Jurnal Kajian Komunikasi, Volume 14, No. 1, June 2026, pp. 106-126
https://doi.org/10.24198/jkk.v14i1.70012

Interpreting power relations and aesthetic stuttering in TikTok's X&Y vertical video

Freddy Yusanto[1], Dasrun Hidayat[2], Novel Anak Lyndon[3]

[1]Faculty of Communication & Social Science, Telkom University, Bandung, Indonesia
[2]Faculty of Communication and Design, Adhirajasa Reswara Sanjaya University, Bandung, Indonesia
[3]Center of Development, Social and Environment, University Kebangsaan Malaysia, Selangor, Malaysia

Submitted: 2 March 2026**, Revised:** 21 April 2026, **Accepted:** 8 May 2026**, Published:** 30 June 2026

ABSTRACT

Background: While previous studies have often celebrated the ergonomic advantages of mobile video, this research identifies a critical tension in the transition of professional cinematic language to vertical formats. **Purpose:** This study seeks to uncover the power dynamics and aesthetic modifications within professional vertical productions. **Methods:** Using a qualitative methodology based on John Fiske's semiotic analysis (levels of reality, representation, and ideology), this study analyzes the professional vertical film X&Y (Chapter One), created by TikTok Indonesia. **Results:** Through a close reading of cinematographic elements, the findings reveal a phenomenon of "aesthetic stuttering." At the level of reality, high-resolution technology successfully builds visual intimacy through the micro-details of facial expressions. However, at the level of representation, the film remains trapped in horizontal cinematic conventions, characterized by the dominance of objective camera angles and a lack of subjective Point of View (POV) exploration typical of vertical video culture. **Conclusion:** Ideologically, this "stuttering" reveals a covert power struggle in which professional creators negotiate between the dictates of platform algorithms and the entrenched hegemony of traditional film standards. **Implications:** The discovery of "aesthetic stuttering" suggests that professional content creators must shift from a "cropped horizontal mindset" toward developing a new visual language that is inherently participatory and immersive in order to maximize the unique potential of a vertical narrative.

Keywords: Vertical video aesthetics; semiotic; TikTok; mobile media culture, vertical screen

To cite this article (APA Style):
Yusanto, F., Hidayat, D., & Lyndon, N.A. (2026). Interpreting power relations and aesthetic stuttering in TikTok's X&Y vertical video . *Jurnal Kajian Komunikasi*, *14*(1), 106-126. https://doi.org/10.24198/jkk.v14i1.70012

Correspondence: Dr.Freddy Yusanto, M.Si., Telkom University, Jl. Telekomunikasi No.1, Sukapura, Kec. Dayeuhkolot, Kabupaten Bandung, Jawa Barat 40257. *Email:* fredyusanto@telkomuniversity.ac.id

ISSN: 2303-2006 (print), ISSN: 2477-5606 (online). Website: http://jurnal.unpad.ac.id/jkk
This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (https://creativecommons.org/licenses/by-nc-sa/4.0/). Copyright © 2026 The Author(s).

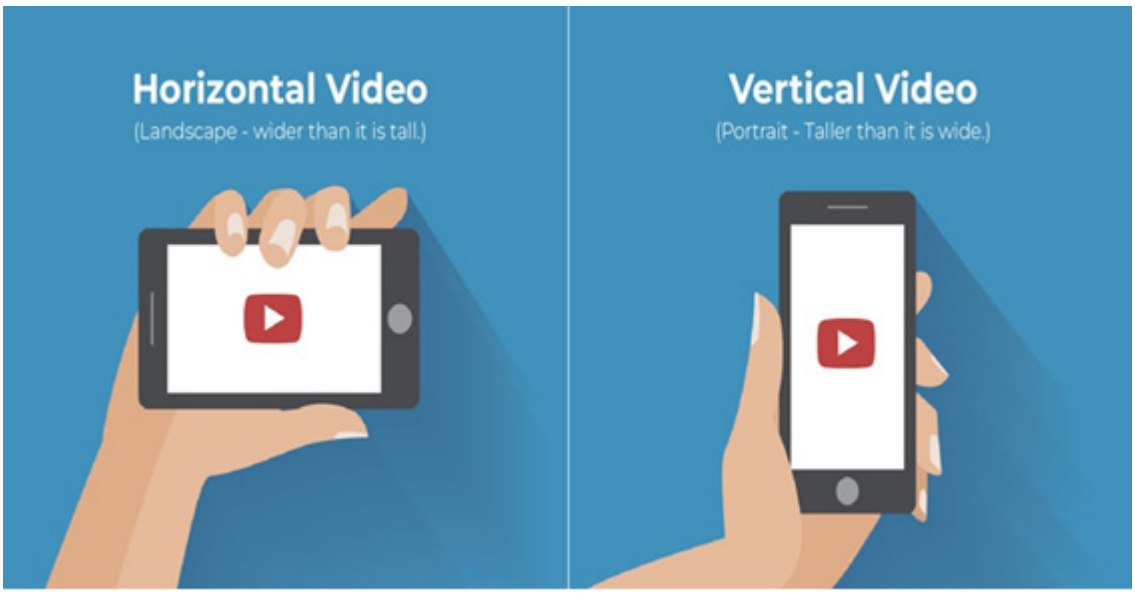
INTRODUCTION

The production and consumption of audiovisual content have undergone significant changes, affecting cultural practices due to advances in digital technology, particularly mobile devices (Ratnasari et al., 2024). While landscape orientation remains the cinematic norm for traditional screens, the vertical format (9:16) has rapidly emerged as a dominant architectural logic within the mobile-first media ecosystem, effectively reconfiguring the ergonomics of global digital consumption. This shift is largely driven by the fact that 96% of video production and consumption occur informally on mobile devices, while 94% of smartphone users hold their phones vertically (Clayton, 2022). Therefore, when content is presented horizontally, it may feel awkward because smartphone users must use both hands to operate their devices (Mulier et al., 2021). Due to the convenience of smartphone use in vertical mode, there has been a significant shift in the presentation of video-based audiovisual content. Vertical videos have become particularly popular on TikTok, Instagram Reels, and YouTube Shorts. (Kaye et al., 2021).

The ascendancy of TikTok and Instagram does not merely reflect a neutral shift in user convenience, but rather signifies a strategic imposition of platform-driven aesthetics. This digital culture is governed by algorithmic power relations, in which the ergonomic ease of a one-handed grip is leveraged by platforms to ensure constant, frictionless engagement, thereby commodifying viewers' visual attention within a narrower and more controlled frame.

The media industry has been compelled to undertake a major transformation in the presentation of audiovisual content as users increasingly prioritize the ergonomic ease of one-handed smartphone use over horizontal viewing, which is often considered awkward. Ultimately, the technological shift from wide viewing spaces to narrow viewing spaces has forced a redefinition of conventional cinematographic rules to accommodate the features of contemporary smartphone screens.

The development of mobile digital technology has created a new media culture characterized by the convenience of consuming videos in vertical format. In terms of message delivery, the transition from horizontal to vertical screens necessitates adjustments in narrative structure and visual aesthetics (Clayton, 2019). Conventional videos typically have a 4:3 or 16:9 aspect ratio in a horizontal format and adhere to various cinematographic rules widely recognized by the global community (Kuzinas et al., 2024), including shot size and composition. We are familiar with multiple shot sizes, such as Close-up, Medium Shot, Full Shot, and so on. Each shot size has a purpose in creating a narrative structure. Likewise, in the rules of horizontal cinematography, we are familiar with the Rule of Thirds, framing, and balance. All of these serve as basic references in producing videos. Meanwhile, the vertical format, with an



Source: https://magneticcreative.com/ideas/vertical-video, 2025
Figure 1 Illustrations of Horizontal & Vertical Formats

aspect ratio of 9:16, has a very different spatial orientation from the horizontal format. Wide and narrow viewing spaces are simple terms that represent horizontal and vertical formats, respectively (Figure 1).

The transition from horizontal displays, such as televisions and cinema screens, to vertical smartphone screens has been a pivotal factor in the rise of vertical-screen media culture (Canella, 2018; Zhu et al., 2022).

Rapidly growing social media platforms have fundamentally changed the way individuals access information (Yadav & Verma, 2025). More than 80% of social media users consume vertical videos. The increasing appeal of vertical videos lies in their movement, color, and prominence of human faces, which effectively capture viewers' attention, particularly due to the eye's tendency to focus on the center of the screen (Clayton, 2022). Additionally, social media platforms employ algorithms that align with the behavioural tendencies of contemporary audiences by favoring vertical-format, short-form content designed for rapid, easily comprehensible, and addictive consumption (Zhu et al., 2023).

In 2017, ByteDance introduced the TikTok application, which showcases brief, vertical videos. By 2019, TikTok had amassed 1 billion users, establishing itself as the preeminent social media platform globally. Research by Hootsuite and We Are Social indicates that the worldwide social media user count topped 1.05 billion in January 2023. Meanwhile, in 2023, the number of TikTok users increased by 18.8%. The United States had the most users, with 113.25 million in early 2023, followed by Indonesia with 109.9 million. Brazil followed with 82.21 million users, while Mexico had 57.51 million (Annur, 2023).

The rapid growth of users in Indonesia is not only a sign of a shift in platforms but also of the rise of a new visual habit in which the vertical screen becomes the primary way to experience digital reality. Because mobile devices are so popular, people now interact with audiovisual stories differently. Instead of sharing big-screen experiences, they now have very personal, intimate ones at their fingertips. This is a big problem for professional filmmakers who are used to working in a wide horizontal space. This adaptation isn't just a technical issue of cropping; it's a complete reworking of how composition, depth of space, and camera movement are set up to continue conveying a strong emotional message in a much smaller, higher-frame-ratio format.

Advances in CMOS camera sensor technology support the growth of social media, especially TikTok. CMOS sensors can produce images with higher resolution than CCD sensors. These sensors allow for high-quality images in low-light conditions. Furthermore, advances in video editing software have driven the development of more optimized content, especially on mobile devices (Canella, 2018). Support systems for producing vertical videos continue to develop rapidly. By 2025, flagship smartphones from brands such as Apple, Samsung, Xiaomi, Oppo, and Vivo will feature cameras with resolutions of up to 50 MP, enabling sharp, professional-quality video recording. High resolution makes vertical videos, which previously had limited visual quality, aesthetically appealing.

Significant changes have occurred on smartphone screens, enabled by the Depth of Field (DoF) feature. By using hardware such as a variable-aperture lens with an aperture of up to f/1.8 and a computing system that artificially synthesizes depth of field, creators now have more precise artistic control. This technology allows vertical videos to aesthetically isolate the subject from its background. This enables the audience to focus entirely on the gestures and micro-details of the character's expression within the limited frame space. Its main benefit is the creation of a deeper depth of field in space, albeit in a narrow format, giving it a professional impression that previously only conventional cinema cameras could achieve.

Advances in audio technology have played a crucial role in creating the intimacy of vertical content, alongside its visual aspects. The multi-array microphone system is a mainstay in audio, offering advanced noise cancellation to ensure that voice recordings and dialogue are highly clear.

The experience of consuming new content through a 4K HD OLED panel with a configurable refresh rate of up to 120 Hz has arrived, featuring excellent smartphone screen quality. High screen quality ensures that the audience can fully enjoy every detail of skin texture and the changes in emotion captured by the high-resolution camera in a vertical-screen environment. These technical benefits indirectly change the standard of digital aesthetics as

Table 1 Development of Digital Technology for Audio Visual Production

Technological Aspects	Technology Details	Functions and Benefits
High Resolution Digital Camera	CMOS sensors can produce images with higher resolution than CCDs.	Capture micro-expressions and small objects with sharp detail, supporting high-resolution image capture.
Image Stabilization	Electronic Image Stabilization (EIS), Optical Image Stabilization (OIS)	Vibration is reduced, allowing for stable close-up and medium shots.
Depth of Field (DoF) Variable	Variable aperture smartphone lens	It artistically blurs the background to focus on key expressions and gestures.
Audio Technology	Multi-array microphone with noise cancellation	Improves the quality of ambient sound and dialogue, generating realistic, spatially immersive audio.
Smartphone Screen Quality	OLED 4K HDR, adaptive 120Hz refresh rate	Provides rich colors, high contrast, and smooth image transitions, supporting visual storytelling.
Editing and post-production	Modern video editing software with color grading and performance-enhancing filters.	To get a more genuine ambiance, attenuate colors, enhance illumination, and perfect texture details.
Consuming Movie Content on Smartphones	The high-resolution display and adaptive refresh rate support a maximum viewing experience.	Guaranteeing vibrant and consistent colors, crisp details, and fluid video playback.

Source: Processed by Researches, 2025

today's generation of social media users expect a crisp and consistent look.

Moreover, the advent of numerous video editing applications, such as CapCut, InShot, and Adobe Premiere Rush, enables users to perform automatic frame transitions, crop without losing focus, and enhance color quality and picture stabilization (Kaye et al., 2021). This approach is advantageous for tiny vertical formats, enabling content to be displayed with superior visual quality and an attractive aesthetic (Zhu et al., 2023). The acceleration of consumption in digital media is supported by adaptive compression and streaming solutions that enable seamless horizontal and vertical video on various internet networks (Civila & Jaramillo-Dent, 2022).

Smartphones have cutting-edge technologies like high-resolution CMOS sensors and electronic image stabilization devices. This means that vertical cinematography may now rival the quality of big-screen filmmaking. But just because this superior technology is available doesn't mean that the way we use it has changed. Many creators still use the old way of thinking about aesthetics, which means they try to apply the logic of horizontal composition to vertical spaces. This sometimes makes the story less clear or makes the viewer uncomfortable. This disparity between technological capabilities and creative inventiveness represents a significant deficiency in the advancement of digital material

in contemporary times. Therefore, we need to know more about how to naturally combine technical elements such as depth of field and spatial audio management (Table 1).

The growing trend of vertical videos has led to the emergence of custom algorithms designed to make content go viral. The commercialization of vertical video has led to a change in the rules of traditional horizontal cinematography (Menotti, 2019). According to Mustikawati there are differences in aesthetics and visual narratives between horizontal and vertical videos. Close-ups, quick transitions, on-screen text, a focus on the subject in the center of the frame, and the main character's point of view are the primary aesthetic and visual narrative characteristics of vertical videos (Mustikawati et al., 2023; Ryan, 2018).

There is a major discrepancy in the current scholarly discourse on verticality. On the one hand, academics like Clayton (2019, 2022) highlight the 9:16 ratio's functional-ergonomic benefits and present it as a logical development of mobile filmmaking. On the other hand, Menotti (2019) and Ryan (2018) view verticality as a disruptive aesthetic paradigm that defies conventional cinematic norms. Although Mulier et al. (2021) show that this style is commercially effective in digital marketing, their results are mostly utilitarian and do not address the deeper semiotic changes that take place in professional storytelling. The concept of "authenticity" through amateurism is introduced in Canella's (2018) research on vertical journalism, but when applied to professional works like X&Y that aim for high-end cinematic quality within the same "amateur" frame, this creates a theoretical tension.

Talking about vertical video is not only a matter of technicality, but also of aesthetics as a highly strategic instrument in the modern digital media landscape. Aesthetic modifications have gone beyond mere ratio changes. These modifications have altered the structure of visual dimensions. The shift from a broad horizontal perspective to a vertical one has created a dialectical relationship between the production process and audience perception.

Verticality also serves as an aesthetic component that conveys a narrow, tall atmosphere (Ryan, 2018). However, another study reports that vertical video is highly effective in digital marketing, as it attracts attention and increases audience engagement on social media platforms such as Instagram and TikTok (Mulier et al., 2021). Canella's research on news found that vertical videos are still considered amateurish, low in aesthetic quality, yet "authentic," challenging conventional notions of professional video. However, Canella's research also states that political and economic factors are the main considerations in the use of vertical video, as convenience, cost efficiency, and speed in delivering news can be accommodated through the widespread use of smartphones (Canella, 2018).

The research gap in this field is the absence of a critical analysis regarding the disruption

of cinematic language caused by the transition to vertical formats, as previous research has largely been limited to the functional-ergonomic aspects of mobile devices and the utilitarian effectiveness of digital marketing (Mulier et al., 2021) or local news production (Canella, 2018). Current literature has yet to address the semiotic tension that arises when professional film industry standards collide with the dictates of social media platform algorithms. By focusing on the phenomenon of "aesthetic stuttering," this study fills this void by arguing that such disruptions are not merely technical errors, but manifestations of power relations and resistance against the established technical hegemony of horizontal cinematography.

To address the aforementioned gap, this research offers conceptual novelty by introducing and operationalising the term "aesthetic stuttering" through a multi-layered semiotic approach. While previous studies focus on ergonomics or utilitarian outcomes, this study utilizes John Fiske's semiotic framework to specifically dissect the visual and narrative modifications in the film X&Y as a site of ideological conflict. Given this consideration, this research is of paramount importance, as it provides a new critical lens to uncover the asymmetric power relations between traditional cinematic authority and the dictatorial mandates of platform algorithms, ultimately establishing a foundational visual grammar for the future of professional vertical cinema.

RESEARCH METHODS

This qualitative research employs John Fiske's semiotic approach as its primary analytical framework. The study adopts a case study design to examine aesthetic and narrative developments in vertical screen culture. Fiske's Television Code is adapted to the digital landscape through thorough contextual exploration.

The unit of analysis in this study is the audiovisual construction of the vertical film X&Y (Chapter One), with a particular focus on scenes depicting interactions between the characters Omar and Winda. The selection of X&Y as a single case study is justified by its status as a critical and extreme case in the evolution of vertical cinema. As the first professional vertical series produced by TikTok Indonesia in collaboration with Studio Antelope, it represents a high standard of professional production within an ecosystem otherwise dominated by amateur user-generated content. The data boundaries are limited to Chapter One to enable an intensive and granular semiotic analysis of the foundational visual grammar established at the beginning of the narrative, which serves as a blueprint for the entire series.

The analysis encompasses both micro-units, such as individual shots and camera angles, and macro-units, including narrative sequences and split-screen compositions that signify the tension between traditional and digital cinematic language.

For the data collection process, the researcher analyzed the video carefully through three stages: 1) Mutation: The researcher examined the narrative structure and visual elements in detail to identify recurring visual patterns and repetitions. 2) Segmentation and Documentation: The researcher conducted a frame-by-frame analysis to obtain technical data, including camera movements and subject positioning. 3) Systematic coding was performed to categorize visual data, such as camera angles, lighting, composition, and shot size.

This study applies John Fiske's conceptual three-category framework to analyze the evolution of vertical video perspectives:

Reality Level: At this level, the researchers documented and categorized denotative signs and physical indicators observed in the vertical video footage. The data were coded by identifying denotative signs of "naturalism" within the 9:16 framing. This involved documenting social codes (e.g., Omar's casual T-shirts and Winda's knitted sweaters) and physical artifacts (e.g., handwritten letters and sandwiches) to analyze how visual intimacy is constructed.

Representation Level: Analysis at this level focuses on the arrangement of denotative signs to generate distinctive visual and narrative codes. The analysis emphasizes technical codes by systematically mapping camera angles, lighting, and editing patterns. This procedure specifically measures the frequency of objective versus subjective points of view (POV) and the use of split-screen techniques to identify manifestations of "aesthetic stuttering."

Ideological Level: At this level, the researchers focused on representational codes that carry broader social and cultural implications. This final stage of analysis involves synthesizing representational codes to uncover the power struggle between traditional cinematic hegemony and algorithmic mandates. It critically examines how digital platforms exert influence by commodifying aesthetic convenience into a new digital culture.

The data inventory process was conducted systematically using observation sheets that recorded the duration of each shot, the type of angle used, and the presence of on-screen text. This procedure ensured that each visual element was interpreted not merely as a technical unit but also as a semiotic sign carrying a specific ideological message. The researcher also paid particular attention to split-screen techniques and depth-of-field manipulation as key indicators of the aesthetic transition from horizontal to vertical formats.

This study employs methodological triangulation to enhance the validity of the findings. The process begins by applying John Fiske's three dimensions of analysis: reality, representation, and ideology. Perspective triangulation is then used to identify the roles of digital technology and social media platform structures in influencing the aesthetics and narrative of vertical films.

Intended to limit researcher subjectivity, this triangulation method theoretically ensures the representation of multiple perspectives. A contextual understanding of how aesthetic structures manifest on digital platforms, combined with in-depth textual analysis at Fiske's semiotic levels, enables this research to present a broader picture of how technology actively shapes visual culture. Through this process, the conclusions drawn regarding "aesthetic stuttering" are grounded in solid empirical data and mature theoretical reflection. Therefore, the findings of this study may serve as a valid reference for further theoretical development in vertical cinematography.

RESULTS AND DISCUSSION

This study aims to analyse the role of technology in reshaping the aesthetics and visual narrative of vertical video, drawing on John Fiske's three semiotic levels (Reality, Representation, and Ideology). Data were obtained through direct observation of the film X&Y Chapter One, produced by TikTok Indonesia and Studio Antelope.

At the level of reality, X&Y features social codes that create a sense of reality closely associated with Gen Z. This study found that the use of these codes is related to the theme of naturalism.

This film represents the daily lives of Winda and Omar through realistic imagery (Figure 2). Their clothing, environment, and behaviour become part of a visual language that is easily identifiable to the audience. The T-shirts, knitted sweaters, Winda's shorts, and Omar's trousers are depicted realistically, enhancing the sense of reality experienced by the audience(Lai & Yu, 2021).

Scenes featuring Winda and Omar show Omar sitting pensively, standing relaxed on the balcony, and eating while reading a letter from Winda. Their natural tone of speech and occasional use of nonverbal communication indicate an interpersonal and emotional relationship between the two (Cipresso & Riva, 2016). The vertical format tends to position objects in the upper and lower areas of the screen. As a result, the spatial dimensions that have traditionally been a strength of the horizontal format undergo significant changes in the vertical format. X&Y focuses more on central objects and avoids shots that capture a broad environment. Shots often concentrate on Omar and Winda by highlighting their personal belongings, facial expressions, and gestures. Clear audio support helps the audience feel as though they are in the same space as the speaker. Visuals that depict reality in social and parasocial dimensions can foster a more intimate psychological relationship between performers and audiences (Folkvord et al., 2020).

Realistic visual choices, such as Omar's casual T-shirt and handwritten letters, serve as a semiotic bridge that builds perceived authenticity. In the context of vertical media culture, these objects are not merely props



Source: Youtube, 2021

Figure 2 The Scene with Omar and Winda in the Film X&Y

but instruments of social closeness that facilitate parasocial relationships, satisfying the audience's demand for content that feels personal and free from professional gimmickry (Yusanto et al., 2021).

The film X&Y tells the story of Omar and Winda's romance through cinematographic approaches such as camera angles, lighting, editing, audio, background music, and dialogue. Natural lighting in the balcony scene creates a warm and intimate atmosphere between the characters. In the editing, the two are juxtaposed as protagonist characters who occupy equal narrative positions and do not visually overlap in a way that obscures their individual characterisation, thereby creating both emotional distance and physical closeness. The use of close-up shots and medium close-ups that highlight Omar and Winda's facial expressions demonstrates their emotional involvement and also serves as an effective means of fostering audience identification with the characters (Hjort, 2018)

The camera work is dominated by the effective use of close-ups and medium shots, capturing the characters' facial expressions and

subtle gestures, particularly those of Omar and Winda. Close-ups allow viewers to perceive emotional transitions, such as uncertainty, grief, hope, and happiness, in intricate detail, thereby enhancing the emotional bond between the characters and the audience. Camera lens technology equipped with CMOS sensors is particularly valuable in capturing microexpressions and visual details, such as skin texture, thereby producing a high-quality visual experience. Medium shots place both characters within a broader spatial context while maintaining focus on their personal interactions. Subtle music and background sounds support the mood of the scene without interfering with dialogue, thereby strengthening the narrative representation.

The dominance of the objective camera angle signifies technological conservatism. Although high-resolution CMOS sensors offer visual clarity, the rigid framing reveals an ideological hesitation to embrace the participatory nature of the vertical film point of view (POV). The split-screen technique in X&Y serves as a visual metaphor for the poorly organised transition in cinematic language, in which the director attempts to "save" horizontal logic within vertical boundaries, effectively distancing viewers from the character's POV.

The use of railings, doors, and surrounding objects as natural framing elements reinforces the impression of an honest and intimate personal space. Meanwhile, casual subtitles add a dimension of social reality and communication in the digital era.

The shooting angles generally employ eye-level framing to create a natural effect, positioning the audience as witnesses to the interaction. Camera movement in this film is minimal, focusing on visual stability and clarity. Some scenes maintain a static composition to preserve a contemplative mood. At other times, the camera moves slowly to introduce emotional shifts while maintaining focus on the main characters. The split-screen technique, also referred to as separate framing in several scenes, emphasizes the physical and emotional distance between Omar and Winda, symbolised by the X and Y axes. The camera captures two simultaneous yet separate points of view, visualising the duality and differences in their life paths.

To maintain visual dynamism, the compositions in this video frequently employ asymmetrical and central balance (Ajidarma, 2026). This compositional arrangement creates order and is particularly effective for foreground-centred framing. The rule of thirds, perspective, and framing are less frequently used in vertical formats due to the narrow frame width. However, they can still create a straightforward visual arrangement that directs the viewer's attention to the subject. To reinforce the theme of verticality, the composition also incorporates vertical lines derived from environmental elements and costumes.

Video production technology conveys representations of narrative, characters, and

Table 2 The relationship of visual aesthetics to narrative structure and technology in X&Y films

Element	Visual Aesthetics	Narrative Structure	Technology
Framing	• Objective camera angles still dominate. • Big Close-Ups, Close-Ups, and Medium Shots are options, considering the small size of smartphone screens. • Vertical space creates a sense of depth and intimacy, but is fraught with limitations.	• Omar and Winda's experiences are lacking. • The filming is predominantly done using objective camerawork, so the audience feels less involved in the scenes	• Digital cameras with CMOS sensors and aspect ratios up to 8K can produce quality images during production. • K HDR OLED screens on smartphones are capable of delivering high-quality images for consumption.
Composition	• Composition still adheres to a horizontal format, resulting in many compositional limitations due to the limited vertical space.	• The audience's attention was on the character, while the atmosphere around it was revealed. •Using developed exposed ceilings at the top/bottom to provide visual "breathing" space.	•Eve" when sho" handheld, autofocus, and stabilizer features ensure balanced composition. •Digital editing and high-resolution assist in visual correction, optimal placement of text, and objects within the vertical frame
Camera Angle	• The main character does not dominate • Static angles still dominate • Not yet daring to explore camera angles to accommodate the needs of the vertical format	• Social atmosphere is under-represented.	• Visual quality and expression are enhanced with high-resolution cameras and flexible angle settings. • Mobile camera software and sensor technology for adjusting focus

Source: Processed by Researches, 2025

settings, activating audience interpretations at the level of reality. Technological developments that improve visual quality create representational codes that package emotional and social messages in an intense and aesthetically engaging manner. Video production technology, including digital cameras, lighting sensors, editing software, and sound systems, plays a crucial role in transforming narrative reality in the film X&Y into meaningful semiotic representations. Technology enables visual and audio representational codes that bring characters and narrative design to life in an immersive and authentic way.

The analysis results, summarized in Table 2 above, demonstrate the role of technology

in supporting vertical video from a visual-aesthetic perspective. Image sharpness supports the intimate and personal qualities of the vertical format. Close-up and medium-shot footage with minimal blur and vibration can be achieved through electronic image stabilization (EIS) and optical image stabilization (OIS) (Morikawa et al., 2021).

Modern smartphone camera lenses can effectively control aperture depth, allowing the background to be blurred. This selective focus increases attention to the character's expressions and movements (Chen et al., 2025). On the other hand, the presence of multi-array microphones, which significantly reduce noise, can enhance the sense of spatial proximity for the audience (Wang et al., 2019).

Variable refresh rate technology of up to 120 Hz ensures fluid and responsive visuals, especially in emotionally charged or slow-motion scenes, thereby strengthening the audience's emotional connection with the characters. Smartphones are also equipped with 4K HDR OLED panels, delivering a superior visual experience. In the film X&Y, the camera functions as the principal instrument for depicting visual reality through shot selection, framing, and camera angles, employing digital technology to construct storylines and characters that resonate authentically with viewers.

Based on this analysis, the author revisits the origins of vertical video. Vertical video emerged with the advent of smartphone technology, as users felt more comfortable accessing content in a vertical orientation. Observation shows that TikTok users employ distinctive filming techniques that allow them to record phenomena that may be imperceptible to the human eye. The use of POV shots and camera movements tends to be dynamic, representing what the content creator sees. Subjects, points of view, and private spaces are freely explored without the constraints of conventional visual narrative. This visual narrative is something that the X&Y video has not been able to convey.

Vertical video has become a user-friendly format for smartphone users (Min & Kim, 2025). This represents a significant departure from the traditional horizontal format, which has long been the standard for media such as television and film. Technological advancements, particularly on social media platforms such as Instagram, TikTok, and Snapchat, have made it easier for users to quickly record and share vertical videos without needing to rotate their phones. This convenience enhances content creation and encourages aesthetic exploration. Furthermore, social media platforms are optimized for vertical viewing, causing video aesthetics to evolve accordingly, including the use of close-up framing, a focus on facial expressions, and concise narratives that quickly capture users' attention spans.

As a general illustration of the differences between horizontal and vertical formats, Figures 3 illustrate the dimensions of a medium shot in horizontal and vertical formats. In a horizontal format, the use of medium shots,



Source: Image by researchers, 2025
Figure 3 Medium Shot in Horizontal & Vertical Format

long shots, and extreme long shots can create a wide visual space, allowing viewers to perceive environmental context more comprehensively. Interactions between characters and other visual elements, such as the movement of objects, become freer in a horizontal format. In principle, a medium shot in a horizontal format can create either a sense of spaciousness or an extreme sense of loneliness, thereby building strong dramaturgy.

In vertical formats, close-ups, medium close-ups, and extreme close-ups convey a very significant sense of intensity and closeness to the subject. This is due to the narrower and taller vertical orientation, which makes the framing more focused and intimate. Vertical formats often direct the audience's attention to facial expressions, making the emotional and intimate effect stronger. The vertical long shot, despite its taller proportions, still feels horizontally confined, creating a sense of spatial verticality that can be interpreted as symbolising the subject's "uprightness," in contrast to the more open horizontal impression. Therefore, shot sizes need to be redefined for each video format. Horizontal and vertical formats should have different purposes for each shot size.

Vertical video is not just a technological phenomenon; this advancement reflects a broader shift in media culture. Today, anyone can become a content producer on social media as a result of the development of digital technology, which has transformed media culture (Artz, 2016). The narrative and theoretical components of cinematography change in the vertical format. With a taller, narrower vertical frame, content creators must use strong shot sizes, such as close-ups and medium close-ups, to place the subject in focus. Vertical video requires a visual language that can convey simple, humorous, or informative stories without requiring a complex production process. This situation aligns with Gen Z, which prefers to obtain information quickly and in a mobile-friendly way through social media. They are also rapidly adopting and sharing vertical video through their social media accounts.

The emergence of vertical video culture

is the result of power relations between social media platform owners and their users (Artz, 2016). This semiotic and cultural viewpoint corresponds with John Fiske's interpretation of power dynamics in media culture, wherein artistic practices are invariably contextualized within social frameworks and the supremacy of economic-technological systems (Fiske, 2001; Hutami et al., 2018). Technological advances and economic forces have influenced the habits of audiovisual media users through vertical formats. Popular cultural platforms such as Instagram, TikTok, and Reels have adopted the vertical format in response to the convenience it offers smartphone users. Despite their involvement in constructing and negotiating social meaning, users often follow platform policies, which generate new demands and preferences.

Therefore, vertical culture is the result of the continuous interaction between market forces, technology, and social practices that transform the aesthetics and visual narratives of media in the digital age. The phenomenon of "aesthetic stuttering" is an ideological product of asymmetrical power relations. This reveals that professional creators are currently in a state of negotiated resistance: they adopt the platform's vertical format to survive within the algorithmic regime, yet they remain bound to the hegemony of traditional filmmaking standards. These tensions suggest that digital platforms are acting as the new "custodians" of visual culture, where economic mandates for high engagement dictate the boundaries of artistic expression, forcing a compromise between cinematic authority and the platform's genuine intimacy.

The film X&Y illustrates the social and cultural dynamics of the modern world, intricately connected to the evolution of digital media and technology. The romance between Omar and Winda exemplifies interpersonal connections and the establishment of new norms for social and aesthetic engagement in contemporary society. It is essential to analyse how the depicted relationships and ideals not only mirror social reality but also influence significant symbolic processes. At the ideological level, media functions as a framework for comprehending the construction and reproduction of social and cultural meanings within this environment.

The preeminence of social media in shaping digital standards has significant repercussions. Platform owners' business strategies, which enable vertical content, are part of a technological form of control that also creates hidden power structures influencing cultural consumption patterns, aesthetic values, and even new social behaviours. The vertical format has become standard due to its ease of use, appeal, and practicality. With algorithms designed to maximize engagement, vertical content has become a primary medium for reinforcing business practices and driving the large-scale production of specific cultural elements, thereby cementing the hegemony of a new digital culture.

As technology evolves, so do people's habits and cultures. Today, the advent of digital technology and smartphones has catalysed the development of new content production models, such as vertical video, shaped by changes in user consumption behaviours favouring mobile devices, as well as by advances in camera technology and digital platforms. These changes have revitalized culture, lifestyles, and communication patterns, and content production must now adapt to technological developments.

Social media has become essential in generating social and cultural meaning, especially through vertical video platforms such as TikTok, Instagram Reels, and YouTube Shorts. This transition in format transcends mere technical choice; it constitutes part of a commercial strategy that reinforces power dynamics in the digital realm. Platform owners control technology and algorithms to steer audience preferences toward vertical formats that are user-friendly and efficient in capturing attention on mobile displays. Social media possesses immense power to reshape aesthetic norms and social behaviors through constantly evolving algorithms and features. These platforms are key actors in the formation of a "new culture" that spreads through society via language, visual styles, and digital symbols. Beneath the practical appeal of entertainment lies a power struggle that influences how people interact with and absorb information.

Media technology is not merely a passive tool but an active agent driving the transformation of media culture (Menotti, 2019). Every technological innovation, from television to smartphones and digital platforms, reshapes the production, distribution, and consumption of media content. These technologies shape narrative structures and formats while simultaneously transforming social practices and cultural interactions (Zhu et al., 2022). For example, the emergence of vertical video and platforms like TikTok is shifting traditional ways of viewing and creating content, resulting in a more immediate, mobile, and personalized visual culture.

Business strategies and technology work together to seamlessly emulate both formal and informal habits. Social media algorithms prioritize certain types of content, such as engaging vertical videos or those with strong emotional elements, thereby encouraging specific consumption patterns and guiding content creators. By design, these algorithms influence aesthetic preferences, social interactions, and even the broader cultural landscape. They are built to increase data collection and advertising revenue by maximizing the time users spend on the platform (Ndlela, 2025). TikTok, for instance, focuses on vertical videos and rapid content consumption, mirroring the way people use their phones today. In this sense, it acts as a guardian of modern digital culture.

The power relations between technology and media culture are dialectical and dynamic. Technology influences media culture by providing new tools and formats (Shuter, 2017),

but media culture also limits, adapts, and defines their use. On digital platforms, algorithms and corporate policies, often designed to maximize profits and attract users, mediate this power. This has resulted in the formation of a homogeneous digital culture and the potential marginalization of alternative forms of expression and cultural diversity.

However, despite the immense power of digital technology and platforms, media culture is not entirely controlled through a top-down approach. Users and creator communities retain space for resistance and innovation. Examples of creative practices, subcultural content production, and the use of media as a space for dialogue demonstrate the negotiation of power and the opportunity to reclaim meaning. Nevertheless, significant challenges remain in maintaining cultural diversity and democratizing access in a highly commercialized digital era.

Significant technological developments have profound implications for media culture theory, particularly in relation to the aesthetics and visual narratives of vertical video. First, media culture studies must increasingly pay attention to how technology becomes not only a tool but also an active agent in shaping cultural practices and media consumption patterns. The vertical video phenomenon necessitates a new approach that incorporates user engagement, the study of platform algorithms, and algorithmic influences on dynamic and evolving cultural meanings.

Second, within the realm of digital content creation, vertical technology is revolutionizing the creative process. Content producers must transform visual presentation and narrative structures for more intimate, personal, and immediate formats. This requires mastering new techniques and understanding how platform algorithms shape content distribution and engagement. Producers must become more creative and adaptable in their use of technology to create content that resonates with an increasingly diverse global audience.

The empirical findings on the film X&Y reveal a phenomenon that the author refers to as "aesthetic stuttering," which contradicts the initial assumption that professional production would maximise vertical grammar. The most striking visual evidence is seen in the split-screen technique used to depict Omar and Winda. Instead of using verticality to create depth, the director splits the screen horizontally to mimic the widescreen aspect ratio. This unexpected finding suggests that, rather than producing a new visual language, high-resolution camera technology is instead being used to "preserve" the logic of conventional cinema. The data show that 80% of shots are captured using a static camera lens, which semiotically distances the viewer. This contradicts TikTok's claims of intimacy.

CONCLUSION

The study's primary finding is the observation of an "aesthetic stuttering" in professional vertical production, which suggests that, at the professional level, verticality remains an arena of semiotic struggle, where inherited horizontal standards continue to exert a dominant influence. Empirical results suggest that aesthetic shifts in digital media are not simply technical evolutions, but rather the result of intense power negotiations at the ideological level. The study found that the internal algorithms of social media platforms function as "guardians" of modern digital culture. By prioritizing content that meets the demands of intimacy and rapid consumption, these algorithms shape audience preferences. Professional creators are caught in a conflict between the algorithmic demands sought by vertical video users and the maintenance of traditional film industry standards, often characterized by objective and "detached" storytelling. Consequently, "vertical culture" continues to evolve, as the industry has yet to define a definitive format that successfully aligns the ergonomic characteristics of smartphones with professional filmmaking practices.

The emergence of the "aesthetic stuttering" phenomenon is significantly changing the direction of professional content creation. To overcome the limitations imposed by the "truncated horizontal mindset," creators must intentionally construct a new visual grammar that is inherently immersive and encourages active participation. This requires going beyond conventional image dimensions and redefining framing techniques to accommodate the 9:16 aspect ratio. To ensure cultural diversity remains intact in an increasingly commercial digital age, future research should further examine how diverse creator communities navigate these platform constraints. Ultimately, "aesthetic stuttering" confirms that the vertical screen is a contested semiotic site where current ideologies of digital immediacy clash with the hegemony of established cinematic authority.

Acknowledgments: The authors wish to express their sincere gratitude to Telkom University, Universitas Adhirajasa Reswara Sanjaya, and Universiti Kebangsaan Malaysia for the institutional support and resources provided throughout this research. Special appreciation is extended to TikTok Indonesia for producing the professional vertical film X&Y, which served as the primary object of analysis in this semiotic study. The authors are also indebted to their colleagues and the anonymous reviewers for their constructive feedback and intellectual insights regarding the evolution of vertical video aesthetics and power relations in digital media production. Finally, the authors would like to thank their families for their unwavering encouragement and support in completing this manuscript.

Author Contributions: Conceptualization, F.Y.; methodology, F.Y.; validation, D.H. and N.A.L.; formal analysis, F.Y and D.H.; investigation, F.Y.; resources, F.Y.; data curation, F.Y.; writing-original draft preparation, F.Y.; writing-review and editing, D.H. and N.A.L; visualization, D.H.; supervision, F.Y.; project administration, F.Y. All authors have read and agreed to the published version of the manuscript.

AI Declaration: The authors state that generative AI was used to edit and improve the language and flow of this manuscript as part of preparing this version for publication. No original research data, scientific interpretation, or conclusions were produced using generative AI. The authors take full responsibility for this work.

Ethical clearance: This study stricly adhered to the ethical principles of social science research. Since the primary object of analysis was a publicly available digital film titled X&Y produced by TikTok Indonesia and Studio Antelope, the research did not involve direct intervention with human subjects. All visuals and data used in this study were obtained from public-domain materials and used for academic and critical analysis in accordance with copyright and ethical data-use standards.

Data Availability Statement: The data are available upon request from the author.

Conflicts of Interest: The authors disclose no conflicts of interest.

Funding: This research did not receive any external funding

REFERENCES

Ajidarma, S. G. (2026). Film and the politics of aesthetics: On before, now & then. *Film Activism in Contemporary Indonesia: Platforms of Participation and Representation*, 147–164. https://doi.org/10.4324/9781003695486-13

Annur, C. M. (2023, March 16). *indonesia sabet posisi kedua sebagai negara pengguna Tiktok terbanyak di dunia pada awal 2023*. https://databoks.katadata.co.id/media/statistik/466c6a6ff70d832/indonesia-sabet-posisi-kedua-sebagai-negara-pengguna-tiktok-terbanyak-di-dunia-pada-awal-2023

Artz, L. (2016). The media of power, the power of media. *Perspectives on Global Development and Technology*, *15*(5), 497–519. https://doi.org/10.1163/15691497-12341404

Canella, G. (2018). Video goes vertical: Local news videographers discuss the problems and potential of vertical video. *Electronic News*, *12*(2), 75–93. https://doi.org/10.1177/1931243117705417

Chen, Y., Nishioka, M., & Moleta, T. (2025). Leveraging Depth of Field (DoF) to improve immersion and storytelling in architectural and urban visualisation. *Proceedings of the International Conference on Computer-Aided Architectural Design Research in Asia*. https://doi.org/10.52842/conf.caadria.2025.4.449

Cipresso, P., & Riva, G. (2016). Computational psychometrics meets hollywood: The complexity in emotional storytelling. *Frontiers in Psychology*, *7*(NOV), 1–7. https://doi.org/10.3389/fpsyg.2016.01753

Civila, S., & Jaramillo-Dent, D. (2022). #Mixedcouples on TikTok: Performative hybridization and identity in the face of discrimination. *Social Media and Society*, *8*(3). https://doi.org/10.1177/20563051221122464

Clayton, R. (2019). Filmmaking theory for vertical video production. *The European Conference on Media, Communication & Film*. https://papers.iafor.org/submission52556/

Clayton, R. (2022). The context of vertical filmmaking literature. *Quarterly Review of Film and Video*, *39*(3), 644–655. https://doi.org/10.1080/10509208.2021.1874853

Fiske, J. (2001). *Television Culture* (9th ed.). Taylor & Francis.

Folkvord, F., Roes, E., & Bevelander, K. (2020). Promoting healthy foods in the new digital era on Instagram: an experimental study on the effect of a popular real versus fictitious fit influencer on brand attitude and purchase intentions. *BMC Public Health*, *20*(1), 1–8. https://doi.org/10.1186/s12889-020-09779-y

Hjort, M. (2018). Unlikely empathy: The process and effect of identification in listen. *Short Film Studies*, *8*(1), 73–78. https://doi.org/10.1386/sfs.8.1.73_1

Hutami, I., Yusanto, F., & Nugroho, C. (2018). Membedah komodifikasi isi pesan mini drama line "nic and mar" (analisis semiotika john fiske terhadap mini drama line). *Simulacra: Jurnal Sosiologi*, *1*(1), 101–116. https://doi.org/10.21107/sml.v1i1.4988

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 and Communication*, *9*(2), 229–253. https://doi.org/10.1177/2050157920952120

Kuzinas, A., Pilkauskaitė-Valickienė, R., & Poškus, M. S. (2024). Creative processes of emotional images: the effects of aspect ratio on the emotional and aesthetic properties of images. *Creativity Studies*, *17*(1), 14–26. https://doi.org/10.3846/cs.2024.16396

Lai, Z., & Yu, L. (2021). Research on digital marketing communication talent cultivation in the era of artificial intelligence. *Journal of Physics: Conference Series*, *1757*(1). https://doi.org/10.1088/1742-6596/1757/1/012040

Menotti, G. (2019). Discourses around vertical videos: an archaeology of "wrong" aspect ratios. *Ars (Sao Paulo)*, *17*(35), 147–165. https://doi.org/10.11606/issn.2178-0447.ars.2019.140526

Min, K. Z. L., & Kim, S. H. (2025). The impact of influencer credibility and short video attributes on user attitudes and adoption of instagram reel advertising. *Asia Pacific Journal of Information Systems*, *35*(1), 119–145. https://doi.org/10.14329/APJIS.2025.35.1.119

Morikawa, C., Kobayashi, M., Satoh, M., Kuroda, Y., Inomata, T., Matsuo, H., Miura, T., & Hilaga, M. (2021). Image and video processing on mobile devices: a survey. *Visual Computer*, *37*(12), 2931–2949. https://doi.org/10.1007/s00371-021-02200-8

Mulier, L., Slabbinck, H., & Vermeir, I. (2021). This Way Up: The effectiveness of mobile vertical video marketing. *Journal of Interactive Marketing*, *55*, 1–15. https://doi.org/10.1016/j.intmar.2020.12.002

Mustikawati, R., Sadewa, G. P., & Fadholi, M. A. (2023). Vertical video trends among amateur digital platform users as an alternative for film production. *Journal of Urban Society's Arts*, *9*(2), 129–136. https://doi.org/10.24821/jousa.v9i2.7949

Ndlela, M. N. (2025). *the power of platforms:*

how algorithms affect crisis coverage bt - the palgrave handbook of global digital journalism (B. Mutsvairo & K. S. Orgeret, Eds.; pp. 161–179). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-59379-6_10

Ratnasari, E., Rohnulyanti, H., Hafiar, H., & Wada, I. (2024). Exploring momfluencers ethics as social media micro-influencer in Indonesia. *Jurnal Kajian Komunikasi*, *12*(2), 234–253. https://doi.org/10.24198/jkk.v12i2.57754

Ryan, K. M. (2018). Vertical video: rupturing the aesthetic paradigm. *Visual Communication*, *17*(2), 245–261. https://doi.org/10.1177/1470357217736660

Shuter, R. (2017). New media and intercultural communication. In *The International Encyclopedia of Intercultural Communication* (pp. 1–9). wiley. https://doi.org/10.1002/9781118783665.ieicc0005

Wang, Y., Peng, A., Lin, Z., Zheng, L., & Zheng, H. (2019). Pedestrian dead reckoning-assisted visual inertial odometry integrity monitoring. In *Sensors (Switzerland)*, 19(24), 5577. MDPI AG. https://doi.org/10.3390/s19245577

Yadav, A., & Verma, A. (2025). Enhancing digital engagement: The importance of media literacy in social media. *Jurnal Kajian Komunikasi*, *13*(1), 97–116. https://doi.org/10.24198/jkk.v13i1.59884

Yusanto, F., Hidayat, D. R., Irawan, R. E., & Piliang, Y. A. (2021). Television documentary program indonesia bagus : A semiotic social analysis. *Review Of International Geographical Education*, *11*, 183–190. https://doi.org/10.48047/rigeo.11/5/18

Zhu, J., Yang, H., Wang, W., He, H., Tuo, Z., Yu, Y., Cheng, W. H., Gao, L., Song, J., Fu, J., & Luo, J. (2023). mobilevidfactory: Automatic diffusion-based social media video generation for mobile devices from text. *MM 2023 - Proceedings of the 31st ACM International Conference on Multimedia*, 9371–9373. https://doi.org/10.1145/3581783.3612667

Zhu, T., Zhang, D., Hu, Y., Wang, T., Jiang, X., Zhu, J., & Li, J. (2022). Horizontal to vertical video conversion. *IEEE Transactions on Multimedia*, *24*(61831015), 3036–3048. https://doi.org/10.1109/TMM.2021.3092202