

---

AI translation · View original & related papers at  
[chinarxiv.org/items/chinaxiv-202507.00112](http://chinarxiv.org/items/chinaxiv-202507.00112)

---

# Sound Design Elements in Digital Publishing: A Preliminary Study (Postprint)

**Authors:** Han Jie

**Date:** 2025-07-09T15:43:20+00:00

## Abstract

**Purpose:** This study investigates the morphology, composition, and current development of sound design elements in digital publishing. **Method:** Through conceptual introduction, it introduces the evaluative elements of “sound design” from the film and television industry into the field of digital publishing. **Results:** The main body of this research is constructed from four aspects: morphological analysis, articulation of general elements, discussion of distinctive strategies, and soundscape. **Conclusion:** As a medium, sound has undergone carrier shifts and transformations in content status within the publishing industry. The boundary of sound in publishing is dissolving, and the creative space for sound design in digital publishing is infinitely vast.

## Full Text

### Preamble

#### A Preliminary Exploration of Sound Design Elements in Digital Publishing

People's Music Publishing House, Beijing 100010

## Abstract

**Purpose:** This study investigates the forms, composition, and current development of sound design elements in digital publishing.

**Method:** Through conceptual introduction, evaluation criteria for “sound design” from the film and television industry are introduced into the digital publishing domain.

**Results:** The main body of research is constructed from four aspects: morphological 梳理, 陈述 of universal elements, discussion of differentiated strategies, and soundscape.

**Conclusion:** As a medium, sound in the publishing industry has undergone

carrier iteration and a transformation in content status. The boundaries of sound in publishing are dissolving, and the creative space for sound design in digital publishing is infinitely vast.

**Keywords:** digital publishing; sound design; soundscape; new reading; panoramic sound

**Classification:** G244

**Document Code:** A

**Article ID:** 1671-0134(2025)04-113-04

**DOI:** 10.19483/j.cnki.11-4653/n.2025.04.023

**Citation Format:** Han Jie. A Preliminary Exploration of Sound Design Elements in Digital Publishing[J]. China Media Technology, 2025, 32(4): 113-116.

## 1. The Evolution of Sound Design in Publishing

### 1.1 The Analog Era: Genesis of Sound Design Thinking

On July 18, 1877, American inventor Thomas Edison recorded and played back his own recitation using the phonograph in his laboratory. Over the subsequent century, advancements in sound recording and amplification technology laid the technical groundwork and hardware support for introducing sound media into publishing, leading to early forms such as books accompanied by disks (tapes). This “text + sound” format reflected initial sound design thinking, where planners and authors consciously introduced sound media into print books to better convey information and enable readers to receive it more intuitively. However, this three-dimensional expression represented a combination of two media, with sound media remaining relatively independent and largely serving a supporting role.

### 1.2 The Digital Era: Comprehensive Participation of Sound Media

The development of digital technology has ushered all fields of human activity into the digital age, and in the publishing industry, analog-based disks (tapes) have gradually been replaced by digital technologies such as optical discs, integrated circuit cards, flash drives, and online publishing [1]. This technological progress has not only improved the representation of sound itself but also created space for sound design to flourish. In terms of sound ontology, the three fundamental parameters—sampling rate, quantization precision, and number of channels—have been comprehensively optimized in the digital era. For publications themselves, sound has become the main expressive element in publications represented by audiobooks, while sound also constitutes an indispensable important element in other rich-media electronic publications.

### 1.3 The 5G Era: Awakening of Sound Design's Subjective Consciousness

From the introduction of VR concepts in the 1992 film *The Lawnmower Man*, Nintendo's release of the Virtual Boy headset product in 1995, Linden Lab's launch of the multiplayer online game *Second Life* in 2003, to the successive market launches of hardware devices such as Oculus, Cardboard, Vive, and PS VR, and the current popularity of the Metaverse concept, the research and development of virtual reality technology and equipment has come a long way [2]. In the 5G era, sound design for VR-based "new reading" has begun to transform from simply "being able to hear" in the analog era and "hearing beautifully" in the digital era toward "being immersed in sound." The spatial sense, presence, and envelopment of sound have given wings to 5G "new reading." Panoramic sound based on VR publishing has made sound an important carrier of expression, and the subjective consciousness of sound design has been fully awakened. Sound creation centered on content has truly come to the forefront, where sound not only serves content but can also expand and sublimate it.

## 2. Universal Elements of Sound Design in Digital Publishing

Sound design in digital publishing basically follows the fundamental process of "production—packaging—analysis—restoration." Beyond this, design elements include four basic aspects: "language, sound effects, music, and sound logos."

### 2.1 Language

Language is the sound symbol through which humans communicate and an important tool for human interaction. Unlike the abstract meaning of music and sound effects, language expresses meaning directly, and its semantic indication function can complete direct connection of sound images. From the perspective of language's expressive attributes, both human language and music share common characteristics of pitch, timbre, rhythm, and emotion. Their "cadence, emphasis, and pacing" sometimes give them expressive power equal to music. Combined with the soft, warm, and highly affable characteristics of human voice, language often becomes the preferred medium in digital publishing. Audiobooks represent this form, and platforms such as Ximalaya, Qingting FM, Litchi FM, Lazy Person's Audiobooks, and Shuqi Novels currently have over 500 million daily active users, becoming the mainstream carrier form with the highest market share in digital publishing.

### 2.2 Sound Effects

Sound effects refer to the effects of sound, including both realistic sounds from nature and non-realistic sounds produced through simulation. In digital publications, they often appear in two forms: first, as prompt elements, such as opening

and closing sound effects, page-turning sound effects, and hyperlink sound effects in audiobooks; second, as auxiliary elements to enhance atmosphere, such as alarm sounds like fire alarms, and sounds of wind, rain, and thunder in plot-driven audiobooks. Together with elements like music and language, they form a “foreground background” perspective relationship to reflect the hierarchy and presence of the entire sound design.

### 2.3 Music

The non-objectivity and non-semantic nature of music art determines its strong inclusiveness, which can express infinitely broad content. Meanwhile, elements such as tonality, harmony, timbre, speed, and rhythm give it strong coloration, providing the best expression for the personalized appeals of publications. Music itself is an extremely common art form with a broad acceptance group and communication foundation, and a large number of readers are willing to read with music. In digital publications, music mostly exists as a “sound background,” presenting three states with language subjects and picture subjects: synchronization, counterpoint, and separation. Synchronization is the high degree of agreement between music and the main form and content, strengthening the main content. Counterpoint is a certain degree of opposition between music and the main content, expressing conflict through certain confrontation. Separation is the distance between music and the main content, guiding readers’ thinking and association through independent statements of both.

### 2.4 Sound Logos

The trend of multimedia convergence in the media industry makes mutual introduction and complementation among visual media, auditory media, and tactile media the norm. Under this concept, logo design gradually introduces auditory elements while strengthening visual experience, becoming the origin of sound logo art. From an artistic attribute perspective, sound logos are sound fragments that use auditory means to express highly condensed brand internal attributes and spiritual connotations. Like visual logos, they have identification functions and represent a combination of art and technology—a form of sound art. Meanwhile, as an art form with strong commercial attributes, the emergence of sound logo art has also benefited from the promotion of commercial trade under the background of globalization. Products need to be promoted and circulated globally, and sound is a promotional carrier that can break national boundaries [3]. In digital publishing, branded audiobooks usually contain sound logo design. For example, the bestselling brand “Miaoboshi” uses a combination of “children’s voice singing + instruments + sound effects” to well outline its positioning for children’s books, making it unforgettable.

### 3. Differentiated Strategies for Sound Design in Digital Publishing

Beyond the four universal elements described above, sound design faces different strategic choices for different publication formats. From the evolution of mono, stereo, and surround sound to the current widespread application of panoramic sound technology in music, film, games, and 5G-based VR publishing, the choice of sound design strategy is determined by product attributes while also being closely related to terminal selection, hardware support, research and development costs, and other elements.

#### 3.1 Mono-Based Sound Design

During the period of integrated circuit carriers, mono design thinking became the main form, determined by hardware limitations. Taking “music cards” as an example, they use transistors to drive piezoelectric ceramic plates to produce sound, with an operating frequency of 300Hz–5kHz. The advantage of this design strategy is that sound signal storage and playback are integrated, achieving high integration of product content and form, with simple and direct communication methods and low production costs. The disadvantage is the lack of low and high frequencies, resulting in thin and harsh sound that lacks aesthetic quality and strong sound design sensibility, only able to transmit basic sound information.

#### 3.2 Stereo-Based Sound Design

The emergence of tape and optical disc carriers brought a qualitative leap to sound expression in digital publishing through stereo (dual-channel). The iteration from tape to optical disc represents a transformation from analog to digital signals. The sampling rate of CD optical discs reaches 44.1Khz, 16Bit, greatly enhancing the spatial sense and clarity of sound and creating a qualitative leap in auditory experience. The widely used “book with disc” format in educational publishing is the main practice of this strategy, where language is the sound subject and other elements are auxiliary. The advantage of this design strategy is that sound elements such as language, music, and sound effects are fused and presented to continuously stimulate readers’ hearing and increase the fun of content transmission, with relatively normal production costs and high cost-effectiveness. The disadvantage is rigid sound space and weak presence. The subsequent stereo format based on online publishing is the online transformation of tape and optical disc carrier content, with no change in sound ontology itself.

#### 3.3 Surround Sound-Based Sound Design

Surround sound design based on 5.1 and 7.1 channels was first applied in the film industry. On the basis of stereo, it creates a two-dimensional sound space by adding 3–5 channels, supplemented by a low-frequency speaker, with frequency

band supplementation making sound thicker and warmer and greatly enhancing listening experience. This sound design scheme has also been applied to concert live broadcasts in recent years, and game products are the most widely applied in digital publishing. As early as 2005, Xbox 360 motion-sensing game products already supported Dolby 5.1 surround sound [4]. In 2017, China's first military game, *Glorious Mission*, used 5.1 surround sound technology to produce game sound effects. To allow players to experience the impact of real guns and live ammunition, its sound design team was stationed at a certain military training base to conduct on-site sound collection [5].

### 3.4 Panoramic Sound-Based Sound Design

Similar to the evolution of visual expression, the expansion from two-dimensional to three-dimensional space is logical. Three-dimensional spatial sound fields represented by Dolby Atmos (Figure 1 [Figure 1: see original paper]) are a revolution in sound technology, supporting up to 34 independent output channels, with typical systems including 5.1.2 (8 channels), 5.1.4 (10 channels), and 7.1.4 (13 channels). The advantage of this strategy is that sound and picture are completely integrated, with extremely strong sound presence. The disadvantage is the need for complete upgrading and reconfiguration of audio-video systems, requiring additional upper-layer speaker layouts and increased total numbers, along with a series of processors capable of supporting three-dimensional immersive sound effect processing capabilities. The domestic panoramic sound solution WANOS also launched in 2017, and by 2022, more than 70 works had applied this system, including *The Wandering Earth*, *Full River Red*, *Hello*, *Li Huanying*, and *The Battle at Lake Changjin* [6]. Wider applications also include games, AR, VR, MR, and other fields, covering terminals including televisions, projectors, PCs, mobile phones, automobiles, headphones, and VR headsets.

## 4. Automation of Sound Creation Methods Under AI Technology

The great development of AI technology has brought changes to sound creation methods. The sudden emergence of IP video traffic requires lower barriers to sound design, while the “5G new reading” model has unprecedentedly high requirements for sound design. All these factors have constructed a new landscape for sound design in digital publishing in the 5G era.

### 4.1 Standardization of Sound Design in the IP Video Context

According to Cisco's Visual Networking Index, global IP video traffic will account for 82% of total network traffic in 2021 [11]. Under these circumstances, sound creation based on IP video presents a different landscape. Most of these videos are drawn from daily life, are short in duration, and are filmed randomly, relying on creativity and curiosity to capture traffic. Therefore, such video plat-

forms provide pre-made sound clips, which can be provided by copyright holders, generated by platforms based on AI technology, or unprocessed sounds recorded by individuals. However, whether such standardized, non-serious creative sound materials can be included in the publishing category, and whether AI-generated music works have copyright ownership issues, pose new challenges to digital publishing.

#### 4.2 Expansion of Sound Expression Space Under “VR+Publishing” Model

In 2015, *Van Gogh Map* published by Electronic Industry Press used VR imaging to restore book content, produced a virtual reality documentary reflecting Van Gogh's life trajectory, and held an exhibition of the same name, representing an early successful case of “VR+publishing” in China. In 2020, “Zhongtu Cloud Innovation” under China Publishing Group, combined with the content characteristics of traditional books, developed related products into ultra-high-definition video products with 3D effects and 720-degree full viewing angles through VR, AR, and other virtual reality technologies, realizing the transformation of books from static pictures and text to dynamic immersive “new reading” scenes, and carried out “5G new reading” cooperation with multiple publishers [12]. Its fully immersive central axis intangible cultural heritage film *Print Central Axis* is a five-fold screen naked-eye 3D immersive film jointly created with the Capital Library. The work takes “historical and cultural resources + panoramic video” as the principle of overall planning, based on content from *Hongxue Yinyuan Tuji*, *Wuyingdian Juzhenban Chengshi*, and *Tangtu Mingsheng Tuhui*, using cutting-edge digital technology as the carrier to highly condense the rich historical and cultural heritage and folk customs of Beijing's central axis [13].

Such innovative cases based on “VR+publishing” provide broad creative space for panoramic sound-based sound design. The resulting sense of sound space, envelopment, and presence has given wings to digital publishing expression. As an independent content subject of publishing, sound media is currently experiencing the gradual decline of optical disc carriers and the rise of online media, yet black vinyl record sales are increasing annually. Sound media as rich media content for digital publishing has undergone carrier iteration and content status transformation from the analog era to the digital and network eras. The boundaries of sound in publishing are dissolving, and the creative space for sound design in digital publishing is infinitely vast.

### References

- [1] Ran Xuemin. Evolution, Boundaries, and Integration: Research on Digital Publishing from a Technical Perspective[J]. *News Research Guide*, 2023(9): 184-187.
- [2] Xinzhi Yuan. From VR to Metaverse: 30 Years of Virtual Reality[J]. *China Industry and Information Technology*, 2022(6): 80-84.
- [3] Han Jie. From Vision to Hearing: Research on Sound Logo Art[J]. *Journal*

of Nanjing Arts Institute (Music and Performance Edition), 2013(4): 52-57+97.

[4] Dolby Digital Surround Sound Brings Cinema Effects to Xbox 360 Games[J]. Computer and Network, 2005(11): 43.

[5] Sina Military. Domestic Game “Glorious Mission” Uses Real Guns and Live Ammunition for Sound Effect Collection[EB/OL]. (2012-06-01)[2024-12-15]. <https://mil.news.sina.com.cn/2012-06-01/1039691981.html>.

[6] WANOS Official Website. Immersive Audio[EB/OL]. (2023-12-01)[2024-12-15]. <https://wanos.cc/home/ImmersiveAudio>.

[7] Yimo Official Website. Understanding “Surround Stereo” at Once[EB/OL]. (2023-01-06)[2024-12-15]. <http://ikmultimedia.100legend.com/news/show/1243.html?agent=ikmultimedia>.

[8] Han Jie, Zhuang Yao. Research on Soundscape Thought in the Context of Electronic Music[J]. Huangzhong (Journal of Wuhan Conservatory of Music, China), 2014(1): 52-57+97.

[9] Machine Heart. Over Ten Thousand Piano Works, More Than One Thousand Hours, ByteDance Releases World’s Largest Piano MIDI Dataset[EB/OL]. (2020-10-27)[2024-12-15]. <https://mp.weixin.qq.com/s/aJhQZ812MgWxC-2gmAb92g>.

[10] Suno Official Website. Work Demonstrations[EB/OL]. (2020-10-27)[2024-12-15]. <https://suno.com>.

[11] Zhao Xiaofang. Preliminary Exploration of Digital Publishing Operation Models in 5G Scenarios[J]. Journal of Hubei Second Normal University, 2023(10): 104-108.

[12] Lin Liying. Practice and Enlightenment of “5G New Reading” in Innovating Publishing Integration Development[J]. New Reading, 2023(2): 14-15.

[13] Zhongtu Cloud Innovation. Zhongtu Cloud Innovation Invited to Participate in 2024 Beijing Cultural Forum Innovation Achievement Exhibition[EB/OL]. (2020-10-27)[2024-12-15]. <https://mp.weixin.qq.com/s/yB9pOt9vIV86xxX9ntgPw>.

## Author Biography

Han Jie (1980–), male, from Qianshan, Anhui, Ph.D., Associate Editor. Currently serves as Deputy Director of Digital Operations Center at People’s Music Publishing House. Research interests include digital publishing and electronic music.

(Responsible Editor: Li Yansong)

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

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