**RESEARCH ARTICLE**

# Creative Paradigms of Handicraft Intangible Cultural Heritage Short Videos: An Integrated Study of Aesthetics, Narration, and Communication in the Digital-Intelligent Era

**Wang Zhenyang[1], Zhang Zeng[2], Hui Yi[3]**

[1] Chengdu Academy of Fine Arts attached to Sichuan Conservatory of Music, Sichuan Province, PRC.E-mail:28460155@qq.com
[2] Sichuan Conservatory of Music, Sichuan Province, PRC.E-mail:theodorezhang@sccm.edu.cn
[3] Johann Wolfgang Goethe-Universität Frankfurt am Main. E-mail:s1628413@stud.uni-frankfurt.de
Corresponding author: Zhang Zeng (e-mail: theodorezhang@sccm.edu.cn).

**Abstract:** Short-video platforms are increasingly important to the communication of craft-based Intangible Cultural Heritage (ICH), but existing research has not sufficiently explained how aesthetic design, narrative organization, and dissemination dynamics work together in this context. This study develops a mechanism-based framework for analyzing craft-based ICH short videos.

A qualitative multi-case study was conducted on 30 Douyin videos produced by three craft-based ICH creators. The corpus was analyzed using multimodal transcription, theory-guided coding, and cross-case comparison.

The results show that effective ICH short videos rely on a recurrent Aesthetics–Narrative–Dissemination chain. Aesthetic strategies enhance procedural visibility and viewer engagement; narrative structuring translates craft processes into culturally resonant meaning; and dissemination mechanisms enable interaction, circulation, and feedback across platform communities. On this basis, the study proposes the A–N–D framework as an analytical model for identifying strong and weak creative configurations in craft-based ICH short videos.

The study extends digital heritage communication research by offering a replicable mechanism for understanding how craft content becomes both culturally meaningful and communicatively effective in platform environments.

**Keywords:** Intangible Cultural Heritage; short-video platforms; multimodal analysis; narrative strategies; digital cultural communication

## 1. Introduction

Craft-based Intangible Cultural Heritage (ICH) represents an important carrier of cultural memory, traditional knowledge, and community identity. In the digital era, the emergence of short-video platforms has profoundly reshaped how such cultural practices are documented, represented, and disseminated. Platforms such as Douyin and TikTok enable creators to present craft processes through highly visual and accessible media formats, allowing traditional craftsmanship to reach broader audiences beyond local communities.

With the increasing popularity of short-video platforms, scholars have begun to explore their role in cultural communication and digital heritage dissemination. Existing studies generally fall into three main strands. The first focuses on audiovisual aesthetics, examining how camera language, editing techniques, and multimodal design contribute to audience engagement and visual attractiveness. The second strand emphasizes narrative strategies, analyzing how storytelling structures, narrative perspectives, and discourse forms translate complex craft knowledge into accessible media content. The third strand investigates platform dissemination mechanisms, focusing on algorithmic recommendation systems, audience communities, and the circulation of cultural content within platform ecosystems (Gao et al., 2023; Gerbaudo, 2024).

Despite these valuable contributions, current research still presents several limitations. Much of the existing literature remains largely descriptive, concentrating on isolated aspects such as visual style or narrative techniques without systematically explaining how different creative elements interact within short-video production and dissemination processes. As a result,

existing studies provide limited insight into how effective creative patterns emerge, how such patterns can be analytically identified, and how they operate within platformized media environments (Kaye et al., 2021).

To address these limitations, this study adopts a theory-building qualitative approach and aims to develop a systematic analytical framework for understanding the creative configurations of craft-based ICH short videos. Rather than treating aesthetics, narrative, and dissemination as separate analytical dimensions, the study explores how these elements interact dynamically within short-video production and circulation.

In this context, the concept of a "creative paradigm" is reconsidered as a configurational mechanism that describes how creative elements are coordinated within specific platform environments. Such an approach shifts the analytical focus from stylistic classification toward the mechanisms that connect visual representation, cultural meaning construction, and media circulation (Quick & Maddox, 2024; Zulli & Zulli, 2022).

Accordingly, the study addresses the following research questions:

RQ1: How are aesthetic strategies configured to trigger viewer attention and immersion in craft-based ICH short videos?

RQ2: How do aesthetic strategies interact with narrative structures to construct cultural meaning in short-video storytelling?

RQ3: How do narrative configurations facilitate dissemination and cross-community diffusion within platformized media environments?

By answering these questions, the study seeks to move beyond descriptive observations and provide a mechanism-oriented understanding of how craft-based ICH short videos function within contemporary digital media ecosystems.

## 2. Literature Review

### 2.1 ICH dissemination in the digital era: from digital preservation to platformized production

Research on the digital dissemination of Intangible Cultural Heritage has gradually shifted from a focus on digital archiving and preservation toward examining how digital media platforms reshape cultural production and circulation. Short-video platforms have emerged as significant infrastructures for cultural communication, enabling creators to document craft processes, present cultural narratives, and interact with broader audiences.

From a platform studies perspective, the concept of platformization highlights how digital infrastructures reorganize cultural production through algorithmic recommendation systems, data-driven distribution mechanisms, and user participation structures (Poell, Nieborg, & van Dijck, 2019; Nieborg & Poell, 2018). Within such environments, cultural content—including ICH—becomes embedded in systems governed by visibility metrics, engagement indicators, and algorithmic curation.

Research on short-video platforms further suggests that algorithmic recommendation mechanisms reshape how cultural content circulates and how audiences encounter heritage-related media. Studies have shown that platforms such as Douyin and TikTok function as interfaces that simultaneously enable cultural participation and shape content visibility through data-driven recommendation systems (Kaye et al., 2021; Paquienseguy, 2025).

Consequently, understanding ICH dissemination in digital environments requires examining not only cultural representation but also the broader platform infrastructures and algorithmic logics that structure cultural visibility and audience engagement.

### 2.2 Short videos as multimodal texts

Short videos are fundamentally multimodal media texts, integrating visual imagery, sound design, editing rhythm, and textual elements within a compact audiovisual format. Multimodal discourse analysis emphasizes that meaning emerges through the interaction of multiple semiotic modes rather than through linguistic text alone.

Within craft-based ICH videos, multimodal design plays a crucial role in translating complex craft processes into visually accessible forms. Techniques such as close-up shots of manual operations, rhythmic editing sequences, and synchronized audiovisual cues enable audiences to perceive craft techniques that were traditionally transmitted through apprenticeship.

Scholars have argued that multimodal analysis provides an effective methodological framework for examining how visual, auditory, and textual elements interact to produce meaning in digital media environments (Geenen et al., 2021; Fazeli et al., 2023; Joy et al., 2024).

In the context of short-video communication, these multimodal strategies function not merely as stylistic devices but as mechanisms that enhance attention, comprehension, and sensory engagement, allowing audiences to access craft knowledge through visual experience.

### 2.3 Narrative structures in ICH short videos

Narrative strategies play a crucial role in transforming fragmented craft procedures into coherent cultural stories. In digital heritage communication, storytelling structures help contextualize traditional practices within broader cultural meanings and social identities.

Research on social media storytelling highlights that narrative construction is closely connected to audience participation. Through comments, reposts, and interpretive engagement, audiences contribute to the negotiation of meaning and cultural interpretation (Schellewald, 2024).

Within short-video environments, narrative structures often rely on process-based storytelling, where sequences of craft production are organized into narrative arcs that highlight craftsmanship, perseverance, and cultural values. These narrative configurations enable audiences to interpret craft practices not merely as technical activities but as expressions of cultural heritage and identity (Darvin, 2022; Mordecai, 2023).

Therefore, narrative mechanisms function as an important intermediary layer linking visual representation with cultural meaning construction.

## 2.4 Platform dissemination and algorithmic circulation

The dissemination of short-video content takes place within algorithmically structured platform environments characterized by interest-based communities and niche publics (Gerbaudo, 2024). Platform algorithms analyze user behavior, interaction patterns, and content metadata to recommend videos to specific audiences.

Although algorithmic distribution can increase cultural visibility, it may also lead to echo chambers, stylistic homogenization, and cultural simplification (Gao et al., 2023). Creators frequently adapt their production strategies to align with algorithmic recommendation logics, shaping video duration, editing rhythm, and narrative presentation in ways that enhance algorithmic visibility (Herman, 2023).

Consequently, understanding the dissemination of ICH short videos requires analyzing not only the content of cultural representation but also the interaction between creative strategies and platform infrastructures (Christin et al., 2024; Lim et al., 2025).

## 2.5 Conceptual framework: the A–N–D mechanism chain

Building upon the literature discussed above, this study proposes a conceptual mechanism framework integrating three analytical dimensions: aesthetics, narrative, and dissemination.

Rather than treating these dimensions as independent variables, the framework conceptualizes them as interconnected components within a dynamic mechanism chain. In this framework, aesthetic strategies function as attention triggers, narrative structures organize cultural meaning, and dissemination mechanisms facilitate audience engagement and cross-community diffusion.

The conceptual structure of this mechanism chain is illustrated in Figure 1, which presents the Aesthetics–Narrative–Dissemination (A–N–D) mechanism chain.

Figure 1. The A–N–D Mechanism Chain for Craft Craftsmanship ICH Short Videos
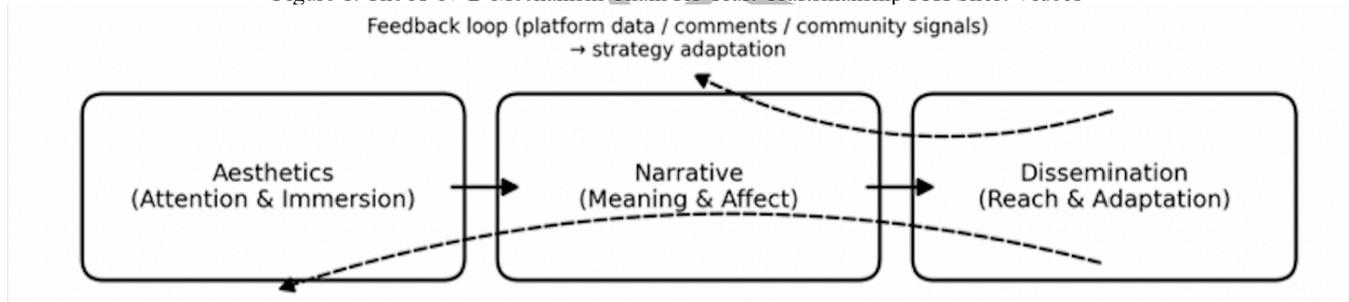


To operationalize this conceptual framework for empirical analysis, the study further develops a set of observable components, indicators, and identification rules. These operational criteria translate the conceptual model into analyzable elements that can be systematically examined within short-video texts.

The operational structure of the framework is summarized in Table 1, which specifies the relationships between analytical dimensions, observable indicators, and identification rules used in the empirical analysis.

Table 1. Components–Indicators–Identification Rules for the Proposed Creative Mechanism Framework

| Mechanism dimension | component elements | Observational indicators (Indicative cues,qualitative) | Identification rules | Strength rating |
|---|---|---|---|---|
| Aesthetics | Micro-visibility | High-frequency close-up and macro photography; key process nodes are "seen"; material texture and force changes are perceivable | Prerequisite A1: At least two types of detailed visualization shots of "key process nodes" should appear; they should correspond to "process logic" rather than being purely decorative | Meeting A1 criteria is considered "up to standard"; if the details are evident throughout >60 % of the key process segments and align with the narrative |

| | | | | information, it is rated as "strong" |
|---|---|---|---|---|
| Aesthetics | Temporal translation | Time delay, omission, and accumulation montage; the long process is compressed yet the logic remains smooth | Prerequisite A2: There exists a clear time compression strategy (such as delaying, omitting, or segmenting) and the audience can still retell the main flow of the process | If the compressed content can still maintain "comprehensible steps + balanced pace", it should be rated as "strong"; if only compression is achieved without logical cues, it should be rated as "weak" |
| Narrative | Narrative agency | Craftsman labor; audience comment participation; implicit presence of traditional spirit and cultural values | Necessary condition N1: At least present both "public subject (craftsman, maker)" and "cultural significance anchor" (subtitle, voiceover, one of scene symbols) simultaneously | The three subjects (public, concealed, absent) can all be identified and mutually supported, rated as "strong" |
| Narrative | Meaning anchoring | Explanatory subtitles and narration; why the craftsmanship is like this; connection with local life and historical memory | Necessary condition N2: There is at least one anchoring point of "cultural significance" (not just a step description), which can be traced back to craftsmanship, community, and memory | If the meaning anchoring runs through the beginning-climax-ending to form a closed loop, it is rated as "strong" |
| Dissemination | Niche entry | Tags, topics, community language; vertical interest matching; stable account persona | Prerequisite D1: Identify target circles (such as handmade crafts, local aesthetics, traditional Chinese style) and use their recognizable symbol systems (tags, discourse, music, etc.) | The use of circle symbols is stable and does not sacrifice cultural information density, rated as "strong" |
| Dissemination | Cross-over triggers） | Universal emotions (healing, stress relief, awe); shareable expressions of "public significance"; cross-domain interaction | Sufficient condition D2: The presence of one of "universal emotions, values" + one of "comprehensible narrative threads" (superposition of two types of elements) | If the content simultaneously meets the criteria of D2 and exhibits cross-circle interpretation and re-narration in the comment section (such as "This is a certain kind of spirit"), it should be |

| Feedback | Platform feedback loop） | Respond to the top comment; adjust the topic and narrative of the next video based on feedback; implement serialized operation | Optional Condition F1: Traceable Feedback - Adjustment of Evidence Chain (Comments, Top-Posting, Serialization) | rated as "strong" The feedback is traceable and can explain the strategy iteration, rated as "strong" |
|---|---|---|---|---|

Through integrating conceptual modeling with operational indicators, this framework provides a structured analytical basis for examining how aesthetic design, narrative construction, and platform dissemination interact in the production and circulation of craft-based ICH short videos.

## 3. Methodology

This study adopts a theory-building qualitative research design combined with an explanatory multi-case study approach to investigate the creative configurations of craft-based Intangible Cultural Heritage (ICH) short videos (Kibler et al., 2025). The research treats short videos as multimodal cultural texts and analyzes how aesthetic strategies, narrative structures, and dissemination mechanisms interact within digital platform environments (Fazeli et al., 2023; Mtisi, 2022).

The methodological design follows a sequential process including theoretical dimension identification, purposive sampling, multimodal transcription, theory-guided coding, and cross-case comparison. This process aims to construct an analytical framework capable of explaining the interaction between aesthetic design, narrative meaning construction, and platform dissemination dynamics.

An overview of the research design and methodological procedures is summarized in Table 2, which outlines the research objectives, analytical strategies, and sequential steps adopted in this study.

Table 2 Research design, methods and process overview

| Research Stage | Research Objective | Method / Analytical Strategy | Data Source | Analytical Output |
|---|---|---|---|---|
| Theoretical dimension identification | Identify key analytical dimensions for studying craft-based ICH short videos | Literature review of multimodal analysis, narrative theory, and platform studies (Geenen et al., 2021; Darvin, 2022; Poell et al., 2019; Nieborg & Poell, 2018; Gerbaudo, 2024) | Academic literature on digital media and cultural communication | Analytical dimensions: aesthetics, narrative, dissemination |
| Case selection and sampling | Construct an empirical dataset of representative craft-based ICH short videos | Purposive sampling based on predefined inclusion criteria | Douyin platform videos from selected creators | Dataset of 30 videos from three creators |
| Multimodal transcription | Transform video materials into analyzable textual data | Multimodal transcription including visual, audio, textual, and interaction layers (Fazeli et al., 2023; Joy et al., 2024) | Archived video materials and platform metadata | Structured multimodal transcripts |
| Theory-guided coding | Identify observable patterns of creative configuration | Structured qualitative coding based on theoretical dimensions | Transcribed video data | Coding categories and operational indicators |
| Cross-case comparison | Compare patterns across creators and videos | Comparative qualitative analysis | Coded video segments | Identification of recurring creative configurations |
| Framework construction | Develop an analytical mechanism explaining ICH short-video creation | Synthesis of empirical findings and theoretical interpretation | Coding results and analytical memos | Aesthetics–Narrative–Dissemination (A–N–D) mechanism framework |

## 3.1 Research Design

The research design follows a theory-building qualitative logic in which empirical analysis is guided by theoretical insights while empirical observations also inform the refinement of analytical categories (Van Hulst & Visser, 2025; Vila-Henninger et al., 2024).

Rather than treating aesthetic, narrative, and dissemination dimensions as independent descriptive categories, the study conceptualizes them as interconnected components within a dynamic analytical structure. The analytical process therefore focuses on examining how aesthetic elements trigger viewer attention, how narrative structures organize cultural meaning, and how dissemination mechanisms shape audience engagement within platformized media environments (Thomas et al., 2024).

## 3.2 Analytical Dimensions and Theoretical Sources

To ensure methodological transparency, the analytical dimensions employed in this study are derived from established theoretical literature rather than constructed ad hoc.

The aesthetic dimension is grounded in multimodal media analysis, which emphasizes how visual composition, editing rhythm, and audiovisual coordination generate meaning within digital media texts (Geenen et al., 2021; Fazeli et al., 2023; Joy et al., 2024).

The narrative dimension draws on narrative analysis in digital media research, focusing on how storytelling structures transform fragmented craft processes into coherent cultural narratives (Darvin, 2022; Mordecai, 2023).

The dissemination dimension is informed by platform studies examining algorithmic recommendation systems and digital cultural circulation (Poell et al., 2019; Nieborg & Poell, 2018; Gerbaudo, 2024).

Together these theoretical perspectives provide the conceptual basis for constructing the analytical dimensions used in the coding framework.

## 3.3 Sampling Strategy and Case Selection

The study adopts purposive sampling to construct a dataset of craft-based ICH short videos published on the Douyin platform. The unit of analysis is an individual short video, including its visual content, audio track, textual elements, and observable audience interaction data.

Three representative creators were selected as case accounts: Peng Nanke, Bai Xingjian, and Gongqiu Tibetan Cultural Heritage. These creators were chosen because their content demonstrates clear craft processes, consistent thematic orientation, and observable audience engagement.

The sample size was determined according to qualitative principles of information richness rather than statistical representativeness. In qualitative inquiry, purposive sampling aims to identify information-rich cases that can illuminate the phenomenon under study in depth, rather than to approximate population parameters. In this study, the research focus is deliberately narrow: craft-based ICH short videos on a single platform (Douyin), analyzed through a clearly specified theoretical lens (the A–N–D mechanism chain). Under such conditions, a smaller but analytically intensive sample is methodologically appropriate, because the required sample size depends on information power, that is, the degree to which the selected material is specific, theoretically relevant, and capable of supporting in-depth analysis(Hennink & Kaiser, 2022; Malterud et al., 2016).

Accordingly, three creator accounts were selected as comparative cases, not to achieve statistical coverage, but to enable cross-case replication and contrast. Each account met the same inclusion criteria—stable craft-oriented production, clear visibility of manual processes, and observable audience interaction—while differing in stylistic configuration and cultural presentation. This design made it possible to examine whether similar mechanism patterns recurred across cases and how specific configurations varied under different creative conditions.

Within each case account, 10 videos were selected, resulting in a final corpus of 30 videos. This number was considered sufficient for three reasons. First, each video constituted a dense multimodal unit containing visual, auditory, textual, and interactional data. Second, the study pursued theory-building through close coding and cross-case comparison rather than broad frequency estimation. Third, during the iterative coding process, later materials mainly reinforced and refined existing categories rather than generating fundamentally new analytical dimensions, indicating that the corpus had reached adequate analytical depth for the purposes of this study. Therefore, the final sample of 30 videos should be understood as a theoretically and methodologically justified corpus for qualitative mechanism construction rather than as a statistically representative sample.

To enhance methodological clarity, the sampling boundaries and selection criteria are explicitly defined and summarized in Table 3.

Table 3 Sample boundaries and sampling criteria

| Sampling Dimension | Criteria | Rationale | Data Outcome |
| --- | --- | --- | --- |
| Platform scope | Videos published on the Douyin platform | Douyin represents one of the most influential short-video platforms in China and provides a mature algorithmic recommendation ecosystem for cultural content dissemination | Platform dataset |
| Content type | Craft-based Intangible Cultural Heritage (ICH) short videos | Focus on videos that present identifiable craft processes or traditional manual techniques | Craft-oriented video corpus |
| Creator selection | Representative creators with stable thematic production related to craft-based ICH | Ensures consistency in content orientation and allows cross-case comparison between creators | Three creators selected |
| Video inclusion criteria | Videos must contain observable craft production processes, clear visual representation of manual techniques, and complete audiovisual sequences | Enables multimodal analysis of aesthetic, narrative, and dissemination features | Valid analytical units |

| Time range | Videos published within the selected research period | Ensures comparability within a stable platform algorithm environment | Temporal dataset |
|---|---|---|---|
| Final sample size | 30 videos (10 videos from each creator) | Allows in-depth qualitative analysis while maintaining cross-case comparison | Final analytical dataset |

## 3.4 Data Collection and Multimodal Transcription

Data collection follows the principle of traceable evidence and reproducible documentation. Each selected video was archived using screen recording and metadata capture to preserve visual, auditory, and textual information (Ohme et al., 2024).

To enable systematic analysis, each video was transcribed into four analytical layers:

Visual layer (camera composition, framing, lighting)

Audio layer (music, ambient sound, voice)

Textual layer (subtitles, titles, tags)

Platform interaction layer (comments and audience responses)

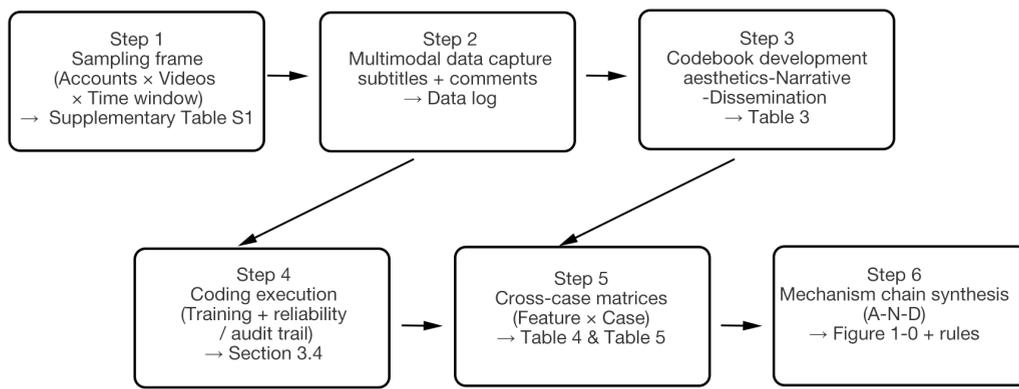The entire analytical workflow and audit trail of the research process are illustrated in Figure 2.



Figure 2 Research workflow

## 3.5 Coding Framework and Codebook Construction

To transform multimodal materials into analyzable data, the study employs theory-guided structured coding.

The coding framework follows a hierarchical structure consisting of:

Analytical dimension

Coding category

Operational definition

Exclusion criteria

A summary of the coding categories and operational definitions is presented in Table 4, which serves as the codebook used during the analysis.

Table 4 Codebook summary for theory-guided structured coding

| Analytical Dimension | Coding Category | Operational Definition | Exclusion Criteria | Theoretical Basis |
|---|---|---|---|---|
| Aesthetics | Micro-visibility of craft processes | Close-up or macro visual presentation highlighting manual operations, tools, or materials used in craft production | General environmental shots without clear craft action | Multimodal analysis (Geenen et al., 2021; Fazeli et al., 2023) |
| Aesthetics | Temporal compression through editing | Editing techniques that condense long craft procedures into short, visually continuous sequences | Real-time documentation without editing condensation | Audiovisual narrative design (Joy et al., 2024) |
| Aesthetics | Multimodal audiovisual coordination | Coordinated use of sound, music, and visual rhythm to enhance sensory engagement | Background music unrelated to craft process | Multimodal discourse theory (Geenen et al., 2021) |
| Narrative | Process-based storytelling | Craft production presented as sequential stages forming a coherent narrative arc | Fragmented shots without process continuity | Narrative analysis in digital media (Darvin, 2022; Mordecai, 2023) |
| Narrative | Cultural contextualization | Narrative elements that explain cultural meaning, tradition, or heritage value | Purely technical demonstration without cultural explanation | Cultural storytelling research (Darvin, 2022) |

| Dissemination | Niche community engagement | Interaction with specific audience communities interested in crafts or cultural heritage | Generic mass entertainment engagement unrelated to craft content | Platform studies (Gerbaudo, 2024) |
|---|---|---|---|---|
| Dissemination | Cross-community diffusion | Evidence of broader audience interaction, sharing, or reinterpretation across different user groups | Interaction limited to creator self-promotion | Platformization theory (Poell et al., 2019; Nieborg & Poell, 2018) |
| Dissemination | Algorithmic visibility cues | Observable signals such as hashtags, trending formats, or engagement indicators suggesting algorithmic circulation | Content lacking any platform interaction indicators | Algorithmic media studies (Kaye et al., 2021) |

## 3.6 Coding Procedure and Reliability Control

The coding process was conducted in several stages. First, videos were segmented into analytical units based on visual transitions and narrative segments. Second, each segment was coded according to the codebook definitions.

To enhance analytical reliability, several procedures were implemented:

Maintenance of an audit trail

Documentation of coding decisions through analytic memos

Peer debriefing with researchers familiar with qualitative media analysis

Negative case analysis to examine instances that do not conform to the proposed framework (Halpin, 2024; Herzog et al., 2025).

Although the study employs a single primary coder, these procedures strengthen the transparency and credibility of the qualitative analysis.

## 3.7 Research Reflexivity and Analytical Transparency

Because qualitative research inevitably involves interpretive judgment, reflexivity was incorporated throughout the analytical process. The researcher maintained detailed analytic memos documenting interpretive decisions and potential biases.

These memos allow readers to trace how analytical interpretations developed from specific empirical observations and coding decisions. Such documentation enhances methodological transparency and supports the credibility of the theory-building process (Vila-Henninger et al., 2024).

## 4. Aesthetic Characteristics of Handicraft ICH Short Video Creation: Refined Design of Audiovisual Language

This chapter focuses on answering one question: In handicraft ICH short videos, which audiovisual aesthetic choices can be stably identified as "creative element configurations" and present relatively consistent organizational patterns across cross-case comparisons? To avoid impressionistic explanatory descriptions, this study operationalizes the aesthetic dimension (A) into three groups of observable elements—camera language, editing organization, and sound-image-text coordination (see the codebook summary in Table 4, 3.5)—and presents the occurrence strength and combination patterns of these elements across different account samples through a cross-case coding matrix (Table 5). Building on this foundation, this chapter further organizes the relationships between aesthetic elements and subsequent narrative functions and dissemination triggers into a traceable evidence chain table (Table 6), enabling readers to clearly see the correspondence between "evidence fragments—coding judgments—mechanism positions," consistent with the direction of the A–N–D mechanism chain (Figure 1).

Therefore, the writing strategy of this chapter adopts a paragraph structure of "functional statement → core finding → evidence anchor (Figure) → coding category (A1–A6) → cross-case reference (Table 5) → mechanism chain positioning (Table 6)." In other words, this chapter not only answers "what the aesthetic characteristics are," but also addresses "how I know it is an identifiable and reviewable characteristic, and what role it plays in the mechanism chain."

## 4.1 Camera Language: Microscopic Gaze and Atmosphere Creation

Macro or close-up gaze (A1) functions less as a stylistic ornament than as a mechanism of procedural legibility. Across the sampled videos, this configuration recurrently appears at key process nodes where tactile operations, material resistance, and micro-level transformations must be made visible to non-expert viewers. The shadow-puppet case, for example, uses close-up framing to render blade angle, force variation, and leather residue visually traceable (Figure 3). The analytical significance of A1 therefore lies not simply in "showing details," but in transforming tacit craft knowledge into readable visual evidence. In the A–N chain, such visibility reduces the audience's comprehension burden and prepares the ground for narrative uptake, because viewers can first identify what is happening before they are invited to interpret why it matters. Cross-case comparison further shows that A1 is not uniformly configured: Bai Xingjian relies on it more strongly as a primary explanatory device, whereas Peng Nanke embeds it within a denser combination of editing and multimodal

coordination, and the Gongqiu case combines close-up visibility with contextual landscape framing. This variation suggests that A1 should be understood as a codable mechanism whose analytical value lies in its contribution to learnability, narrative entry, and subsequent engagement rather than in visual attractiveness alone (see Tables 5–6).



**Figure 3. Macro photography, shadow puppetry prop making**

Corresponding to "detail readability," another common strategy employs "defamiliarized perspective/composition" (A2) to create non-everyday viewing experiences, thereby raising the first-screen attention threshold and forming visual tension that signals "worth continuing to watch" within an extremely short duration. In the sample, the presentation of pigment diffusion patterns captured from an underwater perspective serves as typical evidence of A2: this perspective is difficult to obtain in daily life and thus possesses inherent novelty and visual spectacle attributes (Figure 4). More importantly, this defamiliarization is not equivalent to mere technical showmanship; it is often bound to key procedural nodes. When viewers are attracted by the novelty and enter the video, the scene immediately returns to the causal relationship between action and material, thereby transforming "novelty" into "understanding." At the mechanism chain level, A2 is typically positioned as an "entry device" for A→N, and in some cases further becomes a trigger point for "circle-breaking clues" at the dissemination layer (Table 6), because it provides visual highlights that can be retold and shared, reducing the explanation cost for cross-circle communication.



**Figure 4. GoPro underwater photography**

Beyond focusing on details and creating defamiliarization, camera language also elevates the workspace from "background" to "meaning field" through lighting organization. In the coding framework, "restrained lighting；simple atmosphere" (A3) points to a stable visual orientation: using natural light or quasi-natural light to enhance the sense of labor authenticity, embedding craft activities within a perceptible living world. Typical evidence includes warm-toned light simulating sunlight streaming through windows, with window frame shadows falling on hand movements, creating a warm yet restrained labor atmosphere (Figure 5). At the narrative level, this type of atmospheric cinematography often couples with "pastoral/shangri-la" imagery, providing viewers with an emotional framework for understanding: watching is no longer merely "learning procedures," but also becomes "entering an imagined world of slow living and craftsmanship." Therefore, A3 is frequently classified in the evidence chain under the pathway of "aesthetics—emotional tone—meaning anchoring" (Table 6), providing a visual foundation for the dissemination layer to form recognizable labels such as "healing," "stress-relieving," or "slow living."



**Figure 5. Tibetan culture in lighting design scenarios**

Table 5. Cross-case coding matrix (Aesthetic feature A dimension)

| Case account | A1 Macro Gaze | A2 Defamiliarization perspective | A3 rustic lighting | A4 delay compression | A5 Accumulation Montage | A6 Sound, Picture, and Text Collaboration | Key evidence points to (Figure No.) |
|---|---|---|---|---|---|---|---|
| Peng Nanke (core) | Moderate | Moderate | Strong | Strong | Strong | Strong | Figure 7–12 (editing, music, subtitles) |
| Bai Xingjian (for reference) | Strong | Weak | Moderate | Weak | Moderate | Moderate | Figure 3 (Macro) |
| Gongqiu Intangible Cultural Heritage Tibetan Culture (for reference) | Moderate | Weak | Strong | Moderate | Weak | Moderate | Figure 5 (Lighting) |

Table 6. A→N→D evidence chain table (how aesthetic features lead to narrative function and communication trigger)

| Aesthetic elements (A) | Direct visible evidence | Narrative function (N, abbreviation) | Propagation trigger (D, abbreviated) | Explanation of mechanism chain (corresponding to Figure 1) |
|---|---|---|---|---|
| A1 Macro Gaze | Figure 3 Enlarged Detail | N Skill Visibility, Knowledge Transfer | D Seeking novelty and staying within professional circles | A. Enhancing "learnability and viewability" → strengthening the credibility of process narratives → boosting the motivation for staying and sharing |
| A2 Defamiliariz ation perspective | Figure 4 Underwater image | N Defamiliarization enhances "eventfulness" | D Novelty clicks and diffusion | A. Create differentiated entry points → Enhance homepage attraction → Increase the possibility of breaking through the existing circle |
| A3 Rustic lighting | Figure 5 Warm Light and Window Shadow | N Realism, pastoral imagery | D Emotional resonance, healing-oriented | A. Shaping the atmosphere → Supporting the "labor narrative" → Forming emotional communication tags |
| A4 Delay compression | Figure 6 Delay Paragraph | N Time can bring about changes | D Sense of completion, sense of satisfaction | A. Compressing long cycles → narrative rhythm consumable → enhancing completion and sharing |
| A5 Accumulati on Montage | Figure 8–10 Repeat process | N Accumulate momentum and strengthen the spirit of craftsmanship | D Decompressing the viscosity of the layer | A. Organizational repetition → formation of psychological rhythm → support for "decompression/concentration" communication discourse |
| A6 Sound, picture and text coordination | Figure 11–12 Music and Subtitles | N Meaning anchoring, lowering the threshold of comprehension | D Improved retellability | A. Reduce comprehension cost → Enhance narrative interpretability → Boost comments and favorites |

## 4.2 Editing Characteristics: Time Compression and Rhythm Control

Handicraft ICH typically involves long production cycles, intricate steps, and high information density, whereas short videos, as "fragmented communication carriers", require the completion of a viewing closure within a limited duration. As noted in the original text, editing fulfills the critical function of "time compression and rhythm control"; this study further translates this function into two codable editing strategies: (A4) Time-lapse Compression and (A5) Cumulative Montage. Together, these two strategies address the same methodological question: how creators transform "long-cycle processes that cannot be viewed in full" into "consumable, immersive, and recheckable viewing segments" (Figure 6).

In crafts involving long-duration changes such as fermentation and air-drying, (A4) Time-lapse Compression serves not only to shorten the video length but also to visualize the value logic of "temporal accumulation". Taking the short video "Vinegar" as an example, the work uses time-lapse sequences—such as the rising and setting of the sun and moon, and crop growth—to present changes occurring over several days or even longer. This transforms "time" from an invisible cost into a perceptible visual rhythm (Figure 7). The key to this strategy lies not in simple fast-forwarding, but in using changes in light, shadow, and the environment as temporal markers for process segments, thereby preserving the integrity of the narrative chain: audiences can understand "why waiting is necessary and what changes waiting brings".

Within the mechanistic chain, A4 is often regarded as the structural enabler of the Aesthetics→Narrative (A→N) link: it embeds long-cycle processes into a progressive narrative timeline, and at the dissemination level, it enhances the sense of completion and satisfaction, boosting watch-through rates and sharing behavior (Table 6). The cross-case matrix (Table 7) is designed to examine whether different accounts exhibit consistent reliance on A4 and whether this reliance shows stable co-occurrence with their content themes (e.g.fermentation-based vs. production-based crafts).

Table 7. Evidence Chain Table

| Chain mechanism position | Video ID | Extracts of quotable evidence (verifiable) | Corresponding code | In-paper anchor points (figures, paragraphs) | Mechanism determination (closed-loop discourse) |
|---|---|---|---|---|---|
| | | | | | |

| A→N | V01 | The scene progresses with details of knife tips and skin residue, allowing the audience to gain "step readability" (details-action-result) | A1 + N1 | Figure 3 paragraph | The macro details turn the process into "readable steps", enhancing the narrative capacity of the process |
|---|---|---|---|---|---|
| A→N | V03 | The warm light and the shadows cast by the window frames on the hands create an atmosphere of "rustic labor" and evoke pastoral imagery | A3 + N3 | Figure 5 paragraph | The lighting atmosphere symbolizes the working space, facilitating the establishment of scene imagery narrative |
| A→N | V04 | Sunrise and sunset, growth and time-lapse compress the long cycle into "watchable rhythm segments" | A4 + N1 | Figure 7 paragraph | Translating time cost into narrative progression points with a delay, ensuring the integrity of the process chain |
| A→N→D | V05 | Repeating similar actions 4-5 times creates a gradual transition and a "decompression" rhythm, making it easier for the audience to be drawn in and generate interactive demands | A5 + N2 + D2 | Figure 8–10 paragraph | Accumulated montage aestheticizes repetitive labor and triggers interactive guidance (seeking the entire process and tutorial) |
| A→D | V02 | Underwater blending/extreme perspective creates a "strange spectacle", enhancing the staying and turning engine | A2 + D3 | Figure 4 paragraph | Defamiliarization raises the attention threshold and provides clues for breaking through the original circle |
| A→N→D | V06 | The synchronized sound (firewood chopping, hammering, and insect chirping), ancient-style musical accompaniment, and regular script subtitles jointly anchor the steps and emotions | A6 + N2 + D1 | Figure 10–12 paragraph | Multimodal collaboration enhances immersion and meaning anchoring, and strengthens community identification and sustained interaction |
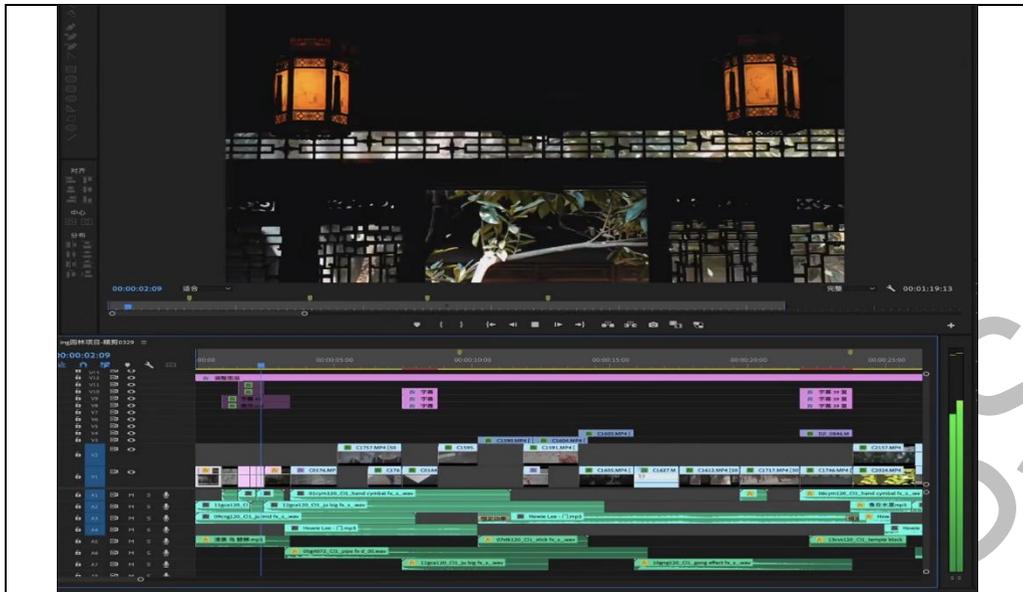
**Figure 6. Operation panel of the audio and video editing software "Jianying"**
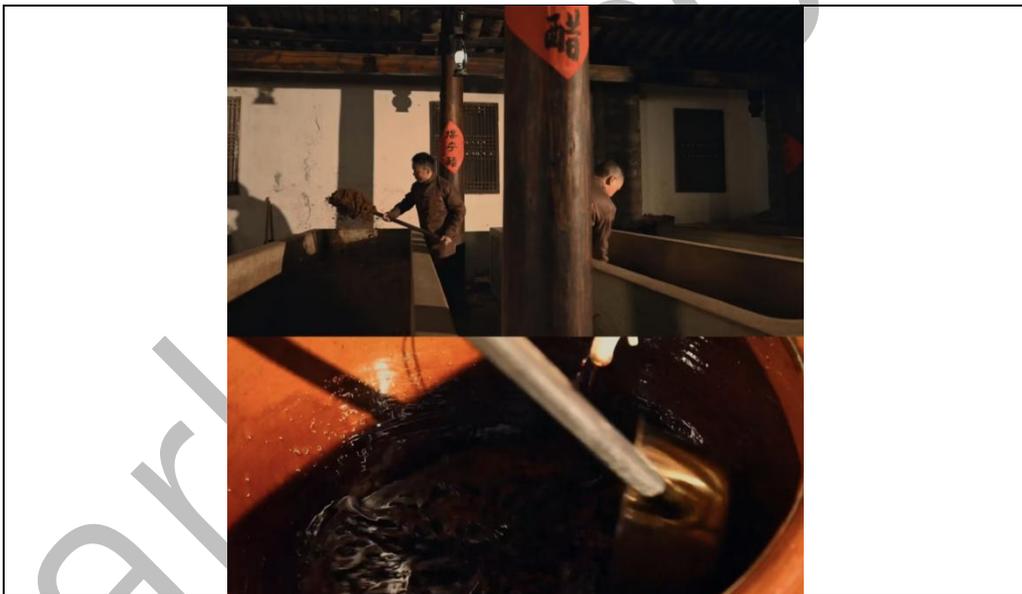


**Figure 7. Time-lapse photography of aged vinegar**

Compared with the time-lapse strategy that emphasizes "macro-compression of time", (A5) "Cumulative Montage" focuses more on "the rhythmic organization of repeated actions". Its aesthetic effect is often experienced by audiences as "stress-relieving or addictive", yet this effect does not arise arbitrarily—it is a verifiable product jointly shaped by shot selection and editing rhythm. Taking "Chaozhou Hand-pulled Teapot" as an example, the video repeats key actions such as clay throwing, shaping and trimming 4–5 times at a consistent rhythm, and builds up momentum through segmented editing (Figures 8–10).

On the narrative level, such repetition does not weaken information delivery; instead, it highlights "the stability and precision of craftsmanship", lending a ritualized rhythm to the viewing experience. On the communication level, repetitive rhythms easily spark interactive demands in the comment section, such as "please upload a tutorial", "show the full process", or "do it again", thus forming circle-based engagement recognizable by platform algorithms.

Accordingly, A5 is identified in the evidence chain as an A→N→D pathway: aesthetic rhythm → narrative ritualization → interaction triggering (Table 6). The cross-case matrix (Table 5) further allows for comparison of whether different creators equally rely on "satisfying rhythm" as their core attraction mechanism.

| Figure 8. Chaozhou hand-pulled teapot: pot-pulling | Figure 9. Chaozhou hand-pulled teapot: plasticity | Figure 10. Chaozhou hand-pulled teapot: scraping the clay |

Beyond camera work and editing, the samples also reveal a more comprehensive aesthetic organization method, namely "sound-image-text coordination" (A6). Your original text has already pointed out that synchronous sound, instrumental music, and explanatory text jointly constitute the immersive experience, citing the soundtrack and regular script subtitles in "Begonia Not Asleep" as examples (Figures 11–12). This study treats this as typical evidence of A6, emphasizing its "coordinative" attribute: synchronous sound provides a sense of authenticity and tactility, the soundtrack supplies emotional rhythm, and subtitles undertake step anchoring and meaning prompting at key nodes. It is precisely through this coordination that viewers obtain both "auditory presence" and "cognitive handles," thereby achieving dual entry into both the procedure and the emotion within a shorter time. At the mechanism chain level, A6 can simultaneously serve A→N (lowering the comprehension threshold, enhancing segment retellability) and N→D (improving retellability and collection motivation), thus it is often labeled as a "synergistic enhancer" in the evidence chain table (Table 6). Meanwhile, Table 5 allows us to further compare whether different accounts exhibit differentiated configurations in terms of "synchronous sound dominance," "soundtrack dominance," or "subtitle dominance," thereby providing a more detailed basis for comparison for subsequent chapters on narrative and dissemination.



| Figure 11. Haitang Weimian: Music Score | Figure 12. White regular script subtitle |

In summary, this chapter does not merely leave aesthetic characteristics at the level of stylistic description, but rather operationalizes them into six types of codifiable aesthetic configuration elements (A1–A6), presenting them in a reviewable manner through two types of result carriers. First, the cross-case coding matrix (Table 5) is used to report the occurrence strength and combination tendencies of each element across different cases. Second, the evidence chain tables (Tables 6-7) are used to mechanistically link aesthetic elements with narrative functions and dissemination triggers, transforming "camera work, editing, or sound-image-text" from scattered descriptions into locatable nodes on the A–N–D mechanism chain (corresponding to Figures 1-12). Within this structure, Figures 3–12 serve not merely as illustrative examples, but as evidence anchors that allow for backtrack coding judgments: each figure corresponds to clear coding categories and mechanism positions, thereby achieving a closed loop of "evidence—coding—mechanism—conclusion."

## 5. Integrated Results and Conclusion: Cross-Layer Evidence of the A–N–D Mechanism Chain and Paradigm Determination

Building on the previous sections, this chapter conducts cross-layer integration and synthesis of the research findings. Section 4 has already identified the key aesthetic configurations (A dimension) of handicraft ICH short videos through multimodal evidence and coding categories, presenting their reviewable audiovisual organization with image evidence anchors (Figures 3–12). On this basis, this chapter advances the research results to the mechanism chain level: on one hand, it explains how these aesthetic inputs are absorbed by narrative strategies and transformed into a meaning structure that is "understandable, empathizable, and recheckable" (N dimension); on the other hand, it reveals how this meaning structure forms dissemination mechanisms of diffusion, interaction, and feedback within the platform's circle-layer ecology (D

dimension). Ultimately, it formalizes the concept of "paradigm or mechanism framework" in the form of determinable configuration rules, thereby completing the closed loop of "evidence—coding—mechanism—determination" (corresponding to Figure 1).

## 5.1 Narrative Mechanism (N): How Aesthetic Input Is Organized into Meaning Structures and Recheckable Discourse

In the A–N–D mechanism chain, narrative (N) serves as a pivotal hub connecting the preceding and following elements. It does not merely reiterate aesthetic effects, but rather organizes the "visibility of the craft process" into the "comprehensibility of meaning." Based on the narratological framework of "story/discourse," this study deconstructs the narrative mechanism into two levels: the first is the story level (what is told: subject, theme, scene), and the second is the discourse level (how it is told: structure, perspective, language). This deconstruction transforms narrative presentation from abstract interpretation into reviewable textual structure analysis, and enables a one-to-one correspondence with the observable aesthetic elements identified in Section 4.

At the story level, the sample narratives exhibit a composite subject structure composed of "overt—covert—absent" elements: the overt subject is typically the craftsman or creator appearing on screen; the covert subject is manifested in the audience participating in meaning negotiation through comment sections and community interactions; the absent subject is embedded in the text and constitutes the value core through "traditional spirit or cultural values" (Schellewald, 2024). The stable establishment of this subject structure relies on the continuous supply of "craft visibility" provided by the aesthetic configurations in Section 4. For example, macro or close-up shots place key actions and material feedback at the visual center (Figure 3), enabling the visual validation of the overt subject's technical authority. Meanwhile, sound-image-text coordination, through the complementary organization of synchronous sound, soundtrack, and explanatory subtitles (Figures 11–12), fills information gaps without disrupting immersion, allowing the value clue of the absent subject to be explicitly prompted and retold by the audience. Thus, the subject is not merely a narrative concept, but a meaning-carrying device continuously supported by observable audiovisual organization (A→N).

At the thematic level, the sample narratives can be summarized into three relatively stable theme types: procedural narrative, ritualized narrative, and memory-reconstructive narrative. Procedural narrative externalizes tacit knowledge through linear processes; ritualized narrative elevates daily production into symbolic experience through rhythmic segments; memory-reconstructive narrative evokes cultural resonance through spatiotemporal compression and digital representation (Henry, 2024; Lin et al., 2023). The ability of these three theme types to form consumable viewing loops in short videos is directly related to the time translation capabilities of the editing strategies discussed in Section 4: time-lapse photography compresses long-cycle changes into perceptible rhythms (Figure 7); accumulative montage generates a sense of "momentum building—completion" through action repetition (Figures 8–10), jointly providing segment closure for procedural and ritual themes; macro details provide "visible evidence" for memory reconstruction, transforming "regeneration" from a value claim into a perceptible craft fact (Figure 3).

At the scene level, the texts repeatedly invoke a binary structure of "natural space—working space": natural space is poeticized into an idyllic image chronotope, providing emotional sustenance; working space, as the core arena for skill demonstration, narrativizes the labor process and generates a sense of authenticity through close-up shots of body-tool interaction (Luo, 2026). This scene mechanism forms a closed loop with the evidence of "simple lighting/atmosphere" (Figure 5) from Section 4: the visual organization of warm-toned light and window shadows falling on hand movements not only enhances immersion but also makes the "idyllic imagery" a reviewable visual code, thereby transforming space from mere background into a field where meaning is continuously generated (A→N).

At the discourse level, narrative structure achieves "clear process logic" and "rhythmic viewing tension" through time compression and segment arrangement. Narrative perspective employs a compound application of internal and external focalization to balance immersion and credibility. Narrative language manifests as the systematic integration of visual, auditory, and textual modalities (Darvin, 2022; Mordecai, 2023). Within this framework, "modal synergy" possesses clear evidence anchors: synchronous sound and soundtrack provide rhythm and emotion, while subtitles undertake step anchoring and meaning prompting at key nodes (Figures 11–12), thereby enhancing the text's comprehensibility and retellability, and providing the necessary mediating mechanism for entry into the dissemination layer (N→D).

## 5.2 Transmission mechanism and feedback loop (D): How does the stratified ecosystem shape the diffusion boundary and inversely shape creation

In the platform ecosystem, the dissemination of handcrafted short videos of intangible cultural heritage is not a linear diffusion, but deeply embedded in the layered structure co-constructed by algorithms, users, and platforms: differentiated creative styles attract different interest groups, algorithms make precise recommendations based on tags and interactive behaviors, user community interactions strengthen internal identity, and platform traffic allocation and commercialization logic further solidify the boundaries of these layers (Huang & Ye, 2024). From the perspective of the A–N–D mechanism chain, layer D is not simply a "result variable", but a set of ecological mechanisms that can inversely influence creative choices: when certain aesthetic symbols are repeatedly rewarded by layered users, they are reused more frequently and drive

content convergence; when narrative themes are tagged by the platform, creators may compress narrative complexity to match recommendation logic, thereby weakening the depth of cultural translation (Christin et al., 2024). This phenomenon corresponds to the emphasis on the "dissemination feedback—strategy adjustment" loop in the method design of Methodology Chapter, and forms a structural closed loop with the feedback arrow returning from D to A/N in Figure 1.

While layering enhances communication efficiency, it also brings about a series of institutional dilemmas: content homogenization and "involution", labeling and context detachment of cultural representation, echo chamber effect caused by layered barriers, emotional trust depletion due to excessive reliance on sales-driven paths, and the digital divide and "aphasia" exacerbated by uneven digital participation. These dilemmas can be understood as weakening factors of "paradigm strength" in Layer D: when homogenization increases and labeling intensifies, even if Layer A has a high degree of completion (exquisite shots/editing), the meaning construction of Layer N may be compressed into consumable stylistic symbols, making it difficult to achieve cross-layer public dialogue and value expansion. Therefore, the communication mechanism not only explains "why it is difficult to break through layers," but also explains "why breaking through layers must be premised on the integrity of the institutional chain" (Bhandari & Bimo, 2022; Hartmann et al., 2025; Lim et al., 2025; Számel et al., 2025).

## 5.3 Formalization of Paradigms: Decision Rules and Strong/Weak Evaluation

This study operationalizes the "creative paradigm" into a verifiable framework of creative mechanisms or a pattern of element configuration, and provides the minimum judgment rules and criteria for strength assessment, ensuring that the "paradigm" moves beyond interpretative descriptions and achieves a reusable and comparable formal expression. This formal expression is guided by the overall structure of Figure 1, with the evidence from Figure 3–12 in Section 4 serving as traceable anchor points for Layer A input.

Firstly, the minimum judgment criteria include four items: (1) Minimum condition for layer A: The work must provide at least one audio-visual evidence of "enhanced visibility of craftsmanship" and be able to point to a clear process node; this condition can be satisfied by any of the following: macro details (Figure 3), time-lapse compression (Figure 7), accumulation montage (Figure 8–10), or sound-image-text synergy (Figure 11–12). (2) Minimum condition for layer N: The work must organize the presentation of craftsmanship into a recognizable narrative structure (at least satisfying two of the three elements: subject-theme-scene, and achieving information complementation and meaning anchoring at the discourse level), otherwise it can only be regarded as "style display" rather than "narrative translation". (3) Minimum condition for layer D: The work must explicitly organize its retellability and accessibility for the platform ecosystem (such as anchoring meaning through subtitles/titles, activating hidden subjects through interaction, forming community entry points through tags, etc.), otherwise dissemination can only remain at the level of sporadic traffic. (4) Chain integrity condition: The work must present at least one traceable path from A to N to D, or exhibit an observable feedback loop (adjustment of D to subsequent A/N strategies). This condition upgrades the "paradigm" from a three-dimensional juxtaposition to a mechanism chain, thereby matching the type of research (framework construction).

Secondly, based on the aforementioned rules, this study proposes a qualitative intensity grading to avoid conceptual generalization: The first level is strong configuration: Layer A has ≥2 pieces of visible evidence; Layer N forms a stable thematic type with sufficient discourse coherence; Layer D possesses observable mechanisms for entering and interacting within the circle; and the chain A→N→D can be traced back in the text. The second level is moderate configuration: Layer A has sufficient evidence but Layer N or Layer D lacks key support (commonly seen in "high aesthetic display but insufficient meaning anchoring" or "complete narrative but weak ecological entry strategy"). The third level is weak configuration: It only possesses stylized Layer A displays or fragmented narrative segments, lacks recheckable organization and ecological entry mechanisms, and is more susceptible to being homogenized and absorbed by the logic of the circle, eventually sliding towards "involution".

## 5.4 Comprehensive conclusions and contributions: closed-loop rebates and research implications

Based on the above results, this study indicates that the effective creation of short videos featuring handcrafted intangible cultural heritage in the digital and intelligent era does not solely rely on "being visually appealing" or "telling stories", but rather depends on a verifiable A–N–D mechanism chain. The A layer translates craftsmanship into viewable evidence through lensing, editing, and multimodal coordination; the N layer organizes the evidence into a meaningful structure through stories and discourse, enhancing its retellability; the D layer achieves diffusion within a layered ecosystem through tagging, interaction, and platform mechanisms, and reversely shapes subsequent creative strategies through feedback loops (corresponding to Figure 1). This closed loop, on one hand, explains the mechanism differences behind "why it becomes popular or why it struggles to break through the circle", and on the other hand, advances the "creative paradigm" from a descriptive concept to an operational mechanism framework and judgment rules, thereby maintaining consistency in the research objectives, paper type, and evidence presentation.

At the practical level, this mechanism framework suggests that creators need to possess threefold abilities: presenting skills through visibility strategies, completing cultural translation through narrative structures, and addressing gerrymandering and

algorithmic logic with ecological awareness. Only by forming synergy in the chain can handcrafted intangible cultural heritage short videos move from gerrymandering interests to public value and living inheritance.

## 6. Conclusion

This study argues that the effective creation of handcrafted ICH short videos in the digital and intelligent era relies on an organic paradigm of the trinity of aesthetics, narrative and communication. The aesthetic layer serves as the foundation of appeal: it translates craftsmanship into perceptible "visual poetics" through sophisticated audiovisual language. The narrative layer constitutes the core of emotional resonance: it accomplishes the in-depth translation from simple craft demonstration to cultural meaning construction by virtue of rich stories and discourses. The communication layer acts as the field of vitality: on the basis of facing up to the circle-based structure, it drives ICH culture from closed interest communities to open public value through systematic circle-breaking strategies.

This paradigm requires creators to integrate three essential identities: a "director" who masters audiovisual aesthetics, a "narrator" proficient in cultural translation, and a "communicator" with an ecological perspective. Only through the synergy of these three roles can handcrafted ICH short videos transcend fleeting traffic attention in the digital wave and truly realize the living inheritance and innovative development of intangible cultural heritage. In doing so, they can provide a sustainable practical path for enhancing cultural confidence, boosting rural revitalization, and enriching the digital cultural ecosystem in the new era.

## Author Contribution Statement

Wang Zhenyang: Conceptualization, Writing—original draft preparation, and manuscript development. Wang Zhenyang was primarily responsible for drafting the article and organizing the overall writing of the manuscript.

Dr. Zhang Zeng: Supervision, Writing—review and editing, and manuscript revision. Zhang Zeng was responsible for proofreading, revising, and improving the manuscript.

Hui Yi: Literature collection, data and material search, and resource support. Hui Yi was mainly responsible for collecting relevant literature and reviewing background materials for the study.

All authors have read and agreed to the published version of the manuscript.

## References

1. Bhandari, A., & Bimo, S. (2022). Why's everyone on TikTok now? The algorithmized self and the future of self-making on social media. Social Media+ Society, 8(1), 20563051221086241.

2. Christin, A., Bernstein, M. S., Hancock, J. T., Jia, C., Mado, M. N., Tsai, J. L., & Xu, C. (2024). Internal fractures: The competing logics of social media platforms. Social Media+ Society, 10(3), 20563051241274668.

3. Darvin, R. (2022). Design, resistance and the performance of identity on TikTok. Discourse, Context & Media, 46, 100591.

4. Fazeli, S., Sabetti, J., & Ferrari, M. (2023). Performing qualitative content analysis of video data in social sciences and medicine: The visual-verbal video analysis method. International Journal of Qualitative Methods, 22, 16094069231185452.

5. Gao, Y., Liu, F., & Gao, L. (2023). Echo chamber effects on short video platforms. Scientific Reports, 13(1), 6282.

6. Geenen, J., Matelau-Doherty, T., & Norris, S. (2021). Visual transcription. A method to analyze the visual and visualize the audible in interaction. Revue Française Des Méthodes Visuelles, (5).

7. Gerbaudo, P. (2024). TikTok and the algorithmic transformation of social media publics: From social networks to social interest clusters. New Media & Society, 14614448241304106.

8. Halpin, S. N. (2024). Inter-coder agreement in qualitative coding: Considerations for its use. American Journal of Qualitative Research, 8(3), 23–43.

9. Hartmann, D., Wang, S. M., Pohlmann, L., & Berendt, B. (2025). A systematic review of echo chamber research: Comparative analysis of conceptualizations, operationalizations, and varying outcomes. Journal of Computational Social Science, 8(2), 52.

10. Henry, A. (2024). Digital ritual: Police–public social media encounters and 'authentic'interaction. The British Journal of Criminology, 64(2), 452–467.

11. Herman, L. (2023). For who page? TikTok creators' algorithmic dependencies.

12. Herzog, C., Handke, C., & Hitters, E. (2025). Digging in or building bridges? A scoping review of thematic analysis. Frontiers in Research Metrics and Analytics, 10, 1617380.

13. Huang, Y., & Ye, W. (2024). 'traffic rewards','algorithmic visibility', and 'advertiser satisfaction': How Chinese short-video platforms cultivate creators in stages. Convergence, 30(1), 659–682.

14. Joy, P., Hammond, B., Hammond, C., Bonardi, O., Wassef, K., & Ferlatte, O. (2024). Pulling at all threads: Reflections on using multimodal critical discourse analysis within arts-based health research. Qualitative Health Research, 10497323241307893.

15. Kaye, D. B. V., Chen, X., & Zeng, J. (2021). The co-evolution of two Chinese mobile short video apps: Parallel platformization of douyin and TikTok. Mobile Media & Communication, 9(2), 229–253.

16. Kibler, E., Vaara, E., & Laine, L. (2025). Cultivating insightful theory-building in qualitative research. Journal of Business Venturing Insights, 24, e00567.

17. Lim, W. M., Koay, K. Y., & Cheah, C. W. (2025). Commercialization of influencer content on social media: The roles of parasocial relationships and sponsorship disclosures. Journal of Strategic Marketing, 33(5), 658–684.

18. Lin, J., Swart, J., & Zeng, G. (2023). Theorising TikTok cultures: Neuro-images in the era of short videos. Media, Culture & Society, 45(8), 1550–1567.

19. Luo, X. (2026). Performing the digital pastoral: Environmental narrative, cultural nostalgia and commodified rurality in li ziqi's videos. Media, Culture & Society, 01634437261417430.

20. Mordecai, C. (2023). # anxiety: A multimodal discourse analysis of narrations of anxiety on TikTok. Computers and Composition, 67, 102763.

21. Mtisi, S. (2022). The qualitative case study research strategy as applied on a rural enterprise development doctoral research project. International Journal of Qualitative Methods, 21, 16094069221145849.

22. Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. New Media & Society, 20(11), 4275–4292.

23. Ohme, J., Araujo, T., Boeschoten, L., Freelon, D., Ram, N., Reeves, B. B., & Robinson, T. N. (2024). Digital trace data collection for social media effects research: APIs, data donation, and (screen) tracking. Communication Methods and Measures, 18(2), 124–141.

24. Paquienseguy, F., & Guo, Q. (2025). Douyin and the digital spread of intangible cultural heritage: Transforming cultural dissemination in the short videos age. Emerging Media, 3(2), 343–366.

25. Poell, T., Nieborg, D., & Van Dijck, J. (2019). Concepts of the digital society: Platformisation. Internet Policy Review, 8(4).

26. Quick, M., & Maddox, J. (2024). Us, them, right, wrong: How TikTok's green screen, duet, and stitch help shape political discourse. First Monday.

27. Schellewald, A. (2024). Discussing the role of TikTok sharing practices in everyday social life. International Journal of Communication, 18, 909–926.

28. Számel, K., Huszár, Á., Horváth, M., Rudas, S., Rakovics, Z., & Koltai, J. (2025). Social class and digital divide: Analyzing digital inequalities on census and digital behavioral data. Socius, 11, 23780231251357646.

29. Thomas, N., Hultgren, A. K., Zuaro, B., Yuksel, D., Wingrove, P., Nao, M., & Beach, D. (2024). Process tracing for applied linguistics. Research Methods in Applied Linguistics, 3(2), 100118.

30. Van Hulst, M., & Visser, E. L. (2025). Abductive analysis in qualitative research. Public Administration Review, 85(2), 567–580.

31. Vila-Henninger, L., Dupuy, C., Van Ingelgom, V., Caprioli, M., Teuber, F., Pennetreau, D., Bussi, M., & Le Gall, C. (2024). Abductive coding: Theory building and qualitative (re) analysis. Sociological Methods & Research, 53(2), 968–1001.

32. Yi, Y. (2023). Negotiating performance between policy and platform—Heritage practice of a Chinese craftsperson on douyin (TikTok). International Journal of Heritage Studies, 29(10), 1089–1109.

33. Zulli, D., & Zulli, D. J. (2022). Extending the internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. New Media & Society, 24(8), 1872–1890.

34. Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. Qualitative Health Research, 26(13), 1753–1760.

35. Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. Social Science & Medicine, 292, 114523.