bilingual Presentation Characteristics of Textual Modality

4.1.1 Bilingual Coverage of Exhibitions

The Six core exhibitions covered by the samples are all equipped with English versions. There are hierarchical differences in the bilingual coverage modules of different exhibitions. Among them, the bilingual completeness of permanent exhibitions is generally higher than that of thematic exhibitions and interactive exhibition items. The exhibition hall navigation module has the highest bilingual coverage rate, while the bilingual coverage rate of exhibit detail descriptions and in-depth cultural interpretation modules gradually decreases. The basic information of bilingual texts for each exhibition is shown in Table 1.

Table 1: Basic Information of Bilingual Texts in Core Exhibitions

Exhibition Name	Chinese Text Length (characters)	English Text Length (words)	Bilingual Coverage Modules	Full Exhibition Bilingual Coverage
<i>The Grand Canal: China's World Cultural Heritage</i>	4286	3952	Exhibition hall navigation, exhibition area description, exhibit introduction, cultural interpretation	Full coverage
<i>Boats on the Canal</i>	3124	2876	Exhibition hall navigation, exhibition area description, core exhibit introduction	Partial coverage
<i>Intangible Cultural Heritage (ICH) of the Grand Canal</i>	3652	2108	Exhibition hall navigation, general overview of exhibition areas, name labeling of ICH items	Partial coverage
<i>Emperor Yang of Sui and the Grand Canal</i>	2897	2643	Exhibition hall navigation, exhibition area description, core cultural relic introduction	Full coverage
<i>The Forbidden City and the Grand Canal</i>	2415	2279	Exhibition hall navigation, exhibition area description, exhibit introduction	Full coverage
<i>The Mystery of the Canal: Supervisor of Canal in the Ming Dynasty</i>	1673	1541	Exhibition item introduction, level description, operation prompts	Partial coverage

4.1.2 Distribution of Translation Types for Culture-Loaded Terms

This study screens 87 culture-loaded terms with typical cultural attributes from the samples, including proper nouns, historical concepts, folk stories, etc. According to translation methods, they can be divided into four categories: transliteration, literal translation, free translation, and amplification with supplementary explanation. The sample quantity and proportion of each category are shown in Table 2.

Table 2: Distribution of Translation Strategies for Culture-Loaded Terms

Translation Type	Sample Quantity	Proportion	Typical Examples
Transliteration	24	27.6%	Shafei craft (沙飞船), Wutou Gate (乌头门), Guazhou (瓜洲)
Literal translation	35	40.2%	The Forbidden City Drifted in from Afar Along the Grand Canal (运河漂来紫禁城), a long parade of red trousseaux (十里红妆), traveling south by boat and north by horse (南船北马)
Free translation	19	21.8%	Water Flows Left and Right, the Youth Yearns Day and Night (河之恋), main shops (正店), Qianzhuang, a local independent Chinese bank (钱庄)
Amplification with supplementary explanation	9	10.4%	Office of the Director-General of Grain Transport (漕运总督部院), ornate bamboo scaffolding-like structures at entrances (彩楼欢门), Zhuazhou, a traditional one-year-old catch ceremony (抓周)

From the perspective of distribution characteristics, literal translation is the most widely used translation method for culture-loaded terms in current exhibitions, followed by transliteration, and amplification with supplementary explanation accounts for the lowest proportion. Most English descriptions of culture-loaded terms only retain core semantics, without additional independent cultural background interpretation paragraphs.

4.2 Image-Text Presentation Characteristics of Visual Modality

This study classifies and analyzes 122 groups of image-text matching samples. According to the presentation weight and functional relationship of image-text information, they are divided into four types: image-dominant with text supplement, image-text balanced, text-dominant with image supplement, and name-only labeling. The distribution of each type is shown in Table 3.

Table 3: Distribution of Image-Text Sample Presentation Types

Presentation Type	Sample Quantity	Proportion	Typical Exhibition Areas
Image-dominant with text supplement	58	47.5%	Ship model display area of <i>Boats on the Canal</i> , scene restoration area of <i>Born from Transportation</i>
Image-text balanced	21	17.2%	Scientific and technological achievements exhibition area of <i>The Grand Canal: China's World Cultural Heritage</i>
Text-dominant with image supplement	11	9.0%	Historical chronology exhibition area of <i>Emperor Yang of Sui and the Grand Canal</i>
Name-only labeling	32	26.2%	ICH exhibit area of <i>Intangible Cultural Heritage of the Grand Canal</i> , urban style exhibition area of <i>World-Famous Canals and Canal Cities</i>

In terms of visual resource types, among the 122 groups of samples, there are 51 groups of real photos of cultural relics, 37 groups of historical scene restoration pictures, 19 groups of digitally created illustrations, and 15 groups of circular screen image screenshots, accounting for 41.8%, 30.3%, 15.6%, and 12.3% respectively.

Visual resources in most exhibition areas occupy the core display position on the page. English text content is mostly arranged below or on the side of pictures, with a font size smaller than that of the original Chinese text.

4.3 Presentation Characteristics of Digital Interactive Modality

The samples include 6 core digital interactive exhibition items, covering five technical forms: VR panoramic roaming, interactive puzzle solving, 5G real-time live broadcasting, circular screen images, and interactive touch screens. The technical characteristics, interaction modes, and bilingual adaptation of each exhibition item are shown in Table 4.

Table 4: Basic Characteristics of Core Digital Interactive Exhibition Items

Exhibition Item Name	Technical Type	Interaction Mode	English Navigation Text	English Interactive Interface	Audio Guide Availability
360° VR Panoramic Exhibition Hall	VR panorama	Passive roaming	Yes	Yes	Chinese only
5G+VR 720° Live Grand Canal View	5G+VR	Passive viewing	Yes	No	Chinese only
View the Grand Canal in 100 Billion Pixels	Ultra-high-definition panorama	Active zoom browsing	Yes	No	Not available
<i>The Mystery of the Canal: Supervisor of Canal in the Ming Dynasty</i>	Interactive puzzle solving	Active level-breaking interaction	Yes	Partial adaptation	Chinese only
360° Circular Screen Experience of Shafei Craft	Circular screen image	Passive viewing	Yes	No	Chinese only
ICH Interactive Touch Screen	Touch interaction	Active click query	Partial coverage	Partial adaptation	Not available

From the overall distribution, all digital exhibition items are equipped with English introduction texts, but only two have realized full English adaptation of the interactive interface. For most exhibition items, operation buttons, pop-up prompts, and interactive content only support Chinese. All supporting audio guide resources only provide Chinese versions, without corresponding English dubbing.

5. Discussion

5.1 Analysis of the Characteristics of Multimodal Intercultural Expression

Existing studies generally point out that digital exhibitions in Chinese museums have common problems such as insufficient multilingual resources, weak cultural context adaptation, and insufficient multimodal collaboration. Meanwhile, they confirm that the collaboration of multimodal symbols can effectively reduce cultural discount in intercultural communication. The characteristics presented by the samples in this study are both consistent with and different from previous conclusions.

5.1.1 Textual Modality

In terms of bilingual coverage, the bilingual coverage rate of core exhibitions at the CGCM is higher than the average level of similar museums in China, reflecting the pioneering role of benchmark museums in international communication construction.

However, in terms of the translation depth of culture-loaded terms, the results of this study are consistent with existing research conclusions: most culture-loaded terms only complete semantic conversion, lacking supplementary interpretation of cultural context, which is likely to cause comprehension barriers for international audiences. The problem of “insufficient decoding of cultural symbols in multimodal translation” pointed out by Wang Honglin also exists here. As is shown in Example 5-1, the translation of “Shafei craft”:

Example 5-1

Chinese: 沙飞船

English: a restored 20-meter-long Shafei craft from the Kangxi Emperor’s reign

This translation only conveys the physical attributes of the artifact, but fails to explain the functional positioning of the “Shafei craft” as a grain transport vessel in the Qing Dynasty and its connotation of sailing as fast as flying. Compared with the optimal relevance principle of multimodal inter-semiotic transformation proposed by Wang Zhiguo, the linguistic modality here fails to form complementary relationship with the ship model image in the visual modality. It only completes the correspondence of symbols, but fails to achieve complete meaning transmission.

Meanwhile, the translation of some exhibition items reflects a strong contextual awareness, which is similar to the conclusion verified by Xin Hongjuan et al. in documentary translation that contextualized translation improves intercultural acceptance. As is shown in the example 5-2, the translation of “Love of the River”:

Example 5-2

Chinese: 河之恋

English: Water Flows Left and Right, the Youth Yearns Day and Night

Instead of literal translation, this translation adapts the classic emotional imagery from *the Book of Songs*, transforming the abstract love into an emotional expression accessible to Western audiences, and realizing the cross-context translation of cultural imagery. However, such high-quality free translation cases only account for 21.8% of the samples and have not yet become the mainstream translation strategy.

5.1.2 Visual Modality

This study finds that the samples have abundant types and sufficient quantities of visual resources, and visual modality undertakes the core ideographic function in exhibitions. As Zhang Delu proposed, visual modality in multimodal discourse can cross language barriers. Liu Jia also pointed out that digital exhibitions reduce communication thresholds with visual symbols.

However, 26.2% of the image-text samples belong to the name-only labeling type. The text only undertakes the identification function and does not carry out cultural interpretation of the visual content. The two form an overlapping state in non-complementary relationships, failing to exert the improvement effect of image-text collaboration on ideography.

This characteristic is different from the mature multimodal collaboration model in the film and television field. Xin Hongjuan and Yang He, in their research on the international communication of *Chang'an*, confirmed that the reinforced complementary relationship between images and subtitles can significantly improve the effect of cultural communication. However, the museum digital exhibitions in this study have not yet established design specifications for image-text collaboration, and visual advantages have not been fully transformed into intercultural communication advantages.

5.1.3 Digital Modality

The wide application of digital technologies such as VR panorama, 5G live broadcasting, and interactive puzzle solving in the samples is consistent with Liu Jia's point that VR technology can break through time and space restrictions and enhance intercultural experience. The digital construction level of the CGCM ranks among the top level of Chinese museums, and the immersion and interactivity of digital exhibition items have reached high standards.

However, the intercultural adaptation of digital interaction obviously lags behind technological construction. Only 33.3% of the core digital exhibition items support English interactive interfaces, and all audio guides have no English versions. Digital technology has built the physical channel for intercultural communication, but insufficient adaptation at the linguistic and cultural levels limits the full release of technological value.

5.2 Motivations for the Formation of Multimodal Expression Status

The unbalanced development pattern of leading digital technology, uneven translation quality and insufficient multimodal coordination in the digital exhibitions of the CGCM is not an isolated case, but a common phenomenon in the digital transformation of cultural venues across China. Such pattern is shaped by the combined effects of policy orientation, development rationale, audience positioning and design mechanism.

Driven by the national cultural digitalization strategy, digital hardware construction has generally been given higher priority than cultural content adaptation in museum development. In recent years, a series of national policies have been rolled out to advance the digital transformation of museums and the development of smart museums, and digitalization has also been identified as a core priority in the construction of the Grand Canal National Cultural Park. As a result, museums have prioritized investment in technical equipment and hardware resources for digital exhibition development, with technologies such as VR, 5G and interactive devices deployed at a rapid pace. In contrast, investment in soft content such as bilingual translation and cross-cultural adaptation has lagged behind. This mismatch between technological progress and content upgrading aligns with the findings of Liu, who noted that digital cultural heritage projects often prioritize technical breakthroughs while overlooking the cultural accessibility of content for overseas audiences.

From the perspective of audience positioning and development rationale, exhibition design has long been centered on local audiences, with the function of international communication placed in a subordinate position. At present, Chinese audiences remain the primary service targets of museum digital exhibitions, and the narrative logic, content depth and interactive design of exhibitions are all tailored to the cognitive habits and cultural backgrounds of local visitors. Bilingual content is mostly added as a supplementary component rather than a core module, and no content production system centered on international audiences has been established. This audience-oriented deviation is consistent with the conclusion of Liao and Chen, who pointed out that local cultural communication practices in China generally lack targeted design for global audiences, leading

to weakened cross-cultural communication effects. Such a positioning directly gives rise to problems including insufficient background supplementation for culture-loaded terms and incomplete English adaptation of interactive interfaces.

In terms of content production, the creation of multimodal content features obvious departmental segmentation, and a collaborative design mechanism has not yet been established. The content production of museum digital exhibitions usually involves multiple departments such as the exhibition department, information technology department and translation team, with visual design, text writing, translation and technical development carried out in separate links. There is a lack of multimodal coordination from the perspective of intercultural communication, where the translation team only takes charge of text conversion without participating in visual and interactive design, while the technical team focuses on function realization without paying attention to the effect of cultural adaptation. This fragmented production model echoes the observation of Wang that multimodal translation practices in China often separate linguistic conversion from other modal design, resulting in insufficient inter-semiotic coordination. Eventually, non-complementary relationships such as overlap and redundancy emerge between different modalities.

In addition, the absence of an effect evaluation mechanism for intercultural communication deprives content optimization and iteration of sufficient data support. At present, the effect evaluation of museum digital exhibitions mostly focuses on indicators such as Chinese visit volume and offline satisfaction, while data collection and analysis of international audiences' access behavior, comprehension degree and communication feedback are largely insufficient. Without precise identification of pain points in cross-cultural communication, it is difficult to promote the continuous optimization of content, causing some common problems to persist for a long time. As Yuan and Nai argued, the current international communication evaluation system in China overemphasizes quantitative traffic indicators and neglects in-depth measurement of audience cognitive effects, which restricts the iterative upgrading of cross-cultural communication practices.

5.3 Optimization Paths for Multimodal Intercultural Expression in Digital Exhibitions

In view of the existing problems, and combined with multimodal discourse theory and the laws of intercultural communication, a systematic optimization framework can be constructed from four dimensions: text system improvement, modal collaboration enhancement, digital interaction adaptation and evaluation mechanism establishment. These pathways not only respond to the common problems identified in previous studies, but also provide targeted solutions tailored to the characteristics of museum digital exhibitions.

The hierarchical adaptation of the bilingual text system serves as the foundation for improving the quality of intercultural expression, with the translation depth of culture-loaded terms as the core optimization focus. Differentiated translation strategies should be adopted for different types of exhibition texts. As for core navigation texts, a combination of literal translation and cultural notes should be applied to ensure information accuracy. While for texts carrying cultural imagery, contextualized free translation strategies should be adopted to reduce cultural discount. And for professional terminology, unified translation specifications should be established to avoid the abuse of transliteration. This hierarchical strategy draws on the optimal relevance principle of multimodal inter-semiotic transformation proposed by Wang, and extends the contextualized translation method verified in documentary translation to the museum exhibition scenario. Drawing on the successful experience of *Love of the River*, core cultural concepts can be translated by combining semantic conversion with imagery adaptation. As is shown in the example 5-3:

Example 5-3

Before optimization: Shafei craft (沙飞船)

After optimization: Shafei fast boat—a cargo with a flat bottom and a passenger vessel, known for its high sailing speed during the Qing Dynasty

By supplementing functional attributes and naming connotations, the translation enables international audiences to grasp the core meaning of cultural symbols through text, which compensates for differences in cultural contexts.

On this basis, it is necessary to establish an image-text collaborative design mechanism to strengthen the complementary relationship between modalities. The functional division of images and texts in different exhibition scenarios should be clarified: for cultural relic display scenarios, an image-dominant and text-supplemented reinforced complementary mode can be adopted, where pictures present the shape of artifacts and texts supplement cultural background; for cultural interpretation scenarios, an image-text balanced non-reinforced complementary mode is applicable, where images construct the scene atmosphere and texts convey deep connotations. This design idea is rooted in Zhang's modal relationship theory, and draws on the proven effect of image-text synergy in film communication. For current ICH exhibition items with only name labeling, concise English cultural descriptions can be added to transform the image-text overlapping relationship into a complementary one, so as to fully unleash the communication advantages of visual modality.

The intercultural adaptation of digital interaction is the key link to unlocking the value of technology empowerment. For digital products such as VR panoramas and interactive exhibition items, basic configurations including English interactive interfaces, English operation prompts and English audio guides need to be completed. For youth-oriented interactive products such as *The Mystery of the Canal: Supervisor of Canal in the Ming Dynasty*, the cultural adaptation in puzzle design can be optimized, with embedded explanations added to Chinese-specific historical concepts to lower the participation threshold for international audiences. This dual adaptation of technology and culture responds to Liu's argument that digital technology can transcend spatial-temporal limitations in cultural communication, and further makes up for the deficiency of existing studies that focus more on technical functions than cultural accessibility. Through the integration of technical upgrading and cultural adaptation, digital technology can truly serve as a carrier rather than a barrier for intercultural communication.

Finally, it is essential to establish an international audience-oriented communication effect evaluation mechanism to form a closed loop of optimization. An international user access data collection module should be added to the backstage of digital exhibitions to track behavioral data of international audiences such as browsing duration, jump paths and interaction completion rates. Audience feedback can be collected through online questionnaires and comment sections. This evaluation mechanism aligns with the precise communication strategy proposed by Liao and Chen, and provides a data-driven approach for the continuous optimization of cross-cultural digital exhibitions.

6. Conclusion

Taking the digital exhibitions of the CGCM as the research object, this study systematically investigates the presentation characteristics of three modalities (text, vision and digital) from the perspective of multimodal discourse analysis. The research shows that the digital exhibitions of the CGCM have formed a relatively complete multimodal ideographic system, with bilingual coverage and digital technology application at the leading level in China's museum sector, laying a solid foundation for intercultural communication. Meanwhile, there is an obvious structural imbalance in the multimodal intercultural expression of current exhibitions, marked by three coexisting deficiencies: insufficient translation depth in textual modality, weak image-text collaboration in visual modality, and lagging cultural adaptation in digital modality. Such overall features of advanced technology, uneven translation quality and inadequate multimodal coordination stem from the combined effects of policy orientation, construction logic, audience positioning and design mechanism, and therefore require collaborative optimization to enhance the overall effectiveness of international communication.

At the theoretical level, this study extends multimodal discourse analysis theory to the intercultural communication scenario of museum digital exhibitions, and verifies the applicability of multimodal complementary relationship theory in digital spaces. Previous studies have pointed out that multimodal translation and intercultural communication research in China is mostly concentrated in the field of film and television, with a clear research gap in museum digital exhibition scenarios. This study fills this gap to a certain extent by supplementing empirical research on the multimodal characteristics of cultural heritage digital exhibitions, enriching the application scenarios of multimodal translation and intercultural communication research, and providing referable analytical pathways and empirical samples for subsequent similar studies. The findings also confirm that Zhang's multimodal discourse analysis framework still has strong explanatory power in digital interactive scenarios, but needs to be adjusted appropriately in combination with the technical

affordance of digital media. At the practical level, this study systematically sorts out the multimodal status and core problems of digital exhibitions at the CGCM. The proposed optimization paths, including hierarchical translation strategies, image-text collaboration mechanisms and digital adaptation schemes, can directly provide practical references for the international communication development of digital exhibitions at the CGCM and similar canal-themed museums. In the context of the Grand Canal Cultural Belt construction, these findings also respond to the call for promoting mutual learning among canal civilizations, and offer actionable solutions for the intercultural digital exhibition design of other cultural heritage museums, so as to facilitate the digital international communication of Chinese culture.

This study still has certain limitations. First, the research samples only focus on the online digital exhibitions of the CGCM without covering digital exhibition items of other museums related to Grand Canal culture, which limits the generalizability of the research conclusions to a certain extent. Second, the research is mainly based on the textual and symbolic analysis of exhibition content rather than empirical research on international audiences, so the judgment of communication effect relies mainly on theoretical deduction and lacks the support of audience feedback data. In addition, the research primarily focuses on Chinese-English bilingual content without involving exhibition resources in other languages, resulting in insufficient comprehensiveness in the analysis of multilingual intercultural communication.

Building on the above limitations, future research can be further expanded in multiple directions. The scope of research samples can be broadened to conduct a comparative analysis of digital exhibitions of different museums themed on canal at home and abroad, so as to extract the common laws and differentiated strategies of intercultural communication. Audience research methods can also be introduced, drawing on the audience-centered research perspective applied in drama translation studies, to collect cognitive data of international audiences through questionnaires, interviews and eye-tracking experiments, and empirically test the impact of multimodal design on communication effects. Furthermore, future studies can explore the application of AI technology in the intercultural expression of museum digital exhibitions, and investigate the empowering pathways of technologies such as intelligent translation, automatic multimodal content generation and personalized recommendation for intercultural communication. This aligns with the development direction of China's national translation strategy in the digital age, and is expected to provide more solutions for cultural international communication in the digital era.

References

- [1] Yuan, X. L., Nai, R. H., A study on the construction of meaning in multimodal discourse in the international communication of “cultural China”. *Foreign Language Education*. 2022; 43(5): 23-29.
- [2] Zhu, Y. S., Theory and methodology of multimodal discourse analysis. *Foreign Language Research*. 2007; (5): 82–86.
- [3] Kress, G. R., van Leeuwen, T., *Reading images: The grammar of visual design*. London: Routledge; 1996.
- [4] Kress, G. R., van Leeuwen, T., *Multimodal discourse: The modes and media of contemporary communication*. Bloomsbury Academic; 2001.
- [5] Jewitt, C., *Multimodality and new communication technologies*. Continuum; 2004.
- [6] Hu, Z. L., Multimodalization in social semiotics. *Language Teaching and Linguistic Studies*. 2007; (1): 1–10.
- [7] Wei, Q. H., On modality, medium and mood in multimodal discourse. *Foreign Language Education*. 2009; 30(4): 54–57.
- [8] Zhang, D. L., On a synthetic theoretical framework for multimodal discourse analysis. *Foreign Languages in China*. 2009; (1): 24–30.
- [9] Halliday, M. A. K., *An introduction to functional grammar* (2nd ed.). London: Arnold; 1994.
- [10] Li, Z. Z., A sociosemiotic analysis of multimodal discourse. *Foreign Languages Research*. 2003; (5): 1–8, 80.

- [11] Zhang, D. L., A systemic-functional synthetic framework for multimodal discourse analysis. *Modern Foreign Languages*. 2018; 41(6): 731–743.
- [12] Wang, H. L., An academic history review of multimodal translation studies. *Chinese Translators Journal*. 2022; 43(6): 106–113.
- [13] Wang, Z. G., Multimodal translation discourse analysis and pragmatic cognition exploration. *Chinese Translators Journal*. 2023; 44(5): 98–105.
- [14] Xin, H. J., Chen, K. X., A study on translation of international publicity documentaries from the perspective of multimodal discourse analysis: Taking Seasons of China as an example. *International Publicity*. 2020; (2): 54–56.
- [15] Lyu, J., Wu, W. Z., Subtitle translation study of the flowers of war from the perspective of multimodal discourse analysis. *Shanghai Journal of Translators*. 2012; (4): 36–38.
- [16] Niu, Q., Experience and enlightenment of subtitle translation of the film Chang'an from the perspective of multimodal discourse analysis. *Movie Literature*. 2025; (5): 145–149.
- [17] Chen, J. R., An analysis of drama translation practice from the perspective of multimodal theory. *Chinese Translators Journal*. 2025; 46(1): 149–155.
- [18] Wu, Y., Mou, Y. W., A study on multimodal national translation strategies for Chinese stories. *Foreign Language Education*. 2022; 43(1): 76–82.
- [19] Chen, X., Pan, H. T., & Pan, L., Multimodal turn in translation studies: present and prospect. *Foreign Language Research*. 2020; (2): 80–87.
- [20] Zou, L., Yu, H. L., Telling Chinese stories of the new era through cultural communication: A review of the symposium “International Communication of Chinese Culture: Discourse System and National Image”. *International Communications*. 2024; (4): 65–67.
- [21] Xin, H. J., Yang, H., Exploration of the international communication logic of the animated film Chang'an going global. *International Communications*. 2024; (8): 37–40.
- [22] Xu, H. J., Zhang, Y. Q., Chen, G., & Ma, X. L., Opening a new chapter of exchanges and mutual learning among canal civilizations in the new era: A review of the “Dialogue on International Communication of China's Grand Canal Cultural Belt”. *International Communications*. 2024; (1): 38–41.
- [23] Liao, X., Chen, J. W., Exploration and innovation of China's international communication narrative system in the new era. *International Communications*. 2025; (3): 71–75.
- [24] Liu, J., Research on the construction of a multimodal corpus platform for digital-enabled international communication of intangible cultural heritage. *Shanghai Journal of Translators*. 2023; (6): 34–39.

Funding

This paper was supported by the project for taking Bridges as “Bridges”: An Exploration of the Translation and International Communication of Bridge Culture along the Eastern Zhejiang Canal (No. 202511646043).

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).