

# The Interactive Significance and Promotion Effect of Urban (Cultural Tourism) Promotional Videos from the Perspective of Multimodal Discourse Analysis

Bingyu Chen<sup>1</sup> & Renjun Pan<sup>2</sup> & Xiaodong Wang<sup>3\*</sup>

1,2,3.School of Arts and Sciences, Northeast Agricultural University, Harbin, China.

\*Corresponding Author Email: wangxiaodong@neau.edu.cn

## Abstract

In the development of modern society, in the face of an increasingly competitive market environment, how to use the film and television medium of city image promotional videos to represent, build a city brand with personalized characteristics, comprehensively enhance the city's competitiveness and well-known brand, and make it stand out in the process of homogenization of cities has become the focus of attention from all sectors of society. Based on the theoretical framework and analysis methods of multimodal discourse analysis, this paper explores the different modal characteristics and construction significance of urban cultural and tourism promotional videos in a certain region. Finally, the regional image and rich emotions of the city are presented based on the artistic symbolism of multimodal.

**Keywords:** *Multimodal Discourse Analysis City Promotional Video Cultural Tourism Modal Characteristics Artistic Symbol.*

## INTRODUCTION

In the information age, human interaction methods have expanded from single-language expression to multi-modal collaborative communication. People not only convey information through language, but also construct meaning with the help of various symbol systems such as music, images and gestures. These diverse expressive resources play a crucial role in contemporary social interaction and information dissemination, especially having significant value in the field of public media communication.

Multimodal discourse analysis, as a cutting-edge field of academic research, focuses on the discourse analysis of the collaborative expression of meaning by two or more modalities, or the combined application of multiple symbol systems in a single modality. Its research scope includes various perceptual modalities such as touch, hearing, vision and smell.

Against this background, as a typical dynamic multimodal discourse, the meaning construction mechanism of urban promotional videos has significant research value. An in-depth analysis of the interaction mechanism of multimodal resources in urban promotional videos and a systematic exploration of how symbols such as text, sound and image collaborate to construct high-quality audio-visual works have dual significance for effectively disseminating the urban image and historical and cultural connotations with Chinese characteristics: On the one hand, it helps international audiences recognize the achievements of China's modernization development; on the other hand, it also provides innovative communication paths for urban cultural and tourism departments. The core objective of this multimodal discourse practice lies in optimizing the interpersonal function of communication effects and constructing a positive social value communication system.

## **1. Multimodal Discourse Analysis**

### **(1) Theoretical Framework**

In the 1990s, some scholars, based on Halliday's meta-function theory, constructed the visual grammatical framework of multimodal discourse for analyzing the meaning construction of images [1]. In systemic functional linguistics, Halliday proposed that language has three meta-functions: conceptual function (used to represent experience), interpersonal function (used for social interaction), and discourse function (used for organizing information). In visual modalities, these three functions correspond respectively to the representational meaning (representing reality), the interactive meaning (establishing the relationship between the viewer and the image), and the composition meaning (the visual organization of the image). With the deepening of multimodal research, scholars have found that the visual grammar framework has certain limitations. Therefore, subsequent studies have expanded and corrected it [2]. For instance, some scholars further divide the representational meaning into character shaping, event relations and background relations to precisely analyze the narrative structure of the image. The study of interactive meaning has been extended to the gaze system, the affect system and visual modality to explore how images affect the emotions and attitudes of viewers. Some scholars adopt the graphic and text layout instead of the original information value framework to explain the synergy between language and images more effectively while retaining the salience features of the images [3].

### **(2) Analytical Methods**

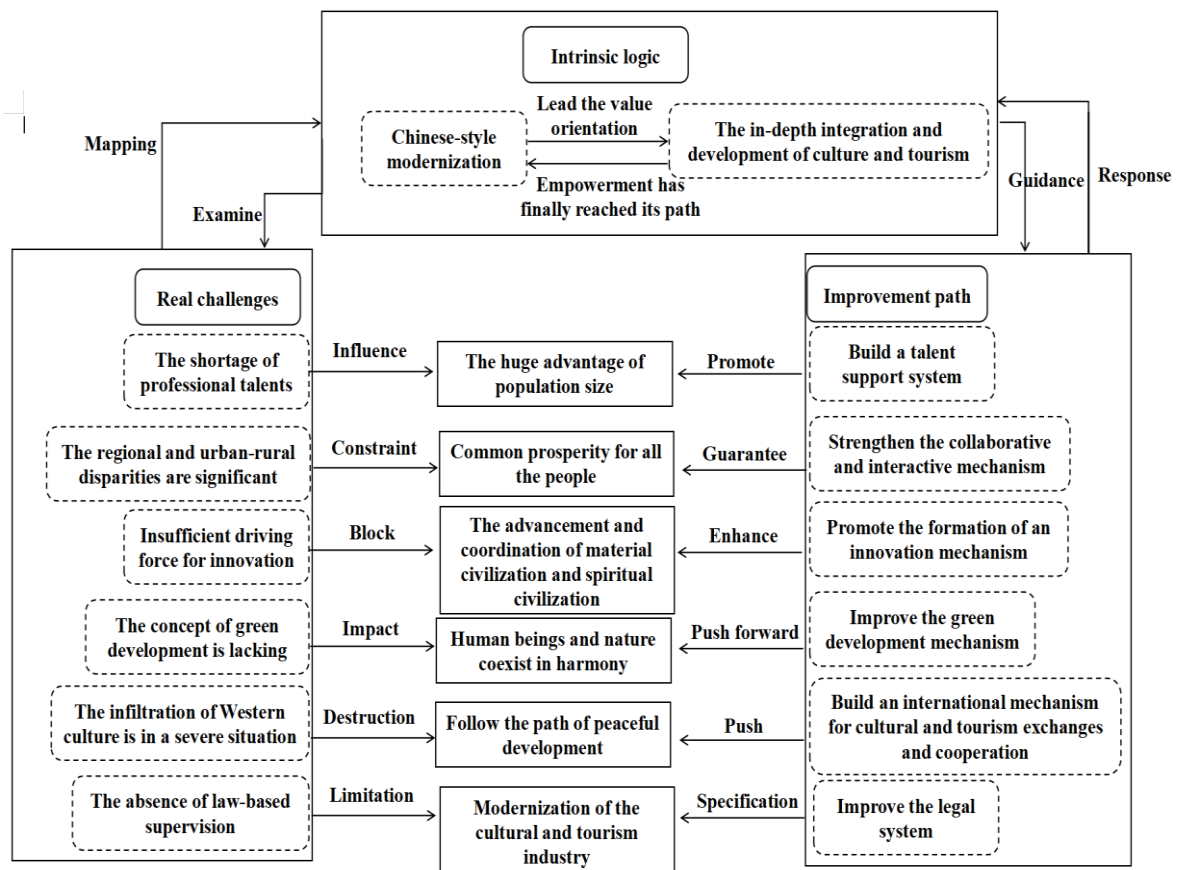
The theory of systemic functional linguistics provides the most suitable theoretical framework for multimodal research. This theory takes the social symbol view as the basic cognitive approach, and multimodal discourse analysis has developed precisely based on this theory. From the perspective of social semiotics, all modalities belong to symbolic resources. The development of different modalities constitutes a network of meaning correlation and selectable options. The interaction between modalities and models has a potential trend of expressing meaning and plays an important role in meaning creation [4]. It can be seen from this that multimodal speech analysis not only studies the important role of all modalities in expressing the meaning of discourse in discourse, but also explores how various modalities collaborate to prompt discourse to generate more meaning [5].

## **2. Construction of the Analytical Framework for Urban Cultural Tourism Promotional Videos**

### **(1) Urban Cultural and tourism development**

At present, the development of urban culture and tourism is steadily advancing in the direction of deep integration, innovation-driven and technological empowerment [6]. The report of the 20th National Congress of the Communist Party of China proposes that we should adhere to shaping tourism with culture and highlighting culture through tourism, and promote the in-depth integration and development of culture and tourism. This not only summarizes the rich experience of the integrated development of culture and tourism in our country, but also clarifies the development direction of the deep integration of culture and tourism in the new era and on the new journey [7]. In the modern social environment where globalization and informatization interweave, cultural and tourism promotion, as an important means to promote local economic development, enhance the city's image and attract tourists' attention, has shown diversity and innovation with the rise of new media technology in cultural and tourism promotion in various places [8].

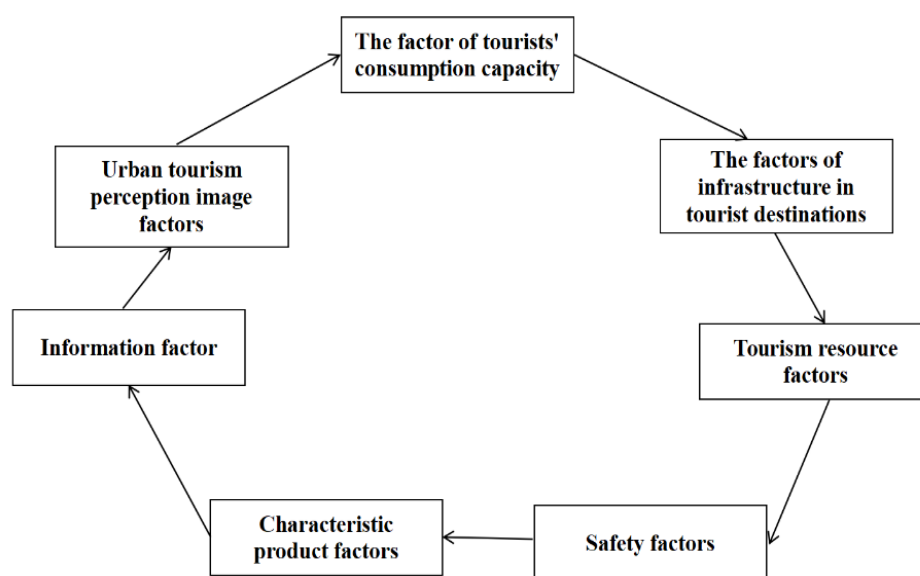
An analysis of numerous cases of cultural and tourism promotion reveals that whether it is an ancient city with profound historical and cultural heritage or a scenic area with beautiful natural scenery, cultural and tourism promotion aims to explore and showcase the local cultural characteristics. The purpose is to enhance the quality of cultural and tourism promotion and display the distinctive traditional culture [9]. From the perspective of the mechanism structure of the deep integration and development of culture and tourism as shown in Figure 1 below, in the process of China's modernization, culture and tourism, as two independent fields, have gradually merged into a whole after the innovation of The Times. We should start from aspects such as professional talents, regional and urban-rural gaps, innovation driving forces, green development, cultural forms, and legal supervision. Only by clarifying the challenges and opportunities faced by the integration and innovation of culture and tourism can theoretical support and effective practice be provided for promoting the development of modernization with Chinese characteristics [10].



**Figure 1: shows the mechanism structure diagram of the in-depth integration and development of culture and tourism**

Taking modern urban facilities as the basic basis and attracting the public based on the rich natural and cultural landscapes and existing service features reserved in the city can further promote the development of urban tourism. Although urban tourism development currently plays a significant role in promoting economic growth and cultural inheritance, in the face of an unstable domestic and international market environment, enterprises and institutions need to deal with challenges such as market competition, infrastructure, and digitalization.

The key to their innovation lies in deeply integrating culture, technology, and tourism to enhance the city's charm and optimize the tourist experience. It is to achieve the sustainable development goals [11]. Based on the development of urban cultural tourism in China in recent years, local tourism development needs to take into account multiple factors, including safety, characteristic products, information, urban tourism perception image, tourists' consumption capacity, infrastructure and tourism resources, as shown in Figure 2 below. On this basis, formulating special strategies to support the digital transformation of the cultural and tourism industry, balancing the construction of regional digital infrastructure, exploring the cultural connotations of cities, and creating attractive cultural and tourism brands based on modern technological means to enhance the popularity and influence of urban tourism can further promote the modernization transformation of the cultural and tourism industry.



**Figure 2: Considerations and Relationships for tourism development**

## (2) Construction of the Analytical framework

For city promotional videos, the relevant modalities are divided into two forms: one refers to the visual modality, and the other refers to the auditory modality [12]. Existing research has found that most scholars mainly apply Kress&Van Leeuwen's vision and method when studying visual modalities. On this basis, the functional grammar of systemic functional linguistics extends to visual modalities. Although this method can be widely used in the study of a single static image, there are still many problems in practical operation. Some scholars have revised and supplemented the above theory, and refined it based on the theoretical framework of visual grammar to ensure its reasonable application in complex visual narratives composed of multiple images, including conceptual meaning, interpersonal meaning and length meaning [13].

The multimodal discourse analysis framework studied in this paper comprehensively applies the concepts proposed by the above-mentioned scholars. When conducting discourse transcription analysis, it is divided into six levels, including the following contents: First, a frame refers to a single image picture; Second, a lens refers to a continuous segment of the picture that remains unchanged when the shooting position of the camera lens does not change.

Thirdly, a scene refers to a series of consecutive performances carried out at one location and one time using two or more shots. Fourth, a stage refers to a series of symbolic modalities that act together. They maintain the consistency of integration within a specific period of time. The meaning expressed by one stage is different from that expressed by other stages in a synchronous work. Fifth, the type steps refer to the beginning, the middle and the end; Sixth, the overall work refers to a specific category, which is fundamentally different from the fact type and the trend type [14].

### **3. Research Design and Result Analysis Based on the Cultural and Tourism Promotional Video of a Certain City**

#### **(1) Research Design**

This paper studies the use of software to take the urban cultural and tourism promotional video produced by a certain place in 2023 as the research object, mainly exploring the following issues: Firstly, what characteristics exist in the different modalities of the local urban promotional video; Secondly, how do the various modalities in the local city's promotional videos construct meanings? Finally, how do various modalities of the local city's promotional videos present the theme [15]? In the multimodal discourse analysis of dynamic videos, two approaches should be chosen.

One refers to the visual grammar theory framework, and the other refers to the system functional grammar theory framework. Moreover, the multimodal corpus technology means are regarded as an auxiliary basis. The audio and video segmentation and annotation tool ELAN6.0 software is selected. After conducting multi-level division and annotation of the city promotional video, a comprehensive discussion is carried out.

After hierarchical processing, it mainly includes the following categories: The first category of marked features is the phonological feature P-, which includes stress (PSt), fast speech rate (PFT), slow speech rate (PST), long pauses (PPL), and short pauses (PPS); The second category belongs to image feature I-, which includes text (IT), image (IP), and text-image (ITP). The third category belongs to Image A-, which includes human images (AP), object images (TP), and human-object images (ATP).

The fourth category belongs to the mode of expression ES-, which includes Narration (ESN), monologue (ESM), and dialogue (ESD). The fifth category belongs to music M-, which includes background music (BGM) and solo music (MS). The sixth category belongs to the shot rate LS-, which includes the slow shot rate (SLS) and the fast shot complaint rate (FLS). The seventh category belongs to the lens movement mode C-, which includes lateral panning (CP), direct panning (CT), zooming/pushing away (CD), panning (CTr), bird 's-eye view (CB), and focus shift (CRF) [16].

#### **(2) Research Results**

##### **1. Visual modality**

Combined with visual grammar analysis, it is found that image semantics contain three forms. Firstly, it refers to the representational meaning. Secondly, it refers to the interactive meaning. Finally, it refers to the compositional meaning. Since urban cultural and tourism promotional videos are usually composed of frame by frame and the theme of the visual simulation mode is images and pictures, the analysis using the ELAN software can yield the annotation results as shown in Table 1 below:

**Table 1: Shows the annotation statistics based on visual modalities**

| mode      | Annotation | Hit | Total annotation duration | Mark the percentage of duration |
|-----------|------------|-----|---------------------------|---------------------------------|
| Animation | AP         | 20  | 73.81                     | 33.573                          |
|           | AT         | 21  | 51.91                     | 23.612                          |
|           | ATP        | 26  | 92.7                      | 42.165                          |
| Image     | IP         | 1   | 0.75                      | 0.341                           |
|           | IT         | 2   | 9.987                     | 4.543                           |

Based on the analysis of the above table, it can be known that this promotional video was created after the city experienced the COVID-19 pandemic. The proportion of video images is lower than that of real images, and the annotation time of both human images and human-object images exceeds 30%, while the appearance time of human images accounts for 75.73%. The research results show that the overall promotional video demonstrates a people-oriented promotional concept. From the perspective of the image, the actual layout starts at the upper left corner and ends at the center. According to the statistics in Table 1, the number of annotations for image text and pictures is smaller. Meanwhile, the promotional video contains a large number of light shows, which illuminate the entire night sky, fully demonstrating the prosperity of the local scenery and the tranquility of the city, and optimizing the meaning expression from different perspectives [17].

## 2. Auditory modality

The auditory modality feature of the entire promotional video lies in the absence of narration, dialogue and monologue, while the English modality plays an important role in the overall meaning and emotional expression [18]. According to the analysis of the annotation statistics based on auditory modalities as shown in Table 2 below, it can be known that the annotation time of musical modalities accounts for 96.24%. It is mainly divided into background music and solo music. The former can choose piano pieces and harmonies, while the latter is specially composed music of the same name. Among them, the lyrics and tracks play a positive role in expressing the theme of the video and constructing the meaning.

**Table 2: Is based on the annotation and statistical results of the auditory model**

| mode   | Annotation | Hit | Total annotation duration | Mark the percentage of duration |
|--------|------------|-----|---------------------------|---------------------------------|
| Music  | BGM        | 1   | 150.64                    | 68.519                          |
|        | MS         | 1   | 60.96                     | 27.728                          |
| Tempo  | FT         | 4   | 99.625                    | 45.315                          |
|        | ST         | 6   | 111.075                   | 50.523                          |
| Pause  | PL         | 1   | 4.07                      | 1.851                           |
|        | PS         | 2   | 3.79                      | 1.724                           |
| Stress | PSt        | 5   | 2.82                      | 1.283                           |

Based on the analysis in Table 2 above, the slow pace is significantly more than the fast pace, and the overall music rhythm is relatively gentle. This proves that the city is developing rapidly and peacefully, reflecting the calm attitude towards life and positive life concept of the local residents. The promotional video is divided into three pauses. One pause represents the end of the video, and the two short pauses respectively represent the transition of music and the end of music for emphasis. Tone, intonation and speed of sound are also the main factors influencing auditory modalities. For instance, the sounds of water, birds singing, cheers and the sound of the zither in promotional videos are more likely to create a realistic and beautiful picture in the minds of viewers [19].



### 3. Modal collaboration

Modal collaborative relationship, as the core content of multimodal functional stylistics research, is of vital importance in the research on the interactive significance and promotion and application of urban cultural tourism promotional videos. From the perspective of logical semantics, there are relationships of reinforcement, extension and detailed expression among different modalities. From the perspective of expressive relationships, complementary and non-complementary analyses can be conducted between them.

Moreover, complementarity has the distinction between reinforcement and non-reinforcement, while non-complementarity possesses concepts such as context exchange, inclusion, and overlap [20]. The data of the auditory and visual modalities of a promotional video in a certain place were studied using the ELAN software, and the results shown in Table 3 below can be obtained:

**Table 3: Shows the statistical results of data based on visual and auditory modalities**

| Modal type      | layer       | Annotation quantity | Total annotation duration | Percentage of annotation duration |
|-----------------|-------------|---------------------|---------------------------|-----------------------------------|
| Auditory mode   | Tempo       | 10                  | 210.7                     | 95.838                            |
|                 | Stress      | 5                   | 2.82                      | 1.283                             |
|                 | Pause       | 3                   | 7.86                      | 3.575                             |
|                 | Music       | 2                   | 211.6                     | 96.247                            |
| Visual modality | Animation   | 67                  | 218.42                    | 99.35                             |
|                 | Image       | 3                   | 10.737                    | 4.884                             |
|                 | Lens Speed  | 10                  | 215.859                   | 98.185                            |
|                 | Camera work | 29                  | 85.784                    | 39.019                            |

Combined with the analysis of the above table, it can be known that the overall promotional video design mainly uses image materials to convey information. Among them, 67 images are marked as A-, and the proportion of music M-marked duration is 96.24%. This proves that music also plays an important role in semantic expression. The opening of the promotion should present the expression theme and Logo image. By combining images and text, it is easier for the audience to have a direct perception.

Other modalities, such as text and music, can complement and optimize the image dissemination. In the early stage of creation, text and images have a complementary relationship. In the effective integration of text and images, the text appears along with slow music, which is more likely to capture the audience's attention and clearly present the promotional theme. From this, it can be seen that the overall design of the promotional video mainly focuses on visual modalities. After effectively integrating auditory modalities, it uses the interaction of symbols to form a harmonious whole, fully expressing rich meanings and unique emotions.

### CONCLUSION

In conclusion, based on multimodal discourse analysis of the interactive significance of urban cultural and tourism promotional videos, clarifying the different modal characteristics and corresponding relationships contained therein, and understanding how cultural and tourism promotional videos in various regions of China in the new era use different modalities to shape excellent urban images and build close connections with the general public, this can not only promote the steady development of the urban cultural and tourism industry, It can also contribute to building the path of socialism with Chinese characteristics.

## References

- 1) Xinwei Huang, Hui Yin. Research on the Construction of Interactive Meaning in German Real Estate Advertising from the Perspective of Multimodal Discourse Analysis [J]. Journal of Zunyi Normal University, 2023, 25(4):80-83.
- 2) Yue Jiao, Guizhen Liu, Xiucan Chen, et al. The Construction of Civilized City Image from the Perspective of Multimodal Critical Discourse Analysis: A Case Study of Cangzhou City's 2021 National Civilized City Creation Promotional Video [J] Cultural Industry, 2023(5):152-156.
- 3) Lu Huang, Lifang Chen. Analysis of Dongguan's City Image from the Perspective of Multimodal Discourse: Taking the Anti-epidemic Promotional Video "Dongguan, We Protect It Together" as an Example [J] Overseas English, 2023(9):73-75.
- 4) Wenqian Yan. The Construction of Changsha's Urban Cultural and Tourism Image from the Perspective of Multimodal Metaphor and Metonymy: A Case Study of "Changsha, The Home of Frequent Visits" [J] Jin Gu Cultural Creativity, 2024(15):91-94.
- 5) Yandi Liu. Multimodal Discourse Analysis of Shanxi's Image Promotional Videos from the Perspective of Visual Grammar: A Case Study of "Multi-Faceted Shanxi, Coming Headfirst" [J] Science Communication, 2024, 16(12):113-116.
- 6) Lihua Li, Yaqian Niu. Multimodal Discourse Analysis of Chengdu City Image Promotional Video from the Perspective of Visual Grammar [J]. Journal of Qiqihar University (Philosophy and Social Sciences Edition), 2024(11):115-119.
- 7) Can Cui. Research on Multimodal Discourse Analysis and International Communication of English News on the "Dual Carbon Goals" of China's International Media [J] Journal of Huangshan University, 2023, 25(4):92-96.
- 8) Zhiying Gao, Jinping Mi, Jingbo Zhao. Multimodal Discourse Analysis of English Promotional Videos: Taking the "Shanxi 8 Minutes" English Promotional Video as an Example [J] Journal of shanxi datong university (social science edition), 2023, 37 (6): 100-104. The DOI: 10.3969 / j.i SSN. 1674-0882.2023.06.020.
- 9) Yixing Zhou, Mei Dong. A Comparative Analysis of the Interactive Significance of City Promotional Videos from a Multimodal Perspective: A Case Study of "Thousand-Year Ancient Capital, Always Come to Chang'an" and "The World of Luoyang" [J] Spread of science and technology, 2023, 15 (16): 59-63. DOI: 10.3969 / j.i SSN. 1674-6708.2023.16.015.
- 10) Jin Yan, Mandy Long. Research on the Translation and Dissemination of Hunan Intangible Cultural Heritage from a Multimodal Perspective: A Case Study of the Documentary "Changsha Hakka Customs and Scenery" [J] Contemporary linguistics, 2023, 11 (7): 2935-2941. The DOI: 10.12677 / ML. 2023.117399.
- 11) Jing Wang, Zhijing Qiu. Multimodal Discourse Analysis from the Perspective of Visual Grammar: A Case Study of the Chengdu Universiade Promotional Video "Shu Road Opens · Universiade Arrives" [J] Overseas English, 2024(13):77-79.



- 12) Weiwei Chen. The Construction of China's National Ecological Image through "Low-Carbon" Promotional Videos from the Perspective of Multimodal Ecological Discourse Analysis [J] Media Forum, 2024, 7(12):47-50.
- 13) Renyan Ji. Multimodal discourse analysis of the poster of 霸王别姬 from the perspective of visual grammar [J]. Comedy World (Mid-month issue), 2023(12):0049-0051.
- 14) Xue Yanyan. Translation Strategies for Cultural Publicity Short Videos from the Perspective of Multimodal Discourse Analysis: A Case Study of Shandong's Publicity Short Videos [J] Jining college journal, 2023, 44 (6): 69-73. The DOI: 10.3969 / j.i SSN. 1004-1877.2023.06.012.
- 15) Douzi Zhu. Research on the Video Dissemination of Chinese Cultural Propaganda from the Perspective of Multimodal Discourse Analysis: Taking the Bilingual Micro-Video "Chinese Style" as an Example [J] Contemporary linguistics, 2024, 12 (12): 456-465. The DOI: 10.12677 / ml. 2024.12121172.
- 16) Heying Hou, Mengjie Zhang, Lihong Gu. Research on Urban Image Communication from the Perspective of Cultural Discourse: A Case Study of Baoding's Youth-Friendly City [J] Communication and Copyright, 2023(22):75-78.
- 17) Ronghan Zhou. Research on Xinhua News Agency's Online Sports Communication from the Perspective of Multimodal Discourse Analysis: A Case Study of "My Asian Games" [J] Editorial and Reporting, 2024(11):64-66.
- 18) Yutong Li. The Meaning Construction of English Translation of Kunqu Opera Scripts from the Perspective of Multimodal Discourse Analysis: A Case Study of the English Translation of "The Everlasting Palace" [J] Modern and Ancient Cultural Creativity, 2023(6):92-95.
- 19) Yue Cheng, Jiayin Chen, Lin Li, et al. Research on Multimodal Meaning Construction and Interaction Mechanism of Ceramic Cultural Landscape Network Platform [J]. Acta Ceramica Sinica, 2023, 44(2):375-381.
- 20) Xindong Wang. Multimodal Discourse Analysis of Movie Posters: Taking the Poster of the Movie "The Big Event in Life" as an Example [J] Comedy World (Second Half of the month), 2023(3):154-156.