The Rhythm Image

Music Videos and New Audiovisual Forms

By Steven Shaviro

BLOOMSBURY ACADEMIC
NEW YORK • LONDON • OXFORD • NEW DELHI • SYDNEY

BLOOMSBURY ACADEMIC
Bloomsbury Publishing Inc
1385 Broadway, New York, NY 10018, USA
50 Bedford Square, London, WC1B 3DP, UK
29 Earlsfort Terrace, Dublin 2, Ireland

BLOOMSBURY, BLOOMSBURY ACADEMIC and the Diana logo are trademarks of Bloomsbury Publishing Plc

First published in the United States of America 2023

Cover photograph: FKA Twigs performs at the Alexander McQueen: Savage Beauty Fashion Gala at the V&A © David M. Benett/Getty Images for Victoria and Albert Museum

A catalogue record for this book is available from the Library of Congress.

Library of Congress Cataloging-in-Publication Data
Names: Shaviro, Steven, author.
Title: The rhythm image : music videos and new audiovisual forms / by Steven Shaviro.
Description: New York : Bloomsbury Academic, 2022. | Series: New approaches to sound, music, and media | Includes bibliographical references and index. |
Summary: "Analyzes recent music videos in depth, and considers how they reflect new technologies and new cultural conditions, from an audiovisual perspective"– Provided by publisher.
Identifiers: LCCN 2022015186 (print) | LCCN 2022015187 (ebook) | ISBN 9781501388552 (hardback) | ISBN 9781501388569 (paperback) | ISBN 9781501388576 (epub) | ISBN 9781501388583 (pdf) | ISBN 9781501388590
Subjects: LCSH: Music videos–History and criticism.
Classification: LCC PN1992.8.M87 S537 2022 (print) | LCC PN1992.8.M87 (ebook) | DDC 780.26/7–dc23
LC record available at https://lccn.loc.gov/2022015186
LC ebook record available at https://lccn.loc.gov/2022015187

ISBN: HB: 978-1-5013-8855-2
PB: 978-1-5013-8856-9
ePDF: 978-1-5013-8858-3
eBook: 978-1-5013-8857-6

Series: New Approaches to Sound, Music, and Media

Typeset by Integra Software Services Pvt. Ltd.

To find out more about our authors and books visit www.bloomsbury.com and sign up for our newsletters.

1

The Rhythm Image

Some Starting Points

"In the beginning, there was rhythm." So sings the British feminist punk band the Slits (Slits 1980). The sentence is also attributed, however, to the nineteenth-century German pianist and conductor Hans von Bülow (Walker 2009, 7). In either case, this declaration is an apt motto for *The Rhythm Image*. I take it for granted that rhythm is fundamental, because—to quote Mark Abel, in turn quoting Victor Zuckerkandl, who himself refers to Henri Bergson—rhythm "is not an event *in* time, but an event *of* time." That is to say, "time happens; time is an event .... Change does not create time; time literally creates change" (Abel 2014, 51; Zuckerkandl 1973, 181–5). Rhythm both enacts change and provides the background of continuity through which change is experienced. Digital audiovisual expression is always, as Maurizio Lazzarato writes, "a duration, a rhythm that enters into relation with other durations and other rhythms" (Lazzarato 2019, 104). Music (moving sounds) and cinema (moving images) are arts of duration, and thereby also arts of rhythm. Their entanglement, not in a rarefied form, but emerging through the hustle and bustle of marketing and publicity, is the basis of the contemporary music video.

In this book, I look closely and intensively at a small number of commercially produced music videos from the English-speaking world, all of them made in 2010 or later. I explore the ways that the visual and sonic rhythms of these videos fascinate and derange us. Music video is a hybrid and impure medium, in several ways. It is the bastard offspring of pop music singles and television advertising. It is a highly collaborative medium as well, since the musician or band usually works with a video director. The director might well have different expressive aims than the musicians do, but their differing perspectives must somehow coalesce. "To make a video," Saul Austerlitz writes, "the performer and director enter into a symbiotic relationship, one best compared to that between an architect and his or her client" (Austerlitz 2007, 8). In addition, both the musician and the director

usually have large crews of technicians and assistants. Often the artist's record company, or management, pays for the video, and demands a say in how it is made and what it shows. All this makes for a lot of people, a lot of perspectives, and a lot of disagreements and crossed signals.

The ostensible purpose of a music video is to promote an already existing song, as well as to establish or embellish the persona of the performing musical artist. The music has already been recorded before the video is conceived. Where filmmakers generally seek sounds that suit and amplify a primary image track, music video directors must seek to create images that match a pre-given soundtrack. And yet, even within these constraints, music videos are often highly inventive and independent artistic expressions in their own right. Their rhythms give rise to different perceptions, feelings, and ideas than the music alone does—or for that matter, than movies and television generally do. As I have written elsewhere, music video is an intrinsically *post-cinematic* medium, with its own tendencies and forms (Shaviro 2010, 2017a). In this book, I seek to describe, in particular cases, both how music videos give us small doses of what Carol Vernallis calls "audiovisual bliss" (Vernallis et al. 2019, 181), and how music videos "work us over completely," as Marshall McLuhan says that new media do, changing our relation to the world by altering the "ratios" of our "sense perceptions" (McLuhan and Fiore 1967, 41).

Rhythm is an informing force, and an organizing principle, for music videos visually as well as sonically. As the brilliant music video director Calmatic puts it in a video interview:

> I treat editing like I'm making a beat. I do all my cuts on the little in-between intricate sounds that people don't even really hear. To me, the edit has its own beat to complement the beat of the actual song.
>
> (Calmatic 2018)

Kodwo Eshun, whose work on music videos—as well as on Black popular music more generally—is absolutely foundational to any discussion of the medium, similarly writes about how sound and image work in entangled counterpoint. In Hype Williams's late-1990s videos for Missy Elliott, Eshun says, "choreography, effects processing and musculature all fuse in an angular fluidity." A music video is "a slice of time that enthralls and appalls in rapid succession." Rhythms cross over between the soundtrack and the image track, so that "new slownesses, new speeds, new fermatas, new polyrhythms, new dynamic events emerge from the elastic reality and softimages of broadcast entertainment culture" (Eshun 2000,

116). Eshun epitomizes these synesthetic transfers in the "sleek equation" of a phrase he takes from George Clinton: "the rhythm of vision is a dancer" (Parliament 1978). Rhythm involves the eyes as well as the ears; and even further, beyond sight and hearing, it resonates proprioceptively through the entire body.

The film theorist Vivian Sobchack argues that this sort of "cross-modal sensory exchange" already takes place in conventional filmviewing. The viewer/listener is a "*cinesthetic subject*," who feels the intensities not just of sight and sound, but of the other senses as well (Sobchack 2004, 69–70). This still applies to music videos, if anything in more concentrated form. Of course, music videos today are usually accessed online, on YouTube and other sites—rather than in large theaters as movies used to be. Also, we encounter them on a much wider variety of screens than was the case in the age of MTV (the 1980s and 1990s), when music video was part of the ongoing flow of cable television (Goodwin 1992, 134). Today, music videos appear on monitors of all sorts and sizes, from gigantic home theater screens all the way down to the tiny viewing areas of mobile phones. And instead of having to wait for a certain video to come up in the course of MTV's rotation, we can watch it whenever we want, and as often as we want.

The ubiquity and everydayness of the circumstances in which we currently watch and listen to music videos might seem to de-intensify them, in comparison both to blockbuster, large-screen cinema, and to the ways that certain videos used to be orchestrated as big media events (like Michael Jackson's "Thriller" in 1983). But in recent years, both the availability of new digital tools and the need to adopt new media strategies have led music video directors to sharpen and heighten their skills, honing their productions to degrees that are never matched in feature-length movies and television series. As I write this chapter, The Weeknd has had to pull one of his videos from being shown in IMAX theaters, due to worries that its "intense strobe lighting" might be triggering for epileptics (Bloom 2021). In contemporary music videos, editing, camera movement, and close-ups all happen on a severely compressed micro-level. As Eshun already said two decades ago, a music video works as "a five-minute hypersyncopated landscape in which everything has rhythmelodic properties," from the "camera angles" to "the limbs, the faces, the gestures, the expressions" (Eshun 2000, 123). One of the aims of this book—and the reason it pays such close, descriptive attention to the second-by-second formal movements of particular videos—is to untangle these new sensory synergies, these changing McLuhanesque ratios of the senses.

My choice of precisely which music videos to discuss at length in the following chapters should be understood as *exemplary*, rather than representative. That is to say, I do not claim to give a comprehensive survey of music video production today as a whole, or to catalog all the important recent trends. Also, I do not discuss fan-made or noncommercial work; I do not look at music videos outside the sphere of the Anglo-American music industry; and I do not discuss the ways that online sites like Twitch and Instagram have recently come to influence video production. The music videos that I discuss here are simply ones that spoke to me in a powerful way, and appealed to my aesthetic sensibility. In all the chapters of this book, I work to analyze the videos in question as closely and carefully as possible. I pay attention to their formal strategies, as well as to the pop culture allusions that they include. I also consider the role the videos play in the overall multimedia construction of the musician's persona. But in addition to such close analysis, I also look for affinities and continuities among the different videos; and I consider ways in which these works are pushing against the limits, or enlarging the affordances, of music video as a medium. Music videos are second only to pornography in their rush to employ the most cutting-edge technologies. And music videos often register, with particular acuteness, the powers and tendencies of digital audiovisual production more generally. Music videos often stand at the most intensive intersections of an overdetermined multiplicity of social, cultural, and technological forces.

When I say that the videos I discuss are exemplary, I am also thinking of Immanuel Kant's aesthetics. Kant says that great works of art are always original; they cannot be imitations of previous works. They are not made by following any preexisting rules; and yet they are sufficiently orderly and nonrandom that they seem to be made according to rules of their own invention. In this way these works seem like powerful examples; but that *of which* they are the examples does not exist anywhere apart from them or prior to them. This means that such works are also exemplary in the sense that they invite us to imitate them in turn. But the only way to create something genuinely powerful and expressive, Kant says, is not to imitate previous great works, but rather to *emulate* the ways in which they reject previous rules and models, and create new procedures of their own (Kant 2000, 193 and passim).

As commercial products of what Horkheimer and Adorno call the culture industry (Horkheimer and Adorno 2002, 94–136), music videos might not seem to be appropriate objects for this sort of high-minded approach. Music videos are fully ensconced within the realm of mass culture, as it is produced

and promoted by large media corporations. They are also rarely self-contained; they most often allude to, parody, and render homage to a plethora of previous works—including other videos, movies, television shows, commercials, internet memes, and fashion trends. As I have insisted before, music video is unavoidably an impure and highly compromised medium (Shaviro 2017a, 7). Still, just as McKenzie Wark argues for the value of *low theory* as opposed to academic high theory (Wark 2021; Webb 2011), so I argue for the low culture of music videos presented on YouTube and commented upon by fans and haters, as opposed to the high culture of video installations in art galleries and museums. As Wark explains,

> High Theory is the grand tradition of philosophy, claiming to legislate for other domains of thought and practice, whereas low theory is the organic concept-forming practices of everyday life, which might borrow from High Theory but really doesn't care about its desires … It is about inventing new practices of knowledge, hopefully more interesting ones.
>
> (Wark 2012)

My claim in this book is that commercial music videos are engaged in inventing new practices of expression, more broadly and exuberantly than conceptualist high-art works are able to do. My invocation of Kant's Romantic aesthetics is unlikely to please either high-minded theorists of modernism and after, or popular-culture scholars and critics. Nevertheless, I think that my invocation of Kant here is relevant and justified, given the ways that—in the words of Lev Manovich—"the genre of music video has been a laboratory for exploring numerous new possibilities of manipulating photographic images made possible by computers" (Manovich 2001, 261; see also a similar passage quoted by Eshun 2000, 116). Formal experimentation and formal innovation, even or especially in the context of a race to rack up the maximum number of views on YouTube, can lead to new, technologized versions of the beautiful and the sublime.

In any case, I am not trying to create a canon of music video masterworks, or to assert that the videos I discuss in this volume are necessarily superior to others that I could equally well have chosen. I only maintain that, more than two decades on from Eshun's and Manovich's observations, music videos have continued to push forward their explorations of the affordances for audiovisual expression that are opened up by new digital technologies. This is as much a process of aesthetic invention as it is one of commercial calculation. Stanley Cavell, writing about film, points out that the affordances of a new technology are

not intrinsically inscribed in the technology. They cannot be known in advance. They need to be both discovered and invented—here, these words are closer to being synonyms than alternatives—in the experimental process of creation. "The aesthetic possibilities of a medium are not givens … Only the art itself can discover its possibilities, and the discovery of a new possibility is the discovery of a new medium" (Cavell 1979, 32). All the music videos I discuss here are engaged in such a process. They all move beyond prior accomplishments, in order to explore new and "extreme possibilities" (to use the phrase made famous by David Duchovny's character Fox Mulder in *The X-Files*).

Scholarship on music videos has only recently caught up with many of the genre's developments. Since the videos themselves are characterized by such rich hybridity, the criticism and scholarship on them has also been scattered among different fields: music and sound studies, film and new media studies, popular culture studies, and so on. This can easily lead to confusion and contradiction. My own background is in film studies; when I have presented my work on music videos at popular music conferences (most notably the Pop Conference, which used to be located in Seattle, but which now takes place, post-Covid, in New York City), I am almost oppressively aware of the contrast between my own cinematically based work, and that of music scholars who probe musical form in ways that I could never manage. This divergence goes back to the very beginning of academic discussion of music videos. Back in the last century, the first Anglo-American book-length study of music videos, *Rocking around the Clock*, was written by the film scholar E. Ann Kaplan (Kaplan 1987). The second, *Dancing in the Distraction Factory*, was written by the pop music scholar Andrew Goodwin (Goodwin 1992). Kaplan writes about music videos from the perspective of feminist film theory, as well as in terms of the then-popular discourse of postmodernism. Goodwin, in contrast, describes his own approach as "rooted in the sociology of popular music, and in musicology itself"; he writes that he "view[s] music television very much from *within* the field of pop consumption" (Goodwin 1992, xviii).

These different starting points lead to strikingly different readings of the same videos. Kaplan sees Madonna's "Material Girl" video (Madonna 1985) as a disruptive, postmodern feminist work in which "the usual hierarchical arrangement of discourses in the classical realist text is totally violated" and deconstructed (Kaplan 1987, 125). Goodwin rather argues that the video works to construct Madonna's "star identity"; in order to accomplish this, the video is in fact "a highly conventional narrative that moves through an initial

lack, via action, to (romantic) resolution" (Goodwin 1992, 99–100). More generally, Goodwin says that Kaplan and other film-studies-based scholars fall short because of the "*visual bias* of [their] theories," which "necessarily neglect the analysis of sound" (Goodwin 1992, 20). They "have tried to analyze music television in iconographic, semiotic, and narrative terms, while paying insufficient attention to the sound track … Strangely enough, very few analysts have thought to consider that music television might resemble music" (Goodwin 1992, 3).

Though my own work is based in film theory, and especially in Gilles Deleuze's two *Cinema* volumes (Deleuze 1986, 1989), I have tried to keep Goodwin's warnings in mind as much as possible. I have drawn upon the often undervalued—but insightful and crucially important—strain of film scholarship that focuses on cinematic sound. And I have endeavored, to the extent that I am capable of it, to engage with academic musical scholarship, as well as musical journalism, as closely as I do with film studies and film-philosophy. Above all, I have attempted to compensate for my unavoidable cinematic bias by attending, as carefully as I can, to the sonic and musical aspects of all the works that I discuss. Of course, it is up to the reader to decide how well I have succeeded in this.

In any case, I am fortunate that academic scholarship with regard to music videos, much of it from a music-based perspective, has expanded and flourished in recent years. Carol Vernallis has led the way with two rich volumes (Vernallis 2004, 2013), and a third one forthcoming (Vernallis 2022b). Mathias Bonde Korsgaard has also published important work on the recent, post-televisual mutations of music video (Korsgaard 2017). There have been a number of rich anthologies as well, containing essays by a wide variety of authors (Arnold et al. 2017; Burns and Hawkins 2019; Vernalllis et al. 2019). I have learned from all of these sources, even when I do not cite them directly in what follows.

All the music videos that I discuss in this book start out from familiar positions in contemporary Anglo-American popular culture. But the directions in which these videos then move, and the aesthetic, emotional, and political perspectives that they propose, are singular inventions. I discuss videos by musical artists who have had massive popular success (The Weeknd), ones who have gained considerable critical respect but less sales (Massive Attack, FKA twigs, Moses Sumney), ones who have established something of a niche following (Dawn Richard, Bonobo), and ones who are just at the start of their careers (Tierra Whack, Tkay Maidza). In all cases, though, these artists' music

videos (sometimes self-directed, but more often made in collaboration with other directors) add new dimensions to the music, expanding its effects, rather than just illustrating it.

Deleuze's two regimes of audiovisual expression

In order to think about the rhythms of music videos, and the relations between sounds and images that they explore, I draw inspiration from Gilles Deleuze's account of cinema in his volumes *Cinema 1: The Movement-Image* (Deleuze 1986) and *Cinema 2: The Time-Image* (Deleuze 1989). Deleuze discusses the ways that time and duration are produced and expressed in the movies, and how this results in the creation of audiovisual worlds. Deleuze distinguishes between the *movement-image*, in which time is expressed through movement and action, and the *time-image*, in which time is freed from movement and subsists "in its pure state" (Deleuze 1989, 17).

Deleuze's two "images" of cinema are both ideal conceptual types. But they also correspond to roughly historical distinctions. The movement-image is characteristic of classical silent and sound cinema, whereas the time-image is grounded in the modernist cinema that emerged after the Second World War (Italian neorealism in the late 1940s, the French New Wave in the late 1950s and early 1960s, the New Hollywood in the early 1970s). Deleuze wrote and published his *Cinema* volumes in the 1980s; so he barely touches on the digital technologies that were only starting to be developed in that decade, and that have radically transformed audiovisual media—in ways he could not foresee—in the years since his death in 1995.

Deleuze says that his overall aim, in the two *Cinema* volumes, is "to isolate certain cinematographic concepts" (Deleuze 1986, ix). Though these concepts seem to reside in images, they have to do, most fundamentally, with *duration*. Even an entirely motionless image in a movie lasts for a particular amount of time; this is part of what differentiates cinema from still photography. I can look at a photographic image for as short or as long a time as I wish; the photograph itself has no intrinsic duration. But as Justin Remes argues, though movies can do entirely without motion, they cannot exist without time (Remes 2015, passim). Maurizio Lazzarato similarly maintains that audiovisual technologies do not unfold *within* the passage of time, so much as they "accumulate and produce duration and time" in the first place (Lazzarato 2019, 73).

Deleuze's typology of cinema—despite his wildly and beautifully poetic language—is actually far more systematic, and even schematic, than it might appear to be at first glance. In the two *Cinema* volumes, he carefully delineates a series of systematic contrasts between the classical movement-image and the modernist time-image. Classical cinema, like classical philosophy, subordinates time to movement, and imagines time only in the linear, serial form that Deleuze elsewhere calls *Chronos* (Deleuze 1990, 77). It is Aristotle who first argues that time can only be conceived as the measure of movement: "when the state of our minds does not change at all, or we have not noticed its changing, we do not think that time has elapsed" (Aristotle 1995, 1:371). The idea that time cannot be experienced, or even noticed, independently of change or movement persists in Western thought for two thousand years thereafter. The corresponding priority of movement in classical cinema is reflected both in the characters' moment-to-moment behavior, and in their progression through the narrative as a whole. Relations in time and space are established through montage, continuity editing, rational (motivated) cuts, and the actors' presentation of character types by means of easily legible bodily motions and facial expressions. Classical cinema unfolds in a Euclidean space, Deleuze says, and its forms mimic the operations of sensory-motor circuits in the brain (Deleuze 1986, 155–9).

All this changes in modernist or postclassical cinema. Or more accurately, while the structures and techniques of the classical movement-image are never abandoned by filmmakers, they are joined, and sometimes replaced, by procedures of a radically different sort. We will never be entirely done with the movement-image, because our sensory-motor circuits remain essential to all mental functioning. But other patterns may arise as well. Yasujiro Ozu's movies are punctuated by *pillow shots*: "seemingly random shots, held for several seconds, of everyday life: of seas and mountains, boats and train tracks, public buildings and private rooms," that have no narrative function of their own, but are interpolated between scenes showing characters and action (Singer 2016). Italian Neorealism leads us to moments when the protagonist is unable to respond to the situation that he or she encounters. In Roberto Rossellini's films with Ingrid Bergman, such as *Europa 51* (1952) and *Journey to Italy* (1954), Bergman's character "discovers something unbearable, beyond the limit of what she can personally bear" (Deleuze 1989, 2). Her paralysis interrupts the time of movement and action, so that we are given over to "time in its pure state," or what Deleuze elsewhere calls *Aion* (Deleuze 1990, 162–8). Italian Neorealism thereby gives us "a cinema of the seer and no longer of the agent" (Deleuze 1989, 2).

This breaking forth of unbearable images also defines the political sensibility of time-image cinema. In *Europa 51*, Bergman's character witnesses the sufferings of the poor in slums, jails, factories, and other such spaces of confinement. This experience transforms her, so that she is no longer capable of worldly action. As a witness, she is compelled to "pass through every state of an internal vision, affliction, compassion, love, happiness, acceptance, extending to the psychiatric hospital where she is locked up" (Deleuze 1989, 2). Her breakdown, culminating in this imprisonment, is also a way of freeing herself from her social role as a wealthy housewife, and more generally from the entire bourgeois-patriarchal order. Rossellini forces us to contemplate both sides of this paradox. The cinema of the time-image gives us harsh visions of social oppression. But its exaltation of time unhinged, or "time in its pure state," is the liberating underside of this oppression (Deleuze 1989, 17).

In Pier Paolo Pasolini's *Teorema* (1968), made some fifteen years after Rossellini's film, the arrival of Terence Stamp's mysterious Visitor similarly awakens all the members of a bourgeois household (father, mother, son, daughter, and servant) from the torpor of their self-imposed confinement. But they are all abandoned to the extremities of existential despair when the Visitor just as mysteriously withdraws. The ruptures of the time-image adumbrate an image of freedom, even though we are then returned to the social reality of capitalist exploitation. While one might describe both *Europa 51* and *Teorema* as movies that negate the principles of post-War capitalist society, it is more accurate (as well as more Deleuzian) to say that they both illuminate a liberating potentiality, indeed a *positivity*, that is implicit within the otherwise oppressive reigning social order. Both movies offer us moments of illumination, against the background of continuing subordination. These moments stand out from the background that immediately re-absorbs them; I am thinking here of Deleuze's lovely description of how lightning "distinguishes itself from the black sky but must also trail it behind, as though it were distinguishing itself from that which does not distinguish itself from it" (Deleuze 1994, 28).

When time thus emerges in its own right, the sensory-motor circuits that underlie action cinema are interrupted, and break down. Instead, we are exposed (along with the characters in the movie) to "pure optical and sound situations" (Deleuze 1989, 9) that cannot be resolved through action. Narrative and character development are suspended. They become open and indeterminate, in the expanded perspective of deep time. In Orson Welles's films, a new articulation of time breaks out from the old, familiar one. Some sequences in

Citizen Kane (1941) still give us the classical passage of time through montage, as in the breakfast-table sequence with Kane and his first wife, which tells us the entire sad story of their marriage and ever-increasing estrangement. But other sequences in *Citizen Kane* convey duration directly, through long takes in deep focus, usually combined with unusual camera placement. Consider the scene in which Kane is reproached by Leland after losing the election, with the camera located almost at floor level. Or even better, consider the scene of Susan's attempted suicide, with the empty glass and medicine bottle right in front of the camera, and the door to the room, through which Kane enters, all the way in the back, in forced perspective. Throughout *Citizen Kane*, Deleuze says, "the images in depth express regions of past as such, each with its own accents or potentials" (Deleuze 1989, 106).

Where continuity editing was a mainstay of the movement-image, it gets disrupted in the cinema of the time-image. Instead of smooth editing that confines action within a homogeneous space, we get irrational cuts, resulting in "false continuity" and "aberrant movement." The jump cuts and deliberate continuity violations in Godard's *Breathless* (1960) are one good example of this (Deleuze 1989, 9). Camera movements also become autonomous and nonfunctional. Think of that moment in Antonioni's *Red Desert* (1964), evoked by Pasolini, in which the camera drifts away from the action, indulging instead in "an insane pan from the bottom up along an electric blue stripe on the whitewashed wall of the warehouse." Antonioni thus "allow[s] himself the greatest poetic freedom, a freedom which approaches—and for this it is intoxicating—the arbitrary" (Pasolini 2005, 178). The movie thus transforms the boredom and neurotic alienation of its protagonist into an unmotivated experience of pure duration.

Deleuze, following Bergson, privileges time over space. But in the time-image, just as time is emptied of action in order to emerge "in its pure state," so space becomes indeterminate as well. Think of the "emptied spaces" in the famous final minutes of Antonioni's *L'Eclisse* (1962), where we see a sequence of shots from the location of the lovers' planned rendezvous, even though neither of them ever shows up. This series of empty spaces, Deleuze says, "might be seen as having absorbed characters and actions, retaining only a geophysical description, an abstract inventory of them" (Deleuze 1989, 5). More generally, in films of the time-image, Euclidean space gives way to Riemanian space, "when the connecting of parts is not predetermined but can take place in many ways: it is a space which is disconnected, purely optical, sound or even tactile." For instance, space becomes tactile or haptic, constructed through gesture rather

than given to us as a whole, in the scene in Bresson's *Pickpocket* (1959) where the camera follows the movements of stolen watches and wallets as the pickpockets transfer them from hand to hand (Deleuze 1989, 129).

In Deleuze's schema, the emergence of the time-image in modernist cinema is correlated with the liberation of time from movement in modern Western philosophy. This liberation is first accomplished by Kant, and then pushed further by Bergson. Kant defines time as "the form of inner sense" (Kant 1998, 163 and passim). This means that time can no longer be understood as the measure of motion, as it was for Aristotle, nor even as an indifferent, external container for events, as it was for Newton. Instead, time "increasingly appears for itself" in the form of duration (Deleuze 1989, xi). This is why we say that rhythms *produce* time, rather than merely taking place *in* time. Kant is the first to associate time with interiority; Bergson radicalizes this, by demonstrating, in Deleuze's words, that "time is not internal to us," so much as "we are internal to time" (Deleuze 1989, 82).

Deleuze's philosophy as a whole—and not only in the two *Cinema* volumes—is largely focused upon time and duration. Being is not substantive for Deleuze, but consists in processes of becoming. Nothing in the world is given all at once; everything is in process. Things continually change. This does not mean that there is no stability or identity. Rather, stability and identity must themselves be produced, and reproduced, through ongoing processes. For Deleuze, as for Whitehead according to Isabelle Stengers, the *endurance* of an entity through time "is an accomplishment and an achievement," and should not be taken for granted (Stengers 2011, 156). In particular, living entities can only maintain themselves by means of unceasing far-from-equilibrium flows and dissipations of energy. When these flows cease, the entity dies. It is only by continually renewing itself that an entity can persist in being—what Spinoza calls *conatus* (Deleuze 1988c, 21 and passim). "Continuity and heterogeneity" are therefore necessarily intertwined, as "the two fundamental characteristics of duration" (Deleuze 1988a, 37).

One shortcoming of Deleuze's theorization of cinema is that he privileges the ocularcentric term *image*, and thereby underplays the degree to which film is also "a sound art," as Michel Chion rightly insists (Chion 2009). This is all the more perplexing when we consider how intimately sound requires duration. You can make a still of an image, but you cannot make a still of a sound. Deleuze does not ignore cinema sound entirely; he attends to it in many discussions of particular directors and films. But even when he does so, Deleuze still refers

to sonic matter, on the model of visual matter, as a kind of *image*. It would be better and more accurate to speak, not of different types of images, but rather of different *regimes of audiovisual expression*.

In any case, Deleuze argues that the emergence of "time in its pure state" also gives rise to what he calls the "autonomous sound-image" (Deleuze 1989, 262). Just as time is liberated from movement, so sound is freed from its subordination to naturalistic speech acts. Deleuze unfortunately says very little about musical soundtracks and other non-diegetic uses of sound. But he does insist that, when time is disengaged from narrative in order to stand out in and by itself, so "the sound must itself become image instead of being a component of the visual image" (Deleuze 1989, 278). In classical cinema, as Michel Chion describes it, the soundtrack is supplemental, giving *added value* to the images while making it seem as if the images contained this value by themselves (Chion 1994, 5). But in the regime of the time-image, the soundtrack separates itself from the image track. In the 1970s films of Marguerite Duras, for instance, we are given "two 'heautonomous' images, one visual and one sound, with a fault, an interstice, an irrational cut between them" (Deleuze 1989, 251ff). Deleuze takes the idea of *heautonomy* from Kant's *Critique of Judgment* (Kant 2000, 28), where it refers to components that are not entirely autonomous from one another, but also not systematically determined by one another. As David Rodowick explains the term, "image and sound are distinct and incommensurable yet complementary" (Rodowick 1997, 145). The principle of the "irrational cut" applies not just to the heterogeneous visual assemblages of modernist cinema, but equally to the ways that this cinema creates a "complex link between the two heterogeneous, non-corresponding disparate [realms of vision and sound]: this new intertwining, a specific relinkage" (Deleuze 1989, 253).

Toward a Third, Digital Regime of Audiovisual Expression

Deleuze does not say much about the rise of digital media. This is not only because he was writing before digital technologies became anywhere near as ubiquitous as they are today. More crucially, digitization *per se* is barely thinkable for Deleuze. Duration, or pure time, is a continuum, which is to say that it is analog. It "cannot be divided without changing in nature at each division" (Deleuze and Guattari 1987, 483). To digitize a continuum, sampling it at regular intervals in order to extract discrete bits or pixels, is therefore to

destroy it. For Deleuze, as for Bergson, the *spatialization of time*—its reduction to homogeneous quantity—is the original sin of Western science and technology. As far as Deleuze is concerned, therefore, we cannot get beyond the time-image, which already frees time to its maximal extent. This is why Alexander Galloway presents Deleuze as "a philosopher of the analog paradigm alone … the analogical philosopher par excellence" (Galloway 2014a, 104).

Nonetheless, it seems increasingly obvious that we cannot understand audiovisual media today without taking digitization into account. Digital tools can readily replicate older analog forms; it has never been easier and cheaper to make either a movement-image film or a time-image film. But these new tools also provide us with new affordances, which allow new media techniques and forms to be invented. We are misled when we think of new technologies as mere extensions of older ones. Marshall McLuhan reminds us that the automobile was originally called the "horseless carriage," and that radio was originally called the "wireless" (telegraph or telephone). But both of these inventions quickly pushed into quite different dimensions than those of the older media that they replaced (McLuhan 1964, 292). The same is true for more recent media forms. When MTV first broadcast music videos on cable television in 1981, one of their advertising taglines was that the service combined "the best of TV with the best of radio" (Speedster 2020). Today, of course, it is evident that music videos are no more a form of radio, or even of broadcast television, than automobiles are a form of the horse and buggy.

Given these developments, a number of theorists have attempted to extend Deleuze's typology of cinema by extrapolating a third type of image—or a third audiovisual regime—beyond the two that he himself describes. Patricia Pisters proposes the *neuro-image* (Pisters 2012); Sergi Sánchez proposes the *no-time-image* (Sánchez 2013); Nick Davis proposes the *desiring-image* (Davis 2013); Cesare Casarino proposes the *life-image* (Casarino 2011); Alexander Galloway proposes the *space-image* (Galloway 2014a, 65; Galloway and Salemy 2013); Elie During proposes the *volume-image* (During 2010); and Steen Ledet Christiansen proposes the *morph-image* (Christiansen 2019). Though I am indebted to all of these formulations, I will adopt none of them, but instead add to the list by proposing what I call the *rhythm image*.

Despite his anti-digital bias, Deleuze does in fact start to consider the radical changes wrought by digitalization in some of his last writings, and especially in his essay "Postscript on Control Societies" (written and published in French in 1990, and translated into English in Deleuze 1995, 177–82). Galloway goes so

far as to suggest that this essay might well be "Deleuze's most lasting legacy." For it proposes "a new image of society and the self that can not simply be reduced to Deleuze's previous tropes like the body without organs, the rhizome, or even the virtual" (Galloway 2014a, 96 and 108). Rather, through an engagement with the work of Michel Foucault—following on from his previous book on Foucault (Deleuze 1988b)—Deleuze begins to speculate, in strikingly new ways, on an emerging third regime both of social organization and of aesthetics.

In much of his work, Foucault traces the transformation of modern Europe from a sovereign society to a disciplinary society. The sovereign society is ruled by a king or despot, and it is largely grounded in prohibitions and public spectacles of power. The disciplinary society is ostensibly democratic, and it is ruled by bureaucratic regulations concerning every aspect of private life. The sovereign society tells you what you *must not* do, and it punishes you spectacularly for violating these prohibitions. The disciplinary society is more concerned with what you *must* do, and it pushes you in all sorts of subtle ways to get you to do it. In Deleuze's words, sovereignty is about "condemning to death," whereas discipline is about "ordering life" (Deleuze 1995, 177). Foucault illustrates this contrast by opening his book *Discipline and Punish* with two descriptions of the treatment of criminals. One is a spectacle: the public torture and execution of a prisoner in 1757. The other is a set of regulations: the minute-by-minute organization of the lives of prisoners, kept away from the public eye, just eighty years later (Foucault 1977, 3–7). Foucault is concerned with the social transformations that led from the former sort of practice to the latter. He shows how the disciplinary society that emerges in the nineteenth century is organized around institutions of confinement—not only prisons, but also factories, schools, military barracks, hospitals, insane asylums, and so on—within which bodies are constrained, and behavior is regimented, on a micro-level.

Deleuze regards the shift from sovereignty to discipline in Foucault's account of European society as congruent with the shift, in his own account of cinema, from the movement-image to the time-image. The emergence of time in itself, first envisioned by Kant, accompanies the disciplinary society's invention of the rigidly determined scheduling of behavior, played out in spaces of confinement, along with its massive expansion of surveillance as figured in the model of the Panopticon (Foucault 1977, 200ff). Deleuze underlines this parallelism, in his "Postscript on Control Societies," by once again citing the nervous breakdown of Ingrid Bergman's character in *Europa 51* at the sight of her husband's factory, which she sees as a place of confinement not much different from a prison

(Deleuze 1995, 177). The disciplinary confinement of populations—their paralysis in space—coincides with the emergence of "pure optical and sound situations" that solicit contemplation, but exceed our powers of action (Deleuze 1989, 9 and passim).

Deleuze extends Foucault's schema, however, by discerning the beginnings of a new, third social formation: the *control society*. In the late twentieth century, Deleuze says in his "Postscript," the "sites of confinement" central to the disciplinary society begin to break down. For instance, mental patients are released from hospitals, and consigned instead to "community psychiatry, day hospitals, and home care" (Delueze 1995, 178). Similarly, workers are no longer assured factory jobs for a lifetime; instead, they are compelled to demonstrate "flexibility" by continually "re-inventing" themselves, in order to get precarious short-term positions. Such changes seem to promise release from confinement, and new degrees of autonomy and self-realization, as neoliberal propagandists incessantly proclaim (Boltanski and Chiapello 2018, passim). But Deleuze argues that these seeming "new freedoms" actually involve new "mechanisms of control as rigorous as the harshest confinement" (Deleuze 1995, 178).

Deleuze schematizes the differences between these two sorts of social regimes. The long-term confinement of the disciplinary society gives way to the continual variations and short-term placements of the control society. Where the logic of the disciplinary society is "analogical," that of the control society is "digital." Where the disciplinary society seeks to concentrate and centralize everything by establishing strong vertical hierarchies, the control society is networked, horizontal, and "dispersive." Where the disciplinary society addresses both specific individuals and large populations, the control society breaks people down into "dividuals," tranches of data that can be aggregated in multiple ways. Where capitalism in the Fordist disciplinary society is "directed towards production," capitalism in the control society is "directed toward metaproduction … What it seeks to sell is services, and what it seeks to buy, activities. It's a capitalism no longer directed toward production but toward products, that is, toward sales or markets" (Deleuze 1995, 181). As Deleuze sums it up in one of his densest formulations,

> confinements are *molds*, different moldings, while controls are a *modulation*, like a self-transmuting molding continually changing from one moment to the next, or like a sieve whose mesh varies from one point to another …. Control is short-term and rapidly shifting, but at the same time continuous and

> unbounded, whereas discipline was long-term, infinite, and discontinuous. A man is no longer a man confined but a man in debt.
>
> (Deleuze 1995, 178–9)

Although Deleuze dwells mostly on the horrors of the new control society, he also hopes that at least it "will not prove worse than its two previous forms" (Deleuze 1988b, 132). His point is that, when these social regimes change, it is not really a question of better or worse; rather, "there's a conflict in each between the ways they free and enslave us" (Deleuze 1995, 178). Every form of oppression also generates its own corresponding potentialities, or "lines of flight" (Deleuze and Guattari 1987, 222ff and passim). The time-image is both a symptom of the disciplinary society, and a possible way out from that society. The question Deleuze leaves us with, but does not answer, is this: What sort of media form, or audiovisual regime, might stand in a similar relation to the digital formations of the control society as the time-image does to the institutions of the disciplinary society?

New Audiovisual Media

A number of media theorists, both pre- and post-Deleuzian, give us clues toward theorizing a new digital audiovisual regime. While the disciplinary society seems to involve the unleashing of deep time—a time that is, in the words that Deleuze likes to quote from *Hamlet*, "out of joint" (Deleuze 1989, 41)—the control society seems to involve the inverse movement, one that is summarized in William Burroughs's call for us to move "out of Time, and into Space" (Burroughs 1987, 252). This inversion is especially worth noting, given that Deleuze takes the very idea of *control* from Burroughs (Deleuze 1995, 178). Many theorists follow Burroughs in seeing new articulations of space as crucial to the globalized capitalism of the late twentieth and early twenty-first centuries. Fredric Jameson observes that, whereas high modernism was obsessed with time, late capitalist or postmodern culture is "increasingly dominated by space and spatial logic" (Jameson 1991, 25). More specifically, Galloway tells us that "the computer is a space-based medium," for which "time is just a variable like any other" (Galloway and Salemy 2013). Lev Manovich similarly notes that computer technology "mak[es] time just one dimension among a number of others"; for Manovich, the aesthetic of the database, will all combinations

simultaneously available, replaces that of the narrative extending linearly through time (Manovich 2001, 147). All these thinkers have contributed to a "spatial turn" in critical theory and the humanities over the past half-century (Soja 1989).

David Rodowick gives us another clue to the new audiovisual regime when he notes that "through digital screens our relation is not to an image, but to function or force." Rodowick means this as a pejorative comment. He rejects digital media because, he claims, they annihilate duration, replacing it "with another conception of time … the time of calculation or computer cycles" (Rodowick 2007, 104). But this position ignores the ways that—as Beatrice Fazi puts it—even within computer science itself "the basic algorithmic model of computation has been challenged by new conceptualizations of the computing machine." Fazi argues for "the contingency of computation" as it interacts with "the unpredictability of the environmental input, so as to bring the indeterminacy of the real world into the computational system" (Fazi 2018, 3). But Rodowick only sees one side of the picture; he ignores the full potentiality of the "function or force" that he attributes to digital media and screens. The control society is indeed grounded in the instrumentalized, bureaucratic "management of information" that Rodowick decries (Rodowick 2007, 141). But when image becomes force, something else also happens. Audiovisual media are no longer experienced in terms of identification and recognition—crucial terms in the modernist film theory championed by Rodowick—but rather in terms of pulsations and rhythms.

This leads to an important philosophical distinction. Bergson describes duration as having "the uninterrupted continuity of melody" (Bergson 1946, 173). But today, we would probably do better to think in terms of a different musical metaphor. Contemporary genres of popular music like hip-hop and EDM (electronic dance music) tend to be organized around "rhythmic and timbral intensity" more than around melody and harmony (James 2015, 33). Rhythms are at the center: as they cross one another in syncopation, they express multiple temporalities at once, encompassing both stability and change. The Marxist musicologist Mark Abel traces the steady meter of modern groove music to the Taylorist rhythms of the assembly line; he agrees with both Adorno and Deleuze that such music "uses as its raw material the abstract, alienated time of capitalism rather than an ideal liberated time." But Abel also insists on something that neither Adorno nor Deleuze is able to

recognize: "polyrhythm and syncopation … suggest alternative pulses and tempos which nonetheless remain integrated within the metrical framework"; these alternative rhythms open up the possibility of "an emancipated, collective temporality" (Abel 2014, 255–6).

What does this mean for Bergson's and Deleuze's "time in its pure state"? As early as the 1930s, Gaston Bachelard, under the influence of the Brazilian philosopher Lúcio Alberto Pinheiro dos Santos, sought to revise Bergson by defining duration in terms of rhythm rather than melody. For Bachelard, "the concept of rhythm is the fundamental concept of time … to have duration, we must entrust ourselves to rhythms" (Bachelard 2000, 20–1). Deleuze himself does not seem particularly interested in this sort of approach; instead, he privileges "the 'nonpulsed time' of a floating music" that he finds in Pierre Boulez and John Cage (Deleuze and Guattari 1987, 262 and 267). But Steve Goodman notes that rhythmic time, as championed by Bachelard, and as featured in the music released by Goodman's own Hyperdub record label, has the Deleuzian virtue of treating time *intensively*, as a force irreducible to spatial extension: "by emphasizing rhythm over melody, Bachelard is emphasizing intensity over duration, arguing in fact that duration is merely an effect of intensity" (Goodman 2009, 86).

Lev Manovich, one of the first theorists to seriously consider the unique traits of digital media, offers us yet another clue to the new audiovisual regime. Manovich describes how the modernist logic of disjunction is replaced, in digital art, by a new "aesthetics of continuity" (Manovich 2001, 135–6). In the formal terms of audiovisual media, this means that the classical cinematic practice of montage (featured in different ways in both the movement-image and the time-image) gives way to the digital technology of compositing:

> Montage aims to create visual, stylistic, semantic, and emotional dissonance between different elements. In contrast, compositing aims to blend them into a seamless whole, a single gestalt … In digital compositing, the elements are not juxtaposed but blended, with their boundaries erased rather than foregrounded.
>
> (Manovich 2001, 136 and 145)

Other theorists have made similar observations about digital technology. Galloway argues that the new digital media present us with "*a multiplicity of vision*," with images "all captured simultaneously, each from its own vantage point" (Galloway 2014a, 66). In philosophical terms, this is not far from Leibniz's *Monadology*, according to which

> The same town, when looked at from different places, appears quite different and is, as it were, multiplied in perspectives. In the same way it happens that, because of the infinite multitude of simple substances, there are just as many different universes, which are nevertheless merely perspectives of a single universe according to the different points of view of each monad.
>
> (Strickland 2014, 25)

Elsewhere, Galloway notes how a computer's "GUI creates spatial continuity through the simultaneous windowing of different spaces …. Fusing cuts within the frame replaces fusing cuts in time" (Galloway 2014b). Elie During, for his part, writes about "the co-existence of images" in digital broadcast media: the most disparate scenes, drawn from the most various times and spaces, are all able to be present on the screen at once (During 2010, passim). These sorts of formulations cannot be found in Deleuze; but they are congruent with his overall claim that the spacetimes of control are "short-term and rapidly shifting," but also "continuous and unbounded." This stands in contrast to the "long-term" but "discontinuous" spacetimes of confinement in the disciplinary society (Deleuze 1995, 181).

All these effects—simultaneity, multiplicity, combinatorial logics, and so on—point up the fact that machines do not see, hear, and feel the world in the same ways that human beings do. Digital computation has its own aesthetic and its own phenomenology, both of which are quite different from ours. One might object to this statement on the grounds that, at least as far as we know, computers are not yet sentient, and therefore do not have an aesthetic or a phenomenology at all. But I think that my extravagance here is justified, because it points up just how *alien* computers are to the anthropocentric modes of perception, thought, and experience that we take for granted.

One reason why computational experience is different from, and incommensurable with, ordinary human experience is that computational media are *procedural* and *generative*. They produce output by running algorithms repeatedly and recursively, with the results each time feeding into the next iteration. Ian Bogost defines *procedurality* as "the computer's special efficiency for formalizing the configuration and behavior of various representative elements" (Bogost 2006, 13). The human mind does not operate in this way. According to Bergson and Deleuze, human perception, or so-called natural perception, is *subtractive*: it "subtracts from the thing whatever does not interest it," creating a simplified and manageable picture (Deleuze 1986, 63). André Bazin praises the analog (chemically based) cameras of his

own time because when they are used "an image of the outside world takes shape automatically, without creative human intervention, following a strict determinism" (Bazin 2009, 7). But this is still ultimately a subtractive process; even without human selection coming into play, the camera's images give us less than the world itself does.

In contrast, computation is ultimately additive. Digital image capture starts out with a reduction, albeit in a different manner from human perception or analog photography. Digital devices *sample* various aspects of the world, taking instantaneous readings at predefined intervals. But this is only the beginning of the process. Computation subsequently works over its sampled material, through iterative processes of mutation and recombination (and sometimes interpolation as well). In this way, digital processing gives us far more information than we started out with. Computational systems accumulate "big data" in multitudinous arrays with vast numbers of parameters. These extend far beyond the ability of any human being, or group of human beings, to parse. Though claims that "the data deluge makes the scientific method obsolete" (Anderson 2008) are silly and overstated, these arguments do point to the vast difference, both in scale and in kind, between computational and human-phenomenological ways of experiencing and understanding the world.

Computational processes also work according to a vastly different temporality than that of human perception. As Mark B. N. Hansen reminds us, new digital media "operate at microtemporal scales without any necessary—let alone any direct—connection to human sense perception and conscious awareness." And yet these "microtemporal processes … despite evading the grasp of our conscious reflection and sense perception, nonetheless impact our sensory lives in significant ways" (Hansen 2015, 37–8). Digital recording and playback devices are now ubiquitously present in our daily lives, in the form of mobile phones, RFID tags, and environmental sensors. These devices interact with us on a subliminal level. They can sense our subjective responses and decisions in their incipient state, as they are just being formed, and before we ourselves become fully conscious of them. Under the control of large corporations, these media can then "nudge" our responses in ways that are quite insidious—even though certain neoliberal thinkers celebrate them (Thaler and Sunstein 2008). Hansen is fully aware of these dangers, but he also expresses the hope that such "twenty-first-century media" can instead be turned to our benefit, through the "the *feed-forward* of data into consciousness," allowing us to access "*the power of potentiality*" (Hansen 2015, 252 and 260). In our current situation of corporate

control, this is a very distant prospect; but it is very much worth cultivating as a Deleuzian line of flight.

Applying these insights about computation to audiovisual media in particular, Shane Denson looks at the ways that "post-cinematic cameras and screens" operate on the basis of "fine-tuned temporal events taking place in intervals fundamentally inaccessible to human perception." The temporality of these media "is no longer tuned to that of their human receivers." This leads to what Denson calls *discorrelation*: "a phenomenological disconnect between viewing subjects and the object-images they view" (Denson 2020, 33). In the rapid, unmotivated edits and jerky camera movements of a Michael Bay action sequence, for instance, we are *affected* by images and sounds without actually being able to track them on a conscious level; they pass by too quickly for full awareness. In the 1980s and 1990s, the use of "fast, nonlinear cuts" and other such devices in feature films and narrative television shows was often disparaged as "MTV-style editing" (Pascual 2021). But such techniques have only become more widespread in the twenty-first century, both in movies and in music videos, thanks to the continuing development of ever-more-powerful digital tools. We should also remember that discorrelation is not just a matter of high velocity. In the *Paranormal Activity* series of horror films, the human attention span is also frustrated by the slow loops and repetitive pans of surveillance cameras, the fixed gaze, complete with timecodes, of webcams, and the nonperceptual abstractions of devices like the Microsoft Kinect system. These movies never let us forget how our world is permeated by audiovisual recording and playback devices (Denson 2020; Shaviro 2017b). All of these technologies—not only depicted within these movies but also used to make them in the first place—display "the affective impact and bypassing of cognitive (as well as narrative) interest through video and computational imaging devices" (Denson 2020, 50).

All these developments effectively lead to the alteration of our sense ratios, as Marshall McLuhan foresaw half a century ago. McLuhan proclaims (even prior to digitization) that electronic media give us "a brand-new world of allatonceness" (McLuhan and Fiore 1967, 63). That is to say, electronic media—even ostensibly image-centric ones—unfold in a space that is no longer predominantly visual, but rather "audile-tactile" (McLuhan and McLuhan 1988, 33 and passim). Allatonceness is apprehended by the ear and by the finger, more than it is by the eye. Sound and touch—each in its own particular manner—are immediate, and intimate, to a far greater extent than vision can be. When sound and touch work together, "space is constituted of resonant intervals, dynamic relationships, and kinetic pressure," all of which are invisible, but directly felt (McLuhan and McLuhan 1988, 35).

Touch requires close proximity, while you have to move a distance away from an object in order to be able to see it clearly. McLuhan tells us that "tactility is the space of the significant bounding line, of pressure, and of the interval." When we explore a surface or a volume through touch, we do not observe these divisions (intervals and boundary lines) and stresses (pressure), so much as we physically feel them. We come as close to Others as possible, but without consuming or assimilating them: "when we touch something, we contact it and create an interaction with it: we don't connect with it, else the hand and the object would become one" (McLuhan and McLuhan 1988, 6).

When Deleuze writes about certain filmmakers (most notably Bresson, as I have already mentioned), and also when he writes about visual art (most notably Cézanne and Francis Bacon), he develops the idea (initially taken from the art historian Alois Riegel) of *haptic space* (Deleuze 2003, 133). As vision becomes extremely close-up, it operates more like touch: "one never sees from a distance in a space of this kind," but only locally, moving from point to point (Deleuze and Guattari 1987, 493). Laura U. Marks, following Deleuze, takes up this notion of *haptic visuality*, and applies it directly to digital audiovisual media (Marks 2000, 2002). Even sight no longer works panoptically, in the manner of what feminist film theorists call the male gaze (Mulvey 1989, 19). Instead, the eyes adapt themselves to a predominantly audile-tactile mode.

This audile-tactile media regime also gives us a new sense of process, becoming, and sensory multiplicity. This is what Steen Christiansen calls *animacy*, and what Deborah Levitt calls the *animatic apparatus*. Audiovisual media today work, as Christiansen puts it, by "articulating our liveliness as a form of mutability," so that "a positive generation of liveliness" extends beyond any traditional notions of the organic, and "disturbs the distinction between animate and inanimate" (Christiansen 2021, 85–7). Levitt similarly notes how "images possess their own forms of vitality" in contemporary audiovisual media (Levitt 2018, 15). Though both Christiansen and Levitt write most explicitly about the visual, their formulations take for granted the reordering of the sensorium that I am evoking here.

I have already noted that Deleuze says less than one might wish about sound in the movies and related media. But the deterritorializing power of sound is evident, since the ear hears from all directions, while the eye is unidirectional. As McLuhan puts it, "we hear from all directions simultaneously; acoustic space has the structure of a sphere in which things create their own space and modify and coerce each other" (McLuhan and McLuhan 1988, 35). Michel Chion similarly insists that hearing is unbounded: while the cinematic image must

appear in a frame, "there is no auditory container for sounds" (Chion 1994, 67). This is what allows movies of all sorts to make such varied, extensive, and sometimes disorienting use of both off-screen sound and non-diegetic sound. Chion also notes that, even though light waves are physically faster than sound waves, in the movies sound tends to precede sight. At such close range, we cannot discern the speed difference; but sound predominates because "the ear analyzes, processes, and synthesizes faster than the eye" (Chion 1994, 10). Sound is experientially immediate and "allatonce"; but we should not forget that sound is also not punctual—it can only unfold in a rhythmic duration, however short this may be. Sound also echoes and prolongs itself through vibration; this is why McLuhan says that "resonance is the mode of acoustic space" (McLuhan and McLuhan 1988, 6).

These new arrangements are foreshadowed in the philosophy of Alfred North Whitehead, who is strongly influenced by Bergson, but not as loyal to him as Deleuze tends to be. The logic of multiple and scattered "digital events," pressed into continuity through compositing, is similar to what Whitehead calls *actual occasions*: finite, singular processes, each with its own specific spacetime location. Every occasion has its own particular duration, but there is nothing like a wider "block of duration" that includes them all. Rather, the individual occasions are each discrete and autonomous. They are densely interrelated, however, through what Whitehead calls the *extensive continuum*. This is a web of connections and contrasts. Each occasion offers a range of potentialities to the others, but each also limits how far the others can go. For any given occasion, the other occasions provide a sort of environment that it has to come to terms with in some way. There is no unity imposed from outside or above, but there are all sorts of overlaps, cross-connections, and influences streaming among the occasions (Whitehead 1978, 51–82). In addition to Deleuze's models of linear, narrative time (Chronos) and of the suspended time of events in their purity (Aion), we have a third model of time as a multiplicity of polyrhythmic, cross-resonating vibrations.

The Rhythm Image

Pulling all these clues together, we get the outline of a third audiovisual regime, correlated both with digital technologies and with the control society. This is what I call the *rhythm image*: the subject for a third *Cinema* volume that Deleuze

never wrote. The rhythm image cannot be described in terms of Bergsonian duration; but it is also different than the *spatialization of time* denounced by both Bergson and Deleuze. The rhythm image does not subordinate time to motion, as the movement-image does; but it also does not liberate time in its pure state, as the time-image does. Rather, the rhythm image renders time as pulsation and force, and as the play of multiple intensities. This new audiovisual regime is the product of digital machines, but it also has its roots in the powerfully embodied polyrhythms of African and Afrodiasporic music. It is linked to the oppressive mechanisms of the control society, but it also offers us potential lines of flight away from those mechanisms.

I have already mentioned how, in the cinema of the time image, camera movements through three-dimensional deep-focus space also provide us with a kind of passage through time. The time-image unfolds for us, Deleuze says, with "all the circles of the past constituting so many stretched or shrunk *regions*, *strata*, and *sheets*" (Deleuze 1989, 99). Think of the camera movement in *Citizen Kane* (actually two shots, but with a hidden dissolve) where we start outside, passing through a neon sign and a rain-soaked skylight in order to descend upon Susan Alexander Kane inside her club, sitting at a table and drinking alone. When the moving camera penetrates deep space, it also passes through these many layers of time, from the present moment of the news reporter trying to interview Susan, and back through her history of failures and humiliations. This is a kind of exploration unique to the time-image.

Movies and videos of the rhythm image offer us a different, but equally unique, mode of spacetime exploration. At a crucial point in the music video director Joseph Kahn's feature film *Detention* (2011)—a film as crucial for the rhythm image as *Citizen Kane* was for the time image—the teenage protagonists are all seated in a circle in the school library. One of the students recounts how he has been stuck in detention, without aging, every day for the past nineteen years. As he explains this, the camera tracks all the way around the circle of seated students. Each time the camera accomplishes one complete 360-degree rotation, we jump in time to a few years earlier. Different students are seated in the same chairs, mostly in the same postures, but wearing different clothes in different styles, and with different music playing on the soundtrack. In this way, we jump back, at several years' intervals, from 2011 to 1993. Time is explored, not through movement in depth, but rather through a superficial or lateral camera movement, accompanied by textual overlays on the screen, and with a continuous circular motion across abrupt cuts, smoothed out through digital

compositing. Different moments of the past coexist, and even share a common space. They all occupy this space in similar ways, but with different surface details each time, as if they had been altered through a series of discrete digital transformations. All these occasions resonate without actually coinciding.

Similar explorations of space and time can be found in many music videos. We see it, for instance, in Marc Klasfeld's 2002 video for Scarface's song "On My Block" (Scarface 2002). The camera circles counter-clockwise around the urban block that Scarface describes as the heart of his neighborhood. As the song proceeds, the camera explores the past by panning through scenes of neighborhood life—pick-up basketball games, police brutality, drug deals, funerals, private homes, and crack dens—along with portraits of Scarface at different ages, from birth to childhood to adolescent sexual experiences to first attempts at rapping, and culminating with his return to the neighborhood as a star. The video goes in a 360-degree circle, ending at the same location as where it began, with a church on the corner, and sneakers slung over a power line. At the start, there is only one pair of sneakers, and the church's billboard reads: "And this too shall come to pass." At the end, there are many pairs of sneakers on the power line, and the billboard reads: "What is in the beginning will be in the end." The movement of the camera, along with the verse-chorus structure of the song, and its repeated piano-riff sample, reconciles linear progression with cyclical repetition, to give a rhythmic account of the lives of the poor Black people who live on the block, where "everything is everything."

This sort of lateral camera movement, in composited long takes—filled with multiple images, and making heavy use of CGI—is common in movies and videos of the rhythm image. Such a movement is used to explore a time that is without depth, but infinite in its extent and in the range of its variations, all of which seem to coexist in a common space. This form of temporality is indeed "short-term and rapidly shifting, but at the same time continuous and unbounded." It makes for a configuration of spacetime that is not a Bersonian-Deleuzian "block of duration," but rather what Rodowick calls "a combination of logically discrete elements completely open and available to changes in value, both perceptual and semantic" (Rodowick 2007, 172).

For yet another example of this configuration, think of Alexander Sokurov's 2002 movie *Russian Ark*, an eighty-six-minute single take in which the camera passes through hundreds of years of Russian history as it traverses the rooms

of the Winter Palace in St. Petersburg. Rodowick complains that *Russian Ark* is not really a film about time passing, because it contains "more than 30,000 'digital events,'" which were added through compositing in postproduction (Rodowick 2007, 165). As "a highly composited artifact," Rodowick says that the movie "is not spatially uniform," and "cannot be considered a single long take" (Rodowick 2007, 173). I quote Rodowick here because I think his descriptions are accurate, even though I reject his pejorative judgment of the film. Sokurov's camera movement through time is indeed of a radically different sort than those we find in time-image movies like *Citizen Kane*. In the world of the control society, time is rhythmic rather than melodic; it is not a single, unified duration, but a pluriverse composed of scattered polyrhythmic "digital events" that interact in unpredictable ways.

And finally, for one more example, consider Grant Singer's music video for Lorde's song "Green Light" (Lorde 2017). The song is about a breakup; Lorde expresses both sadness and anger about being dumped by her former boyfriend, but looks for a "green light" that will allow her to move on to something new. The song starts in a melancholy manner, with Lorde's plaintive voice backed only by sparse piano chords. In the course of the song, however, the instrumentation becomes thicker and more percussive, and also more upbeat. Lorde pivots from mournfulness in the verse, and reproach in the refrain, to hopefulness and anticipation in the pre-chorus, and then to an ecstatic release in the chorus. The second time through each of these sections of the song, the cycle is repeated more emphatically, and at a higher energy level (in terms both of Lorde's voice and the instrumentation). The music video is set mostly at night; Singer gives a distinctive look to each section of the song, not only in terms of Lorde's actions, but also in terms of lighting and color, as well as cinematography and editing. We see Lorde staring into the bathroom mirror, dancing in a club, riding in an Uber, dancing on the car's roof, and strutting and dancing down the street (see Figure 1.1).

As the song proceeds, and as Lorde gets more energetic and positive, the editing becomes quicker as well. In the latter parts of the song, shots from all the earlier sections of the video are rapidly and continually intercut with newer shots. This creates a polyrhythmic visual mix, in which the different sections of the song seem to coexist. Lorde's changing moods, and her actions, are somehow simultaneous, as well as successive. In the course of the video, the upbeat shots gradually predominate over the melancholy ones, but this feels more like a

Figure 1.1 Lorde, "Green Light." Dir. Grant Singer, 2017.

statistical redistribution than like a linear progression. The video ends with the sky getting light behind Lorde, as the dark night, lit only in garish neon, gives way to the orange, purple, and pink layers of the horizon from which the Sun is about to rise. Change does happen; Lorde will not be stuck in her misery forever. But this change is not the result of any sort of linear or even dialectical progression. Rather, joy arises through the accumulation of multiple, discontinuous pulses of energy, all resonating at different frequencies and in rapidly shifting intervals.

The structure of the music video, including the ways that its images respond to and interact with the music, can be summarized in the form of a table (see Table 1).

All in all, the song has a four-part structure, with the four-part sequence repeated twice. The first and second parts are emotionally subdued, but the third and especially the fourth parts are more upbeat. The Chorus sounds positively jubilant, even though the lyrics still imply hesitation (Lorde sings that she is still waiting for a green light, not that she has gotten one). The second iteration of the four parts is musically denser, and thereby more emotionally intense, than the first. The video responds to this musical structure in the way that it moves from relatively longer takes to much shorter ones as the song proceeds. There is also a certain thickness or lack of definition to the images, because the video was initially shot on 16-mm film (Hogan 2017). The video exemplifies what I have elsewhere called a *post-continuity* style, appropriate to "the time of microintervals and speed-of-light transformations, that are characteristic of globalized, high-tech financial capital" (Shaviro 2016), but which here endeavors to transfigure this temporality in personalized, affective terms.

Table 1: Breakdown of Lorde, "Green Light" (dir. Grant Singer, 2017).

Time	Section	Music	Lyrics	Scenes	Visuals	Colors
0:00	Verse 1	Sparse piano	I do my makeup …	Close-up on Lorde's face in mirror	Single long take; cut at very end to side view	Bluish background
0:30	Refrain	Piano and percussion	But you're not in love …	Lorde's face as she dances; background blurry	Extreme close-ups, with several jump cuts	Bluish, a few greens, strobing
0:50	Pre-Chorus	Piano and percussion, louder and more upbeat	But I hear sounds …	Lorde exits club, enters car, leans out of car	Longer shot of entering car; close-up in car; close-up of driver; background blur; shot from outside car as Lorde leans out	Naturalistic nighttime lighting; Lorde's red/purple dress
1:18	Chorus	Heavier instrumentation, choral singing	That green light …	Lorde leans out of car, then she dances and lies on top of car	Camera is further away; Lorde's upper torso or whole body; camera circles ecstatically around her or pans down the length; some longer shots showing whole scene, with driver standing at distance from car	Reddish light on Lorde's body; urban night background
1:35	Verse 2	Thicker orchestration	Sometimes I wake up …	Lorde walks down street; back to dancing on car	Close-up of Lorde's face as she walks down street, with blurry background; similar when she is back on top of car	Naturalistic lighting as she walks; back to reddish glow on top of car

(continued)

Table 1 *(continued)*

Time	Section	Music	Lyrics	Scenes	Visuals	Colors
1:48	Refrain	Percussion and synthesizers	But you're not in love …	Lorde dancing down the street; turns around pole, twirls her hair, grabs pay phone	Two medium traveling shots, one from behind Lorde, one from front and side; we see more of her body; background in focus	More naturalistic night-street lighting; Lorde's dress stands out
2:08	Pre-Chorus	Back to piano and percussion	But I hear sounds …	Lorde dancing in street, enters door, bathroom as she looks in mirror and Jack Antonoff plays piano in background	Cuts between longer shots of Lorde dancing on sidewalk and in bathroom	Bathroom has contrasting sickly green and bluish haze
2:35	Chorus	Thickest orchestration, choral singing	That green light …	Lorde dancing and running, highest energy so far	Fast montage of all the scenes seen previously: street, in car, on car, dancing in club, dancing in bathroom	Rapid alternation of all lights we've seen before: blue, reddish, green
3:05	Outro	Repeated thick sound	That green light …	Lorde dancing in multiple venues, also walking over highway bridge; from 3:47, she is motionless on the highway bridge, looking toward the camera	Continued rapid montage, often blurry, moving camera, tracking shot right-to-left over highway bridge; camera moves in on her for a few more seconds after music ends	Continued color alterations; shots over highway show more naturalistic lighting of sunrise in background

Summary

The distinctions among the three audiovisual regimes, or "images," can be summarized in this way:

Movement Image	**Time Image**	**Rhythm Image**
Sovereign society	Disciplinary society	Control society
Fixed social roles	Individuals	Dividuals
Servitude	Confinement	Debt
Time via movement	Time in its pure state	Time as pulsation/force
Sensory-motor circuits	Pure optical-sound situations	Machinic microperception
Euclidean space	Riemannian space	Audile-tactile space
Continuity editing	False continuity	Compositing
Rational cuts	Irrational cuts	Concealed cuts
Normal movement	Aberrant movement	Suspended movement
Time in montage	Durations in deep focus	Occasions in lateral pan
Chronos	Aion	Polyrhythms
Classical cinema	Modernist cinema	Music videos

I take Deleuze's typology, and my extension of it, as a useful heuristic device, rather than as some sort of definitive formulation. It is rare to find a complete example of any one type of image; earlier formations persist in the midst of later ones, and later formations are anticipated in earlier ones. A movie like Tony Scott's *Domino* (2005) includes elements of the movement-image (in its action sequences), of the time-image (in its nonlinear chronology and its isolation of privileged objects and critical moments), and of the rhythm image (in its aggressive, speed-freak editing style, which divides even the simplest actions into clashing and overlapping multiple shots, often distorted in color and angle as well). Moreover, the temporal transitions that Deleuze layers on top of one another—among social formations, philosophical arguments, and cinema formalisms—do not necessarily coincide. For instance, the transition from the sovereign society to the disciplinary society according to Foucault took place before the invention of the cinema, so it cannot literally coincide with the passage from the movement-image to the time-image. Indeed, as Jacques

Rancière observes, Deleuze sometimes uses the very same films and sequences to exemplify both the movement-image and the time-image, so that "the passage from one book to the other would not mark the passage from one age of the cinematographic image to another but the passage to another point of view on the same images" (Rancière 2006, 107–23).

Despite all these reservations, however, I still find Deleuze's classifications both useful and informative. They allow us to draw crucial distinctions (even if these are unstable, and ultimately subject to deconstruction). And most importantly, these three "images," or regimes, allow us to notice and pick out important features of audiovisual works that would be much harder to discern otherwise. In spite of all its shortcuts and ambiguities, Deleuze provides us with a powerfully generative framework for thinking about audiovisual forms and how they have changed over the course of the past century. He provides many valuable suggestions for linking formal and technical devices to larger meanings and feelings. He also provides—albeit reluctantly—a broad sense of how historical changes in film production and reception are related to, and embedded within, larger social, political, and technological developments. If film study tends to be divided between formal analysis and ideology critique, Deleuze moves along a diagonal, offering insights and methods of approach that are reducible to neither of these approaches, but relevant to both of them.

In this book, I try to avoid over-generalizing my claims about the "rhythm image" as a third audiovisual regime in the history of cinema. My approach is by design tentative and experimental. Whatever can be said about twenty-first-century cinema, made with digital tools in general, I try to restrict my focus to characteristics to the particular digital music videos, produced in the context of the largely Anglophone music industry, that I analyze and discuss in depth. I am ultimately more interested in singular expressions and effects than I am in larger trends, even though I remain aware that context always matters. Obviously all the works I discuss in this book are defined to a great extent by the musical and cinematic conventions that they assume, refer to, and (to a large extent) adhere to. But I still seek to search for the points at which these works deviate from such contexts and frameworks, in order to express novel feelings and ideas.

In other words, my concern throughout this book is less to give a specifically Deleuzian account of music video as a genre or form, than to understand how music videos function as particularly dense nodes of expression for a series of transformations in contemporary social and cultural life. These transformations include:

1. aesthetic changes, or changes in the sensorium, having to do with our experience of relations of space and time, of sound and image, and of attention and distraction;
2. social and cultural changes having to do with the ways that things like race, gender, sexuality, and relative affluence or impoverishment are both lived in individual experience, and structured transpersonally in ways that are simultaneously constraining and liberating;
3. technological changes (in such things as screens and devices, data gathering and, digital processing) having to do with new modes of media production and distribution, and with new habits of personal consumption;
4. affective changes having to do with the ways that people (and particularly younger people) negotiate many of the issues that have traditionally been central to pop music (sex and romance, money and fame, questions about personal identity, and so on).

These are all, of course, issues that have been extensively discussed in recent years by philosophers, social scientists, media theorists, and scholars of popular culture. The driving idea behind my book is that we can *also* approach all these issues *intensively* (rather than *extensively*), in terms of the ways that they are compressed and enfolded in singular cultural productions like the videos I discuss. Most broadly, I seek to consider the ways that popular music culture in the Anglophone world today negotiates and explores the spacetime (and especially the temporal structures) of our contemporary world, and how it finds ways to express, within this world, new sorts of feelings, desires, and frustrations.