



Visual Metafunctions of Drone Videography in the Presidential Secretariat's YouTube Public Communication: A Social Semiotic Analysis

Eko Agung Sugiyarto^{1*}, Marlinda Irwanti Poernomo²

¹Visual Communication Design, Dharma Indonesia University, Banten, 15710, Indonesia

²Institute of Communication and Business LSPR, Jakarta, 12120, Indonesia

Abstract. This study examines how drone videography constructs visual meaning in government public communication through a social semiotic perspective. While drone technology has become increasingly prevalent in governmental media production, existing studies have largely emphasized its technical and aesthetic advantages rather than its role in generating communicative meaning. Addressing this gap, the study analyzes drone videography featured in the Presidential Secretariat's official YouTube channel using the visual social semiotic framework developed by Kress and van Leeuwen. A qualitative approach was employed by examining visual data through three metafunctions namely representational meaning, interactive meaning, and compositional meaning. The analysis focuses on how aerial camera movements, framing, viewing angles, distance, salience, and information value collectively construct relationships between political actors, national space, and audiences. The findings indicate that representational meanings portray the President as an active national leader while simultaneously depicting Indonesia's territorial landscape, infrastructure development, cultural diversity, and strategic national events. Interactive meanings are established through elevated viewing angles, long shots, and dynamic aerial movements that position viewers as observers of state achievements while fostering credibility, authority, and national pride. Compositional meanings are created through the integration of visual salience, spatial organization, and cinematic sequencing, producing coherent narratives that reinforce government legitimacy and national identity.

Keywords: *Unmanned Aerial Vehicle (UAV); Drone Videography; Public Communication; Social Semiotics; Government Media Content*

1. Introduction

The increasing adoption of digital media has transformed how governments communicate with citizens. Rather than relying solely on traditional mass media, government institutions now disseminate information directly through social media platforms that combine textual, visual, and audiovisual modes of communication (Hofmann et al., 2013; Hubert et al., 2020; Sáez Martín et al., 2015; Sanders et al., 2011).

*Corresponding author's email: ekoagung@undhi.ac.id, Telp. +6282136294692



Among these platforms, YouTube has become an important channel for documenting governmental activities and shaping public understanding of state policies and political leadership (Agustina, 2023; Charolin & Wedhasmara, 2024; Chatfield & Brajawidagda, 2013). In this environment, visual communication is no longer merely a supporting element but has become a strategic resource through which governments construct institutional legitimacy, national identity, and public engagement.

Wang et al. (2019) and Laghari et al. (2023) argue that recent developments in Unmanned Aerial Vehicle (UAV) technology have significantly expanded the visual repertoire available for audiovisual communication. Similarly, Pinem & Zarkasi (2025) found that the drone videography enables dynamic aerial perspectives, cinematic camera movements, and panoramic spatial representations that cannot easily be achieved through conventional filming techniques. Previous studies have primarily examined UAVs from technological, operational, and cinematic perspectives, highlighting their efficiency, flexibility, and capacity to enhance visual aesthetics (Hermanus et al., 2023; Vujičić et al., 2022). Other research has investigated drone journalism, focusing on newsroom practices, ethical considerations, and transformations in news production (Livberber et al., 2024; Pathak, 2024; Ramirez et al., 2025). While these studies demonstrate the growing importance of drone technology in media production, they predominantly conceptualize drones as technological tools rather than semiotic resources for constructing meaning.

This limitation is particularly evident in studies of government communication. Existing scholarship has extensively discussed digital government communication, public relations, transparency, and citizen engagement through social media (Berger et al., 2016; Fajri et al., 2024; Kencana et al., 2023; Laenens et al., 2018). However, relatively little attention has been devoted to how visual meaning is constructed through specific audiovisual techniques employed in government-produced media. More specifically, previous studies rarely examine how drone-generated visual perspectives contribute to the representation of political actors, the establishment of relationships with audiences, or the organization of visual narratives within institutional communication. Consequently, the semiotic functions of drone videography remain theoretically underexplored despite their increasing prominence in official government media.

This study addresses this gap by integrating visual social semiotics with government communication studies. Drawing upon the visual grammar framework developed by Kress & van Leeuwen (2020), which extends Halliday's metafunctional theory into the visual domain, this research conceptualizes drone videography as a multimodal semiotic resource rather than simply a recording technology. Within this framework, visual meaning is constructed through three interconnected metafunctions: representational meaning, which depicts participants and social actions; interactive meaning, which establishes relationships between images and viewers through gaze, distance, and camera angle; and compositional meaning, which organizes visual elements into coherent communicative structures. This theoretical perspective enables a systematic examination of how aerial cinematography participates in the production of institutional meaning.

The Presidential Secretariat's official YouTube channel provides an appropriate empirical context because it functions as the Indonesian government's primary audiovisual platform for documenting presidential activities, national ceremonies, infrastructure development, diplomatic events, and public policy communication. Beyond serving as an archival repository, its audiovisual content is frequently referenced and redistributed by mainstream news organizations, thereby extending its influence within



Indonesia's digital public sphere. Consequently, the channel represents a significant site for examining how government institutions strategically employ visual communication to shape public narratives.

Accordingly, this study investigates how drone videography constructs meaning in the Presidential Secretariat's YouTube public communication through a visual social semiotic analysis. Specifically, the study examines how representational, interactive, and compositional meanings are realized through aerial camera perspectives and visual composition. By positioning drone videography as a semiotic resource within multimodal government communication, this research contributes to three areas of scholarship. First, it extends visual social semiotic theory into the emerging field of drone-mediated visual communication. Second, it enriches multimodal discourse studies by demonstrating how aerial cinematography functions as a meaning-making system rather than merely an aesthetic production technique. Third, it contributes to government communication research by explaining how visual strategies participate in constructing institutional legitimacy, political representation, and national identity in contemporary digital media.

2. Method

This study employed a qualitative research design using visual social semiotic analysis to examine how meaning is constructed through drone videography in government public communication (Bertotti, 2024; Fongesta & Aulia, 2023; Ghazvineh, 2024). The qualitative approach was selected because it enables an in-depth interpretation of multimodal meaning-making processes embedded in audiovisual texts rather than measuring frequencies or audience responses (Goundar, 2025; Johnson et al., 2020; Stratton, 2024). The analytical framework was adapted from Kress & van Leeuwen (2020) visual grammar, which conceptualizes visual communication through three metafunctions namely representational, interactive, and compositional meanings.

The primary data consisted of the official YouTube video entitled "*President Jokowi Invites ASEAN Leaders Aboard a Phinisi Ship, Labuan Bajo, May 10, 2023*" published by the Presidential Secretariat (<https://www.youtube.com/watch?v=mxt-K4N2hfo>). The video was selected through purposive sampling (Etikan, 2016) based on four criteria: (1) extensive use of drone videography throughout the narrative; (2) representation of official government public communication; (3) high public engagement indicated by views, comments, and dissemination across mainstream media; and (4) its strategic role in documenting an international diplomatic event. Rather than aiming for statistical generalization, this study adopts an information-rich case approach to explore how visual meaning is constructed within a significant example of government audiovisual communication.

Secondary data were obtained through semi-structured interviews with [number] informants consisting of communication scholars, visual journalism practitioners, and content production personnel from the Press, Media, and Information Bureau of the Presidential Secretariat. Participants were selected purposively according to their expertise and direct involvement in government audiovisual communication. Each interview lasted approximately [duration] minutes and explored three themes: (1) the rationale for employing drone videography, (2) visual production strategies, and (3) intended communicative messages. The interview findings were not analyzed as independent data but were used to support interpretation and provide source triangulation.



The analysis followed four sequential stages. First, the selected video was repeatedly viewed to obtain a comprehensive understanding of its narrative structure. The video was then segmented into meaningful analytical units based on changes in scenes, camera movements, and visual sequences. Each segment became the basic unit of analysis. Second, every segment was coded according to the three visual metafunctions proposed by Kress & van Leeuwen (2020). Representational meaning focused on participants, actions, settings, and symbolic representations. Interactive meaning examined camera angle, social distance, gaze, perspective, and viewer positioning. Compositional meaning analyzed information value, salience, framing, spatial organization, and the integration of visual elements across sequences.

Third, coded segments were grouped into broader interpretive categories to identify recurring visual patterns throughout the video. Particular attention was given to the relationship between aerial camera movements and the construction of institutional representation, political authority, national identity, and public engagement. Fourth, interpretations derived from the visual analysis were compared with interview data and relevant documentary sources to enhance analytical credibility through triangulation. Rather than treating interviews as confirmatory evidence, the researchers used them to contextualize production decisions and strengthen the interpretation of visual meanings.

To enhance the trustworthiness of the findings, this study adopted several qualitative validation strategies (Soysal & Türkmen, 2024). Credibility was strengthened through data source triangulation involving visual data, interviews, and official government documents. Dependability was supported by maintaining a transparent coding procedure and documenting each stage of analysis. Confirmability was enhanced through continuous comparison between empirical observations and theoretical concepts within the visual social semiotic framework. Throughout the analytical process, the researchers practiced reflexivity by acknowledging that the interpretation of visual meaning is inevitably influenced by theoretical perspectives and therefore consistently grounded interpretations in observable visual evidence. The purpose of this study is analytical rather than statistical generalization. Accordingly, the selected case is intended to provide theoretical insights into how drone videography functions as a semiotic resource in government public communication rather than to represent all government audiovisual productions.

3. Results and Discussion

The opening sequence adopts a flashback narrative by presenting interview excerpts before introducing the visual journey of the ASEAN leaders. Rather than functioning merely as an introduction, this narrative structure positions viewers to interpret the subsequent aerial footage through the participants' personal experiences. Within Kress & van Leeuwen (2020) representational metafunction, the interview statements operate as verbal resources that frame the visual narrative, guiding audiences to understand the voyage not simply as a diplomatic event but as a shared experiential activity. The comments by Lee Hsien Loong, Ferdinand Marcos Jr., and Abdul Mateen establish experiential meanings that are later reinforced through the visual representation of the journey.

The subsequent drone shot presents a panoramic view of the Phinisi vessel sailing toward Bidadari Island during sunset. From a representational perspective, the image depicts participants, setting, and action through a narrative process in which the vessel functions as the primary vector connecting political actors with the surrounding maritime



landscape. The aerial perspective extends beyond documenting the voyage by integrating Indonesia's natural environment into the diplomatic narrative. In visual social semiotics, landscape is not interpreted merely as scenery but as a semiotic resource that contributes to the construction of institutional and national meanings (Alkateeb, 2023; O'Hagan & Serafinelli, 2022).

The interactive metafunction is realized through the use of a high-angle aerial shot and an extreme long shot. These camera choices create a relatively distant social relationship between viewers and participants while simultaneously emphasizing the scale of the surrounding environment. Rather than inviting interpersonal intimacy, the drone perspective positions viewers as observers of an official state event. This finding supports previous studies arguing that aerial cinematography frequently produces institutional authority and spatial legitimacy by presenting events from perspectives unavailable to ordinary human vision (Bonatti et al., 2020; Passalis & Tefas, 2020). Consequently, the drone functions not only as a recording device but also as a semiotic resource that shapes audience positioning.



Figure 1 Analysis of Language Metafunction in Drone Aerial Footage 1

Source: YouTube Account of the Presidential Secretariat, 2025 (<http://www.youtube.com/@SekretariatPresiden>)

Figure 1 illustrates how the three dimensions of visual metafunction are simultaneously realized through drone aerial footage in the Presidential Secretariat's YouTube communication. Representationally, the image depicts the Phinisi vessel navigating Indonesia's maritime territory, thereby constructing an experiential meaning associated with mobility, connectivity, and diplomatic engagement within an archipelagic context. The interactive metafunction is established through the high-angle aerial perspective and long shot composition, which position viewers as observers of a broader geopolitical landscape rather than as participants in direct interpersonal interaction. Meanwhile, the compositional metafunction emerges from the balanced arrangement of



the vessel, surrounding islands, and horizon, with the Phinisi functioning as the primary visual salience against the expansive seascape that contextualizes Indonesia's strategic maritime identity. The warm color modality produced by the sunset enhances visual coherence and cinematic continuity across the sequence, although its meaning should not be reduced to a symbolic representation of peace or regional solidarity. Instead, the semiotic significance of the scene is generated through the interaction between color modality, spatial organization, and the broader narrative context established throughout the video. Consequently, the drone footage operates not merely as an aesthetic device but as a multimodal semiotic resource that contributes to the construction of institutional narratives linking political leadership, national territory, and diplomatic activity.

The compositional metafunction is constructed through the balanced arrangement of the Phinisi vessel, the surrounding islands, and the horizon. The vessel occupies the visual salience of the frame, while the expansive seascape provides contextual information that reinforces the strategic importance of Indonesia's maritime geography. The warm color modality generated by the sunset contributes to visual coherence and cinematic continuity across the sequence. However, the sunset itself should not be interpreted as directly symbolizing peace or ASEAN solidarity. Instead, the semiotic significance emerges from the interaction between color modality, spatial composition, and the narrative context established through participants' interviews. Within this multimodal configuration, the visual sequence constructs an experiential representation of diplomacy embedded within Indonesia's maritime landscape rather than relying solely on symbolic interpretations of natural scenery.

These findings indicate that drone videography performs more than an aesthetic function in government public communication. By organizing representational, interactive, and compositional meanings simultaneously, aerial cinematography contributes to the construction of institutional narratives that associate political leadership with national territory and diplomatic engagement. This extends previous scholarship on drone media, which has primarily emphasized technological innovation and visual aesthetics, by demonstrating its role as a multimodal semiotic resource in government communication.

3.1. Narrative Representation: Constructing Action Through Drone-Based Visual Grammar

Figure 1 depicts a drone-generated aerial view of a Phinisi vessel navigating the waters of Labuan Bajo during the ASEAN Summit. Within Kress & van Leeuwen (2020) framework, the scene represents a transactional narrative process because it contains a clearly identifiable actor, vector, and goal. The Phinisi functions as the principal Actor, while its forward movement generates an implicit vector directing viewers' attention toward the destination. Although the destination is not fully visible within the frame, the orientation of the vessel establishes a directional process that visually connects the participants with a future point of action. Rather than simply depicting movement, the transactional process organizes the event as purposeful political mobility, in which the voyage itself becomes the central narrative activity.

Unlike conventional ground-level cinematography, the drone perspective enables the vessel to be observed within a broader spatial context. The aerial framing integrates the moving ship, surrounding islands, and maritime landscape into a single visual composition. Consequently, the representational meaning is not limited to the actors alone but extends to the relationship between political participants and national territory. This finding illustrates how aerial cinematography expands the narrative potential of



visual grammar by embedding political action within geographical space, thereby reinforcing previous arguments that drone imagery reconstructs spatial experience rather than merely documenting events (Ejaz & Choudhury, 2024; Li et al., 2025).

The situational context further clarifies how representational meaning is constructed. Following Halliday's contextual dimensions, the field concerns the diplomatic voyage undertaken during the ASEAN Summit, the tenor involves ASEAN heads of government participating in the event, and the mode is realized through professionally produced drone videography disseminated via the Presidential Secretariat's official YouTube channel. These contextual dimensions do not themselves generate symbolic meaning; rather, they shape how viewers interpret the visual representation within the institutional context of government communication.

The cultural significance of the Phinisi should therefore be understood through the interaction between visual representation and broader cultural discourse rather than through symbolism alone. As Indonesia's recognized maritime heritage, the Phinisi carries established cultural associations rooted in Bugis-Makassar shipbuilding traditions (Suardi et al., 2023). Within the video, however, its semiotic function arises because this culturally recognizable object is positioned as the central participant in an official diplomatic event. The representational meaning is therefore produced through the selection of a culturally significant artifact as the primary visual actor, rather than by assuming that the vessel automatically symbolizes ASEAN solidarity.

Similarly, the surrounding seascape and sunset should not be interpreted independently as signs of peace, wisdom, or regional harmony. Within visual social semiotics, these environmental elements function primarily as circumstances that provide spatial and temporal context for the narrative process. Their contribution to meaning emerges from their relationship with the moving vessel, the institutional setting of the ASEAN Summit, and the accompanying verbal discourse presented through interviews and official narration. The impression of diplomatic cordiality is therefore constructed multimodally through the interaction of visual composition, participant actions, and contextual framing rather than through the landscape itself.

This interpretation also refines the concept of transactional narrative processes proposed by Kress and van Leeuwen. In this case, the transactional process extends beyond depicting physical movement by organizing political actors, cultural artifacts, and national space into a coherent institutional narrative. Drone videography thus functions not merely as a recording technology but as a semiotic resource that structures political action within a geographically expansive visual field. This finding extends previous applications of visual grammar by demonstrating how aerial cinematography contributes to the construction of government communication through the integration of movement, spatial representation, and cultural context.

3.2. Interactive Meaning: Audience Positioning and Visual Power Relations

The interactive meaning of Figure 1 is constructed through the combined use of gaze, camera angle, and aerial perspective rather than through direct interpersonal interaction. None of the participants establish eye contact with the viewer, indicating what Kress & van Leeuwen (2020) describe as an offer image. Instead of demanding engagement, the image presents the diplomatic event as an object of observation. However, this absence of gaze should not be understood merely as a formal visual characteristic. Within the context of government communication, the "offer" configuration contributes to an institutional



mode of address in which audiences are positioned primarily as witnesses to state activities rather than active participants in political interaction.

The use of an oblique camera angle further reinforces this relationship. Rather than inviting viewers into the event through frontal alignment, the oblique perspective maintains symbolic distance between the audience and political actors. This distance does not necessarily reduce audience engagement; instead, it establishes the event as an institutional performance unfolding within a governmental sphere. Consequently, viewers are encouraged to observe the official narrative from an external position, lending the representation an appearance of objectivity while preserving the ceremonial authority of the event. Similar observations have been reported in studies of political visual communication, where compositional distance contributes to perceptions of institutional credibility and legitimacy rather than interpersonal intimacy.

The bird's-eye perspective generated through drone videography introduces a further ideological dimension that extends beyond conventional visual grammar. While aerial viewpoints appear to position viewers physically above the depicted participants, the visual authority produced by the image does not translate into audience dominance over political actors. Instead, the elevated perspective functions as a state-controlled mode of vision through which governmental institutions determine what becomes visible and how public space is represented. The drone therefore operates not only as a recording technology but also as a mechanism of visual governance, organizing territory, political actors, and national space into a coherent institutional narrative.

This finding suggests that power within aerial government communication is relational rather than hierarchical. The audience receives access to perspectives that would otherwise remain unavailable in everyday experience, creating an impression of transparency and privileged observation. Simultaneously, however, this visual access remains carefully curated through institutional decisions regarding camera placement, movement, framing, and editing. As a result, viewers experience a controlled form of visual participation in which they are invited to observe state activities while the government retains authority over the production and circulation of visual knowledge.

These findings extend Kress and van Leeuwen's interactive metafunction by demonstrating that aerial perspectives in government communication perform ideological work beyond interpersonal positioning. Rather than simply establishing social distance or viewer perspective, drone imagery contributes to the construction of institutional authority through what may be understood as strategic visual surveillance: a communicative practice in which the state presents itself as transparent, technologically capable, and territorially comprehensive while simultaneously directing how audiences perceive political reality. This interpretation also complements previous studies of drone media (Montes-Romero et al., 2020; Tilak, 2020), which emphasize the capacity of aerial imagery to reorganize spatial perception, by showing that such visual reorganization simultaneously reshapes political relationships between the state, territory, and citizens.

3.3. Compositional Meaning: Organizing Visual Information in Government Communication

The compositional meaning of Figure 1 is produced through the systematic organization of visual elements rather than through aesthetic qualities alone. Following Kress & van Leeuwen (2020), compositional analysis examines how information value, salience, and framing integrate individual visual components into a coherent communicative structure. In this sequence, the drone shot does not merely function as a transitional image between interview excerpts and subsequent scenes. Instead, it



reorganizes verbal and visual information into a unified multimodal narrative that guides viewers' interpretation of the diplomatic event.

Visual salience is primarily established through the placement of the Phinisi vessel at the center of the frame. Although the surrounding sea occupies a larger proportion of the image, the vessel attracts immediate attention because of its contrast in scale, movement, and spatial isolation. The expansive maritime background therefore functions less as decorative scenery than as contextual information that situates political actors within Indonesia's maritime territory. This compositional arrangement directs viewers to interpret the diplomatic activity as inseparable from the nation's geographical identity.

Information value is further constructed through the spatial relationship between the foreground, middle ground, and background. The open sea in the foreground creates visual depth, while the moving vessel occupies the central informational position. The islands positioned in the background complete the geographical context without competing for visual prominence. Rather than presenting these elements independently, the composition integrates them into a coherent representation of political space. The visual organization therefore communicates not only where the event takes place but also how national territory becomes part of the institutional narrative.

Framing also contributes to ideological meaning. The absence of strong visual boundaries between the vessel, surrounding waters, and islands produces an integrated composition in which political actors appear embedded within the national landscape. This integration reduces the separation between governmental activity and territorial representation, encouraging viewers to perceive state leadership as naturally situated within Indonesia's maritime environment. The compositional structure thus performs ideological work by visually connecting political authority with national space instead of presenting government actors as isolated individuals.

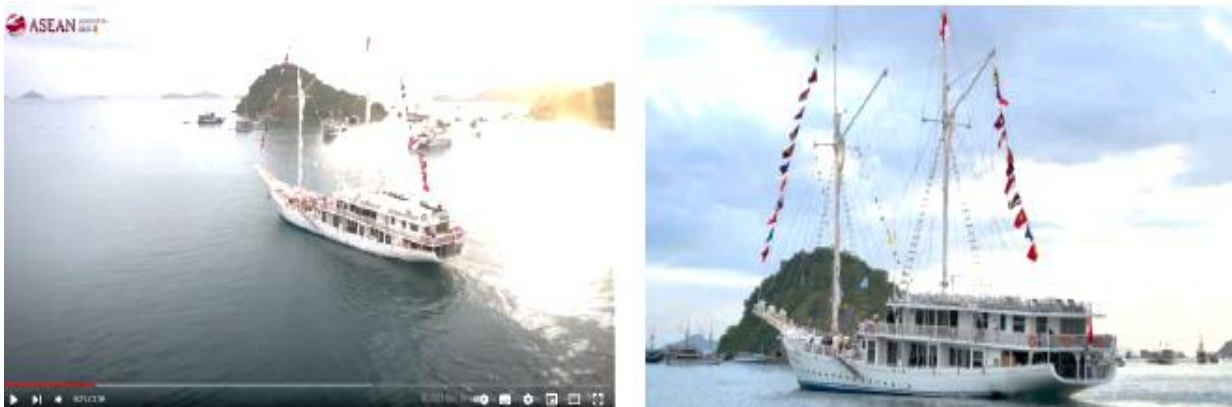


Figure 2 Comparison of Aerial Images and Images Taken by a Ground Based Camera

Source: YouTube Channel of the Presidential Secretariat, 2025 (<http://www.youtube.com/@SekretariatPresiden>)

Figure 2 demonstrates the distinctive semiotic functions of aerial imagery in comparison with images captured by a ground-based camera within the Presidential Secretariat's public communication strategy. While ground-based footage primarily provides proximity, facial expressions, and interpersonal engagement that support the interactive metafunction, the drone imagery expands the representational scope by situating political actors and diplomatic activities within a broader spatial and national



context. From a multimodal perspective, the aerial sequence should not be interpreted merely as a visually impressive shot or as a technique of continuity editing; rather, its communicative significance emerges through its relationship with the surrounding verbal and visual modes. The interview excerpts preceding the drone footage establish an experiential framework for the diplomatic journey, whereas the panoramic aerial composition spatially materializes this experience by visualizing movement across Indonesia's maritime landscape. Consequently, meaning is produced through the interaction between verbal discourse and visual organization rather than through either mode independently. The findings further indicate that compositional metafunction extends beyond the arrangement of visual elements, functioning instead as a strategic semiotic resource through which institutional narratives are spatially organized and politically legitimized. By integrating political actors, cultural symbols, and national territory within a coherent visual hierarchy, drone videography contributes to the construction of state legitimacy and national identity, demonstrating that aerial composition operates not merely as a cinematic technique but also as an ideological mechanism through which governments shape public perception in digital media.

The drone sequence should not be interpreted simply as a visually impressive shot or as an example of continuity editing. Its communicative function emerges through its relationship with the preceding interview excerpts and the subsequent narrative sequence. The verbal testimonies establish an experiential frame for the diplomatic journey, while the aerial composition spatially materializes that experience through a panoramic visualization of the voyage. Meaning is therefore generated through the interaction between verbal discourse and visual organization rather than by either mode independently.



Figure 3 Analysis of Language Metafunction in Drone Aerial Footage 2

Source: YouTube Account of the Presidential Secretariat, 2025
(<http://www.youtube.com/@SekretariatPresiden>)

Figure 3 further demonstrates how the three dimensions of visual metafunction are realized through drone aerial footage in the public communication content produced by



the Presidential Secretariat. At the representational level, the aerial perspective constructs an experiential meaning by positioning political actors, cultural symbols, and geographical landscapes within a unified national space, thereby transforming physical territory into a communicative resource for political storytelling. The interactive metafunction is established through the distant aerial gaze and elevated camera angle, which encourage viewers to adopt an observational position associated with authority, overview, and strategic vision rather than interpersonal intimacy. Meanwhile, the compositional metafunction emerges through the careful organization of visual salience, framing, and information value, where institutional actors are embedded within broader spatial configurations that emphasize national scale and territorial continuity. These findings indicate that composition extends beyond the mere arrangement of visual elements and functions instead as a strategic semiotic resource through which institutional narratives are spatially organized and politically legitimized. By integrating political leadership, cultural identity, and national territory within a coherent visual hierarchy, drone videography contributes to the production of state legitimacy and national identity. Consequently, aerial composition should be understood not only as a cinematic technique for enhancing visual aesthetics but also as an ideological mechanism through which governments organize public perception and communicate institutional authority within digital media environments.

These findings suggest that compositional metafunction extends beyond the arrangement of visual elements. In government public communication, composition functions as a strategic semiotic resource through which institutional narratives are spatially organized and politically legitimized. By integrating political actors, cultural symbols, and national territory within a coherent visual hierarchy, drone videography contributes to the production of state legitimacy and national identity. This finding extends previous applications of visual grammar by demonstrating that aerial composition is not merely a cinematic technique but an ideological mechanism through which governments organize public perception in digital media.

3.4. Conceptual Representation: Landscape as an Institutional Representation of National Space

Unlike the previous sequence, Figure 3 does not depict a clearly identifiable vector or directional movement. Consequently, the image is more appropriately interpreted as a conceptual representation rather than a narrative representation within Kress & van Leeuwen (2020) visual grammar. The scene organizes visual elements into a stable classificatory structure in which the maritime landscape, anchored Phinisi vessels, surrounding islands, and sunset collectively constitute an integrated representation of place rather than an unfolding action. Meaning therefore emerges from the spatial relationship among these visual components instead of from participant movement.

Within the compositional structure, the expansive seascape occupies the largest visual area, while the vessels appear relatively small within the broader landscape. This distribution of visual salience shifts attention away from individual political actors toward the territorial environment itself. Rather than functioning merely as an attractive natural backdrop, the landscape becomes the principal representational resource through which national space is constructed. The image therefore encourages viewers to perceive Indonesia's maritime territory as an institutional setting for regional diplomacy rather than as a neutral geographical location.



The frequent use of panoramic aerial imagery throughout the video should not be interpreted simply as an effort to display natural beauty or promote tourism. Such an interpretation overlooks the ideological function of visual composition within government communication. By repeatedly presenting wide aerial views of Labuan Bajo before and during diplomatic interactions, the video embeds political events within carefully curated representations of national territory (Kodir et al., 2020). Drone aesthetics consequently perform a dual communicative function: they produce visual attractiveness while simultaneously associating governmental leadership with Indonesia's maritime identity and territorial legitimacy.

This finding also illustrates how aerial imagery contributes to political image-building. The absence of crowds, urban complexity, or everyday social activities produces an idealized representation of space in which the landscape appears orderly, peaceful, and visually coherent. Such compositional choices selectively privilege particular aspects of reality while excluding others, thereby constructing an institutional narrative consistent with official government communication. The resulting image is therefore not a neutral documentation of place but a strategically composed visual representation that supports the government's preferred portrayal of Indonesia within the context of international diplomacy.

The use of warm sunset lighting should likewise be interpreted cautiously. Rather than assuming that the golden-hour atmosphere inherently symbolizes harmony, hospitality, or cultural openness, its semiotic significance arises from its interaction with framing, salience, and narrative context. The soft color modality contributes to visual coherence and emotional tone, but the ideological meaning emerges because these aesthetic resources are integrated into an official state-produced audiovisual narrative. In this sense, aesthetic choices become communicative resources that reinforce institutional legitimacy without relying on explicit political messaging.

These findings extend previous studies of drone media by demonstrating that aerial landscape imagery in government communication performs representational work beyond tourism promotion or cinematic spectacle. Through conceptual representation, the landscape is transformed into a semiotic resource that naturalizes the relationship between political authority, national territory, and regional diplomacy. Drone videography therefore functions not merely as an instrument for producing visually attractive images but also as a strategic mechanism for organizing political meaning within contemporary government public communication.

3.5. Compositional Meaning: Organizing Visual Information Through Spatial Integration

Within the compositional metafunction, Figure 3 organizes visual information through the spatial integration of maritime elements rather than through the prominence of a single focal object. The sea, islands, anchored Phinisi vessels, and horizon are distributed across the frame in a manner that produces a relatively balanced composition. Instead of establishing a dominant visual hierarchy, these elements function collectively to construct a coherent representation of place. Following Kress & van Leeuwen (2020), meaning is therefore generated through the relationships among visual components rather than through any isolated element.

Color modality also contributes to this compositional structure. The warm tonal quality created by the late-afternoon light provides visual coherence across the image, linking the sea, sky, and surrounding landscape into a unified scene. Within a social semiotic perspective, the lighting should not be interpreted as inherently symbolizing



warmth, hospitality, or particular emotional states. Rather, it functions as a semiotic resource that shapes the visual modality of the image and contributes to its overall representational coherence. Any emotional interpretation remains a potential reading that emerges from the interaction between color, composition, and the broader narrative context, rather than a demonstrable audience response.

Similarly, the inclusion of multiple small islands should not automatically be understood as representing Indonesia's archipelagic identity. Instead, their semiotic significance derives from their compositional relationship with the maritime setting and the diplomatic event portrayed in the video. By integrating territorial features into the visual frame, the composition situates governmental activities within a geographically recognizable environment. The image therefore constructs national space as the context in which diplomacy takes place, rather than merely providing an attractive natural backdrop.



Figure 4 Analysis of Language Metafunction in Drone Aerial Footage 3

Source: *YouTube Account of the Presidential Secretariat, 2025*
 (<http://www.youtube.com/@SekretariatPresiden>)

Figure 4 illustrates how compositional metafunction in drone aerial footage organizes spatial relationships between political actors, cultural artifacts, and territorial environments to produce a coherent institutional narrative. Through the strategic arrangement of visual salience, framing, and spatial positioning, the aerial composition constructs the meaning potential through which government communication represents national territory and diplomatic activity. The analysis does not attempt to determine how audiences interpret these visual resources, as audience reception lies beyond the scope of this study. Instead, it focuses on how drone videography functions as a multimodal semiotic resource that structures and communicates institutional meanings within digital public communication.



These findings indicate that compositional metafunction in drone videography operates by organizing spatial relationships that connect political actors, cultural artifacts, and territorial environments into a coherent institutional narrative. The analysis does not claim how audiences actually interpret these visual resources, as audience reception falls outside the scope of this study. Rather, it demonstrates how the arrangement of visual elements creates the meaning potential through which government communication represents national territory and diplomatic activity.

3.6. Narrative Representation: Concluding the Diplomatic Narrative Through Visual Closure

Figure 4 functions as the concluding visual sequence of the video. Unlike the opening scenes, which emphasize movement and diplomatic activities, this image presents a relatively stable composition in which the setting itself becomes the principal representational resource. The low position of the sun near the horizon, the anchored Phinisi vessel, and the surrounding maritime landscape together establish a visually coherent conclusion to the audiovisual narrative. Rather than depicting a new action, the scene provides narrative closure by linking the preceding interviews with a final panoramic representation of the diplomatic experience.

From the perspective of visual grammar, the scene is more appropriately interpreted as a conceptual representation than as a narrative process. The absence of clear vectors connecting participants to specific actions shifts attention from movement to symbolic organization. The maritime landscape, the Phinisi, and the sunset are arranged as complementary visual elements that collectively frame the diplomatic event within Indonesia's coastal environment. Meaning therefore emerges through the spatial integration of these elements rather than through participant interaction.

The concluding aerial perspective also reinforces the institutional framing of the event. By presenting the diplomatic gathering within a broad maritime landscape, the video situates ASEAN cooperation within a recognizable representation of Indonesian territory. This compositional strategy aligns governmental activities with national space and cultural identity without requiring explicit verbal explanation. The visual sequence therefore contributes to the representation of Indonesia as the host of a regional diplomatic event through the integration of landscape, cultural heritage, and official state activities.

Importantly, this analysis does not interpret the sunset as an inherent symbol of peace, political wisdom, or regional solidarity. Such meanings cannot be established solely from the visual evidence. Instead, the sunset functions as a modality resource that contributes to the temporal closure and visual coherence of the narrative. Any broader political interpretation must be supported by additional verbal, textual, or contextual evidence beyond the image itself.

The relationship between visual representation and diplomacy is therefore constructed through the multimodal interaction of interview excerpts, official narration, and aerial imagery rather than through individual visual symbols. Within this multimodal configuration, the drone footage supports the institutional narrative by providing a spatially expansive representation of the diplomatic setting while maintaining consistency with the broader communicative objectives of the Presidential Secretariat.

These findings suggest that the primary contribution of drone videography in this sequence lies not in communicating abstract geopolitical messages but in organizing visual resources that reinforce institutional representation. Through compositional coherence, aerial perspective, and multimodal integration, the concluding sequence



strengthens the government's visual narrative of diplomatic engagement while remaining grounded in observable semiotic evidence.

3.7. Interactive Meaning: Audience Positioning and Institutional Authority

The interactive configuration of Figure 4 shares several visual characteristics with the opening aerial sequence, including the absence of direct gaze, the use of an oblique perspective, and a distant aerial viewpoint. However, rather than simply repeating the earlier interactional pattern, the concluding sequence performs a different communicative function. At this stage of the narrative, the aerial perspective no longer introduces the setting but provides visual closure by positioning viewers as observers of a completed diplomatic event.

The absence of eye contact constructs what Kress & van Leeuwen (2020) describe as an offer image, in which viewers are invited to observe rather than directly participate in the represented event. Nevertheless, within the institutional context of government communication, this visual configuration also establishes a distinctive relationship between the state and the audience. Political actors are presented as objects of public observation, yet the terms of that observation remain institutionally controlled through camera placement, framing, editing, and aerial movement. Consequently, visual access is provided without diminishing governmental authority over the representation itself.

The elevated drone perspective further contributes to this institutional relationship. Unlike conventional ground-level imagery that may encourage interpersonal proximity, the bird's-eye view produces emotional distance by emphasizing spatial overview rather than individual interaction. The audience is positioned to observe diplomatic activities from a panoramic perspective that is unavailable in ordinary everyday experience. This visual strategy encourages viewers to engage with the event primarily as an institutional spectacle rather than as participants in interpersonal communication.

The aerial perspective also illustrates how drone videography functions as a technology of visual governance. By presenting political actors, cultural symbols, and national territory within a single panoramic frame, the state determines what becomes visible and how spatial relationships are organized. The resulting representation creates an appearance of transparency by offering viewers privileged visual access to official activities while simultaneously maintaining institutional control over the production of visual knowledge. In this sense, drone imagery performs a dual communicative function: it expands public visibility while reinforcing governmental authority over the framing of political events.

Rather than claiming that audiences necessarily experience particular emotions or interpret the image in a specific way, this analysis demonstrates how interactive visual resources construct the potential for particular readings. The combination of social distance, elevated perspective, and institutional framing makes available a representation in which diplomacy is understood through spatial overview, technological sophistication, and state-managed visibility. These findings extend previous applications of interactive metafunction by showing that drone aesthetics in government communication organize not only interpersonal relationships but also broader power relations between state institutions, public audiences, and the visual representation of national territory.

3.8. Compositional Meaning: Spatial Organization and Institutional Representation

The compositional metafunction of Figure 4 is realized through the hierarchical organization of visual information rather than through the technical characteristics of



drone cinematography. The sunset occupies the most visually salient position within the frame, while the surrounding sea, Phinisi vessel, and islands function as complementary elements that stabilize the composition. This visual hierarchy directs attention toward the overall spatial organization of the diplomatic setting instead of emphasizing individual participants.

Within Kress & van Leeuwen (2020) framework, information value is constructed through the spatial relationship among these elements. The central placement of the horizon and the balanced distribution of maritime features create a coherent visual structure in which landscape functions as an integral component of institutional representation. Rather than serving merely as an aesthetically pleasing background, the surrounding environment contextualizes the diplomatic event within Indonesia's maritime territory.

The compositional arrangement also contributes to the visual coherence of the government's communication strategy. The aerial perspective enables political activities, cultural symbols, and geographical space to be integrated within a single visual frame. Importantly, this integration should not be interpreted as providing a complete or objective representation of reality. From a social semiotic perspective, every visual representation involves processes of selection, framing, and emphasis. Drone imagery therefore constructs a particular institutional perspective by determining which aspects of the event become visually prominent and which remain outside the frame.

This finding highlights an important communicative function of drone videography. Rather than simply improving visibility or producing more attractive images, aerial composition organizes spatial information in ways that support the government's institutional narrative. The panoramic perspective presents diplomatic activities as embedded within an expansive maritime landscape, thereby reinforcing the association between state authority, national territory, and official public communication.

The analysis therefore moves beyond technical explanations of aerial filming. Drone videography is understood here as a multimodal semiotic resource through which governmental institutions strategically organize visual information, construct institutional visibility, and shape the public representation of political events. In this sense, the communicative significance of aerial composition lies not in its technological novelty but in its capacity to structure meaning within contemporary government visual communication.

4. Conclusion

This study examined how drone videography functions as a multimodal semiotic resource in the Presidential Secretariat's YouTube video documenting the ASEAN Summit in Labuan Bajo. Using a social semiotic approach, the findings show that representational, interactive, and compositional metafunctions work together to construct an institutional visual narrative rather than merely recording diplomatic activities. Drone imagery organizes political actors, cultural symbols, and maritime landscapes into a coherent representation that situates government communication within Indonesia's territorial and cultural context.

The study contributes to the development of social semiotic and visual communication research by extending the application of visual grammar to drone-based government communication. The findings demonstrate that aerial imagery is not only a technological or cinematic feature but also a strategic semiotic resource through which governments organize institutional visibility, represent national territory, and frame



diplomatic events within official communication. From a practical perspective, the study suggests that government institutions can utilize drone videography not merely to enhance visual attractiveness but also to strengthen institutional narratives and public representation in digital media.

This study has several limitations, including its focus on a single case from the Presidential Secretariat's YouTube channel and its emphasis on textual meaning rather than audience interpretation or communication effects. Future studies could compare drone usage across different governmental institutions and political events while incorporating audience reception, production studies, or computational visual analysis. Such approaches would provide a broader understanding of how drone aesthetics operate in the production, representation, and interpretation of contemporary government digital communication.

Acknowledgments

The author would like to express appreciation and gratitude to all parties who provided support for the conduct of this research. Special thanks are extended to the Press, Media, and Information Bureau of the Secretariat of the President of the Republic of Indonesia for providing access to the data and content that formed the subject of this study.

Conflict of Interest

The authors declare no conflict of interests.

References

- Agustina, I. R. (2023). Implementasi Open Government Indonesia melalui Saluran Youtube Resmi Sekretariat Presiden. *Jurnal Komunikasi Profesional*, 7(1). <https://doi.org/10.25139/jkp.v7i1.5674>
- Alkateeb, H. A. (2023). The British Council's role in nourishing the English language teaching industry in the Gulf Cooperation Council region: a visual social semiotic perspective. *Social Semiotics*, 33(2). <https://doi.org/10.1080/10350330.2020.1833686>
- Berger, J. B., Hertzum, M., & Schreiber, T. (2016). Does local government staff perceive digital communication with citizens as improved service? *Government Information Quarterly*, 33(2). <https://doi.org/10.1016/j.giq.2016.03.003>
- Bertotti, V. (2024). Choice, marketing and subjectivities: A discursive-semiotic analysis of six Montessori websites. *Discourse Studies*, 26(6). <https://doi.org/10.1177/14614456241242924>
- Bonatti, R., Wang, W., Ho, C., Ahuja, A., Gschwindt, M., Camci, E., Kayacan, E., Choudhury, S., & Scherer, S. (2020). Autonomous aerial cinematography in unstructured environments with learned artistic decision-making. *Journal of Field Robotics*, 37(4). <https://doi.org/10.1002/rob.21931>
- Charolin, N., & Wedhasmara, A. (2024). Evaluation of community engagement on government youtube: Duration, title, dialogue, content, emotion. *Journal of Business and Information Systems* (e-ISSN: 2685-2543), 6(1). <https://doi.org/10.36067/jbis.v6i1.234>
- Chatfield, A. T., & Brajawidagda, U. (2013). Political will and strategic use of youtube to advancing government transparency: An analysis of jakarta government-generated youtube videos. *Lecture Notes in Computer Science (Including Subseries Lecture Notes*



- in Artificial Intelligence and Lecture Notes in Bioinformatics*), 8074 LNCS. https://doi.org/10.1007/978-3-642-40358-3_3
- Ejaz, N., & Choudhury, S. (2024). Computer vision in drone imagery for infrastructure management. In *Automation in Construction* (Vol. 163). <https://doi.org/10.1016/j.autcon.2024.105418>
- Etikan, I. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, 5(1). <https://doi.org/10.11648/j.ajtas.20160501.11>
- Fajri, C., Najih, M., Longani, K. D., Novianto, F., & Kusuma, A. (2024). Government Digital Communication Model in Corporate Social Responsibility Programs (Study in Gunung Kidul, Yogyakarta-Indonesia). *Komunikator*, 16(2). <https://doi.org/10.18196/jkm.23765>
- Fongesta, B. P., & Aulia, S. (2023). Analisis Semiotika Iklan Pariwisata Labuan Bajo. *Kiwari*, 2(1). <https://doi.org/10.24912/ki.v2i1.23067>
- Ghazvineh, A. (2024). An inter-semiotic analysis of ideational meaning in text-prompted AI-generated images. *Language and Semiotic Studies*, 10(1). <https://doi.org/10.1515/lass-2023-0030>
- Goundar, P. R. (2025). Researcher Positionality: Ways to Include it in a Qualitative Research Design. *International Journal of Qualitative Methods*, 24. <https://doi.org/10.1177/16094069251321251>
- Hermanus, D. R., Suhono Harso Supangkat, & Fadhil Hidayat. (2023). Web-Based Anomaly Detection for Smart Urban Living: Drone Photography and Videography. *Jurnal Sistem Cerdas*, 6(2). <https://doi.org/10.37396/jsc.v6i2.330>
- Hofmann, S., Räckers, M., Beverungen, D., & Becker, J. (2013). Old blunders in new media? How local governments communicate with citizens in online social networks. *Proceedings of the Annual Hawaii International Conference on System Sciences*. <https://doi.org/10.1109/HICSS.2013.421>
- Hubert, R. B., Estevez, E., Maguitman, A., & Janowski, T. (2020). Analyzing and Visualizing Government-Citizen Interactions on Twitter to Support Public Policy-making. *Digital Government: Research and Practice*, 1(2). <https://doi.org/10.1145/3360001>
- Johnson, J. L., Adkins, D., & Chauvin, S. (2020). A review of the quality indicators of rigor in qualitative research. In *American Journal of Pharmaceutical Education* (Vol. 84, Number 1). <https://doi.org/10.5688/ajpe7120>
- Kencana, W. H., Situmeang, I. V. O., Meisyanti, M., Rahmawati, K. J., Nugroho, H. S., & Lubis, A. Y. (2023). Digital Communication Transformation in Sukabumi City Government. *Komunikator*, 15(1). <https://doi.org/10.18196/jkm.17206>
- Kodir, A., Tanjung, A., Astina, I. K., Nurwan, M. A., Nusantara, A. G., & Ahmad, R. (2020). The dynamics of access on tourism development in Labuan Bajo, Indonesia. *Geojournal of Tourism and Geosites*, 29(2). <https://doi.org/10.30892/gtg.29222-497>
- Kress, G., & van Leeuwen, T. (2020). *Reading Images*. Routledge. <https://doi.org/10.4324/9781003099857>
- Laenens, W., Van den Broeck, W., & Mariën, I. (2018). Channel choice determinants of (Digital) government communication: A case study of spatial planning in flanders. *Media and Communication*, 6(4Theoretical Reflections and Case Studies). <https://doi.org/10.17645/mac.v6i4.1652>
- Laghari, A. A., Jumani, A. K., Laghari, R. A., & Nawaz, H. (2023). Unmanned aerial vehicles: A review. In *Cognitive Robotics* (Vol. 3). <https://doi.org/10.1016/j.cogr.2022.12.004>



- Li, W., Xiao, L., Yao, S., Hou, C., Wen, Z., & Ren, D. (2025). ED-YOLO: an object detection algorithm for drone imagery focusing on edge information and small object features. *Multimedia Systems*, 31(3). <https://doi.org/10.1007/s00530-025-01783-9>
- Livberber, T., Ayvaz, S., & Yalman, A. (2024). Perspectives on drone journalism in Turkiye: Journalists' views on the use of drones. *Heliyon*, 10(21). <https://doi.org/10.1016/j.heliyon.2024.e39790>
- Montes-Romero, Á., Torres-González, A., Capitán, J., Montagnuolo, M., Metta, S., Negro, F., Messina, A., & Ollero, A. (2020). Director tools for autonomous media production with a team of drones. *Applied Sciences (Switzerland)*, 10(4). <https://doi.org/10.3390/app10041494>
- O'Hagan, L. A., & Serafinelli, E. (2022). Transhistoricizing the Drone: A Comparative Visual Social Semiotic Analysis of Pigeon and Domestic Drone Photography. *Photography and Culture*, 15(4). <https://doi.org/10.1080/17514517.2022.2116899>
- Passalis, N., & Tefas, A. (2020). Continuous drone control using deep reinforcement learning for frontal view person shooting. *Neural Computing and Applications*, 32(9). <https://doi.org/10.1007/s00521-019-04330-6>
- Pathak, K. (2024). Global trends and status of drone journalism: a systematic analysis of research constituents and key themes. *Global Knowledge, Memory and Communication*. <https://doi.org/10.1108/GKMC-09-2023-0323>
- Pinem, C. D., & Zarkasi, I. R. (2025). Metafungsi Videografi Unmanned Aerial Vehicle Pada Konten Komunikasi Publik Pemerintah. *WACANA: Jurnal Ilmiah Ilmu Komunikasi*. <https://doi.org/10.32509/wacana.v24i1.4651>
- Ramirez, F., Mari, W., & Martingayle, D. (2025). Let Us Fly Our Drones: An Examination of Student Newspapers' Coverage of Drone Journalism. *Journalism Practice*, 19(4). <https://doi.org/10.1080/17512786.2023.2218332>
- Sáez Martín, A., Haro de Rosario, A., & Caba Pérez, M. D. C. (2015). Using Twitter for Dialogic Communication: Local Government Strategies in the European Union. *Local Government Studies*, 41(3). <https://doi.org/10.1080/03003930.2014.991866>
- Sanders, K., Crespo, M. J. C., & Holtz-Bacha, C. (2011). Communicating governments: A three-country comparison of how governments communicate with citizens. *International Journal of Press/Politics*, 16(4). <https://doi.org/10.1177/1940161211418225>
- Soysal, Y., & Türkmen, S. (2024). Reinterpreting the Member Checking Validation Strategy in Qualitative Research Through the Hermeneutics Lens. *Qualitative Inquiry in Education: Theory & Practice*, 2(1). <https://doi.org/10.59455/qietp.19>
- Stratton, S. J. (2024). Purposeful Sampling: Advantages and Pitfalls. In *Prehospital and Disaster Medicine* (Vol. 39, Number 2). <https://doi.org/10.1017/S1049023X24000281>
- Suardi, S., Rahmat, A., Setiawan, W., Pawara, M. U., Alamsyah, A., Nugraha Arifuddin, A. M., & Hidayat, T. (2023). Design a Phinisi-Type Tourist Ship to Increase Tourist Interest in Vacationing at Taka Bonerate National Park. *International Journal of Marine Engineering Innovation and Research*, 8(2). <https://doi.org/10.12962/j25481479.v8i2.16616>
- Tilak, G. (2020). Drones and media industry. *RUDN Journal of Studies in Literature and Journalism*, 25(2). <https://doi.org/10.22363/2312-9220-2020-25-2-360-366>
- Vujičić, M. D., Kennell, J., Stankov, U., Gretzel, U., Vasiljević, Đ. A., & Morrison, A. M. (2022). Keeping up with the drones! Techno-social dimensions of tourist drone videography. *Technology in Society*, 68. <https://doi.org/10.1016/j.techsoc.2021.101838>



Wang, B. H., Wang, D. B., Ali, Z. A., Ting Ting, B., & Wang, H. (2019). An overview of various kinds of wind effects on unmanned aerial vehicle. *Measurement and Control (United Kingdom)*, 52(7-8). <https://doi.org/10.1177/0020294019847688>

