Fundamental Workflow and Techniques of Audio Post-Production for Film and Television: A Case Study of 'The Incarnation' (Post-print)

Authors: Wan Zesi

Date: 2023-10-08T00:00:00+00:00

Abstract

With the development of modern science and technology, film and television works have gained increased opportunities for advancement. Film art plays a non-negligible role in enriching public life and conveying Chinese values. How to immerse viewers in a scene through a single film constitutes the core and vitality of sound in film and television. This article will commence with the fundamental workflow of sound post-production for film and television, introduce the composition of sound and basic post-production techniques, and expound in detail upon the function of music and sound effects in film and television works.

Full Text

Basic Process and Techniques of Film and Television Audio Post-Production: A Case Study of "Incarnation"

Abstract: With the development of modern science and technology, film and television works have gained unprecedented opportunities for growth. As an art form, cinema plays a vital role in enriching public life and transmitting Chinese values. The core and vitality of film sound lie in its ability to immerse viewers in the narrative world through a single work. This paper examines the fundamental workflow of film and television sound post-production, introduces the composition of film sound and basic production techniques, and elaborates on the functions of music and sound effects in film works.

Keywords: film and television; audio production; workflow; sound effects

As film and television production output continues to rise annually and cinematic art advances rapidly, film has become a crucial medium for disseminating

culture, promoting Chinese values, creating rich living environments, and fostering comprehensive, coordinated, and sustainable economic and cultural development. A film work comprises diverse expressive elements, including visuals, sound, special effects, and other artistic treatments. Whether in animated films, science fiction, or live-action productions, sound constitutes a critical element affecting overall work quality. In outstanding film works, effective sound not only reflects social life and conveys genuine emotions but also elevates themes, enhances atmosphere, creates artistic conception, and strikes directly at the heart. Immersing viewers through a single work represents the core and vitality of film sound. This paper explores the fundamental workflow of film and television sound post-production, introduces sound composition and basic techniques, and details the collection and editing methods of music and sound effects while analyzing their functions in film works.

1. Basic Workflow of Film and Television Sound Post-Production

As an audio-visual art form integrating drama, cinematography, painting, music, dance, literature, sculpture, and architecture, film captures images and sound using film stock, videotape, or digital media to deliver exceptional audio-visual experiences. With societal development, film and television have permeated every aspect of public life, becoming an indispensable part of daily existence. Sound represents one of the fundamental elements of the film medium, transforming works from purely visual to audio-visual sensory experiences. Since the release of the first sound film *The Jazz Singer* in 1927, diverse forms of film and television sound have emerged over the past 92 years, establishing unified national industry standards. Currently, China adopts the SMPTE standard for film industry loudness: -20 dBFS, while France and Japan use the EBU standard: -18 dBFS. SMPTE (Society of Motion Picture and Television Engineers) also introduced the timecode concept for temporal synchronization between devices, a counting system with parameters including Hours, Minutes, Seconds, and Frames. The film industry typically references SMPTE 24 Film Sync, representing 24 frames per second playback. Signal formats and recording specifications for film are 48kHz, 24-bit, with file formats using BWF or WAVE.

Most film and television sound revolves around three main components: human voice, natural sound effects, and music. These three categories can be subdivided into dialogue, crowd sounds, background sounds, action sounds, friction sounds, footsteps, special sound effects, and music. In sound post-production, these primarily consist of ADR, Foley, SFX, Music, and BG. Drawing from personal experience in sound production for the thriller short film *Incarnation*, this paper introduces the composition of film audio and its basic production workflow and techniques.

1.1 ADR

ADR refers to Automated Dialogue Replacement, meaning the re-recording in a studio of dialogue that was poorly recorded during production or subsequently modified, to replace the original location sound. Before voice actors enter the studio, the recording engineer must prepare a dialogue script containing each character' s lines with corresponding start timecodes, calculated from the onset of the first phoneme. The script may also note delivery tone, speech rate, and actors' specific linguistic habits to better match the emotional context of the location recording. During formal recording, the shotgun microphone used in location recording is employed, simulating the microphone-to-source distance according to shot requirements to match the timbre and volume of the production sound. The author compared widely-used condenser microphones for vocal recording (such as the Neumann U87) with shotgun microphones used in location recording (such as the Sennheiser 416). Conventional condenser microphones produce exceptionally clear sound quality; as transducers relying on capacitance variation with extremely thin diaphragms, they exhibit high sensitivity and excellent transient response, resulting in pronounced sibilance that lacks the sonic perspective appropriate to on-screen scenes. Shotgun microphones, composed of interference tubes on both sides, cause off-axis sounds to reach the diaphragm with varying phase relationships, resulting in partial cancellation that suppresses lateral sound energy and reduces ambient noise. Employing an appropriate shotgun microphone, with its strong directionality and effective lateral sound energy suppression, enables recording from greater distances than conventional microphones, capturing vocals with high signal-to-noise ratio and clarity.

During recording, voice actors reference timecodes and beep tones, mimicking characters' emotions and delivery for line replacement. Each line should retain adequate breath sounds to create natural transitions. Since sound propagation naturally attenuates high and low frequencies, producing certain coloration, while studio recordings with close microphone placement suffer minimal frequency loss and lack appropriate spatiality, resulting in a "dry" sound that appears too close to the listener and fails to integrate with the narrative. Therefore, post-production requires equalization to reduce certain high and low frequencies, simulating natural propagation characteristics, combined with early reflections and reverb to enhance spatiality. After these effects, the sound becomes more scene-appropriate, eliminating the "two layers" sensation of disconnected audio.

1.2 Foley

Foley refers to the artificial simulation of required sound effects in film or the post-processing of recorded action sounds to achieve effects that cannot be directly recorded. Foley comprises three components: cloth (clothing friction sounds), steps (footsteps), and props (action sounds). Foley pursues "subjective reality" in perception rather than "objective reality" in actual life. For

example, when someone in cloth shoes walks on a wooden floor, the sound in reality is extremely faint and barely noticeable. Due to the masking effect and cocktail party effect, people selectively attend only to information they deem important and valuable. Consequently, real-life footsteps often go unnoticed as very subtle sounds. However, in film and television works, footsteps serve as crucial sound elements; when 特写 shots of feet appear, footsteps are often much louder than in real life. Numerous such examples exist. Since viewers concentrate intensely when watching films, everyday life details in cinema are amplified, requiring corresponding amplification of sound compared to daily life. Additionally, "subjectively real" Foley must match the emotional atmosphere of the scene. For instance, when a killer enters an elevator, the actual door closing sound from location recording is very faint and thin. Yet in the film, the elevator door sound is heavy and muffled, like the gates of hell closing—this results from post-production Foley processing. Viewers do not perceive this processed sound as jarring, unnatural, or incongruous; on the contrary, using the "real" location sound would make audiences feel it "unrealistic and disconnected." This occurs because viewers have been immersed in the film' s atmosphere and accept its emotional framework, thus accepting sounds that align with that mood.

Foley also demands precision and accuracy in action sounds, requiring not only keen intuition for natural sound timbres but also thorough familiarity with scene rhythm to ensure every minute action synchronizes perfectly without misalignment. Foley content primarily consists of detail sounds, so recording must first ensure adequate signal-to-noise ratio. During Foley recording, engineers must avoid cable friction sounds, chair creaking, air impact noises from movements, and even stomach growling from hunger to maintain clean sound. In post-processing, action sounds and footsteps similarly require equalization to simulate sound coloration during propagation, attenuating certain high and low frequencies while adding early reflections and reverb to create more natural spatial integration and avoid the "two layers" effect. Some sounds cannot be directly simulated or lack appropriate sound effects; in such cases, separate sound fragments can be created and then layered. For example, for the "baseball bat hitting the ground" sound in the film, one might first simulate the "heavy impact," "bat bouncing," and "tail sustain" separately, then layer these fragments to create a similar timbral effect. During post-production, new Foley elements not present in the original can also be added according to film details and the director' s vision.

1.3 SFX

Sound effects refer to using sound to create effects that enhance scene realism and atmosphere. Sound effects can be used individually, layered, combined with location sound, or overlaid with Foley or music. Music and sound effect editing tends to be fragmented and labor-intensive. When searching for materials, sound characteristics should be described concisely and specifically, using divergent thinking to find potentially similar sounds. The collection scope should

be broad and wide-ranging, while application should be precise and accurate. Sound effects fall into two categories: those with distinct pitch and those without. Pitched sound effects have relatively narrow application and are not easily layered with music, only usable alone or 叠加 with other non-pitched sounds to enhance effects.

1.4 Music

In music collection, any material matching the intended mood—whether in style, instrumentation, or tempo—should be gathered as richly as possible, with copyright considerations paramount in selection. In music editing, due to the short film duration (5 minutes), a complete piece cannot 贯穿 the entire work. Therefore, the primary approach involves selecting developmental passages, namely the main melodic sections, for collage, with specific methods detailed later.

1.5 BG

BG refers to background ambient sound, typically recorded in stereo during location recording. If location ambient sound is flawed and unusable, appropriate ambient sounds can be sourced from libraries and layered to create natural-sounding environments. At the film' s beginning, the sisters' conversation in the studio appears 平淡 and everyday, yet tension must be conveyed to the audience. Therefore, background sounds with substantial low-frequency content are layered to create a 压抑 and tense atmosphere. Background sound also plays a crucial role in scene transitions and emotional shifts. Different scenes require completely different ambient sounds; marking cue points during scene changes allows for 剪辑 ambient sounds accordingly, then 拼接 with the next scene' s ambient sound and applying appropriate fade-ins creates clean environmental sound transitions that make scene changes more distinct. After adding BG, both ADR and Foley gain significantly enhanced spatiality, with timbres sounding more realistic and eliminating the "in your face" sensation for viewers.

1.6 5.1 Surround Sound Track Building and Mixing

Create tracks for L, C, R, LS, RS, and LFE, with output settings configured for 5.1 surround. For greater operational convenience, L and R can be simplified as front tracks, LS and RS as rear tracks, with panning adjusted to desired positions. For automated panning, simply change the track header' s "read" mode to "latch"mode, then click and hold the pan position to write corresponding automation, creating desired effects.

During export, first export six separate track files from the 5.1 mix, then import them into Sonic Foundry Soft Encode software for encoding. In the software, assign each audio file to its corresponding channel, verify all parameter settings, and export an AC3 file. Finally, place the AC3 audio file into MKV TOOLS software to mux with the video file into an MKV container.

2. Techniques and Functions of Sound Effects and Music

2.1 Sound Effects Usage Techniques and Functions

Sound effects can enhance atmosphere and create artistic conception. When using sound effects, emphasis often lies not on specific sounds but on intentionality—the atmospheric tendency. This tendency manifests in two ways. The first is positive reinforcement: using sound effects matching the narrative and musical atmosphere to strengthen mood. For example, in the scene where the male killer enters the elevator, a faint low-frequency humming sound blends with the elevator' s inherent sound, seemingly present yet elusive, enhancing the 诡异 atmosphere and gradually building audience tension. The second is reverse contrast: using sound effects contrasting with narrative and musical atmosphere to create strong emotional juxtaposition that highlights the creator' s intended emotion. For instance, after the climactic "baseball bat murder," absolute silence is punctuated by a water droplet falling and splashing with several hollow echoes, introducing subtle dynamics and variation into complete stillness and highlighting unease within silence—proving more effective than absolute quiet alone.

Sound effects also serve to bridge music, capture audience attention, and facilitate transitions. Sound effects can be 配合 used when plot emphasis shifts, rhythm changes, or music transitions. In one scene, the camera follows the sister' s gaze from the table where she' s searching to her sister' s suitcase. The film adds a "whoosh" sound similar to a flashback effect, capturing audience attention and 暗示 that the focus is changing, creating a sense of movement that shifts audience attention along with the actor while subconsciously generating curiosity: the focus has changed, but to where? This makes it easier to establish the suitcase' s key 地位 in the narrative. In the elevator scene, sound effects also bridge music and aid transitions. Combined with music, sound effects add greater dynamics and tension. During the confrontation between the male killer and the sister, music 叠加 with progressively intensifying sound effects enhances the musical atmosphere, continuously escalating 恐怖 emotion and pushing the film toward its climax.

Sound effects can also supplement characters' psychological activities and enrich scene environments. During the confrontation scene between the male killer and the sister, a heartbeat sound is inserted to express the sister' s fear and tension. At the ending, the sound of traffic passing outside the window emphasizes post-accident 恐怖感.

2.2 Music Usage Techniques and Functions

Music plays a crucial role in creating artistic conception, with good 配乐 comprising half of a film' s success. Music usage, like sound effects, relies on intentionality. Using a single piece of music throughout would be monotonous, while changing music too frequently disconnects emotional continuity and risks pulling audiences out of the narrative. Music can be changed according to major

plot segments, with different styles and emotions interwoven to avoid monotony.

Music creates artistic conception and establishes atmosphere. The film' s first half uses music with 压抑 emotion to establish atmosphere, while the latter half selects 恐怖-themed music to advance the plot. In the corridor scene where the male killer walks, the film employs 恐怖 music with music box timbres; the higher pitch combined with wind chime sound effects creates an ethereal, supernatural feeling that intensifies 恐怖气息 and prepares the narrative for the upcoming climax. The ending music features not only 诡异 melodies but also supernatural vocal chants that, as the mystery unravels, fully release the film' s 恐怖 atmosphere, driving it to a second climax. To avoid abrupt gaps during music transitions, continuous sound effects can be added to maintain emotional coherence.

Music also controls narrative rhythm. In the same corridor scene, the film employs a relatively 舒缓 and light musical tempo that temporarily eases tension without completely relaxing it, maintaining emotional continuity. This gives audiences breathing room while building toward the climax. When the male killer enters the room, the background music simultaneously shifts to more tense and aggressive music that rapidly elevates the 恐怖意境 and atmosphere, 调动 audience emotions and building anticipation for subsequent events. Alternating between 舒缓 and 紧张 music allows audience emotions to rise and fall with the score, preventing excessive tension or numbness, effectively regulating atmosphere and rhythm.

Every film and work possesses its own strengths, weaknesses, and supporting elements. Understanding the basic workflow and techniques of sound production and mastering the functions of music and sound effects enables films to better convey creators' intentions, making works more vivid, soulful, and moving. Continuously exploring how to create better film and television sound remains the relentless pursuit every film creator should uphold.

References

[1] Zhou Xiaodong. *Handbook for Recording Engineers* [M]. Beijing: China Radio and Television Press, 2006: 13-14.

[2] Wang Jue. *Film and Television Recording Technology and Techniques* [M]. Beijing: Communication University of China Press, 2016: 25-27.

[3] HOLMAN, TOMLINSON. *Multichannel Surround Sound Technology* [M]. Beijing: People' s Posts and Telecommunications Press, 2011: 131-137.

(Author affiliation: School of Music and Recording Arts, Communication University of China)

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