

Generative Montage: An Analysis of AI-based Video Editing Software Effects on Montage

MADALENA MIRANDA, UNIDCOM IADE – Universidade Europeia, Portugal;
email: madalena.miranda@universidadeeuropeia.pt

ABSTRACT

This chapter explores how AI-based video editing software is transforming the grammar of cinematic montage. As artificial intelligence becomes embedded in creative workflows, particularly through features such as automated scene detection, transcription, and generative extension and rhythm analysis, traditional notions of editing as a human-driven, but tech-led expressive practice are being redefined. The study addresses two questions: how do these tools alter foundational principles of montage, and what are their implications for the editor's cognitive and creative agency?

The chapter employs a theoretical and critical approach, drawing on film theory, media studies, and software analysis. It makes an exploratory comparison between the early evolution of montage in Soviet film theory – primarily the work of Eisenstein, Kuleshov, Pudovkin, and Vertov – through the digital turn described by media and film theorists such as Manovich, Casetti, Pearlman and Furstenau, and the present-day incorporation of AI in editing software. It reflects how the aesthetics, temporality, and ontology of montage shift when decision-making is partially automated.

Attention is paid to the role of the software's interface as a site of negotiation between human intention and algorithmic suggestion. Ultimately, the chapter argues that AI-enabled editing does not eliminate montage, but reconfigures it: from a dialectical grammar to a generative, semi-automated, hybridized form of media composition. The chapter contributes to a broader understanding of how cinema's expressive and representational possibilities are being reshaped in this moment of algorithmic aesthetics.

KEYWORDS: Automated Collage, Cinema Montage, AI-based Video Editing Software

INTRODUCTION

Over the past decade, artificial intelligence has emerged as a key agent of transformation across multiple media industries. In audiovisual production, AI-driven tools have rapidly evolved from peripheral enhancements to central elements of the creative workflow. This shift has been evident in the domain of video editing, where AI-based software is reshaping not only how content is assembled but also how cinematic language itself is conceived and practised.

This chapter addresses some implications of this transformation for one of cinema's foundational elements: montage. Traditionally conceived as the deliberate selection and arrangement of moving images to produce narrative, rhythm, or ideological effect, montage has historically demanded human judgment, intuition, and creativity to create meaning. However, as editing platforms nowadays integrate machine learning algorithms capable of automatic scene detection, rhythm-based cutting, and emotion-driven sequencing, the once-manual process of montage is becoming increasingly automated. These developments raise fundamental questions about creativity and authorship, decision-making, and aesthetic agency in the age of computational cinema.

The central concern of this chapter is to understand how AI-based video editing software begins to alter some points and practices of the grammar of montage. More specifically, the chapter asks: What happens to the disruptive, the unexpected “montage of the attractions” and its expressive potential when non-human systems settle pre-structured decisions? How do these technologies influence the editor's cognitive processes and creative freedom? And to what extent can such systems be seen as co-authors in the making of cinematic forms?

With experience as a filmmaker and editor across different montage exigence stages – from television series to institutional, experimental, and authorial fictional and documentary films – along with training and research in the area, the work pre-

sented here combines practical experience and theoretical questioning.

To explore these questions, the chapter adopts a theoretical and critical analysis that combines film and media theory, software studies, and a historical reflection on montage practices. Drawing on foundational montage theorists – such as Sergei Eisenstein, Dziga Vertov, and Lev Kuleshov – and contemporary authors – such as Lev Manovich, Francesco Casetti, Marc Furstenau, and Karen Pearlman – the analysis considers the evolution of montage alongside that of media technologies. Particular attention is paid to the interfaces and affordances of current AI-based editing platforms, which not only support but also shape users' decisions. The methodology used in this chapter will focus on content analysis of reviews of editing AI tools on professional YouTube channels, using a sample of professional and semi-professional YouTube channels focused on video editing software and trends.

Ultimately, this chapter situates montage within a broader transformation of media grammar in the digital age. The chapter argues that the integration of AI into editing workflows demands a rethinking of montage as a grammar, but with an internal, opaque, and automated dynamic, rather than an external one in which the cinematic and visual culture of the editor propelled the grammar's changes and neologisms. It also problematizes montage as a mutable, negotiated practice – one in which human and machine inputs collide, converge, and co-compose. The result is not only a technical shift, but also a cultural and conceptual one: the automation of montage may simultaneously constrain and expand the possibilities of cinematic production in the twenty-first century.

HISTORICAL AND TECHNICAL OVERVIEW OF VIDEO EDITING SOFTWARE

The evolution of video editing software reflects broader shifts in media production and digital aesthetics. In the 20th century, film editing was a labor-intensive, linear



FIGURE 1. Elizaveta Slilova, “The Man with the Movie Camera”,
© Dziga Vertov, 1929.

process, reliant on mechanical tools and physical splicing. Some of this intensive labor was also gendered, as seen in “The Man with the Movie Camera” (Vertov, 1929), where the “monteuse” is a woman, as was the case in many films in film history, where the detailed, hard, meticulous work was a woman’s job. This was the core action of filmic language, the site of its expressive uniqueness since the early days of cinema. As Marc Furstenu highlights, early in his “Aesthetics of Digital Montage”, editing functioned as “a plastic art,” whose plasticity lay in manipulating artistically cinematic material directly on the editing table, the historical moviola. In this sense, montage was not only a narrative tool but a tactile, material engagement – one that established cinema’s distinct artistic ontology.

This was the heart of avant-garde cinema, as Russian authors both experimented with and theorized about. As Sergei Eisenstein wrote in “The essence of montage is conflict: conflict between graphical directions, between volumes, between spatial movements... conflict between an object and its spatial nature.” In “Film Form” (p. 45, 1949). The principles of opposition and dialectics became the foundation of cinematic meaning-making, distinguishing montage not as continuity, but as conceptual rupture and synthesis.

Such were the examples of *Battleship Potemkin* (1925), or Vsevolod Pudovkin’s articulation of montage as a means of building narrative logic and emotional resonance – “the linkage of shots, not their collision,” as he emphasized in *Film Technique and Film Acting* (1929). Likewise, Lev Kuleshov’s experiments demonstrated how context reshapes meaning: through the juxtaposition of identical images, he showed that viewers construct emotion and narrative from the relational placement of shots, an idea foundational to theories of *affective editing*. This is well known as the *Kuleshov effect*.

The foundations of dialectical montage, or film editing, were established. Although different types of editing were developed during the 20th century, notably

among modern authors and contemporary filmmakers such as Nouvelle Vague cineasts and independent American authors, and names like Stanley Kubrick, variations in the understanding of the editing process were an early achievement of Soviet avant-garde film theory. To that, also, was added the steadiness of the 35 mm film editing process until linear video stations arrived.

The introduction of Non-Linear Editing (NLE) in the late 1980s and 1990s, such as Avid Media Composer and later Adobe Premiere and Final Cut Pro, marked a profound transformation. These digital tools allowed editors to manipulate footage with unprecedented freedom, enabling faster workflows, iterative processes, and experimentation with narrative structures, reorganising the tools and reinventing the language and its grammar.

In between the newest modes of montage available and the most classical and material ones, the flow of conscience and creativity was there. The ease of the process, the convenience, the higher cutting speed, and the experimental capacity of the first digital software did not compromise the editor’s creative flow. This new workflow departed from what William Warner famously called “dealing with the same pain from the linear editor,” a phrase he used to describe his motivation for developing Avid Media Composer after struggling with the limitations of linear editing.¹

The modular structures of cinematic grammar quickly changed with this ontological transformation, as a more fluid mode of expression emerged from the editing tables. This shift resonated with broader sociocultural paradigms such as Zygmunt Bauman’s notion of *Liquid Modernity* (Bauman, 2000), which describes a world marked by flexibility, change, and uncertainty, or Manuel Castells’ vision of the *Network Society* (Castells, 1996), where information technologies reconfigure production, communication and identity. The

1 “The Untold Story: Interview With Avid Media Composer Inventor”, <https://www.youtube.com/watch?v=bthKAx707UQ&t=1023s>



FIGURE 2. Screen still frames from “Avid 1/Media Composer – 1990 Avid Experience Tour Promo Video”, https://www.youtube.com/@The_Rough_Cut <https://www.youtube.com/watch?v=yDkTHOYneZ0>

transformation in editing practices thus mirrored these broader transitions, aligning aesthetic modularity with societal fluidity and networked logic.

The emergence of non-linear editing (NLE) systems marked a turning point in the history of audiovisual post-production. Before the 1980s, editing involved physically cutting and splicing reels of film – a process constrained by linearity, irreversibility, and the material limits of celluloid. This changed with the introduction of the first commercially available NLE systems, notably the CMX 600 in the early 1970s (developed by CBS and Memorex), and later, the revolutionary Avid Media Composer in 1989. Avid was among the first systems to allow digital, frame-accurate, real-time editing of broadcast-quality video, and it quickly became the industry standard in film post-production workflows across various countries.

In the United States, high-profile productions initially adopted it as advertised. An interesting fact in the testimony about the editing process in the video is the acceleration of the process. Internationally Avid gained prominence through its use in European cinema, particularly in the UK and France, where it was employed in independent productions and national film schools. Filmmakers valued its ability to offer precise control over pacing and structure, essential to both narrative and experimental editing strategies. Its adoption in film schools like *La Fémis* in Paris and the NFTS in London further cemented Avid's role in shaping a new generation of editors and cinematic grammar. Its precision, speed, and ability to integrate image and sound in real-time made it the dominant choice throughout the 1990s for many editors working in both cinema and emerging digital formats.

The idea of finishing the load of the material relation with tapes in linear editing was an important step to the adherence Avid Media Pro had among the community, hence *moviola* nostalgia. Its widespread adoption by editors signaled a decisive break from analogue practices and helped

redefine professional editing as a digital craft, before Adobe Premiere and Final Cut Pro became most popular. These platforms digitized footage and enabled real-time, random-access manipulation of image and sound. Editors could now preview, rear-range, and modify clips with less degradation and image quality loss and an easier process. This development accelerated the workflow and redefined the conceptual architecture of editing: filmic time *gained plasticity*, sequences became potentially iterated, and the montage process became more transposable and fragmented.

These characteristics capture what Lev Manovich describes in *The Language of New Media* (2002) as key features of new media: modularity, variability, and working according to the logic of the computed database. Editing becomes not just a technical process but an emergent artistic and cultural practice shaped by the aesthetics and affordances of computational media.

Adobe Premiere, launched in 1991, brought NLE capabilities to personal computers, while Apple's Final Cut Pro, released in 1999, further democratized access to professional-level editing tools. The advent of these systems also aligned with the rise of digital cinematography, completing the transition to a fully digital production workflow. Although it raises some paradoxes, both in image capture as well as in digital editing, as Francesco Casetti introduces in "The Lumière Galaxy":

Throughout its history, cinema has often met with changes that seemed to push it beyond its technological, linguistic, and institutional boundaries, and at the same time it has always sought to remain itself, also negotiating or appropriating novelties (Casetti 2015:4).

The new status quo relates both material and cultural conditions of the cinematic experience, and the editing process is a part of this ecology. This editing environment was shaped by technological developments

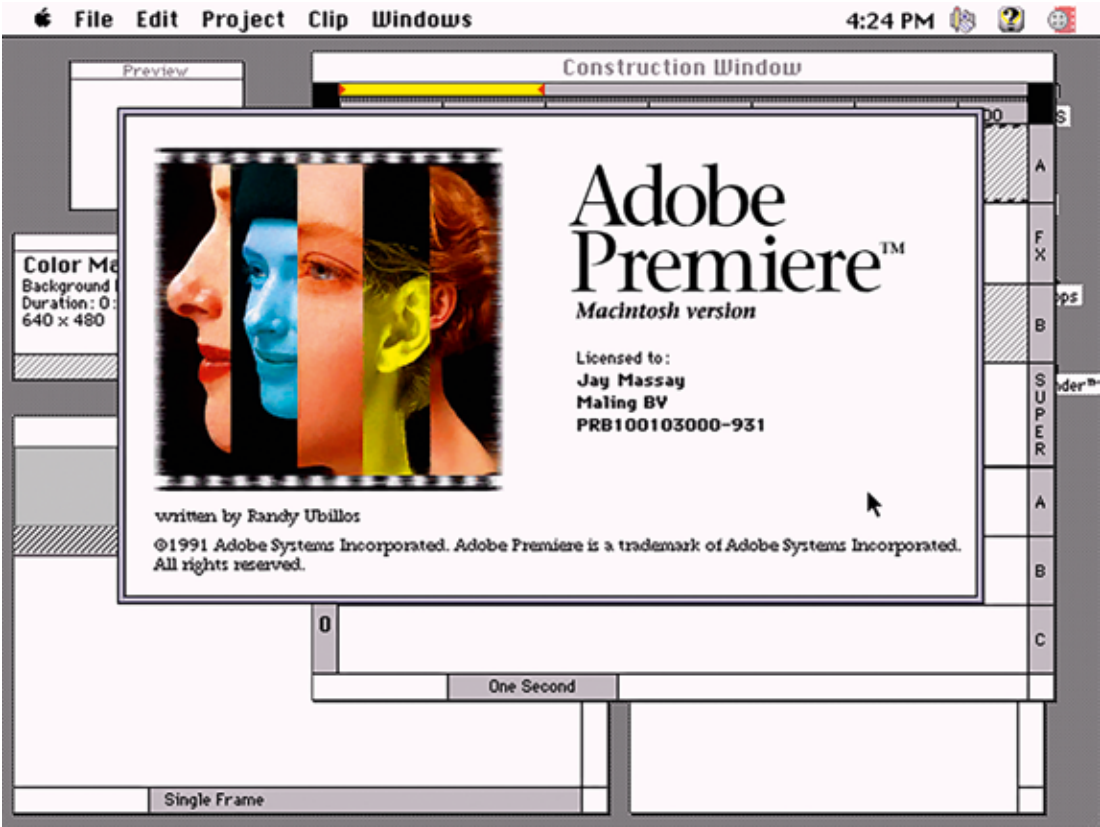


FIGURE 3. Screen still frame from the interface of Adobe Premiere 1.0 (Mac) in <https://winworldpc.com/product/adobe-premiere/10-mac>

and also by the integration of digital image capture, transforming filmmaking into a process increasingly embedded in computational inscription and manipulation. The transformation of editing, through the rise of digital montage as the primary locus of computational manipulation, directly reflects this shift. Editing became a site of aesthetic consequences, where digital enhancement tools created limitless possibilities for image construction that shaped a tendency and a time. As Furstenau (2024:11) emphasizes, “It is the new editing technology that is understood, in the most direct sense, to have transformed the basic identity of the cinema.”

Although this profound shift, while non-linear editing redefined temporal and spatial manipulation, its core logic remained *user-directed*, focused on human intervention. Editors selected, arranged, and trimmed clips according to the narrative, the rhythm or the expressive authorial intentions. This shift toward digital tools made cinema probable as a “database”, the main argument in Soft(ware) Cinema, Manovich’s project. In his essay, *Database as Symbolic Form* (1998), Manovich further conceptualizes the database as a new aesthetic form that replaces traditional narrative logic.

This transformation, aligned with broader cultural trends towards modularity, remixability, and interactive media, became a new paradigm of audiovisual and cinematic production and consequently impacted visual culture. In this context, Jonathan Beller (2006) describes the emergence of a contemporary “cinematic mode of production”, a socio-technical system in which the visual regime of cinema extends beyond the screen and becomes embedded in everyday acts of watching, interacting, and producing. For Beller, this logic organizes the attention economy, where image circulation and engagement become sites of labor and value creation. And reclaims to the editing table the place where it also takes place:

Materially speaking, industrialization enters the visual as follows: early cinematic montage extended the logic of the assembly-line (the sequencing of discreet, programmatic machine-orchestrated human operations) to the sensorium and brought the industrial revolution to the eye. Cinema welds human sensual activity, what Marx called ‘sensual labour’, in the context of commodity production, to celluloid. (Beller 2006:95)

Although this theoretical extreme proposal of enframing montage as an act of navigating and recognizing or mapping from a collection of indexed possibilities and database materials, Cinema did not disappear, and montage and editors continued to work. The continuity of film editing has shifted in its coordinates, from temporal to spatial; the film strips and the photogrammes were reorganized with new affordances in vectorial terms, and editing creativity kept its challenges.

The cinematic work emerges from a second layer of interaction, after the interface interplay between selection, categorization and recombination. In contemporary practice of film editing, the *data set* becomes an important part of the user experience, as both the interface and the interaction. Although the new media object that is digital cinema is perceived mostly as a narrative form, what the current montage paradigm shows are changes also in the cultural forms: “What we encountered here is an example of the general principle of new media: the projection of the ontology of a computer onto culture itself” (Manovich 1998:5).

Some decades later, with the evolution of the NLE systems, new challenges emerge. With the advent of Generative Artificial Intelligence, this process is undergoing another paradigmatic shift. AI-based editing software, such as Adobe Sensei, Runway ML, and DaVinci Resolve’s Neural Engine, now incorporates machine learning



FIGURE 4. Screen still frames of Final Cut Express 4.0 commercial,
<https://www.macintoshrepositary.org/9277-final-cut-express-4>

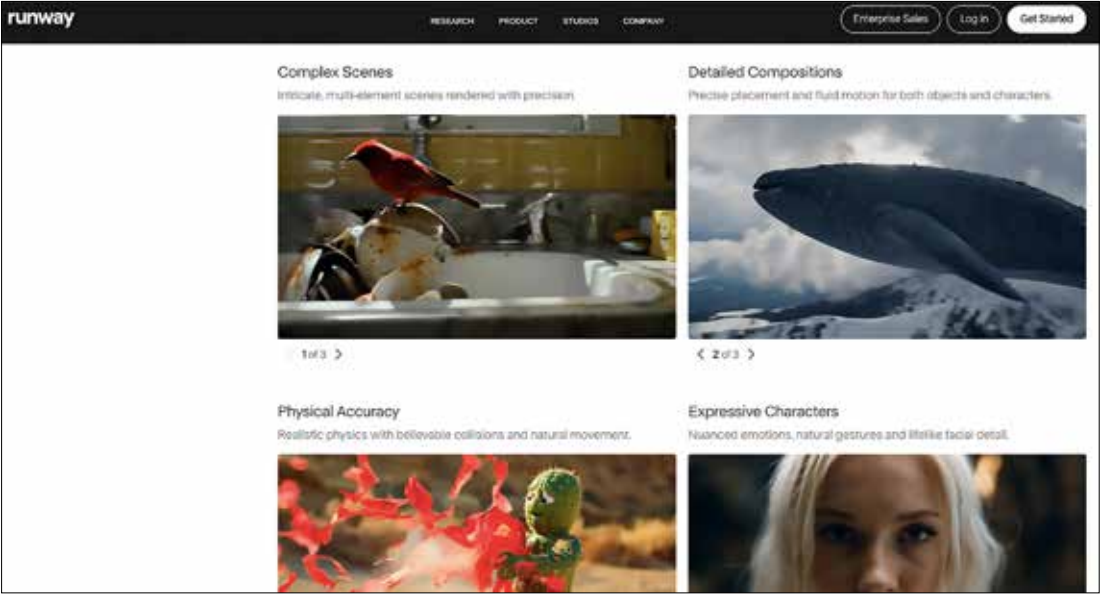


FIGURE 5. Screen still frame about new video generation model Runway Gen-4.5 from <https://runwayml.com/research/introducing-runway-gen-4.5#features-section>

systems that automate decisions and processes previously performed by humans. These tools offer features like automatic scene detection, rhythm-based cuts, and even style imitation. Rather than merely executing commands, the software begins to *pre-select, predict, suggest*, and edit based on learned data and pre-configured grammar. In this context, the editor's role and relevance are redefined: a creative decision-maker can become a curator of algorithmic options.

This growing automated autonomy has been called *Alaesthetics* (2022, Arielli & Manovich), in which computational systems assume an active role in shaping visual culture and artistic production. Editing in this landscape integrates a process of verification, validation, and supervision of the software's outputs. These developments raise critical questions: How does this shift affect the cognitive process of editing? Does it narrow or expand creative possibilities? And does the editor have a *performative conscience* of the whole process? What remains of montage as an expressive, political, or dialectical act when decisions are filtered through algorithmic logic? To address these questions, it is necessary to reflect upon editing software, not merely as a technological progression, but as a cultural artefact – one that encodes, standardises, and perhaps even dictates the possibilities of cinematic expression.

A PRACTICE-BASED APPROACH TO ARTIFICIAL INTELLIGENCE EDITING TOOLS USES

As a practical examination of how AI tools modify foundational principles of montage and influence the editor's cognitive agency, this section analyses generative AI features in contemporary editing software by drawing on the most frequently discussed reviews and the features highlighted by professional editors on their YouTube channels. The choice of this methodology concerns the newness of the AI editing tools, and the necessary practice-based in-depth of their use. If some discourse is already available in the software business, and

some professional and industry editors are using some tools, the mainstream reviews and access to these are on active editors from the audiovisual scene, working with different levels of engagement, who are usually the tech ambassadors.

Some of the most referred-to tools are audio tools, such as Voice Conversion and Ripple Delete Silence in Da Vinci Resolve, which are among the features presented². Audio manipulation is easier and less resource-intensive to render, and audio voice cloning is becoming increasingly advanced. Another tool, variously referred to by different video editors, is a transcription tool, where, in the clips, the AI can create a file with the spoken transcription, and the editor can check the moment for each clip. It is referred to both in Da Vinci Resolve³ or Avid Media Composer⁴. Also, the conversion of the transcription to subtitles is transversal to both editors. This can be a clue to understanding advances in language-learning models that range from word and sound transcripts to written languages. The image effects are more modest and still limited.

For example, the AI Speed Warp⁵ is a reference; once, it changes the quality of the speed-duration change, liberating the temporality of editing choices from the quality of the shot material. In Adobe Premiere, one of the most impactful AI tools is Generative Extend, which can generate 2 more seconds at the beginning or end of a clip with AI. As seen on Justin Brown's Primal Video YouTube Channel, the Generative Extend works particularly well on ambient, nature, or catalogue-style clips. There is a difference between reimagining a cinema scene and redoing it in the editing suite. Also, an image-based AI tool and a commonly used feature is scene edit

2 These features belong to a Greg Edits Video Channel, with around 12 k followers. <https://www.youtube.com/watch?v=D33F1tktokg>

3 Idem

4 Janus Edit Master is another YouTube channel, with an analysis of Avid Media Composer – <https://www.youtube.com/watch?v=l6Gf4N-dLT4>

5 Also from Greg Edits: <https://www.youtube.com/watch?v=D33F1tktokg>



FIGURE 6. Screen still frame about AI automated scene detection tool from Justin Brown Primal Video <https://www.youtube.com/watch?v=PjObX9XQvgI&t=648s>

FIGURE 7. Screen still frame about automated dialogue transcription from Janus Edit Master: <https://www.youtube.com/watch?v=I6Gf4N-dLT4>

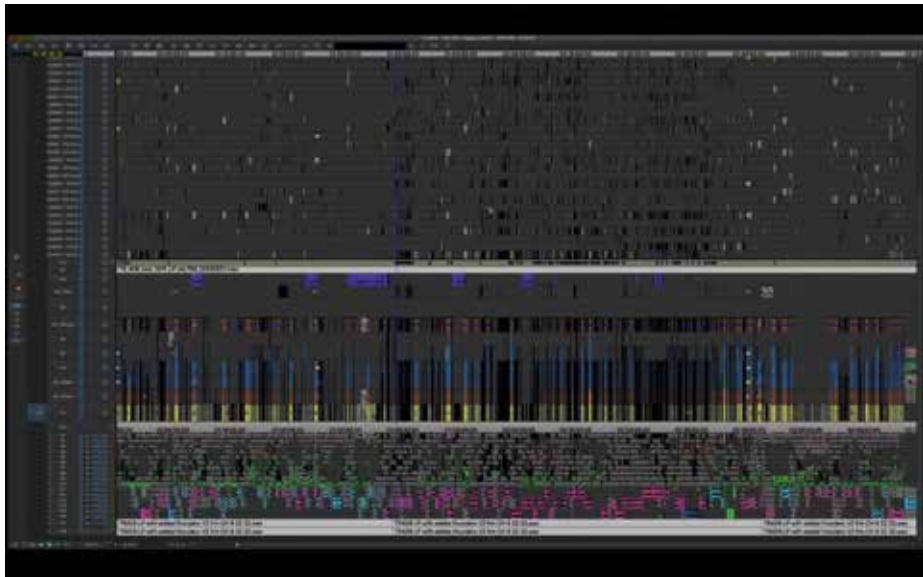
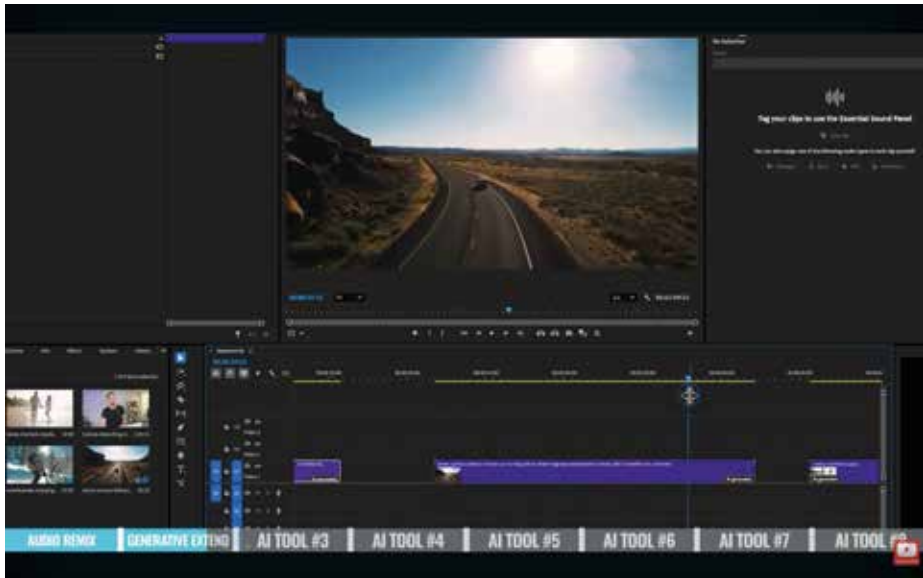


FIGURE 8. Screen Still Frame about generative extend tool from Justin Brown Primal Video
<https://www.youtube.com/watch?v=PjObX9XQvgI&t=648s>

FIGURE 9. Still frame from Learning the Art of VFX Editing – Stranger Things – The Rough Cut,
<https://www.youtube.com/watch?v=zMVSqerhLj4&t=936s>

detection, which identifies previous edits and reorganizes them in a new timeline. Sometimes there is no previous project file to work upon. Also referred to are the AI reframe, which enables the transmediation of formats much more easily; the Morph Cut, a transition using morphing images to make the cuts; and the colour correction AI-assistant, which automates the colouring processes.

Compared with Rough Cut⁶, a YouTube channel that defines itself as “Exploring the tips, techniques and careers of the top editors and postproduction pros working in film and television”, no special video on AI uses has been released. The channel’s strategy is to have the testimony of top editors showing and explaining their editing processes, without focusing on the software’s features in terms of technicality or user experience. It rather focuses on the purpose of Cinema and the film editing or editing workflow it has.

MONTAGE THEORY: DIALECTICS, CONFLICT, AND FORM

Montage is not merely a technique of cinematic construction; it is the very grammar of cinema, a language of assembling meaning through juxtaposition. Since its earliest theorization, montage has served as both an artistic strategy and a representational gesture. In the Soviet avant-garde of the 1920s, it became a tool for shaping consciousness, producing ideological effects through the collision of images.

Sergei Eisenstein developed a theory of montage that foregrounded conflict as a generative force. In *Film Form* (1949), he argued that “the essence of montage is conflict: conflict between graphical directions, between volumes, between spatial movements... conflict between an object and its spatial nature” (p. 45). This approach to editing did not aim at continuity or realism but at creating a synthesis through contradiction. For Eisenstein, montage was a dialectical process grounded in the

philosophical and Marxist notion of thesis-antithesis-synthesis. He believed that it was through the clash of shots, each bearing its ideological charge and pathos, that new ideas and emotional responses would emerge. In the Odessa Stairs editing, each cut and collage corresponds to an effect and an affect, through the narrative and the sequence.

Also, Lev Kuleshov, a contemporary of Eisenstein, provided experimental evidence for his theory through what later became known as the “Kuleshov Effect.” By showing the same neutral facial expression intercut with different contextual images – a bowl of soup, a child in a coffin, a woman reclining – Kuleshov demonstrated that meaning in cinema arises not from individual images, but from their sequential arrangement. The viewer actively constructs emotional and narrative coherence through the montage. And with the power of association, the relations and inner forces of montage as the grammar of film became evident and structured.

These Soviet theorists established montage as a foundational language of cinema. Their theories differ in their emphasis and conclusions about conflict, construction, and revelation or catharsis, yet they share a belief in the transformative potential of editing. Montage is not neutral; it is a place to create meaning, rational, sensorial and emotional engagement. Understanding this is essential for evaluating the role of AI in contemporary editing. The transition to generative and automated montage challenges the centrality of human intention in sequencing and juxtaposition. In the next section, we will argue how AI-based tools interfere with these foundational models of cinematic form.

NEW MEDIA AND THE AUTOMATED SPECTACLE

The rise of artificial intelligence in editing interfaces can be interpreted through the lens of media archaeology and early film theory. In particular, Tom Gunning’s seminal concept of the cinema of attractions framework (1986). Gunning argues that early

6 The Rough Cut https://www.youtube.com/@The_Rough_Cut



FIGURE 10. Film still frames of “Potemkin’s Battleship” (1925),
<https://marxist.com/soviet-cinema-montage-revolution-and-the-fight-for-artistic-freedom.htm>

FIGURE 11. film still frame of the Kuleshov effect in
https://www.imdb.com/title/tt3736946/?ref_=mv_close

cinema, before the dominance of narrative continuity, was primarily a visual spectacle arousing curiosity and surprise. This cinema did not attempt to immerse viewers in a story world, but instead engaged them through direct sensations display – a cinema of showing over telling.

Wanda Strauven (2006) revisits and expands Gunning's argument, suggesting that the attraction never truly disappeared but remained latent, periodically resurfacing in moments of technological innovation. With the advent of digital media, and particularly in interactive and computational interfaces, the logic of attraction returns. The emphasis shifts back to the loop, the immediacy, and installation aesthetics. Something that resonates with remediation concepts, as in Bolter and Grusin (1999). The film continuum becomes programmed – not only what we see, but how we see it, how fast, with automated feedback. The premonition of the loop becomes the now-overwhelming infinite scroll.

If we follow this argument, making an important shift from the reception of the spectator to the reception of producer, in this case, the editor, interacting with the software interfaces, the AI-based video editing software can be analyzed as a new “engine of attractions”, once the editors are the first spectators of the “software attractions”, hence some of these current technological responses are much more ways to fasten audiovisual medium editing, in an overwhelmed embodied environment. These editing systems with AI features, both generative and extended tools, such as Adobe Sensei, Runway ML, and DaVinci Resolve's Neural Engine, offer real-time responses, predictive editing, and automated audio and visual generation.

Its usability *attracts* users with its technical prowess: pre-calculated cuts, rhythm predictions, and media reconstruction responses. The editor's experience becomes an interaction with a responsive, interactive surface – an interface that simulates a frictionless efficiency, with a focus on the editing tool usability.

The speed and surprise with which these tools operate establishes a parallel with the logic of the attractions of early cinema and early digital image days. Much like early cinema's fascination with mechanical motion and illusion, AI NLE lies not only in the final results but in the performativity of the editing process itself. Watching the system “learn,” “detect,” or “generate” in real time evokes a contemporary wonder, a form of computational magic. This can be particularly relevant when the spectator who was the subject of reception in the Cinema of Attractions is now the “prosumer”, meaning the producer, of the majority of the circulating digital imagery made with AI.

The notion of immediacy is crucial here; hence, interfaces are designed to reduce latency and to deliver creative results at a new pace. In doing so, they reinforce a culture of instant gratification, aligning with the logic of platforms and feeds. And this also influences editing, which now has more of a posture, a performance of direct action and reaction, almost like a live event or streaming. It evolves within the user experience ideology of intuitive, immediate dialogue between the machine and the human author.

This techno-aesthetic recalibration transforms the editor's relationship to time, labor, and form. If classical montage theory emphasized choice, rhythm, and dialectical tension, the AI-enhanced interface underlines speed, ease, predicted cuts and pre-organized timelines, a dive into flow. In other words, what is gained in ease and access may be lost in an intentional cut. *Montage as usability* replaces *montage as collage*.

AI editing as “attraction engines” opens questions about perception, cognition, and visual pleasure⁷ in the age of automation, that are now present also at the beginning of the cinematic transformation in the editing process, hence digital culture prosumers (Jenkins, 2016) are impacting audiovisual content creators that also shape the professional tools as

7

Laura Mulvey's essay (2006) on “visual pleasure” concept can be reenacted considering the principle of AI image generated pleasure.

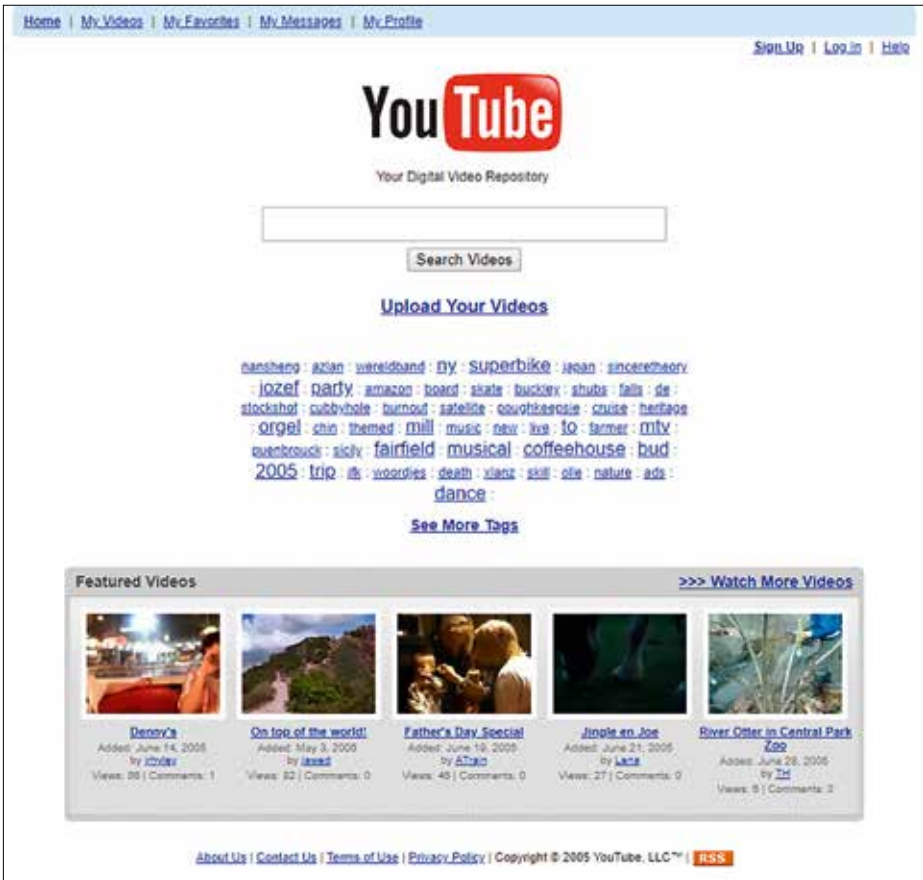


FIGURE 12. YouTube frontpage in 2005
(<https://www.webdesignmuseum.org/gallery/youtube-2005>)

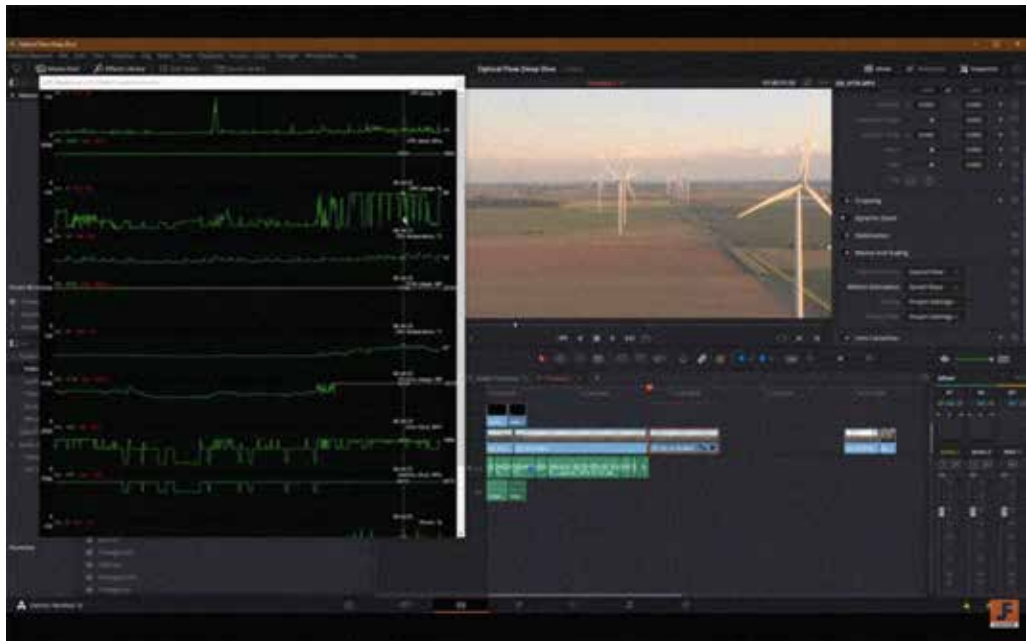


FIGURE 13. Screen still frame of Da Vinci Resolve Neural Engine in John's Films' YouTube channel (https://www.youtube.com/watch?v=u_4KG8ljliY)

editing software and consequently the movie industry practices. Recent features being implemented in editing software will change the structures of professional cinema, too. Even if, as Casetti affirms, cinema will always be cinema. The attraction engine of AI software reflects a broader epistemic shift: from narrative and conflictual dialectics to interface dynamics, from human gesture to algorithmic suggestion. These changes must be critically examined as part of a larger reconfiguration of audiovisual creation in the 21st century.

In the following section, we will explore how this new logic manifests in the cognitive and aesthetic consequences of generative editing workflows, asking how far these tools transform not only what we edit, but also how we think about montage itself.

AESTHETIC AND COGNITIVE IMPACT OF GENERATIVE EDITING INTERFACES

The advent of AI-assisted editing interfaces has not only restructured workflows but also redefined the ontological foundations of cinema itself. As Marc Furstenau (2018, 2024) argues, the digital reconfiguration of montage introduces a new paradigm in which the computational logic of software reshapes the material and temporal coherence of the cinematic medium. Montage, once the locus of human expression in Cinema, decision-making and rhythm, is now shared, if not encompassed, by algorithmic processes that simulate or predict the editor's human and artistic agency.

In classical editing environments, the editor operated as "*a sculptor of time*"; a reference to filmmaker Andrei Tarkovsky – making deliberate choices based on intuition, narrative structure, and aesthetic rhythm. Furstenau (2018) emphasizes that the shift to digital montage erodes this model, introducing a new form of editing practice in which decisions are no longer solely intentional but increasingly reactive to what the digital interface proposes. With this shift, the editor becomes what we might call a *curator of suggestions* – a verifier and adjuster, a supervisor of computa-

tionally generated outputs. But is this not an emergent, broad way of interaction with Artificial Intelligence, as LLMs and Generative AI systems are also used in other domains of cultural production?

This transformation has profound cognitive implications. The editor confronts an interface already populated with options, and the labor becomes one of curatorial selection. The psychological and pedagogical dimensions of editing, formerly rooted in trial and error, in watching and feeling through footage, shift toward automated evaluation and preselection's of the material, which may not be entirely screened. This reframes cognition from generative intention to a critical filter. The produced imagery is now independent, pre-exists the editor, the director and at last, the film itself.

Furstenau (2024) refers to this shift as an ontological rupture in the aesthetics of montage. If film was once defined by its capacity to manipulate time and form through manual sequencing, AI-based tools turn editing into a database-driven interaction. The software operates through a logic of affordances: what it allows the editor to do, and what it does *before* the editor acts.

Moreover, the visibility of editing interfaces themselves plays a crucial role in shaping decisions. Contemporary software often visualizes options through previews, auto-arranges sequences, or recommends structures. These interface shapes and forms subtly guide the user's sense of what is appropriate, effective, or even cinematic. They not only facilitate decisions but also influence them and the user experience design of the interface, which Galloway (2012) defined as an autonomous zone of interaction. This also creates a new aesthetic condition in which machine affordances train human culture.

What is at stake is not merely the speed or convenience of editing, but its conceptual grounding. As the editor role shifts, a new form of cinematic authorship emerges – distributed, contingent, and mediated. The montage is no longer an act of composition but a networked negotiation

between human and nonhuman agents, as Bruno Latour's Actor Network Theory predicted.

The process is not devoid of potential. Editors may reuse or creatively misuse these affordances. What can emerge is a hybrid montage: a form in which pre-structured logic meets improvisational decision. The result can again be a conflictual space where montage becomes differently constructed, and new significance can emerge, as new dialectical forces.

The implications of this transformation extend beyond aesthetics into broader questions of media literacy, training, and artistic autonomy. If AI tools offer new forms of access, they also impose new constraints. As such, understanding the aesthetic and cognitive impact of generative editing is essential not just for theorizing cinema but for imagining the future of audiovisual creation. In the next section, we conclude by showing how these developments redefine the grammar of montage, interrogating the tension introduced by computational reconfiguration.

FROM COLLAGE MONTAGE GRAMMAR TO GENERATIVE MONTAGE

The grammar of montage, forged through decades of artistic experimentation and theoretical elaboration, is undergoing a radical reframing. The introduction of AI-based editing tools does not merely optimize existing workflows – it redefines the rules and logics through which cinematic meaning is constructed. The shift from intentional sequencing to algorithmic suggestion implies not only a new technique but a new syntax, one that emerges from the computational infrastructures underlying digital media.

In classical montage theory, as seen in Eisenstein or Kuleshov, the cut was an intentional disruption, a chosen point of transformation where meaning, emotion, or rhythm was reconfigured. Today, cuts can be generated by machine vision, structured by automated pattern recognition, and suggested through templates designed

by developers. The logic of montage has become partially embedded in code, shifting the axis of decision from artistic to operational. This evolution signals a movement from montage as composition to montage as configuration. Instead of manually assembling sequences based on expressive goals, the editor interacts with system-generated options: timelines pre-populated with recommended edits, transitions auto-inserted for emotional pacing, or rhythm adjusted to music through automated synching. Generative montage does not erase the editor's agency, but it displaces it, replacing authorship as a negotiation with pre-structured possibilities.

Such changes invite us to consider whether montage remains a site of dialectical tension or becomes a field of *optimized continuity*. What happens to conflict, cut and the disruptive power of editing when the software privileges flow, coherence, and audience prediction? The dialectical montage described by Eisenstein, full of friction and contrast, could give way to a flat affective experience. This does not suggest that algorithmic montage cannot be expressive or politically charged. But it does reorient the grammar: from dialectics to data flow, from disruption to effectiveness. Editing becomes more likely to be interface navigation than material confrontation. The editor now works within a spectrum of algorithmically viable options, rather than from a blank canvas of imaginative potential.

Moreover, the generative system itself becomes a co-author of form. Automated montage extends this logic: the software not only stores and retrieves, but also arranges and proposes. It could become a procedural machine of aesthetic assembly – an apparatus that renders cinematic language as a programmable set of outcomes. However, this transformation also opens critical space. If the grammar of montage is being redefined, it can also be reprogrammed. Editors, artists, and theorists are beginning to explore ways to subvert the defaults of generative tools – misusing templates, breaking recommended structures, or training models on subversive

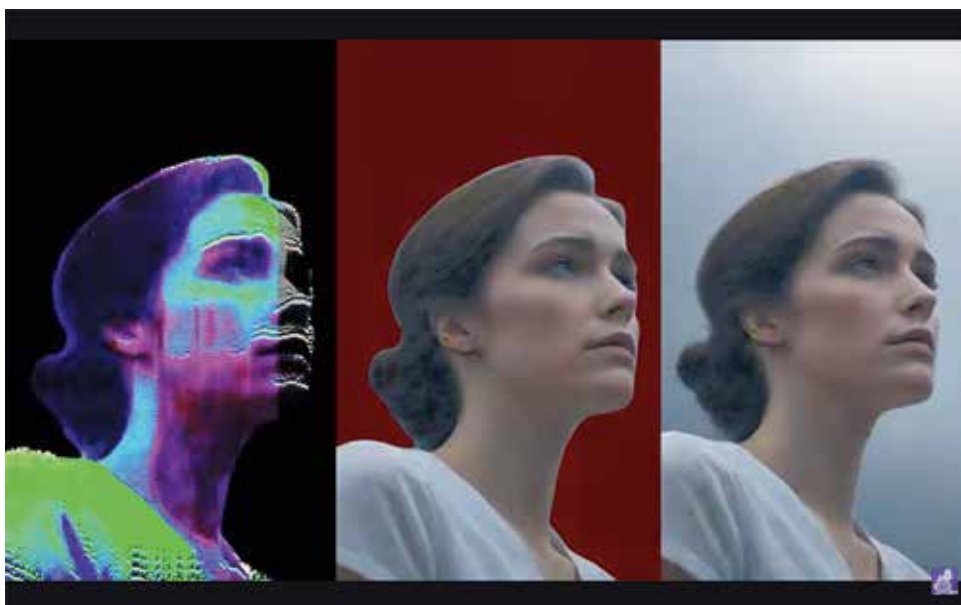


FIGURE 14. Screen still frame from YouTube Channel YOVOZOL
(<https://www.youtube.com/watch?v=ZGE4pnXxTcI>)

aesthetics. This practice reclaims montage as resistance within the very tools that seek to automate it.

In this emergent phase, generative montage can be understood not as the end of cinematic grammar but as its mutation. A new hybrid syntax is forming – part pre-structured logic, part improvisational play. Understanding this transformation is crucial for assessing the future of film form, and for ensuring that montage remains a site of creative and critical intervention in the age of automation.

CONCLUSIONS

We have traced how non-linear editing systems introduced modularity and temporal flexibility, and how AI systems are now pushing further, transforming the editor from an authorial agent into a supervisor of algorithmic proposals. This shift affects not only the workflow but also the cognitive, affective, and conceptual dimensions of editing. The *cut*, once an intentional act, now often arrives pre-assembled, guided by invisible and opaque logics of data aggregation and pattern recognition.

Revisiting foundational Eisenstein's dialectical montage, Kuleshov's affective sequencing, we can assess what is at stake in this transition. Their commitment to montage as an expressive and political force contrasts with the seamlessness of optimized, and often opaque, structures of generative editing interfaces. The shift from synthetic montage to *montage as a system* invites a broader critique of contemporary media production and its entanglement with automation. At the same time, this transformation can open new creative and critical possibilities.

Generative montage can become not necessarily a closure of form, but a mutation, a hybridization of programmed suggestion and human improvisation. Editors and artists are already finding ways to challenge the defaults, inject contingency, and reclaim montage as a space of rupture, irony, and experimentation.

To understand the future of montage is to understand the shifting relationships between humans, machines, and media grammars. As AI continues to expand its role in audiovisual production, the need to theorize, critique, and reshape these tools becomes ever more urgent. Only through such engagement can we ensure that cinematic language evolves not as a passive product of automation, but as an active terrain of artistic negotiation and cultural intervention.

REFERENCES

- Arielli, E., & Manovich, L. 2022. AI-aesthetics and the Anthropocentric Myth of Creativity. *Nodes* (19–20):91–97, <https://doi.org/10.57633/NODES-19-20/1-ENG>.
- Bauman, Z. (2000). *Liquid Modernity*. Polity Press.
- Beller, Jonathan 2006. *The Cinematic Mode of Production: Attention Economy and the Society of the Spectacle*. Dartmouth College Press.
- Eisenstein, Sergei [1949] 1977. *Film Form: Essays in Film Theory*. Ed. and trans. Jay Leyda. San Diego: Harcourt.
- Casetti, F. 2015. *The Lumière galaxy: Seven key words for the cinema to come*. Columbia University Press.
- Castells, M. (1996). *The Rise of the Network Society: The Information Age: Economy, Society, and Culture*. Oxford: Blackwell.
- Furstenau, Michael 2018. 'Film Editing, Digital Montage, and the "Ontology" of Cinema'. – *Cinemas: Journal of Film Studies* 28, 2, 29–49.
- Furstenau, Michael 2024. *The Aesthetics of Digital Montage: Film Editing and Technological Change*.
- Bolter, J. David, and Richard A. Grusin. *Remediation: Understanding New Media*, MIT Press, 1999.
- Galloway, Alexander R. *The interface effect*. Polity, 2012.
- Jenkins, H., & Ito, M. 2015. *Participatory culture in a networked era: A conversation on youth, learning, commerce, and politics*. John Wiley & Sons.
- Kuleshov, Lev 1974. *Kuleshov on Film: Writings of Lev Kuleshov*. Ed. and trans. Ronald Levaco. Berkeley: University of California Press.
- Manovich, Lev 2002. *The Language of New Media*. Cambridge, MA: MIT Press.
- Manovich, Lev 2013. 'Database as a Symbolic Form'. – Ross Parry (ed.), *Museums in a Digital Age*. London: Routledge, 64–71.
- Mulvey, Laura. "Visual pleasure and narrative cinema." *Media and cultural studies: Keywords* (2006): 342–352.
- Strauven, Wanda (ed.) 2006. *The Cinema of Attractions Reloaded*. Amsterdam: Amsterdam University Press.