

# Postprint: A Study on the Construction of China's Urban Image through Travel Vlogs on Bilibili in the Context of Audiovisual Symbols

**Authors:** Liu Jing

**Date:** 2023-03-24T00:00:00+00:00

## Abstract

The success of urban image communication bears directly upon a city's core competitiveness and sustainable development. In recent years, audio-visual communication has emerged as the principal vehicle for urban image dissemination, with travel Vlogs representing a novel modality for city image construction. Characterized by a grassroots perspective, these Vlogs present and shape urban imagery, thereby engendering profound city impressions among viewers. This study examines domestic travel Vlogs on Bilibili, China's premier video platform, classifies and analyzes the audio-visual symbols contained therein to investigate how such symbols construct Chinese urban images, and proposes development recommendations for leveraging travel Vlogs in the construction of China's urban image, grounded in practical realities.

## Full Text

### Preamble

**Abstract:** The success of urban image dissemination is directly related to a city's core competitiveness and sustainable development. In recent years, audio-visual communication has become the primary means of urban image transmission, with travel Vlogs representing an emerging method for shaping urban image from a civilian perspective, thereby fostering profound urban impressions. This paper examines domestic travel Vlogs on bilibili (hereinafter referred to as "Bilibili"), China's leading video platform, conducting a systematic classification and analysis of their audio-visual symbols to investigate how these symbols construct China's urban image, and offering practical recommendations for the future development of urban image construction through travel Vlogs.

**Keywords:** urban image; semiotics; audio-visual symbols; Vlog

The renowned American sociologist Lewis Mumford defined urban image as a composite formation resulting from the combined influences of mass media, personal experience, interpersonal communication, memory, and environmental factors[?]. In today's globalized environment, urban image and the effectiveness of its dissemination have become critical metrics for evaluating urban competitiveness. In recent years, audio-visual communication has enriched urban image representation through multi-dimensional audio-visual language, establishing itself as a primary mode of urban image dissemination. Among these, travel Vlogs have emerged as a new method for documenting and disseminating urban image. Unlike official promotional frameworks that tend to be grandiose, travel Vlogs construct urban image from a civilian perspective.

Bilibili serves as the primary distribution platform for travel Vlogs in China. Consequently, this paper conducts a semiotic analysis of the audio-visual symbols in travel Vlogs on Bilibili to explore the mechanisms and pathways through which these Vlogs construct China's urban image.

## **I. Urban Image and Communication Media**

### **(I) Urban Image and Urban Image Dissemination**

The concept of urban image was first proposed by scholar Kevin Lynch in the 1960s, who conceptualized it as comprising paths, edges, districts, nodes, and landmarks[?]. This definition emphasizes the material aspects of urban image while neglecting subjective human impressions. Research on urban image in the mid-20th century focused primarily on urban architecture, landscape, and functional design. Since the 21st century, with the continuous development of media technology, research on urban image dissemination has proliferated. Studies on urban image and its dissemination both domestically and internationally encompass paradigm research, disciplinary construction, strategy research, communication effects research, and research on media technologies in communication.

The maturation of media technology and forms has injected new connotations and aura into urban image. As technology iterates and upgrades, imaging technology has significantly advanced and expanded, with production, processing, and dissemination techniques gradually maturing. Audio-visual communication has broken the traditional pattern of text and image content dissemination, characterized by multi-sensory, multi-dimensional, digital, and contextual features, becoming one of the primary methods of urban image dissemination.

### **(II) Travel Vlogs: A Civilian-Perspective Medium for Urban Image Dissemination**

Vlog (Video Blog) first gained popularity on the foreign video platform YouTube in 2012, but only became trendy in China in 2018. Vlogs involve creators recording videos from a first-person perspective to present their lives, thoughts, and

ideas, featuring strong personal expression and typically lasting 5-10 minutes. As a genre of Vlog, travel Vlogs document the creator's experiences, environments, and reflections during travel, employing numerous audio-visual symbols that carry rich meanings beneath their surface representations. Notably, these symbols not only convey the creator's personal feelings and thoughts but also construct China's urban image and facilitate its dissemination.

In the traditional media era, the subjects of urban image audio-visual dissemination were official, with governments and mainstream media constructing and disseminating urban image primarily through city promotional videos and documentaries. From this official perspective, narratives tended to be grandiose, presenting macroscopic views of cities and embedding urban style and spirit within the content. In the new media era's UGC (User-Generated Content) production mechanism, ordinary citizens can also independently create content, becoming one of the subjects of urban image dissemination. Travel Vlogs construct urban image from a civilian perspective, typically presenting microscopic and fragmented perspectives imbued with the creator's subjective impressions of the city. Consequently, the construction of China's urban image is no longer confined to official discourse, demonstrating characteristics of diversified expression forms, pluralistic discourse, and multi-dimensional angles. Universal participation in urban image construction also greatly benefits the sharing of urban memories and the co-construction of urban impressions.

## II. Classification of Audio-Visual Symbols in Domestic Travel Vlogs

### (I) Classification of Visual Symbols

Visual symbols are non-linguistic symbols composed of line, light, color, force, expression, balance, form, and other symbolic elements used to convey various information[?]. Visual symbols include image symbols and text symbols. In urban image construction, visual symbols directly affect human senses at the first level, presenting the "appearance" of urban image; at the second level, visual symbols contain abstract imagery of objects and can enable viewers to comprehend the connotations of urban image through isomorphic associations between symbols.

**1. Image Symbols** Image symbols in domestic travel Vlogs published on Bilibili consist of urban natural landscapes, cultural landscapes, and people. The *Modern Dictionary of Geographical Science* defines natural landscape as a natural complex formed by the interconnection of natural landscape elements, representing the general term for both natural and human-modified natural aspects. Within this scope, natural landscapes presented in domestic travel Vlogs include natural heritage such as mountains, rivers, farmland, orchards, gardens, as well as nature reserves and parks. For example, the Vlog *Suzhou VLOG/The Jiangnan Water Town is So Beautiful!* by content creator "Miwivon" presents

Suzhou's river and garden scenery from a first-person experiential perspective, showcasing the city's unique natural landscape.

Cultural landscapes, in contrast to natural landscapes, refer to landscapes altered by human activities. Urban cultural landscapes in domestic travel Vlogs comprise material and non-material cultural landscapes. Material cultural landscapes primarily include urban architecture, streets, landmarks, etc., while non-material cultural landscapes include historical culture, folk customs, and lifestyle concepts. The Vlog *A Sophomore Girl's First Solo Trip: Recording the Super Healing Summer of Qingdao* by content creator "Peach Bead" presents Qingdao's streets, architecture, the famous landmark University Road, and the relaxed and leisurely nightlife of locals, imbued with a strong sense of daily life.

Image symbols also include human symbols. The most frequently appearing figures in travel Vlogs are the Vlog creators themselves, as well as local residents and tourists. Through facial expressions and body movements that convey different emotions and feelings, the city is not presented as purely material but is imbued with vitality. As components of the urban living organism, people also constitute an important element of cultural landscapes and influence urban image.

**2. Text Symbols** Text symbols in domestic travel Vlogs published on Bilibili primarily include subtitles and text within urban landscapes. Subtitles refer to text added during post-production, frequently appearing at the beginning, middle, and end of travel Vlogs with the creator's unique personal style, serving functions of introduction, explanation, and interaction, helping viewers understand the content and evoking emotional responses. Text as components of urban landscapes typically appears on shop signs, scenic spot plaques, building billboards, etc. In the Vlog *After Leaving, I Started Traveling Alone • Suzhou* by content creator "Shi Ruo'a," the camera records the plaque "Lotus Wind on Four Sides" on the Lotus Wind on Four Sides Pavilion in the Humble Administrator's Garden, supplemented by the subtitle "Every place's name corresponds to its scenery," before cutting to images of red water and lotus leaves on all four sides of the pavilion. This maximizes the effectiveness of text symbols in video, presenting a leisurely and tranquil Suzhou urban image.

## (II) Classification of Auditory Symbols

Auditory symbols are symbols that convey meaning at the auditory level, including linguistic auditory symbols and non-linguistic auditory symbols. Linguistic auditory symbols refer to human speech. Non-linguistic auditory symbols consist of paralanguage and other sound symbols, where paralanguage refers to variations in pitch, tone, volume, and speed of human vocal sounds, while other sound symbols are composed of musical instruments and other natural sounds[?].

**1. Linguistic Auditory Symbols** In travel Vlogs, linguistic auditory symbols primarily include synchronous sound language and post-production

voiceover language. Synchronous sound language refers to linguistic auditory symbols recorded by Vlog creators during filming, serving functions of introduction, explanation, and emotional expression. During travel, synchronous sound language typically represents immediate, relatively authentic reactions and expressions. In the Vlog *Romance and Solitude by the Sea in Winter Like a Movie/Aranya Healing Journey VLOG* by content creator “Lü Zhengmao MaoMao,” which documents a trip to Aranya in Qinhuangdao, extensive use of synchronous sound language systematically introduces local architectural styles, the famous landmark Lonely Library, and seaside scenery, focusing on a corner of the city and showcasing Qinhuangdao’s unique urban charm.

Post-production voiceover language refers to linguistic recordings made by Vlog creators during video post-processing, which can run through the entire video or serve as supplementation and embellishment beyond synchronous sound. Travel content creator “Wild Wang Xiaotu” uses extensive post-production voiceover language in the “3 Days 2 Nights” column Vlogs, adopting a third-person perspective and using concise language to introduce the city’s characteristic natural and cultural landscapes along a chronological travel timeline, presenting urban image multi-dimensionally.

Whether to use linguistic auditory symbols in travel Vlogs depends on the creator’s preferences, style, and the characteristics of the urban image. Additionally, non-linguistic auditory symbols also play supporting roles in travel Vlogs.

**2. Non-Linguistic Auditory Symbols** Non-linguistic auditory symbols in travel Vlogs primarily include background music and non-linguistic synchronous sounds. Background music refers to music played to create a certain environmental atmosphere[?]. In travel Vlogs, background music is mainly used to complement atmosphere and environment, render emotions, and express feelings, harmonizing with video content and rhythm. Non-linguistic synchronous sounds, as opposed to synchronous sound language, refer to non-linguistic auditory symbols recorded by video creators during filming. In travel Vlogs, non-linguistic synchronous sounds primarily include paralanguage, musical instrument sounds, and natural sounds. For example, content creator “A ASKA” in the Vlog *Healing Travel Manual·Say Goodbye to All Troubles/Incomplete Dali Travel Guide Vlog* matches different scenes with different styles of background music and records natural sounds such as lake water sounds, bird calls, and container collision sounds, fully embodying Dali’s natural and humanistic atmosphere.

### III. Construction of China’s Urban Image through Audio-Visual Symbols in Travel Vlogs

#### (I) The “Signifier” and “Signified” of Audio-Visual Symbols

In Saussure’s semiotic theoretical framework, signs are divided into signifier and signified, where signifier refers to the phonetic image of a sign and signi-

fied refers to its intrinsic meaning. When signs and sign systems function, they produce two signifying sequences: “denotation” and “connotation”[?]. Roland Barthes, as the founder of urban semiotics, applied post-structuralism to urban studies and innovatively applied semiotic theory to visual communication, developing non-coded iconic messages and coded iconic messages based on linguistic messages. Non-coded iconic messages refer to direct signification, which is denotative images; coded iconic messages refer to connotative signification, which is connotative images[?].

Domestic travel Vlogs combine visual and auditory symbols to directly present urban landscapes at the signifier level, while conveying urban temperament, customs, and cultural heritage at the signified level. For instance, in the previously mentioned Vlog *Healing Travel Manual· Say Goodbye to All Troubles/Incomplete Dali Travel Guide Vlog*, the creator records Dali’s natural scenery of Cangshan Mountain and Erhai Lake, the low-key and unadorned style of Xizhou Ancient Town, and local specialties such as rice noodles and tea in clay pots, interspersed with natural water sounds, bird calls, and diners’ conversations, complemented by suitable background music. Beyond introducing urban landscapes, subtitles such as “The scenery is so beautiful,” “Eating in the courtyard is particularly pleasant,” and “Returning to the garden is not easy, maintaining simplicity is even more rare” express the creator’s affection for Dali and reflect the city’s characteristics from the side. This Vlog combines multiple audio-visual symbols to vividly express Dali’s simple, tranquil, and healing urban temperament.

## (II) Simile and Metaphor in Audio-Visual Rhetoric

Rhetoric refers to techniques that enhance the persuasiveness or artistic effect of words or sentences[?]. Audio-visual rhetoric, within the rhetorical framework, can be understood as techniques that use audio-visual symbols to enhance the persuasiveness or artistic effect of audio-visual works. In constructing China’s urban image, travel Vlogs primarily employ two rhetorical devices: simile and metaphor. Simile in audio-visual rhetoric refers to techniques that express metaphorical meaning by using words such as “like” or “as,” or by connecting the tenor and vehicle. In the Vlog *Qingdao Travel Vlog/Summer Means Going to the Beach with Good Friends/The Sea Never Gets Old* by content creator “Shi Dabi Qing Ya,” one scene shows the content creator and friends walking hand-in-hand on the grass of Qingdao’s Wheat Island, with blue sky and white clouds in the distance connecting to the azure coastline. The corresponding subtitle reads, “This moment is truly so beautiful, it feels like we’re walking on clouds, indescribably beautiful.” The combination of non-linguistic and linguistic symbols more explicitly clarifies the simile relationship, thereby shaping urban image.

Metaphor in audio-visual rhetoric, also known as “metaphor,” refers to rhetorical techniques that use copulative conjunctions or omit connective words while using tenor and vehicle. A metaphorical relationship exists between travel Vlogs and urban image, where famous landmarks and local cultural symbols serve as

synonyms for a city—for instance, Suzhou gardens and Su embroidery typically serve as synonyms for Suzhou, while hot pot and “mountain city” typically serve as synonyms for Chengdu. In travel Vlogs, creators often emphasize these visual symbols through cinematography, supplemented by linguistic auditory symbols for explanation, using suggestive and relatively concealed techniques to connect distinctive urban landscapes (objects) with urban image, naturally constructing urban image.

### (III) Myth: The “Naturalization” of Urban Cultural Image

In the semiotic perspective, myth refers to “ways of thinking about people, products, places, and ideas constructed to convey particular messages to audiences.” It is both an expression of the relationship between humans and the world and an “anonymous ideology”[?]. In audio-visual communication practice, the completion of cultural ideology construction and dissemination requires both communicators and audiences to share a common space of meaning—that is, to share common cultural backgrounds and cultural significance. Travel Vlogs use numerous audio-visual symbols in combination to shape urban landscape appearances, predominantly employing first-person perspectives that generate immersive experiences for viewers, achieving “bodily presence.” Based on presenting urban cultural symbols, audio-visual language further reinforces urban cultural image. Through the “repeated emphasis” of increasingly similar content, urban cultural image becomes constructed. For example, in recent years, besides cuisine, the most frequent theme in Chongqing travel Vlogs has been the city’s complex, vertically interwoven urban roads. In the Vlog *Chongqing Vlog/First Trip with Friends/Hot Pot, Light Rail, Night View, Store Exploration, Food Tour/Truly So Happy* by content creator “Liu Chongchong,” the difficulty of hailing taxis in Chongqing is documented, with the comment “Chongqing’s map needs to be made in 3D.” The presentation of many such videos has increasingly embedded the image of Chongqing as an “8D magical city” in public consciousness, attracting numerous tourists to experience it and creating a new “myth.”

## IV. Development Suggestions for Constructing China’s Urban Image through Travel Vlogs

### (I) Enhancement: Employ Diverse Audio-Visual Symbols to Highlight Urban Cultural Connotations

Limited by video creators’ personal capabilities and qualities, domestic travel Vlogs on Bilibili generally suffer from low content quality, monotonous use of audio-visual symbols, and often remain at superficial levels of signification without conveying deeper urban culture. In response, Vlog creators should collect as many natural and cultural images with urban characteristics as possible during the material acquisition stage, while also paying attention to collecting synchronous sounds such as mountain and forest wind sounds, river and lake water

sounds, and other natural sounds, as well as human voices and sounds from other objects. During video shooting, attention should be paid to camera movement techniques and the selection of shot sizes and distances. In the video production stage, high-quality subtitle copy should be produced, and suitable background music should be selected for different landscapes. By employing diverse audio-visual symbols and applying appropriate audio-visual rhetoric, urban culture and connotations behind urban landscapes should be reflected, injecting the “soul” of value and spirit into the “skeleton” of urban landscapes to present vibrant urban images.

### **(II) Collaboration: Joint Official-Civilian Narratives to Enrich Perspectives on Image Construction**

Content expression in civilian-produced travel Vlogs tends to be fragmented, with relatively one-sided narrative angles focusing only on one or several urban characteristics, such as only natural or cultural landscapes, making it difficult for viewers to form holistic impressions of urban image based solely on these scattered contents. This necessitates collaboration between civilian and official subjects, where official subjects provide opportunities, rules, resources, and guidance, utilize video platforms to create topics for disseminating positive urban image, and civilian subjects conduct video creation based on these foundations to expand narrative perspectives on urban image. During the dissemination stage, official subjects should leverage their influence and guiding power to amplify the spread of high-quality videos and present urban image in diversified ways.

### **(III) Innovation: Excavate New Urban Symbols and Guard Against Urban Stereotypes**

Currently, travel Vlogs on platforms exhibit homogenization tendencies in presenting the same city’s image. Creators tend to select urban symbols with high popularity for sightseeing and check-ins, causing urban image to become fixed and “internet-famous,” such as Chengdu’s hot pot, Xiamen’s Gulangyu Island, and Tianjin’s “Tianjin Eye” Ferris wheel. Over time, this reinforces people’s stereotypes about cities and weakens the diversity of urban image. In response, video creators should focus on urban heterogeneity, deeply experience urban natural and cultural landscapes, excavate new urban symbols, explore their cultural and historical characteristics, and emphasize the culture, spirit, and values embodied by urban symbols in Vlogs, allowing urban image to be “de-labeled” and presenting authentic and vivid urban image in multi-dimensional and three-dimensional ways.

Cities are not only areas for human survival and activity but also carriers of human civilization development. Currently, travel Vlogs have become a new method for recording urban features, landscapes, and culture, as well as constructing and disseminating urban image. However, problems such as superficial symbolism, fragmented content, and homogenized image still exist in the

process of constructing urban image. This urgently requires people to delve into cities, observe, experience, and comprehend urban internal culture and values with heart, depict urban image with diverse symbols and multiple perspectives, and better promote the construction and dissemination of China's urban image.

## References

- [1] Lewis Mumford. *The City in History: Its Origins, Its Transformations, and Its Prospects* [M]. Translated by Song Junling, Ni Wenyan. Beijing: China Architecture & Building Press, 2005.23
- [2] Kevin Lynch. *The Image of the City* [M]. Translated by Xiang Bingren. Beijing: China Architecture & Building Press, 1990.
- [3] Ying Ming, Luo Yi. Cross-cultural Communication in the Context of Visual Symbols [J]. *Contemporary Communication*, 2007(06):22-24.
- [4] Li Yuqin. The Relationship Between Auditory Non-linguistic Symbols and Emotion in Television Communication [J]. *Art Panorama*, 2005(02):55-56.
- [5] Wang Sen, Yu Langtian. *Introduction to Music Communication Studies: The Intertextual Construction of Music and Communication* [M]. Chongqing: Southwest China Normal University Press, 2008.
- [6] Wu Huifan. Signification and Narrative: A Semiotic Interpretation of Urban Culture Communication [J]. *Contemporary Communication*, 2018(03):31-34.
- [7] Zhang Jing. On Roland Barthes' Semiotics and Visual Communication Theory [J]. *Humanities Magazine*, 2013(09):70-74.
- [8] Zhao Yiheng. The Main Form of Rhetoric Revival: Semiotic Rhetoric [J]. *Academic Monthly*, 2010, 42(09):109-115.
- [9] Roland Barthes. *Mythologies: Critique and Truth* [M]. Translated by Tu Youxiang et al. Shanghai: Shanghai People's Publishing House, 2009:175.

*Note: Figure translations are in progress. See original paper for figures.*

*Source: ChinaXiv — Machine translation. Verify with original.*