

Embodied Meaning in Motion: A Videographic Approach to Image Schemas in Film Studies

Maarten Coëgnarts^{1,2,3,*,†}, Miklós Kiss^{4,†}

¹University of Antwerp, Prinsstraat 13, 2000 Antwerpen, Belgium

²LUCA School of Arts, Paleizenstraat 70, 1030 Brussel, Belgium

³University of the Free State (UFS), Bloemfontein Campus, South Africa

⁴University of Groningen, Oude Boteringestraat 34, 9712GC Groningen, The Netherlands

Abstract

This paper explores the potential of videographic methods to represent and analyse dynamic image schemas in film. Building on John Bateman’s keynote at The Eighth Image Schema Day 2024, which emphasized the need for embodied formalisations of image schemas, we respond to the call for more dynamic and temporally sensitive approaches to their representation. While past efforts in cognitive linguistics have proposed various diagrammatic models, they often remain static and struggle to capture the inherently dynamic nature of schemas like CONTAINMENT. Drawing on the practice of videographic criticism and two recent video essays by the authors, this paper develops a method that aligns with the temporal, embodied, and experiential character of image-schematic structures in film. Focusing on Robert Dewell’s work on containment, we animate four core patterns—entry, exit, inclusion, and exclusion—and demonstrate how their sequential combination contributes to the experiential rhythm of cinematic scenes. Two case studies illustrate how this method can enhance our understanding of embodied meaning in film.

Keywords

Containment, embodied cognition, film analysis, image schemas, videographic criticism

1. Introduction

In his keynote address at The Eighth Image Schema Day 2024, John Bateman [1] emphasized the crucial role of image-schematic relationships in audiovisual communication, particularly in film. Looking ahead, he proposed two key directions for further development: (a) that filmic representations may offer a particularly suitable medium for depicting image schemas, and (b) that there remains a pressing need to formalise image schemas in ways that engage more directly with first-person embodiment and simulation.

This call to rethink formalisation connects with a broader body of work in cognitive linguistics and related fields, where various authors have proposed visual representations of image schemas to clarify their structure and facilitate interpretation (e.g., Langacker [2]; Talmy [3]; Mandler [4]; Hedblom and Kutz [5]). The most recent of these efforts, by Hedblom et al. [6], introduces a diagrammatic image schema language (DISL) designed to overcome the complexity of earlier models by emphasizing compositional, diagrammatic clarity. However, as Bateman [1] points out, despite these advances, such representations remain static—at odds with the inherently dynamic nature of image schemas. From this perspective, he argues, any attempt to represent image schemas —“whether those are filmic, static pictures, diagrams, verbal expressions, comics, tangible interfaces, or even logics, can be evaluated by asking how well they function as embodied ‘simulators’ for the phenomena they are attempting to capture” (Bateman [1, p. 6]).

Proceedings of the Joint Ontology Workshops (JOWO) - Episode XI: The Sicilian Summer under the Etna, co-located with the 15th International Conference on Formal Ontology in Information Systems (FOIS 2025), September 8-9, 2025, Catania, Italy

*Corresponding author.

†These authors contributed equally.

✉ maarten.coegnarts@uantwerpen.be (M. Coëgnarts)

ORCID 0000-0003-2813-4847 (M. Coëgnarts); 0000-0002-9136-8200 (M. Kiss)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

This paper explores the potential of videographic methods to animate image schemas dynamically in film. Drawing on the growing practice of videographic criticism in film studies (e.g., Grant [7]; Keating [8]; Kiss [9]; Scott [10])—which fuses audiovisual montage with analytical reflection—it proposes a graphic animated method that resonates with the temporal, embodied, and experiential character of image schemas. The insights presented here build upon the practical process of making two video essays (Coëgnarts [11]; Kiss and Coëgnarts [12], which served as both methodological experiments and critical explorations of image-schematic structures in cinema.

The structure of this short paper is as follows. In the first part, we build on Robert Dewell’s [13] work on dynamic patterns of containment to animate four key patterns: entry, exit, inclusion, and exclusion. These patterns, grounded in the conceptual primitives of path, container, and object, function as building blocks that shape cinematic motion. We argue that film can combine these patterns sequentially to generate complex temporal structures (i.e., syntagmas of dynamic patterns), thereby shaping the experiential rhythm of particular scenes. Two case studies will be used to illustrate this argument.

2. Dynamic Patterns of Containment in Film

In his chapter “Dynamic Patterns of Containment”, Robert Dewell [13] argues for a renewed emphasis on Mark Johnson’s foundational insight that image schemas are not static forms but inherently dynamic patterns of experience. As early as in *The Body in the Mind* ([14], p. 29), Johnson made it clear that image schemas should not be understood as fixed, pictorial representations, but rather as recurring, embodied patterns of sensorimotor activity. To illustrate this point, Johnson offered a detailed analysis of the CONTAINMENT schema—one of the most frequently discussed and intuitively accessible image schemas in cognitive linguistics. Dewell builds on this by emphasizing the dynamic nature of CONTAINMENT and elaborating its structure through two primary manifestations: ENTRY and ENCLOSURE. Both are grounded in the same set of conceptual primitives—container, object, and path—but they differ in how these elements are configured and in which component is in motion. In the ENTRY pattern, the object is the mobile figure that moves along a path into a relatively stationary container (e.g., a person walking into a room). In contrast, ENCLOSURE reverses this dynamic: here, the container itself is the mobile figure that moves around and closes in on a stationary object (e.g., a door closing around someone, or arms wrapping around an object). These two patterns are not merely variations in spatial configuration; they reflect different experiential logics and affordances. By distinguishing between them, Dewell underscores the importance of capturing the temporal and embodied structure of image schemas—something that static representations often fail to do.

The literature presents a variety of diagrammatic strategies for representing dynamic image-schematic patterns, with CONTAINMENT being one of the most frequently explored (for a good review, see Hedblom et al. [6]). Broadly speaking these visualisations tend to fall into two broad categories. The first includes summary scans—static images that condense a dynamic spatial pattern into a single representation. Examples of this approach include Johnson’s [14] diagrams using circles and arrows to depict conceptual primitives like container, path, and object, as well as Mandler’s [4] use of varied arrow types to indicate different forms of movement. Dewell [13] builds on these earlier efforts by explicitly including moving objects in the diagrams, adding a layer of specificity. The second category comprises visualisations that aim to reflect sequential state transitions, thereby incorporating a temporal dimension. A notable example is Langacker’s [2] panel-based format, which represents dynamic events as a series of frames, each capturing a different stage of a process. This method makes the temporal progression of image schemas more explicit. Building on Langacker’s approach, Hedblom et al. [6] recently introduced the Diagrammatic Image Schema Language (DISL). Drawing inspiration from both panel-based linguistics and the narrative structure of comic strips, DISL uses sequential “panels” to portray how image-schematic states unfold over time. This format offers a more compositional and temporally aware way to model experiential patterns. For instance, as illustrated in Figure 1, DISL can represent a full scenario of an object entering a container, capturing not just spatial relations but also the dynamic sequence of the event.

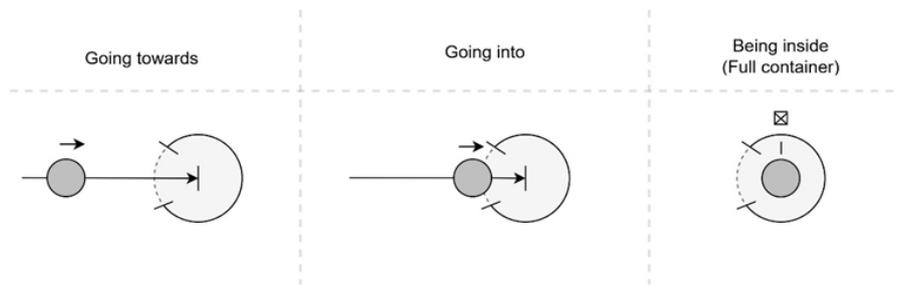


Figure 1: Recreated example of Hedblom et al.'s [6, p. 159] DISL representation of ENTRY.

In his work, Coëgnarts (e.g., [15], [16]) extends Dewell's theoretical insights by proposing that dynamic containment patterns are intrinsic to the dynamic logic of cinematic form. These patterns can be defined with greater precision: ENTRY occurs when an object or character moves into the frame while the camera remains static, whereas INCLUSION refers to the camera itself bringing a previously unseen object into view, either by panning, tilting, or tracking. Conversely, EXIT denotes a static-camera situation in which an object leaves the frame, while EXCLUSION describes the camera moving away from the object until it is no longer visible. Other patterns—ENCLOSURE (camera moves closer to an object), EXPOSURE (camera moves away from an object while keeping it in focus), DISTANCING (an object moves away from the camera), and APPROACHING (an object moves toward the camera)—occur primarily along the z-axis, articulating depth and spatial proximity in film. Movements along this axis also inherently engage the SCALE family of spatial primitives and image schemas, as changes in perceived depth are often accompanied by qualitative and contextually significant shifts in the apparent size of objects. Such transformations may manifest as GROWING or SHRINKING, depending on whether the object appears larger or smaller over time (see also Hedblom et al. [6, p. 154]). In practice, these patterns rarely occur in isolation—filmmakers often combine them, for example, INCLUSION + ENTRY, where camera movement and subject movement into frame coincide, generating layered spatial dynamics. Together, these basic motion patterns can be sequenced or interwoven, either within a long take or through editing, to construct rhythmic structures that align closely with the scene's emotional and narrative fabric.

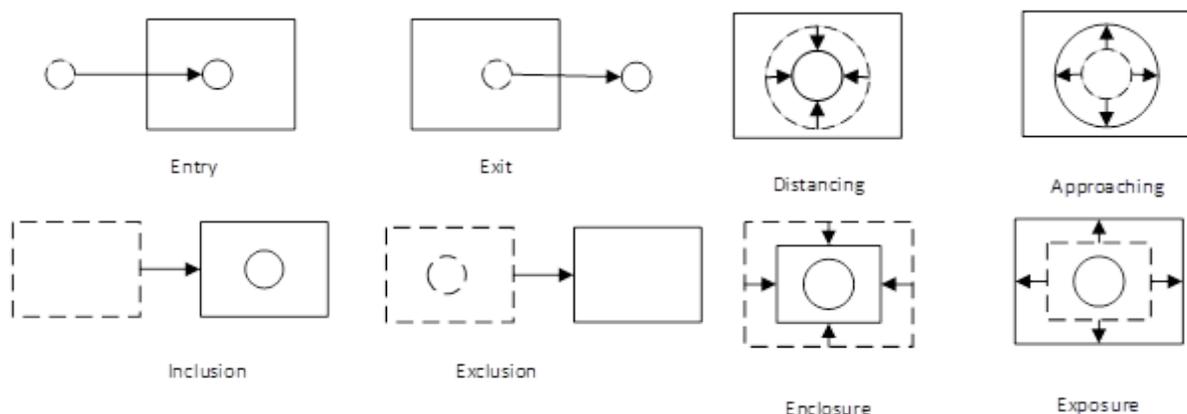


Figure 2: Dynamic patterns of containment (after Coëgnarts [15]).

More recently, these dynamic tools have been expanded and animated through the videographic format itself. The video essays “Embodied Visual Meaning in Motion” [11] and “Predictable Unpredictability” [12], both published in the peer-review journal [in] Transition—represent the first practical outcomes of this theoretical approach, offering a novel method of analysing embodied visual meaning in film. Using standard video editing software (such as DaVinci Resolve or PowerDirector), conceptual

primitives are animated as abstract figures and brought in relation to film excerpts (either sequentially, side-by-side or in overlay). In “Predictable Unpredictability” [12] objects are represented through Picture-in-Picture (PiP) overlays—simplified geometric shapes or icons layered atop the footage and animated through position and scale. Containers are visualized through added frames or borders that enclose the shot, evoking a sense of spatial containment. Paths are rendered via keyframe animation, which enables controlled transformations in position, scale, or opacity over time. When combined, these schematic elements articulate dynamic spatial relations that are often implicit in cinematic storytelling.

What becomes particularly apparent through this videographic analytical method and illustration is the affective variety embedded within these otherwise generic patterns. While the underlying schemas—such as ENTRY, EXIT, INCLUSION and EXCLUSION—are abstract and broadly applicable, their experiential texture varies dramatically depending on how they are rendered in a specific cinematic context. The pacing of a camera movement, the rhythm of an edit, the proximity of a character to the lens—all inflect the schematic pattern with distinct emotional resonance. For instance, a slow enclosure of a character might evoke intimacy or menace, whereas a rapid exit could suggest urgency or disorientation. These affective nuances are difficult to convey through static diagrams or verbal description alone, but they become palpable and analyzable in the videographic format. By animating these patterns directly onto the film’s temporal flow, the videographic format not only reveals the embodied structure of cinematic meaning but also highlights its expressive richness—showing how formal repetition can lead to emotional diversity.

In the following, we demonstrate the analytical potential of this videographic approach through two brief case studies. Each example highlights how a specific configuration of dynamic patterns—shaped by particular film-stylistic choices—generate embodied visual meaning that resonates with the scene’s narrative and emotional texture.

3. Videographic Analysis of Image Schemas: Two Cases

3.1. Taxi Driver (Martin Scorsese, 1976)

The first case analyzes the iconic phone call scene from Martin Scorsese’s *Taxi Driver* (1976), also included in “Embodied Visual Meaning” [11, 05:28–06:09]. In this well-known long take, Travis Bickle (Robert De Niro) is shown on the phone, attempting to arrange a second date with Betsy after a failed first encounter. The shot unfolds through a sequence of three dynamic containment patterns. While it is clear that Travis has no success in arranging another date, the camera excludes Travis along the x-axis, visually suggesting the emotional discomfort of the moment (see Figure 3, A1-A2). The camera moves laterally to the right, finally resting on the converging depth lines (z-axis) of an empty hallway. This empty frame not only emphasizes absence but also anticipates the emergence of the two subsequent spatial patterns. The emptiness of the hallway invites Travis’s re-ENTRY from the left edge of the frame once the call ends (see Figure 3, A3). As he moves into the frame, the strong z-axis perspective cues his subsequent motion from the foreground of the frame to the background; a pattern of DISTANCING, that also activates a NEAR-FAR schema (see Figure 3, A4). In the video essay, these dynamic patterns were presented sequentially: first as animated schematics, then followed by the corresponding film fragment. In the video referred by the public DOI code in the caption of Figure 3, we have reinterpreted the original videographic diagram using the overlay method further developed in “Predictable Unpredictability” [12], with two key modifications. First, the circle representing the character is now positioned directly over the actor’s face. This adjustment provides a more accurate reference point for determining shot size—an essential cinematic parameter typically gauged by the perceived distance between the camera and the subject. Second, we adjusted the scale of the circle to account for the change along the z-axis depth, enhancing the diagram’s spatial precision.

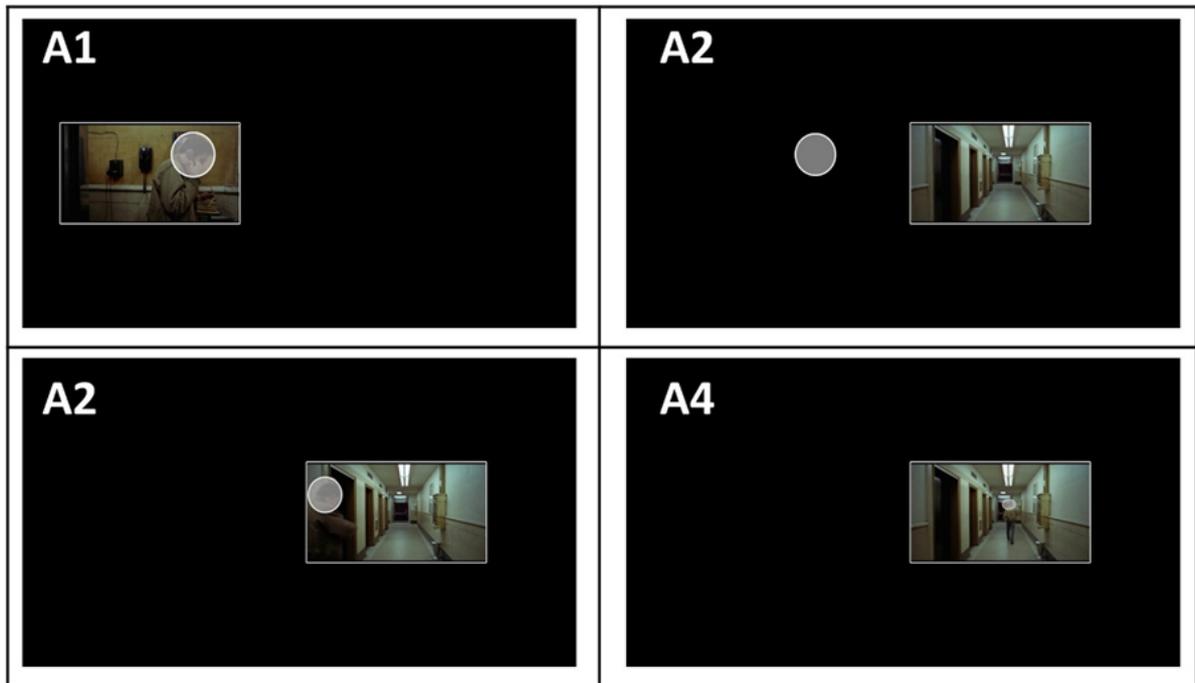


Figure 3: Four film stills from the videographic analysis of the dynamic patterns of containment underlying the phone call scene from *Taxi Driver* (1975). To access the video, go to <https://doi.org/10.6084/m9.figshare.29936117>.

3.2. *La Notte* (Michelangelo Antonioni, 1961)

The second case examines an excerpt from Michelangelo Antonioni's *La Notte* (1961), also featured in "Embodied Visual Meaning" [11, 03:58–06:09]. *La Notte*, the second film in Antonioni's "trilogy of alienation," explores the strained relationship between Giovanni (Marcello Mastroianni) and his wife Lidia (Jeanne Moreau), which is further complicated by the alluring presence of Valentina (Monica Vitti). In the selected excerpt, these interpersonal dynamics are conveyed through a carefully orchestrated sequence of dynamic containment patterns.

The excerpt begins with Lidia initiating a movement that sets the camera in motion while also predicting a pattern of INCLUSION: as she walks toward Valentina, the camera tracks her, eventually framing Valentina within Lidia's spatial "container" (see Figure 4, A1-2). As Coëgnarts [17] explains from the cognitive perspective of Predictive Processing, the widescreen aspect ratio affords the potential to fill the empty space on the left side of the frame (see Figure 4, A3). It establishes an anticipatory structure that is soon fulfilled as Giovanni enters the frame from the left, completing a triangular composition among the three characters, with Valentina positioned at the center (see Figure 4, A4). Symbolically, she becomes enclosed by the couple, suggesting a sense of emotional or relational confinement. Yet this triangular arrangement is unstable and short-lived. It anticipates Valentina's EXIT from the frame. At this point, Antonioni cuts to a new shot that reorients the perspective—shifting focus from the couple's backs to a frontal view. Now placed in the foreground, Valentina exits to the right, disrupting the compositional triad and returning the focus to the couple, once again framed in isolation as a dyadic unit.

Also here we have decided to reinterpret the original videographic diagram using the overlay method (see the video referred by the DOI code in the caption of Figure 4).

4. Discussion and Future Work

The videographic method outlined in this study offers a novel tool for both film analysis and image schema research. By operationalizing foundational concepts from cognitive linguistics—such as image

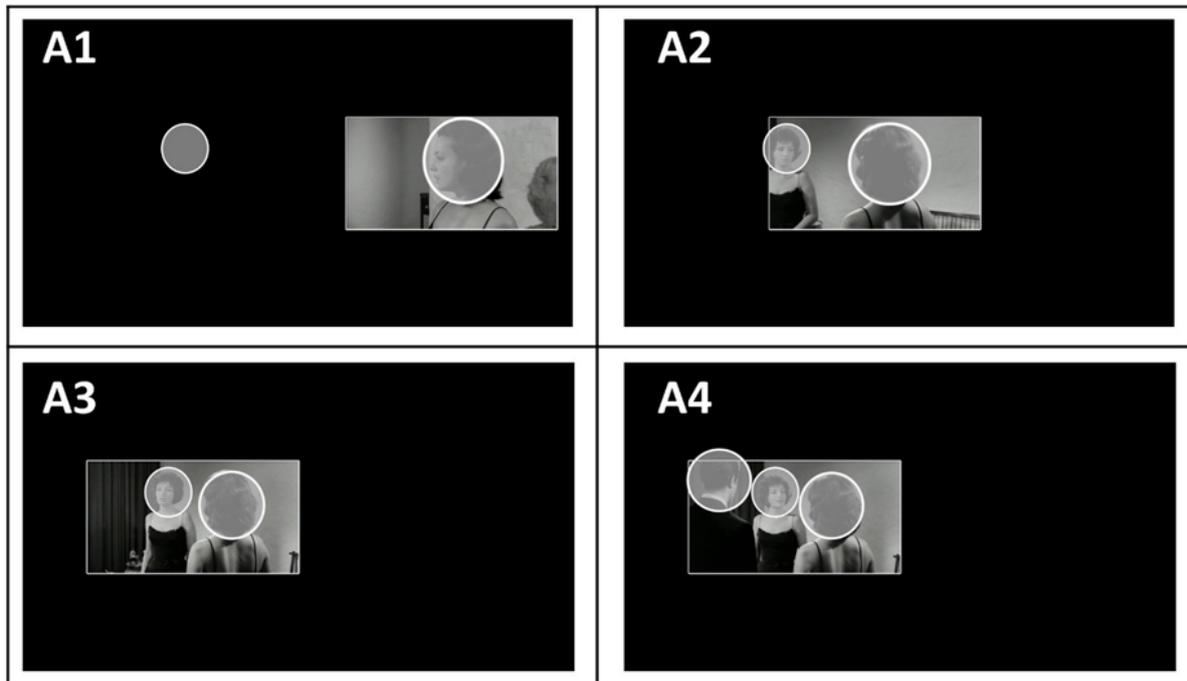


Figure 4: Four film stills from the videographic analysis of containment patterns in a scene from *La Notte* (1961). To access the video, go to <https://doi.org/10.6084/m9.figshare.29936144>.

schemas, conceptual primitives, and embodied metaphor— within a temporal and audiovisual format, the approach bridges the gap between abstract theoretical models and the perceptual experience of cinema. It enables scholars to not only identify dynamic patterns such as containment, path, and force, but to render them experientially and analytically visible through motion and rhythm.

For film analysis, this method enhances our ability to track the expressive function of formal elements—camera movement, editing, framing—not as isolated stylistic choices but as carriers of embodied meaning. It foregrounds how narrative and emotional effects are grounded in recurring schematic structures, yet always modulated by their particular spatiotemporal execution. This opens new pathways for analysing affect, style, and narrative coherence, especially in scenes where visual meaning is central but difficult to verbalize. For image schema theory, the videographic approach introduces a complementary method for testing and visualizing theoretical claims. Rather than relying solely on verbal or static diagrammatic representations, scholars can now model, simulate, and iterate schematic phenomena in motion—capturing their compositionality, variability, and affective resonance. In this way, videographic work may serve as a form of “embodied experimentation,” opening up interdisciplinary dialogues between cognitive science, film studies, and digital media.

Future research could expand this videographic approach to visualizing image schemas by incorporating additional schemas (such as force dynamics, balance, or verticality), or exploring cross-cultural cinematic uses of schematic patterns. Moreover, collaborative projects between cognitive linguists and film scholars could further explore how cultural conventions and genre norms shape the expression and perception of these embodied structures. The integration of videographic techniques promises to enrich both fields by offering a medium that is not only analytical but also performative—revealing how we come to understand film not just with the mind, but with the body in motion.

Declaration of Generative AI

The author(s) have not employed any Generative AI tools.

References

- [1] J. A. Bateman, Image schemas, cognitive metaphor, and film: Bridging discourses, in: The Eighth Image Schema Day co-located with The 23rd International Conference of the Italian Association for Artificial Intelligence (AI* IA 2024), Bozen-Bolzano, Italy, November 27-28, 2024, CEUR-WS.org, 2024.
- [2] R. W. Langacker, Foundations of cognitive grammar: Volume I: Theoretical prerequisites, volume 1, Stanford University Press, California, 1987.
- [3] L. Talmy, Towards a cognitive semantics. cambridge, ma-london, england: Bradford book, 2000.
- [4] J. M. Mandler, How to build a baby: II. Conceptual primitives, *Psychological review* 99 (1992) 587–604.
- [5] M. M. Hedblom, O. Kutz, Conceptual Puzzle Pieces: an image schema experiment on object conceptualisation, in: G. Bella, P. Bouquet (Eds.), *Modeling and Using Context (CONTEXT)*, LNCS, Springer Nature, Cham, Switzerland, 2019, pp. 98–111.
- [6] M. M. Hedblom, F. Neuhaus, T. Mossakowski, The diagrammatic image schema language, *Spatial Cognition and Computation* 25 (2025) 138–175. doi:PATH10.1080/13875868.2024.2377284.
- [7] C. Grant, The audiovisual essay as performative research, *NECSUS. European Journal of Media Studies* 5 (2016) 255–265.
- [8] P. Keating, The video essay as cumulative and recursive scholarship. the cine-files fall 15, 2020. URL: <http://www.thecine-files.com/the-video-essay-as-cumulative-and-recursive-scholarship/>.
- [9] M. Kiss, Videographic criticism in the classroom: Research method and communication mode in scholarly practice. the cine-files 15, 2020. URL: <https://www.thecine-files.com/videographic-criticism-in-the-classroom/>.
- [10] D. Scott, Videographic criticism and the art of creative thinking, *Media Practice and Education* (2024) 1–23.
- [11] M. Coëgnarts, Embodied visual meaning [in] motion, [in] *Transition* 10 (2023). doi:PATH10.16995/intransition.11280.
- [12] M. Kiss, M. Coëgnarts, Predictable unpredictability: Dynamic embodied patterns in miklós jancsó's film style'(creators' statement), [in] *Transition* 12 (2025). doi:PATH10.16995/intransition.17622.
- [13] R. Dewell, Dynamic patterns of CONTAINMENT, in: B. Hampe, J. E. Grady (Eds.), *From perception to meaning: Image schemas in cognitive linguistics*, Mouton De Gruyter Berlin, Germany, 2005, pp. 369–393.
- [14] M. Johnson, *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*, The University of Chicago Press, Chicago and London, 1987.
- [15] M. Coëgnarts, *Film as embodied art: Bodily meaning in the cinema of Stanley Kubrick*, Academic Studies Press, 2019.
- [16] M. Coëgnarts, M. Slugan, Embodying meaning visually: from perceptual dynamics to motion kinematics, *Art & Perception* 10 (2022) 137–158.
- [17] M. Coëgnarts, The predictive embodied mind: a case-based encounter with film aesthetics, *Frontiers in Neuroscience* 19 (2025) 1583107.