

From Horizontal Screen to Vertical Screen: Visual Logic Reconstruction and Aesthetic Adaptation in Micro-Dramas

Xu Zhang*

Huanghe S & T University, Zhengzhou, 450000, China

*Corresponding author: m15837300919@163.com

Abstract: *The deep integration of mobile internet and the proliferation of handheld terminals have made the vertical screen format the dominant mode of presentation for micro-dramas, posing a fundamental challenge to the film and television language system rooted in horizontal screen aesthetics. The vertical screen is not a simple cropping of the horizontal screen; rather, it reconstructs visual grammar with the vertical frame as its logical starting point. This study examines this paradigm shift from the perspectives of visual logic reconstruction and aesthetic adaptation. The vertical screen employs central composition to shift the visual unit from spatial setting to human figure, with shot scales gravitating toward medium and close-up shots, thereby enhancing the intensity of visual focus. Spatial compression curbs the perception of depth, fostering a tendency toward flatness; visual information assumes a vertical hierarchy, and the trajectory of attention capture shifts to a top-down flow. The vertical frame achieves a dynamic balance between rapid transitions and the Gestalt of vertical movement, with editing reconstructing coherence along the vertical axis as the new baseline. Ultimately, the visual grammar of the vertical screen shapes an aesthetic form of immediacy, in which the visual intensity of the present moment becomes the core criterion of aesthetic judgment-revealing the mechanism through which media properties profoundly reshape visual language.*

Keywords: *vertical screen micro-dramas; visual logic; aspect ratio aesthetics; spatial compression; immediacy aesthetics; shot rhythm*

Introduction

The evolution of visual art has always been closely intertwined with the physical properties of its medium, and each shift in aspect ratio has been accompanied by a reconstruction of visual language and an adaptation of aesthetic paradigms. In recent years, as mobile terminals have become the primary carriers of image consumption, the vertical screen format has transformed from an expedient for playback media into an active choice in content production. The rise of micro-dramas epitomizes this shift in media landscape. However, existing film and television theories have long been constructed around the horizontal screen, making their discussions of composition principles, spatial staging, and editing logic difficult to apply directly to vertical-screen imagery. The horizontal screen follows the visual inertia of lateral extension, pursuing spatial exposition and depth layering, whereas the vertical screen confronts a contraction of information capacity and a reshaping of the visual trajectory along the vertical dimension—a fundamental rupture in visual logic between the two. As the aspect ratio shifts from horizontal to vertical, how the basic units of visual narrative are reorganized, how the mechanism of visual attention allocation is redirected, how shot rhythm and editing logic are adapted, and ultimately what new aesthetic forms emerge—these questions constitute the core issues in understanding vertical-screen imagery. This study focuses on the reconstruction of visual logic and the process of aesthetic adaptation in vertical-screen micro-dramas. By systematically examining the transformation across three dimensions—compositional paradigm, spatial perception, and visual rhythm—it seeks to reveal the shaping mechanism of vertical-screen imagery on visual language and its aesthetic implications.

1. Vertical Screen Compositional Paradigm and the Shift of Visual Focus

1.1 Paradigm Shift from Horizontal Screen Panoramic Narrative to Vertical Screen Central Composition

Horizontal screen film and television creation has long followed the visual logic of panoramic narrative, as the widescreen nature of the format provides ample visual capacity for spatial staging and environmental exposition, while the extension of the horizontal field allows the relationship between setting and character to be fully presented. The emergence of vertical screen micro-dramas has broken this established paradigm, as the vertical frame fundamentally alters the basic units of image composition and their mode of organization. The horizontal axis, which dominates the horizontal screen, is replaced in the vertical screen by the vertical axis, and the arrangement of visual elements is no longer oriented toward lateral extension but is reorganized around the vertical direction^[1].

The use of the long shot in the vertical screen is significantly constrained, as the limited frame height prevents the horizontal information contained in the long shot from being fully presented, thereby compelling creators to turn to a visual strategy centered on central composition. The human subject is positioned at the center of the frame, the background space is substantially compressed or even omitted, and the visual focus becomes highly concentrated on the character's face and its subtle expressions. This shift in compositional paradigm is not merely a simple adjustment of aspect ratio; rather, it signifies that the basic unit of visual narrative has shifted from spatial setting to human figure, and from environmental relationships to micro-expressions. Consequently, the core elements of visual language undergo a fundamental displacement.

1.2 The Generative Logic of the Visual Guidance Mechanism within Vertical Space

The vertical extension of the vertical screen format introduces a new technical pathway for the reconstruction of the visual guidance mechanism. When viewing horizontal screen content, the viewer's gaze typically moves in the horizontal direction, as the arrangement of elements within the frame follows the reading habit of moving from left to right or from right to left, resulting in a relatively balanced distribution of visual focus. In contrast, the vertical space of the vertical screen forces a shift in the trajectory of eye movement, making top-to-bottom or bottom-to-top vertical scanning the primary mode of visual reception. This transformation requires creators to establish a corresponding visual guidance mechanism to ensure that key information can be effectively captured during the brief viewing process. The placement of the character's face at the center constitutes the visual starting point, while factors such as the direction of the gaze, hand gestures, and the positioning of props together form a vertical visual path. Within the frame, techniques such as light distribution, color contrast, and focal plane manipulation are employed to construct a hierarchical structure of visual gravity along the vertical dimension, allowing the viewer's gaze to move up and down the frame along a predetermined trajectory. The swipe-based transition unique to the vertical screen further reinforces this vertical guidance mechanism, as the habitual action of swiping upward with a single finger becomes isomorphic with the direction of eye movement. With visual guidance and gestural guidance thus aligned along the vertical axis, the coherence of the viewing experience is thereby secured at the technical level.

1.3 Frame Constraints and the Reorganization of the Visual Information Hierarchy

The physical constraints of the vertical screen format exert a profound influence on the manner in which visual information is presented. Compared with the horizontal screen, the effective visual field of the vertical screen is significantly reduced, and its capacity to accommodate horizontal information is greatly diminished, necessitating a fundamental restructuring of the traditional methods of organizing on-screen information. Compound information commonly found in horizontal screen imagery—such as multiple characters in a single frame, environmental details, and depth layering—cannot be clearly presented simultaneously within the vertical screen, thereby significantly increasing the risk of information overload. To adapt to these constraints, micro-dramas have developed a hierarchical logic for organizing visual information. The foreground layer carries the core narrative information, typically consisting of the character's face and key actions, occupying the central area of the frame and receiving the highest visual priority. The middle-ground layer provides supplementary information, including hand gestures, handheld props, and close-range environmental elements, distributed across the midsection of the frame. The background layer serves an atmospheric function, being highly simplified

due to frame cropping and retaining only essential color or light cues. These three layers are arranged sequentially along the vertical direction, ensuring the prominent presentation of core information while avoiding visual clutter caused by horizontal information crowding. This reorganization of the information hierarchy endows vertical-screen imagery with a distinctive visual grammar, in which information reception no longer depends on the breadth of horizontal scanning but rather on the depth of vertical layering, thereby effecting a fundamental transformation in the allocation of visual attention^[2].

2. Spatial Compression and the Shift in Visual Perception

2.1 Shot Scale Transformation and Visual Focus Intensity under Frame Contraction

The physical constraints of the vertical screen format directly trigger a systematic adjustment of the shot scale system. In horizontal screen production, the choice of shot scale serves the panoramic presentation of narrative space and the complete expression of environmental relationships, with the alternating use of long shots, medium shots, and close-ups forming the basic framework of visual rhythm. In the vertical screen, due to the narrowing of the horizontal field of view, the expressive power of the long shot is significantly weakened, as the spatial relationship between characters and their environment becomes difficult to present fully within the limited frame. This situation compels creators to shift the emphasis of shot scale toward medium shots, close-ups, and extreme close-ups.

This shift is not merely a simple change in focal length selection; it signifies a fundamental enhancement of visual focus intensity. When the character's face occupies the main area of the frame, the viewer's gaze is forcibly locked onto micro-expressions and subtle movements, while redundant information that might otherwise disperse attention within the frame is substantially reduced. The direct effect of shot scale contraction is an increase in visual focus intensity, allowing the core information carried by each frame to be presented with greater density per unit of time. Close-up shots, which in horizontal screen production are often used as a means of emotional intensification, become a conventional narrative unit in the vertical screen. Subtle twitches of facial muscles, fleeting changes in the eyes, and slight openings of the lips all become legible visual signs, thereby shifting the language of imagery from macro-level spatial narration to micro-level expression-based narration.

2.2 The Suppression of Depth Perception by Vertical Composition and the Tendency toward Flatness

Horizontal screen imagery has long relied on depth arrangement to construct spatial layers, with the longitudinal arrangement of foreground, middle ground, and background providing viewers with the visual illusion of three-dimensional space. The vertical composition of the vertical screen fundamentally limits the scope within which depth arrangement can be deployed. Given the limited frame height, the longitudinal distance between foreground and background cannot be fully realized, thereby diminishing the expressive power of deep-focus shots. This suppression of depth perception gives rise to a tendency toward flatness in visual presentation, in which the relationship between character and background shifts from vertical separation to vertical superimposition^[3].

Background elements no longer exist as independent spatial layers; instead, they are attached to the periphery of the character in a vertically arranged manner, forming a visual structure akin to planar composition. Lighting also shifts from three-dimensional rendering to contour delineation, with an increased use of side-backlighting and rim lighting to compensate for the loss of volumetric sensation caused by depth compression. This tendency toward flatness does not represent a regression in visual expression; rather, it constitutes a new visual grammar engendered by the medium specificity of the vertical screen. It requires creators to reconstruct a sense of spatial order on the two-dimensional plane through means such as vertical relationships, size contrasts, and color superimposition. The weakening of depth perception directs the viewer's attention more toward the flow of information across the surface of the frame, rather than immersing it in an illusion of deep space.

2.3 The Mechanism of Capturing and Sustaining Visual Attention within the Vertical Field of View

The vertical field of view defined by the vertical screen format imposes new demands on the operational patterns of visual attention. When viewing horizontal screen content, viewers' eye movements are primarily characterized by horizontal scanning, with visual attention distributed relatively evenly, granting both the left and right sides of the frame equal opportunities to be noticed.

The vertical field of view of the vertical screen alters this balance: the distance traveled by the gaze in the vertical direction increases, while its range of movement in the horizontal direction narrows, thereby shifting the manner in which visual attention is captured. The central area of the frame, aligning with the natural landing point of the gaze, becomes the primary position for capturing attention, and creators concentrate core information in the midsection of the frame to ensure that it is received at the moment of viewing.

The sustenance of attention, in turn, depends on the design of visual flow along the vertical direction. By employing such means as the vertical extension of character movements, the upward or downward displacement of props, and the vertical scrolling of subtitles, creators guide the gaze to maintain continuous motion within the frame, thereby preventing the loss of attention caused by stagnant information. The vertical swipe-based transition unique to the vertical screen further reinforces this sustenance mechanism: each swipe constitutes a visual refresh, as a new frame enters the field of view from the vertical direction and forms vertical visual continuity with the preceding frame, ensuring the continuity of attention through technical means. Within the vertical field of view, the operation of visual attention no longer follows the inertia of horizontal scanning; instead, it forms a vertical flow trajectory moving from top to bottom or from bottom to top. This trajectory aligns closely with the operational habits of handheld devices, thereby constituting the unique physiological basis for viewing vertical-screen imagery^[4].

3. Vertical Screen Visual Rhythm and the Formation of Aesthetic Forms

3.1 The Shaping Effect of the Vertical Frame on the Shot Rhythm of Micro-Dramas

The vertical constraints of the vertical screen format exert a profound shaping effect on the shot rhythm of micro-dramas. In horizontal screen production, the determination of shot rhythm often takes into account the coordination among spatial staging, character movement, and environmental exposition, with the horizontal field of view provided by the widescreen format allowing perceptual space for relatively moderate shot transitions. In the vertical screen, due to the contraction of horizontal information capacity, the visual elements available for sustained gaze within the frame are significantly reduced; therefore, maintaining information density per unit time must rely on faster shot transitions. The vertical frame simultaneously alters the trajectory characteristics of intra-shot movement, making characters' vertical actions and vertical displacements the focus of visual attention, with the duration required by these movements directly influencing the length of the shot.

When a character performs a top-to-bottom movement within the vertical screen, the trajectory runs through the entire frame, and the visual Gestalt requires sufficient duration to be properly perceived—an intrinsic demand that imposes constraints on overly rapid transitions. The shot rhythm of the vertical screen thus achieves a dynamic balance between rapid transitions and the Gestalt of vertical movement, maintaining both a high-frequency output of information and ensuring the complete presentation of movement within the vertical field of view. The distribution of shot durations exhibits a bipolar tendency: some shots last as briefly as a few seconds to sustain rhythmic intensity, while others are relatively extended to accommodate the full execution of vertical movement, granting rhythmic elasticity a new form under the constraints of the vertical frame. This shaping effect is also manifested in the alignment between shot transitions and the vertical visual path; when the frequency of transitions resonates with the physiological rhythm of upward and downward eye movement, the fluency of viewing is enhanced, thereby establishing a comfort zone for rhythm. The vertical frame not only alters the distribution of shot lengths but also reshapes the logic of matching rhythm with content: intense emotional confrontations tend to correspond to faster transition frequencies, whereas nuanced emotional expressions require longer shot durations to be effectively conveyed, making rhythm itself a significant vehicle of meaning^[5].

3.2 The Adaptation Logic of Visual Continuity in Fast Editing for the Vertical Screen

Fast editing stands out as a prominent formal feature of micro-dramas, and maintaining visual continuity within the vertical frame becomes the core issue in adapting editing logic. Horizontal screen editing has long relied on horizontal motion matching and gaze alignment to sustain fluidity; when the direction of a character's movement aligns with the horizontal axis of the frame, the coherence of transitions is naturally achieved. Vertical screen editing must reconstruct this logic of continuity by establishing the vertical direction as the dominant axis for motion matching. When a character's gaze or

action extends vertically, the editing point is positioned at a vertically continuous point along the motion trajectory, enabling the viewer's gaze to transition smoothly between the upper and lower frames.

The stability of central composition provides a visual anchor for fast editing: with the character's face consistently positioned at the center of the frame, even during rapid shot transitions, the fixed location of the visual focus reduces the sense of abruptness caused by switching. Similarities in color, lighting, and compositional structure serve as implicit links that connect different shots, ensuring that the continuity of these visual elements maintains overall perceptual unity during rapid transitions. The adaptation of editing logic for the vertical screen is also reflected in the avoidance or transformation of horizontal movement. Horizontal traversal within the vertical screen tends to cause visual defocus; therefore, editing favors breaking down horizontal actions into vertically continuous presentations or reconstructing the movement path through combinations of multiple shots, thereby ensuring that the direction of movement consistently aligns with the vertical orientation of the frame.

At the same time, editing rhythm establishes a corresponding relationship with the information hierarchy within the vertical field of view: the switching frequency for core information is relatively slower, while that for auxiliary information can be appropriately accelerated, thereby guiding the allocation of the viewer's attention through variations in rhythmic density. The maintenance of visual continuity in fast editing also relies on the precise control of gaze direction; when the direction of a character's gaze remains consistent across consecutive shots, the viewer's sense of immersion is sustained, whereas sudden reversals in gaze direction require buffering through the insertion of transitional shots. The editing logic adapted to the vertical screen is, in essence, a reconstruction of spatial perception: it transforms spatial continuity in the horizontal direction into visual continuity in the vertical direction, establishing an intrinsic coordination mechanism between rapid transitions and vertical composition.

3.3 The Visual Grammar of the Vertical Screen and the Aesthetic Experience of Immediacy in Micro-Dramas

The formation of the visual grammar of the vertical screen has given rise to a distinctive aesthetic experience of immediacy specific to micro-dramas. The aesthetic habits cultivated by long-term exposure to horizontal screen imagery emphasize immersion and continuity, as viewers invest extended periods of time to enter the narrative world, with the accumulation and release of emotions following a gradual rhythm. In contrast, the visual grammar of vertical screen micro-dramas is organized around immediacy: the highly focused composition, the sustained contraction of shot scales, and the rapid advancement of rhythm collectively constitute an expressive system suited for instantaneous reception and quick response. Each frame pursues the immediate effect of visual impact, with close-ups of characters' faces directly conveying emotional signals and the climactic moments of actions compressed into short shots for presentation. The reception of information no longer depends on the accumulation of context across preceding and subsequent frames but relies instead on the visual intensity of a single frame.

This aesthetic experience of immediacy forms an isomorphism with the operational mode of the vertical screen: the replacement of images brought about by swiping up or down corresponds to the rapid alternation of emotions, allowing viewers to complete a full cycle of emotional consumption with each swipe. The simplification and intensification of visual grammar significantly increase the density of the aesthetic experience per unit of time, as the triggering of emotion, the arrival of climax, and the attainment of satisfaction are compressed into a matter of seconds, with the extensiveness of the aesthetic process replaced by instantaneousness. The aesthetic experience of immediacy is also reflected in the directness of emotional triggering: the visual grammar of the vertical screen weakens the roles of narrative groundwork and emotional buildup, instead employing the concentrated presentation of high-intensity visual symbols to place viewers in a state of high emotional arousal within an extremely short span of time. This aesthetic form imposes new demands on the encoding methods of visual language: rhetorical devices such as symbolism and metaphor, which require cognitive processing, are diminished, whereas direct presentation and immediate triggering become the dominant expressive strategies. At the same time, the aesthetic experience of immediacy does not simply consist of a patchwork of emotional fragments; rather, it establishes a frequency of emotional resonance through rapid alternation. When multiple short shots converge around the same emotional core, each transition serves as an accumulation of the same emotional intensity, ultimately coalescing into a high-density aesthetic sensation.

The aesthetic form of immediacy shaped by the visual grammar of the vertical screen reflects the profound restructuring of aesthetic modes by media properties, as the temporal dimension of images is redefined within the vertical screen, making the visual intensity of the present moment the core criterion of aesthetic judgment. This aesthetic experience further gives rise to new viewing habits: viewers gradually become accustomed to and anticipate such high-frequency emotional stimulation, thereby raising the threshold of aesthetic expectation, which in turn drives the evolution of visual grammar toward even greater density and heightened immediacy.

Conclusion

The transition from horizontal screen to vertical screen is not a simple adjustment of aspect ratio; rather, it involves an aesthetic reconstruction of the underlying logic of visual language. With its vertical constraints, the vertical screen challenges the visual paradigm long dominated by the horizontal screen: at the level of composition, it effects a shift from panoramic narrative to central composition, as the fundamental unit of imagery moves from spatial relationships to the human figure itself. Spatial compression redirects the shot scale system toward micro-level focus, depth perception is supplanted by flatness-oriented visuals, and a distinctive mechanism of attention capture emerges within the vertical field of view. Shot rhythm achieves a balance between rapid transitions and the Gestalt of vertical movement, while editing reconstructs visual continuity along the vertical axis as its new baseline, ultimately giving rise to an aesthetic form centered on immediacy. This transformation reveals the fundamental shaping of visual language by the physical properties of the medium: the temporal dimension of imagery is redefined, and the visual intensity of the present moment becomes the core criterion of aesthetic judgment. The formation of the visual grammar of the vertical screen does not constitute a negation of the horizontal screen tradition; rather, it represents an expansion of the possibilities of visual language under new media conditions. In the future, as display technologies evolve and viewing habits become further entrenched, the visual grammar of the vertical screen will continue to diversify and refine, and the mutual exchange and integration of the two visual logics will jointly propel the ongoing evolution of visual language.

References

- [1] Yuan Laichen. "A Study on Visual Metaphors in AI Micro-Dramas from the Perspective of Cultural Semiotics." *Vision*, no. 16 (2025): 17-20.
- [2] Zheng Jialin. *A Study on the Integration and Innovation of Cultural Tourism Micro-Dramas from the Perspective of Reception Aesthetics*. 2025. Southwest University, MA thesis.
- [3] Niu Muzhi. *A Study on Audience Guidance Strategies for Vertical Screen Micro-Dramas*. 2025. Southwest University, MA thesis.
- [4] Yuan Jingwei. "An Exploration of the Visual Narrative Innovation and Issues in Internet Vertical Screen Micro-Dramas." *Journal of Tonghua Normal University*, vol. 45, no. 11 (2024): 92-97.
- [5] Zhang Xing. "A Study on Communication Strategies Based on the Visual Characteristics of Vertical Screen Micro-Dramas." *Media Review*, no. 04 (2024): 39-41.