The Perception Machine

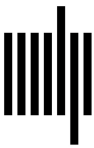
Our Photographic Future between the Eye and AI

By: Joanna Zylinska

Citation:

The Perception Machine: Our Photographic Future between the Eye and AI
By: Joanna Zylinska
DOI: 10.7551/mitpress/14471.001.0001
ISBN (electronic): 9780262376631
Publisher: The MIT Press
Published: 2023

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open



The MIT Press

6 Can You Photograph the Future?

Figure 6.1
Screenshot from *Devs*, 2020. Dir. Alex Garland. FX on Hulu.

Ten seconds into the future

A small group of people are sitting around a table in what seems like a cross between a corporate meeting and university lab seminar, looking at screens big and small. Their attention is on a moving image of a worm, overlaid with the graphic visualization of its movements. The two moving lines converge nicely (figure 6.1). The group are looking ten seconds into the future—the worm's future, that is. This is a restaging of the WormBot project,[1] a digital simulation of the behavior of a nematode worm represented in the form of time-lapse images with a view to predicting the worm's death, and it is happening as part of the first episode of a TV series *Devs*, created and directed by Alex Garland (2020, FX on Hulu). The idea is to recreate a living

organism in a computer system by rendering all of its neural connections in a digital form. After ten seconds the correlation between the simulated screen movement and its predicted graphic visualization is lost, but the professor-lookalike boss is impressed enough to elevate the young Russian AI programmer to the next phase of his company's project: the eponymous Devs. The project lab, hosted in a vacuum-protected Faraday cage in the middle of a forest, contains a hugely powerful quantum computer whose goal is to achieve for the billions of neurons in the human brain what the WormBot-lookalike did for nematode's 302. The (mad) professor-boss, a *deus*-like character called Forest, has built his whole enterprise to generate an alternative version of reality online, one in which his wife and young daughter do not die in a car accident (inadvertently caused by him). In other words, Forest wants to overcome the causality of the physical world by creating one in which a different future can unfold—a future he will be able to inhabit together with his digitally restored family. The *Devs* series deals with the age-old problem of determinism versus free will, packaging it as a fairly conventional story of a heterosexual white nuclear family (and its multiethnic romantic counterparts) experiencing life's travails.

What is of particular interest to me about *Devs* in the context of this book is its visualization of the problem of predicting the future by capturing an image of it. The first step involves the mobilization of the power of the quantum computer to render detailed images from a remote past. Borrowing from the past's "multiverse" simulations allows the Devs developers to obtain a rich enough dataset to generate adequate images of the past: Jesus on the cross or prehistorical people making cave art. The next step will involve using the technology and data from Forest's more recent past to render a new version of the future for him and his kin. I would like to propose we bracket for now the determinism debate and the rather implausible quantum frippery of the series' plot to focus on the very idea of rendering the future as an image. For me, *Devs* serves as a framing device for the key question that will drive this chapter: Can you photograph the future?

To explore this question, I will look at the temporality of the photographic image while engaging with the latest developments in CGI, computational photography, and machine vision—technologies that are premised on forecasting an image and then making it look real. Drawing on neuroscience research on the role of prediction in human consciousness, and the application of machine learning to developing machine consciousness,

I will then move to a brief overview of the mobilization of photography and photo-rendering in predictive technology, in areas such as weather forecasts, stock markets, epidemiology, and consumer behavior. I will also consider the sociopolitical consequences of the wide implementation of predictive technology in our lives. Last but not least—and in line with the overall ambitions of this book—I will explore the possibility of imaging and imagining *better* futures.

Time is just change when there is no one to track it

The question about the possibility of photographing the future may seem both absurd and out-of-date—Marty McFly already videoed it back in 1985![2]—yet the premise behind it underpins many of the current developments in imaging technology, computation, and AI. Indeed, photography poses us a temporal conundrum, on the level both of physics and common sense. As a medium, photography has typically been associated with the past: it has been constituted as a quintessential medium of memory, of time gone by, or even, as discussed in chapter 1, of death. Logically, it is only "the present" that can ever be photographed, but the present itself is an illusive concept, unable to be captured in an image or thought about without simultaneously receding to the past. Photographic technology has always played with logic and physics to create a spectral ambience around its practices. We can think here of long-exposure photography, where several hours were required to obtain an image trace on the not-so-sensitive substrate in the early days of the medium, or about stop-motion imagining, intended to explicitly show the passage of time, be it over the period of one day or several years. There are also photographic dispatches from a prehuman past, such as the visual rendering of the data captured by a radio telescope, in June 2005, of solar dust cloud radiation in the Taurus Molecular Cloud, where the data represented an event that "took place in 1585, or thereabouts,"[3] or even a "baby picture" of our universe, produced by NASA between 1990 and 1992 and showing the map of the cosmic microwave background radiation supposedly released as heat during the Big Bang.[4] The temporal determination of a photograph depends on the point of reference: specifically, on the positioning of the observer within the system. Time is what it (supposedly) is only for an observer whose life is grasped as a timeline; outside the (human) observer there is just change.

In its computational guise, photography can therefore be said to have warped the arrow of time, displacing temporality as we know and sense it with our bodies and minds with the machinic logic of multiplication and nonlinear choice. Reflecting on Google's Pixel 3 smartphone and its sophisticated camera, Sy Taffel explains that a photograph is now an extraction of an image from a sequence, coupled with an averaging of that sequence. Or, as developer Vasily Zubarev puts it, "When you tap the 'take a photo' button, the photo has actually already been taken."[5] Taffel clarifies further that, unlike in traditional cameras,

> where the photographer pressing the shutter-release triggers the recording process, hitting the record button signals the mid-point within a 15-frame image stream the camera sends for computational processing. The significance is not just that the camera composites these images, the practice of taking a photograph is transformed from a human operator capturing a moment of choice, to delineating a sequence of images to begin working backwards and forwards in time from. Using data from those 15 frames, the Pixel 3 generates a composite image containing HDR, greater resolution and lower levels of noise than would be possible with a single exposure.[6]

In the computational images of the world as deployed by the latest smartphone cameras, such as Google's Pixel or Apple's iPhone, the future literally influences the past.

Even though it reaches beyond human scales, photography gains a symbolic value in our human-centered universe as a practice of stopping time—but also, perhaps more strongly, a practice of making time. Photography allows us to make cuts in the optic flow unfolding around us. Yet it is not so much the already well-studied relationship between photography and time that interests me here but rather, and more specifically, the relationship between photography and the future. I want to approach it by looking at the problem of prediction—and the mobilization of photography and other forms of imaging in constructing predictions in different disciplines—from cognitive psychology and neuroscience through to AI, machine learning, and computer vision. I am mindful here of Matthew Cobb's suggestion in *The Idea of the Brain* that our scientific horizon at any given time depends on the technological metaphors we use, but also that those very metaphors reflect the technical knowledge of the time, its discourses and vocabularies. "Metaphors can flow both ways," says Cobb.[7] As discussed in chapter 2, since the medium's invention in the early nineteenth century photography

has been intrinsically linked with neuroscience, both enabling the neuroscientific imaging practice and shaping the understanding of the working of neurological processes, from memory through to learning. Indeed, it was in terms of "imprinting" on the brain's tissue and its neural extensions that early research in neurology explained perception and action in humans.[8] This perhaps goes some way toward explaining the key role of the concept of "the mental image" in cognitive psychology and neuroscience. Yet since the middle of the twentieth century the photographic discourse has been supplemented by concepts and metaphors from computing, describing the working of the neural system with reference to "feedback loops, information, codes and computation."[9]

Interestingly, Cobb points out that, even though we now "stand on the brink of understanding how patterns of activity in networks of neurons create perception," some scientists "sense we are approaching an impasse in how we understand the brain."[10] The computer metaphor in particular is said to have reached its limit, with the proliferation of brain data being accompanied by frequent resignation on the part of its human interpreters to its voluminous excess and unexplainability. It is therefore worth probing whether current changes to the photographic medium through its encounter with computation, network technologies, and AI are altering in any way the neuroscientific horizon, and, in particular, our understanding of neural processes related to imaging and perception. What *is* the mental image today—and what can it become? Can we learn to see ourselves and our world better once we have changed our metaphors? And, picking up on this bidirectional flow of metaphors, can the reframing of key concepts lead to a different understanding of photography and imaging?

We need to go deeper

The concept of the mental image, referring to quasi-perceptual experience which is prelinguistic, and which forms the basis of what we now understand as consciousness, has a long history in Western epistemology. An aspect of representationalist thinking unfolding across centuries, it originally posited that ideas were just pictures in our mind, things we saw "with our mind's eye." The concept has now lost its literal connotations without challenging the implicit primacy of visuality in our understanding of knowledge production. Indeed "'imagery' has become the generally accepted term

amongst cognitive scientists for quasi-perceptual experience in any sense mode," with the term embracing auditory, olfactory, kinesthetic, and haptic images alongside visual ones.[11] The fact that the image still constitutes a cornerstone of the neuroscientific rhetorical register delineates a particular conceptual trajectory for the discipline: it prescribes what can be seen and said, and how.

In what follows, I want to look at the use of the concept of the image in the current theories of consciousness as outlined by two of its leading researchers who are also known for writing for wider audiences: neuroscientists Antonio Damasio and Anil Seth. Damasio and Seth are physicalists (aka materialists), yet in their respective theories they go beyond single-organ functionalism. This is to say, they recognize the role of the body and its situatedness in the world in the making of consciousness, a shared assumption that differentiates them from a variety of cognitive scientists currently involved in modeling human intelligence and human vision in machines. But they are also interesting for me because of the different rhetorical registers through which they articulate their positions, coming, respectively, from image theory and information science. In the further part of the chapter I will look at the role of images in Damasio's and Seth's respective theories of perception; at the link between perception, prediction, and imaging in those theories; and, last but not least, at the possibility of *conceptualizing consciousness as a person's orientation toward the future, which involves making images of that future*. The automaticity of this process, be it in mental image-making or in the production of mechanical—or, to use Flusser's term, technical—images of the future, will allow me to take further steps in my own investigation of the possibility of photographing a future.

Readers may have assumed all along that this this idea about photographing a future is a metaphor. I would not deny this, but my understanding of the metaphor in this chapter, and in the book as a whole, positions it as an agentic entity which enacts something, rather than just being an articulation of an otherwise stable "something" in a purposefully figurative sense, for poetic or rhetorical effect. I am following here Jacques Derrida's intimation, outlined in his well-known essay, "White Mythology," that metaphors are philosophemes (i.e., concepts), that they are foundational to philosophy (even if many philosophers would deny this state of events), and that they also exceed the boundaries of any philosophical system.[12] The best we can therefore do with metaphors is "use" them with more awareness

as to their modus operandi, while giving up on the fantasy of any pure, original, uncontaminated meaning "before the metaphor." In other words, we can try to make *better* metaphors, by way of challenging the "white mythology which reassembles and reflects the culture of the West," the white man's *mythos* that passes for universal Reason.[13] An attempt to imagine photographing the future is therefore also an attempt, metaphorical and hence also ontological—as well as, inevitably, political—to figure out ways of making better futures, beyond the fixed frames of our present-day images and imaginations.

For years avoided by scientists as exclusively focused on an individual self and hence impossible to study objectively, consciousness has recently gained a new respectability as a domain of scientific inquiry, partly thanks to the encounter between neuroscience and the multidisciplinary field of artificial intelligence. (Some go so far as to argue that work on artificial intelligence should in fact be recognized as an attempt to create artificial consciousness.)[14] Defined by Damasio as "an organism's awareness of its own self and surroundings,"[15] consciousness is broadly understood as a first-person phenomenon which unfolds as part of a private, first-person process to which we give the name "mind." Damasio distinguishes between *core* consciousness, which is a sense of the here and now, and is also likely to be found in other animals, and *extended* consciousness, which spans one's lifetime, requires memory and attention, and is enhanced by language. Interestingly, it is through references to the metaphorical apparatus of image-making that Damasio explains the working of those two types of consciousness. He proposes that "core consciousness includes an inner sense based on images,"[16] and that those are images of feelings. Responding to stimuli, which are experienced by us *as* feelings, the brain makes patterns in the circuits of its neural cells. In line with the metaphoricity of the wider approach of this chapter, we could call these patterns "impressions" or "imprints."

Damasio concludes that to understand consciousness we primarily need to understand the problem of mental images—i.e., of "how the brain inside the human organism engenders the mental patterns we call, for lack of a better term, the images of an object,"[17] where that object may be material (a person, a place) or immaterial (a melody, a toothache). An image, in turn, stands for "a mental pattern in any of the sensory modalities, e.g., a sound image, a tactile image, the image of a state of well-being."[18] But

then a shift happens in Damasio's narrative, where he is getting his media purposefully yet somewhat unexpectedly mixed. He says: "the first problem of consciousness is the problem of how we get a 'movie-in-the-brain,'" relying on what he himself describes as a "rough metaphor."[19] He also highlights that "the movie has as many sensory tracks as our nervous system has sensory portals—sight, sound, taste, and olfaction, touch, inner senses."[20] Some interesting forms of mediation are activated here, with the brain positioned as the cinematic screen (which we already encountered in chapter 3, via Deleuze). Also, an odd circularity creeps into Damasio's argument as part of the mediation process, with images functioning as stand-ins for his (and everyone else's) inability to clearly explain the sequence of events in the production of consciousness. Damasio goes so far as to acknowledge that we have to explain not only how the movie-in-the-brain is generated (assuming that there is a movie, of course) but also how the brain "generates the sense that there is an owner and observer for that movie."[21] If consciousness does behave like a movie, it is like a retro silent one, as language is only believed to arise at a "later" stage.

The brain-as-screen is perhaps less like a conventional cinema screen and more like the 1950s Circarama, renamed Circle-Vision a decade later, with screens surrounding the viewer from all angles but also requiring the viewer's corporeal movement to appreciate the experience fully. The viewer only *becomes a viewer* through their 360-degree mobility in the image space. This sense of the image envelope that requires mobile attention from the viewer returns in VR and 3D gaming, with the viewer's/player's positioning resembling an M. C. Escher drawing in which the movie dreams itself up as a movie, playing in a theater with an audience. The viewer is thus both inside the image and watching the image. That sense of the self-generativity of experience from within the image envelope is being used in current experiments with machine learning and AI that involve an attempt to recreate the human experience of perception—or rather the externally perceptible aspects of that experience—in machines.

Google's DeepDream algorithm developed in 2015 allowed for the production of images in which nonexistent objects were "discovered" by running previously trained neural network models on those images multiple times.[22] This resulted in infamous images showing spaghetti with eyes, camel-birds, and psychedelic squiggles covering familiar landscapes. This aesthetic, short-lived as it was, became known as "Inceptionism," a

term derived "from the Network in network paper by Lin et al. in conjunction with the famous 'we need to go deeper' internet meme."[23] Lin et al.'s 2013 paper proposed a deep network structure called "Network In Network," which was to be "used to enhance model discriminability for local patches within the receptive field."[24] The operations of recurrent neural networks are meant to recreate the (posited) experience of human perception and human learning, producing images and then feeding them into the network to create more—and more detailed, or "deeper"—images. As we already discussed in chapter 4, the computational model of perception in AI research can be described as reductionist, as it brackets off both the human body and its movement in the environment as not so vital aspects of the perceptive process. This (il)logic is replicated in the current design of neural networks—algorithms whose name is strictly, and restrictively, metaphorical, as there is nothing "neural" about those networks. Yet what is modeled more accurately from the human organism is the sense of recurrence, or circularity, whereby the elements produced as part of the work of the machine's posited "intelligence" are themselves producing that intelligence. (As mentioned before, consciousness would be a more apposite term here, or, more accurately, behavior that presents as consciousness to a human observer.)

Bringing together all the constitutive elements, Damasio puts forward the following schema for the image-based operations of consciousness in humans:

> The wordless narrative . . . is based on neural patterns which become images, images being the same fundamental currency in which the description of the consciousness-causing object is also carried out. Most importantly, the images that constitute this narrative are incorporated in the stream of thoughts. The images in the consciousness narrative flow like shadows along with the images of the object for which they are providing an unwitting, unsolicited comment. To come back to the metaphor of movie-in-the-brain, they are within the movie. There is no external spectator.[25]

A number of things are intriguing about this schema. First is the perceived need to split the process of the emergence of consciousness into two phases: the wordless narrative of images (even though it is still called a *narrative*, its role consisting in having to *describe* the object perceived or *comment* on images of that object) and a subsequent timeline of thoughts (which are not yet words). Second is the distinction between "images in

the consciousness narrative" and "images of the object." The first kind of images function like cybernetic mirrors: they produce consciousness by providing an image of the self (or, more precisely, of the self's body) that is then being incorporated into the picture of the self. I am calling this mirror *cybernetic* because of the feedback loop effect required for its operation. Yet it is not a straightforward device that forms images by reflection. The cybernetic mirror producing consciousness by way of making images is more like the *Magic Mirror* envisaged by Escher, where the reflection becomes part of the scene being reflected. The two strands and two timelines of those images seem to overlap and merge, collapsing the odd Platonism of Damasio's theater of consciousness in which shadows appear before objects into a more contemporary version of the screening venue: a movie theater. The supposed cinematic experience is staged by "the brain" for us, an experience that is also said to generate "us" in the process.

Yet it is worth delving a little deeper into the mixed media metaphors here and realize that consciousness does not really have the designed smoothness of a VR or immersive gaming experience. Damasio himself acknowledges that "core consciousness is generated in pulselike fashion, for each content of which we are to be conscious."[26] Indeed, we flit in and out of consciousness, both during our sleep (when it is only in the relatively short time of deep sleep that our consciousness is suspended) and in daytime. To continue with the media metaphor, consciousness is therefore perhaps more like an experimental modernist short, one that is made up of still photographs, multiple exposure frames, slow-mo animations, and material scratches to the film's surface. The role of the body, foregrounded in the formal experiments of modernist filmmakers such as Sergei Eisenstein, Hans Richter, or Aleksandr Rodchenko, is crucial in the creation and, of course, perception of that movie. We could say that there is no movie without a body. As Damasio puts it, "the images you form in your mind always signal to the organism [which stands for a material bodily entity] its own engagement with the business of making images."[27] The human body therefore presents itself as a self(ie)-camera. The images are needed to generate us, but this sense of "us" emerges through our "action." We become us by imaging and hence imagining ourselves into being, we could say. The image of "the organism" that we internally create for ourselves to map out (not yet quite) our external boundaries, and hence help maintain our life,

becomes a blueprint for the imaging of the self that is subsequently seen as living that life.

Importantly, this ur-image of the organism needs to be followed by many others. Indeed, in order to know how to act, we need what Damasio refers to as "good images." He writes: "Images allow us to choose among repertoires of previously available patterns of action and optimize the delivery of the chosen action—we can, more or less deliberately, more or less automatically, review mentally the images which represent different options of action, different scenarios, different outcomes of action. We can pick and choose the most appropriate and reject the bad ones."[28] Yet it is not only through somewhat confusing spatiality in which a reflection becomes an image that Damasio conveys the working of consciousness to us. His model also entails a warped temporality. While core consciousness, as he figuratively puts it, "does not illuminate the future,"[29] extended consciousness involves "holding an image over time,"[30] which is another name for the process of memory. Drawing on the famous experiment conducted in the 1980s by Benjamin Libet on the relationship between stimulus, reaction, and consciousness, which showed a delay in our consciousness catching up with our body,[31] Damasio points out that "we are always hopelessly late for consciousness,"[32] the lateness amounting to about five hundred milliseconds. This presumed lateness has led philosophers and neuroscientists to interrogate whether our actions are indeed caused by brain activity or whether they are preprogrammed all along. In the latter scenario, consciousness only catches up with our activity five hundred milliseconds later and thus only provides a report on the activity, rather than being its source of origin. The varied responses in the science community to the relationship between consciousness, brain, and body reflect the diverse set of positions on consciousness today, which are in turn underpinned by diverse approaches to the age-old problem of determinism and free will.

Devs could be read as a parable for illustrating this problem. The orchestrator of the Devs enterprise, mad-professor-cum-entrepreneur Forest, believes there is no such thing as free will. As mentioned previously, Forest is mourning his family's death in a car crash, which was caused by his wife while she was on the phone to him—making him a contributor to the crash. Yet for Forest we all run on "tramlines." This means that his wife and daughter died in that car crash because they were always going to die in it,

with the tracks of causality extending well beyond and ahead of any individual's actions or acts of will. The belief in the determinism of the universe provides solace to his guilt but also a recipe for trying to recreate another universe in which a different set of events would be guaranteed to unfold, one that would be more satisfactory to its designer's wishes. Through a trick of smoke and mirrors involving a lot of misuse of the term "quantum," things both do and do not go to plan. A young programmer called Lily—a heroine on a quest to revenge the death of her Russian-spy boyfriend and save the world from Forest's divine machinations, and one whose heroic adventure has already been predicted by Forest's all-seeing computer—at the last moment throws out the gun she was supposed to use to kill the evil Devs-master, thus messing up the prediction algorithm. Yet both Lily and Forest do die in the altercation anyway once the cage that held them drops down, joining the rendered universe on the other side of the screen in which Forest is reunited with his family while Lily is dating her other (good) boyfriend. The philosophical dilemma as to whether they are still themselves in that rendered universe, or whether they are mere digital copies of themselves that just *look like* them to an outside observer, is closed off by the foundational assumption that drives the show (as well as much current AI research) that both perception and consciousness are computational. For *Devs* it also means that it is possible for the characters to experience themselves as "continuous and unified" not just across time (which we all do, every day, on waking up)[33] but also across platforms and media.

Commenting on the science behind the show, computer science professor Scott Aaronson is rather skeptical about the show's overuse of the term "quantum" with reference to what is complexity. Indeed, complexity is the reason computer developers from the scene discussed at the beginning of this chapter are unable to see more than ten seconds into the nematode worm's future. Aaronson then adds, acerbically: "Predicting the weather three weeks from now might be forever impossible."[34]

The perception machine as a prediction machine

A number of successful predictions do nevertheless happen everyday, albeit at smaller scales and temporal ranges, and they do have profound significance not just for how we can live our lives but also for how we can understand ourselves better. Yet their operating model is somewhat different than

claimed by Libet and taken up by *Devs*. Neuroscientist Anil Seth draws on more recent research by Aeron Schurger to provide an interpretation of Libet's experiment which suggests that the assessment of the brain's "readiness potential" to initiate a given action was itself an artifact of the process of measuring this potential, and something that could only be captured retrospectively. According to Seth, "you will see something that looks like a readiness potential if you look back in time from moments of fast responses."[35] This allows him to interpret the brain's supposed readiness as an accumulation of sensory data from which the brain can make the best guess. The brain is therefore not ready for a *specific* action; it has to create this action by envisaging the external state of events—and those events' unfolding. The act of envisaging itself is a simulation of probabilities, a prediction or "best guess," not a foretelling of what will most definitely happen. This statement supports Seth's account of consciousness as the organism's ability to make predictions. We will return to this point shortly. But, for now, I want to focus on another aspect of Seth's argument, namely, his conclusion that free will is a "perceptual experience," a statement that allows him to reconcile our inner sense of agency with the causality of the laws of physics as we know them.[36] This is a stronger claim than it might initially seem, because for Seth the only way we can access anything, internally (in ourselves) or externally (in the world), is through our experiences of it, which are clusters of perceptions. We could go so far as to say, as Seth indeed does, that our sense of self is produced through the process of future-forwarding, i.e., looking at (what becomes) ourselves, at regular intervals, while trying to guess the causes of the sensory signals being received. For Seth, Descartes's "I think, therefore I am" becomes "I predict (myself), therefore I am." This argument is encapsulated by his proposition that our brains are (Bayesian) prediction machines, attempting to guess the causes of sensory signals from the perceptual inferences they receive.[37]

I am interested in this formulation and the idea behind this concept of the prediction machine for two reasons, both related to core aspects of this book: the posited role of perception in the constitution of the self, and the figure of the machine in explaining consciousness. In Seth's framework, prediction and perception are inextricably interwoven: "The brain is continually generating predictions about sensory signals and comparing these predictions with the sensory signals that arrive at the eyes and the ears—and the nose, and the skin, and so on."[38] His post-Cartesian quip

could thus also be articulated as "I perceive (myself), therefore I am." This theory confirms the constitutive role of perception in the emergence of both our self and what we call the world. Yet, even though multisensory perception is such a strong aspect of his framework of thought, Seth is far less inclined than Damasio to resort to media metaphors, whether based on still or moving images, to explain it. Instead, his rhetorical register, with its signals, inputs, outputs, inferences, controls, and Bayesian guesses, is primarily drawn from cybernetics, information theory, and computer science. For Seth we are "beast machines—self-sustaining flesh-bags that care about their own persistence."[39] He traces his use of the concept of the machine to *L'Homme machine* (Machine Man) published in 1747 by Julien Offray de La Mettrie, explaining the workings of both the body and the mind solely in materialist terms, and thus overcoming Descartes's infamous dualism that still shapes much of metaphysics, as well as most religious beliefs. Yet Seth's model of consciousness and selfhood is not computational in the sense used in *Devs*, or applied in the machine vision research discussed in chapter 4. For him consciousness is phenomenological; it is always a consciousness of something, which means it is embodied and embedded in the world.

Echoing the sentiment well known to Marxists, visual culture theorists, and other humanities scholars for decades—namely, that "how things seem is a poor guide to how they actually are"—Seth's theory redefines perception as "controlled hallucination." The notion of systemic control is important to it, with conscious experiences being described, to use Giulio Tononi and Gerald M. Edelman's term, as "informative and integrated."[40] This sense of integration is needed for the execution of the ultimate function of perception: to enable, via a series of prediction-driven actions, our survival in the world. We could perhaps suggest that the role of perception is to enable control over (what we might call, contra Damasio) our self-image, or to go all the way down the imagistic metaphorical whirlwind, to photograph ourselves into existence. If memory, for Damasio as much as for Bergson, involves holding an image over time, selfhood—and self-consciousness, which is its foundational aspect—requires an incessant production of such images. Another way of putting it, in line with the structuring philosophical approach of this book, would be that *we photograph ourselves into being*. The model that emerges here is not that of a movie-in-the-brain, an avant-garde silent film, or even a poorly edited and choppy YouTube video, but rather that of a Polaroid print (albeit one updated for the digital age). But

we also *make the world* by *making images of it*. This is to say that not only do we photograph ourselves into being, we also make the world into what it is (for us) by making images of it, over and over again. It needs to be acknowledged, as I have several times previously in this volume, that "stuff" (aka "matter") does exist "out there" without our perception or other form of intervention, but for it to become what we see as and call the world, it needs this imagistic process to be instantiated. Needless to say, the outcome of the process could look very different for different species.

Seth explicitly rejects the idea of the brain as a kind of computer in the skull, "processing sensory information to build an inner picture of the outside world for the benefit of the self" through feature detection. Nor does he partake of the image register to explain what happens, only resorting to the metaphor of "pictures" in order to dismiss the imaging model. So the brain is not an imaging machine for him, transmitting results from features detected in the world and thus producing their mere reflections. In Seth's model, perceptions do not come directly from outside, i.e., from sensory impulses or inputs; they come from the brain's *predictions about* the causes of these sensory signals. Building on the notion of perception as inference by nineteenth-century physicist and philosopher Hermann von Helmholtz, Seth puts forward his idea of a prediction machine, which is a creative and interpretive device, not just a recording one.

Yet this is perhaps too rigid an opposition, given that prediction and imaging cannot be so easily decoupled, or that they are in fact being increasingly linked in various imaging technologies, from computational photography and CGI through to what has become known as "predtech." We are of course dealing here with an expanded understanding of imagining, and of photography in particular, one that involves a shift from seeing photography as a passive transmission of what is out there in the world through to an active creation of "the world" through the conjoined human-machinic apparatus. Even if we are to agree that perception is not a reading of the world out there and that it is a "controlled hallucination," a way of dreaming up the world on the basis of the "data" received from it, that proposition surely calls for an expanded sense of media through which such a hallucination can occur and present itself to us. Although Seth explicitly rejects the model of the computer, his preference for information theory as the organizer of his metaphorical horizon effectively ends up turning the brain into a high-level processing unit that transforms sensory signal inputs

into abstract predictions, and then spits them out as perceptions. What format and what medium these predictions come in remains unanswered in his theory. Seth emphasizes that "brains are not computers made of meat" but rather "chemical machines," and that every brain is "part of a living body, embedded in and interacting with its environment—an environment which in many cases contains other embodied brains."[41] Putting the image of meat at the center of his argument, he does not follow through on its consequences. The linguistic paradigm, which is inadvertently introduced into Seth's information-driven model—as evidenced, for example, in his claim that "we never experience sensory signals themselves, we only ever experience interpretations of them"[42]—misses out on the creative possibility of the imagistic apparatus. What is more concerning for me is that his model also plays into the hands of those who want to posit prediction as a medium-independent process that can be done better and faster *by* machines, especially those of nonhuman variety. With this we enter the weird world of predtech: predictive technology.

Predtech and the capitalization of perception

This development, long predating advanced research in machine vision and machine learning, was already envisaged by Virilio in his *Vision Machine*. Virilio posited that

> the act of seeing is an act that [precedes] action, a kind of pre-action partly explained by Searle's studies of "intentionality." If seeing is in fact foreseeing, no wonder forecasting has recently become an industry in its own right, with the rapid rise of professional simulation and company projections, and ultimately, hypothetically, the advent of "vision machines" designed to see and foresee in our place. These synthetic-perception machines will be capable of replacing us in certain domains, in certain ultra high-speed operations for which our own visual capacities are inadequate, not because of our ocular system's limited depth of focus, as was the case with the telescope and the microscope, but because of the limited *depth of time* of our physiological "take."[43]

Virilio's analysis was premised on developments taking place in "artificial intelligence," a term curiously yet not wrongly placed by him in quotation marks, in the 1980s, long before the emergence of deep learning and machine vision premised on "Inceptionism," i.e., recursive neural nets. The angle of Virilio's prophetic analysis was sociopolitical. He warned us

against "the automation of perception,"[44] coupled with the production of synthetic imagery that had no indexical reference to anything we recognized as reality, beyond its verisimilitude. Virilio thus knew already in 1988 (the year the French edition of his book came out) that prediction was in fact a creation. The obfuscation of this fact by the photorealism of the media through which the forecasting is delivered has serious consequences for ourselves as both individual selves and political subjects.

Virilio analyzed the use of prediction in military technology, with war maneuvers and wars' outcomes increasingly being simulated on geographically remote screens, where action on the ground serves as a singular actualization of the virtual possibilities run on prediction machines in a different part of the globe. Today predtech is widely deployed in areas such as finance, weather prediction, epidemiology, cancer detection, consumer behavior, and crime prevention. Combining data mining, neural network analysis, and visual rendering, it stages virtual scenarios while also making realities. To explain the working of predtech, marketing technology company Bluecore uses the example of a fuzzy picture which is gradually filled in with more data. Comparing traditional marketing to "taking a very zoomed in view of a photograph,"[45] they pride themselves on using AI to provide "a bigger picture." Yet the process of rendering this picture not only creates an image of their client's customers but also *creates customers*: shaping their desires, choices, and actions while simultaneously rendering the very idea of a desirable customer base for the client using Blucore's predictive technology. For example, "you can use it to calculate a customer's predicted lifetime value, helping your team focus on customers that will drive the most revenue long term."[46] This process is also in operation in other areas where the use of resources, especially those of human nature, is to be optimized. "As of 2013, Hewlett-Packard was predictively scoring its more than 300,000 workers with the probability of whether they'd quit their job—HP called this the Flight Risk score, and it was delivered to managers."[47] Predtech therefore not only predicts what is going to happen but also *makes things happen*. Not only can it create "bad workers" by reporting on them as potentially less reliable and steady in advance of anything they actually do while also shaping punitive management tactics, it also designates certain areas and certain groups of people as less desirable and less valuable, as poorer, weaker, and more prone to crime.

And thus machines photograph the future for us, imaging into existence through publicity and simulation both our desires and the world to come. It was recently reported that AI will not "just help us find things, it will generate what we're actually seeking."[48] Even though the generation of new images through compositing has shaped photographic practice since its inception, with AI-enhanced computation the technique has shifted to the center of image production, while not requiring any expertise from the user anymore. Scott Prevost of Adobe Sensei, an AI and machine learning section of the creative imaging software company, explains that soon the search function "will be able to apply machine learning to blend assets and create an image that never existed before—an image that's exactly what you had in your imagination."[49] The more recent "text-to-image" generators such as Open AI's DALL·E have already been able to generate what we ourselves cannot yet quite imagine. Although presented as a useful feature for all sorts of creatives, it is easy to envisage this technology being utilized to enhance the (mal)functioning of platform capitalism. Following Bernard Stiegler, Taffel describes this process, "whereby knowledge is displaced from humans into machines designed to commodify, privatise and monetise data, information and experience,"[50] as a proletarianization of the user. The impending metaverse as envisaged by the tech companies will perhaps be a final installation in this attempt to not only fully automate our perception but also commodify it.

Of course, things do not have to be this way. Prediction need not be seen as a passive enterprise, an impassionate observation of what will have happened. Having buried the indexical fantasies of photography, we can instead explore the idea of photographing the future as a form of creation, an imaginative rendering of a future we would like to see. This open-ended concept of the future is already (albeit implicitly) embraced by Adobe, Meta, and other giants of the tech industry, although, in supporting a particular form of popular and populist data exchange on their platforms, they end up consolidating the contrary belief that "the future is inevitable, something that can at best be predicted," a belief that of course works in the service of those companies. Yet, as Nick Montfort phrased it, "The future is not something to be predicted, but to be made."[51] We must not therefore just respond to the tech companies' visions and vistas but offer a vision our own—an approach Montfort describes as "future-making." John Norton, a technology writer for the *Observer* and a self-confessed recovering utopian,

recognizes that "if we want to make things better, our focus has to be on changing the machine's purpose and obligations"[52]—with "the machine" standing for him not just for the tech companies of Silicon Valley, but also for their financial enablers, political endorsers, and infrastructural supporters. We must therefore change the key parameters of the prediction machine, while making use of its warped temporality. Photographing the future can become a way not just of seeing *what will happen* but of *creating images of what we want to happen*—and of how we want the world to look—and making better choices from within the image stack. This shift is to take cognizance of the fact that, as discussed in this chapter, computational imaging is always a form of buffering, with a singular photograph being just a cut in a range of possibilities. The next chapter will offer a singular attempt to make better cuts in the image flow—and, with this, to envisage ourselves a better future.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in ITC Stone Serif Std and ITC Stone Sans Std by New Best-set Typesetters Ltd.

Library of Congress Cataloging-in-Publication Data

Names: Zylinska, Joanna, 1971- author.
Title: The perception machine : our photographic future between the eye and AI / Joanna Zylinska.
Description: Cambridge, Massachusetts : The MIT Press, [2023] | Includes bibliographical references and index.
Identifiers: LCCN 2022050332 (print) | LCCN 2022050333 (ebook) | ISBN 9780262546836 (paperback) | ISBN 9780262376624 (epub) | ISBN 9780262376631 (pdf)
Subjects: LCSH: Photography—Philosophy. | Digital images—Social aspects. | Image processing—Digital techniques. | Artificial intelligence.
Classification: LCC TR183 .Z955 2023 (print) | LCC TR183 (ebook) | DDC 770.1—dc23/eng/20230405
LC record available at https://lccn.loc.gov/2022050332
LC ebook record available at https://lccn.loc.gov/2022050333

10 9 8 7 6 5 4 3 2 1