MAST

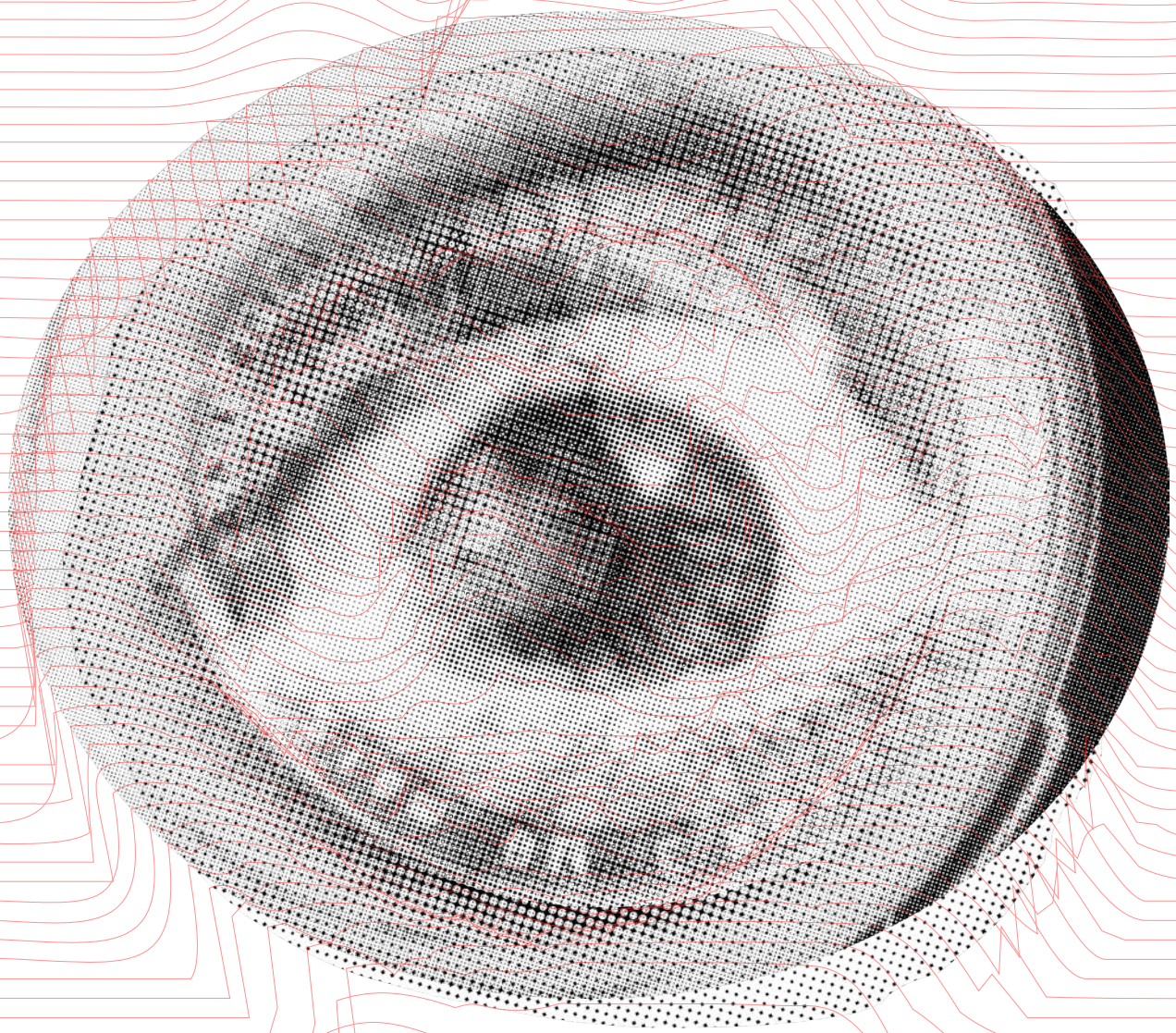
The Journal of Media Art Study and Theory

Special Issue:

Automating Visuality

Edited by Dominique Routhier, Lila Lee-Morrison and Kathrin Maurer

Volume 3 | Number 1 | April 2022



ISSN 2691-1566
mast-journal.org

MAST

The Journal of Media Art Study and Theory

ISSN 2691-1566

Author Benefits

- ✓ MAST is an open-access journal and provides free access to the full texts of all its contents immediately upon publication. No fees are charged to either readers or authors.
- ✓ MAST uses a double-blind peer-review process. The editors perform an initial review of all submissions and may reject papers that are clearly outside the scope of the journal. Accepted submissions within the scope will be sent for blind peer-review and will be reviewed by at least two internal/external reviewers.
- ✓ Published articles are licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Authors retain copyright and once published, they can post a copy of their articles in their websites and in an Institutional repository providing acknowledgment is given to MAST as the original source of publication.
- ✓ MAST is part of Radical Open Access Collective.

Twitter: twitter.com/MAST_NeMLA

Websites: mast-journal.org

NeMLA NORTHEAST MODERN LANGUAGE ASSOCIATION

MAST The Journal of Media Art Study and Theory

University at Buffalo The State University of New York

Aims and Scope

MAST is an online, open-access, and double-blind peer-reviewed journal featuring interdisciplinary scholarship in the domain of media studies. *MAST* stands for "Media Art Study and Theory" and aims to publish innovative research, writing, and work by artists and scholars who present new methods, approaches, questions, and researches in the field of media studies in theory and practice. The journal is relevant to academics, artists, researchers, theorists, and art curators with an interest in artistic research, theory, and praxis of media, introducing works that demonstrate creative engagements with current debates in media studies. *MAST* is housed in and sponsored by NeMLA (The Northeast Modern Language Association) at the State University of New York at Buffalo. *MAST* issues are published digitally twice a year (Spring and Fall).

Contents

Special Issue: Automating Visuality
Edited by Dominique Routhier, Lila Lee-Morrison, and Kathrin Maurer

The Journal of Media Art Study and Theory
Volume 3, Issue 1, 2022
Automating Visuality
www.mast-journal.org
MAST

Automating Visuality: An Introduction

Dominique Routhier, Lila Lee-Morrison, and Kathrin Maurer
University of Southern Denmark

Intelligent machines and self-learning algorithms increasingly determine not only *who* and *what* we see—what pops up on our screens or bursts into our social media bubbles—but also *how* we see and perceive the world around us. In drawing attention to the "automation of visuality"—by which we mean the automatic sifting, framing, and structuring of everyday life (and death) through machinic forms of vision—we want to examine not only how machines see but what it means when visuality itself, the social fact of seeing, is "curated" by machines.

Today, machines sort and shape what we see in all kinds of seemingly innocuous ways. We are so accustomed to having our personal news feed, social media, or Netflix suggestions served to us automatically that we rarely realize the extent to which our feeds, our "windows" to the world, are custom tailored by unobtrusive algorithms. As the authors of a recently published book, *The Age of AI*, notice: the algorithmic interventions in everyday life have reached a point where "the individual comes to rely, often *instinctively* or *subconsciously*, on software processes to organize and cull necessary and useful information" (Kissinger, Schmidt, and Huttenlocher 105, emphasis added). This observation, for better or worse, extends to "visuality" more broadly as new machinic forms of vision increasingly structure perceptions on a societal scale.

In a sense, the problematic at the heart of this special issue reaches all the way back to Walter Benjamin's celebrated essay "The Work of Art in The Age of Mechanical Reproduction" from 1935. Benjamin's essay, which centered on cinema's social and political implications, has since spawned a

range of critical perspectives on the nature of machinic vision technologies and their impact on society at large. On the following pages we bring together new scholarly approaches to the evolution and current state of automation in the realm of vision technologies with the aim to critically examine the machinic "filtering, decrypting, and pattern recognition" that artist Hito Steyerl rightly identifies as defining traits of our contemporary "post-cinematic" visual culture (47).

The contributions to this issue explore different aspects of the automation of visuality and can be roughly divided into three distinct but interrelated themes: forms of automated surveillance, particularly outside the military nexus (Søilen; Wellendorf et al.); ruptures in the regime of representation and in the correlation between images and their supposed truth content (Uliasz; Lehmuskallio and Meyer); and questions of power inherent to advanced imaging technologies both historically and in the present (Buel; Amin). Much of this new research builds on, extends, and enters into critical conversation with existing scholarship on image technologies and the effects of these technologies on visual culture at large.

A key point of reference for this special issue is Trevor Paglen's idea that the "automation of vision"—including the technology of smart bombs, drone sensors, intelligent surveillance cameras, automatic facial recognition, automatic license plate readers, etc.— challenges entrenched understandings of visual culture as narrowly tied to human interests and human sense-making ("Invisible Pictures"). Following the German filmmaker Harun Farocki, Paglen describes this historical change in terms of a "tectonic shift" in visual culture: a shift from a human-centered ecology of images to a near autonomous feedback loop of machine-readable or "operative" images, of which its aesthetics are not intended.

Farocki's artistic exploration and theorization of what he termed as "operational images" was formulated against the backdrop of US-led warfare in Iraq and based on visual analysis of missile footage from military "targeting" operations ("Phantom Images"). Media theorist Jussi Parikka notes in a forthcoming book on operational images that, although Farocki's founding work on operative images speaks to this military "logic of targeting," his work in this field also encompassed "architectural

modeling, traffic control systems, construction of affective environments such as malls, and other examples that have also paved the way toward current topics of AI culture."

While Paglen and others have also highlighted how forms of operative images are encroaching on various cultural aspects beyond the military realm, much of the discussion about machine vision remains centered on the analysis of overtly militarized aspects of everyday life, most notably contexts of policing, surveillance, and border control. While these remain pertinent contexts to study, recent scholarship has highlighted that operative imaging techniques are increasingly relevant also outside of such para-military contexts (for instance, Aud Sissel Hoel and Frank Lindseth's "Differential Interventions"). This special issue also contributes with new perspectives on these alternative contexts of machine vision.

This partial shift of focus relates to the fact that most AI-assisted imaging systems serve so-called "dual" purposes. The automation of visuality, in other words, cuts across from the military contexts into the less spectacular, quotidian technologies under AI-saturated "surveillance capitalism" where both gendered and racialized hierarchies of power can be evidenced and extended through the visualizing processes of machine vision (Zuboff). As surveillance scholars have pointed out, the automation of surveillance continues, and in some cases even *reinforces*, existing forms of social inequities, racialization, and marginalization (Browne). Faced with new machine vision technologies—automated facial recognition systems are a case in point—scholars are providing more nuanced understandings of surveillance that challenge the traditional dichotomies between target and observer, surveyor and surveillé, enemy and friend.

The figure of the big brother of panoptic surveillance is changing in the context of big data tracking. One important feature of today's imaging systems is that their workings defy the classical military "logic of targeting" so central to the theory of operative images. Today's advanced imaging techniques form part of a contemporary set of data-based targeting procedures that, as the French philosopher Grégoire Chamayou argues in his article "Patterns of Life," not merely tracks and surveils

pre-existing individuals but rather *produces* its subjects across "fields as diverse as policing, military reconnaissance and marketing."

Distinctions between civilian and military realms are disintegrating, calling for analyses that transgress the military-civilian divide and look closer at the ways that surveillance machines operate today. Many users of social media and smart home tech are now surveilled voluntarily. We invite technologies that track, follow, and monitor us into our houses—think about Google Nest or Alexa here—and we carry "smart" devices everywhere in our pockets, on our wrists, or strapped to our bodies. We are physically, affectively, and materially entangled with these new forms of "big data" surveillance. However, this does not necessarily make these technologies less intrusive or violent. On the contrary. But the automated gaze is no longer the piercing camera eye or the all-seeing panopticon but rather a form of sensorial 24/7 surveillance that most of us have gladly welcomed into our domestic life (Andrejevic).

The automated gaze of the robotic vacuum cleaner is a prime example of this type of affective big data surveillance and an excellent entry-point for discussing the automation of visuality in the context of the Western "modern home." These co-bots are often sold under the pretense of helpful household companions, as they can eliminate dirt while you are gone, overcome any surface obstacles, and duly recharge and park themselves after their services. Often these robo-cleaners are equipped with laser based and infra-red cameras and can use artificial intelligence software and deep learning for better mapping (or mopping).

Karen Louise Grova Søilen explores in her contribution, "The Haunting of the Automated Gaze," the robovac-gaze in an aesthetic context analyzing the video installation *Modern Escape* (2018) by the artist duo Hanne Nielsen and Birgit Johnsen. Their artwork exhibits the type of uncanny atmosphere that the non-human cleaner (and observer) creates and how it is able haunt the house with its eerie presence. This haunting triggered by the affective encounter of the robovac makes explicit the military optics and technologies of automatization and its logics of safety, security, and elimination.

The harmless looking vacuum cleaner penetrates our private sphere as a military spy in domestic camouflage.

The authors of "Calm Surveillance in the Leaky Home" (Kassandra Wellendorf, Karen Louise Grova Søilen, and Kristin Veel) also discuss the robotic vacuum cleaner as an example of the automation of visuality. They do not focus on an artwork but rather follow an ethnographic and empirical method, interviewing selected users and owners of a robovac. According to their findings, people who own a robotic vacuum cleaner have mixed affectivities about it. They anthropomorphize it at times as a "benign pet" but also perceive it as a "part-surveillant intruder that sees more than its keeper" (Wellendorf et.al.). In fact, most users highlighted that it is quite a hassle to have such a thing and that the industry's promise of calm co-habitation and harmonious dwelling with it is a myth. Rather robovacs and automated dirt elimination technologies embody a data collection machine with "always on" surveillance that becomes a permanent member of your family.

The robovac is just an example how automated visual technologies can terrorize—or in Søilen's words "haunt"—our private and intimate surroundings. Social media with algorithmic modes of observation, automated generation of deep fake images, facial recognition, and drone imaging – all embodiments of an automated gaze discussed in this theme issue—surveil us often in ways that we do not even know about. They track, map, listen, and observe us in ways that are invisible and inaudible. However, their modes of networked surveillance make us nevertheless complicit with the military origins of these technologies. As Søilen states, referring to the works of Annie Ring, our complicity manifests itself in the "intimate entanglement and shared material histories between the military and consumer technologies." Moreover, our robo-friends can create databases that can be sold to governmental and non-governmental institutions and companies as was the case with the facial recognition company Clearview, which harvested data from social media and sold them to, for example, law enforcement, security firms, and police departments. As these examples of the blurring between the military and the civilian realms suggest, we are confronted with new forms of what drone scholar Caren Kaplan and her colleagues aptly call "everyday militarism" (Kaplan, Kirk, and Lea).

But it is not only the nature of surveillance that changes as it becomes automated, but also the nature of images. Just as it is necessary to think outside the dichotomies of traditional surveillance studies, so too is it crucial to think beyond the dichotomy of "visible and invisible" that still structures contemporary image theory. While it is tempting to think of our automated visual culture in terms of an opposition between visibility and non-visibility, such an approach risks, as Luciana Parisi notes, "to re-inject the truth of representation back into the deep learning of automated networks" (5). The defining feature of visual culture today is not so much that it is "invisible," but rather that the social fact of seeing—or "visuality" in short—passes through opaque proprietary algorithms that perpetuate the structural inequities of capitalist society in coded form.

More important than the accelerated dissemination of discrete images or the overwhelming of the senses by a constant stream of images is the fact that the components of the visual architecture of our shared reality, the image form, are mutating into something else. Notably, this change includes the fact that most images now form a part of mass aggregates, or "image ensembles," that are better understood as statistical entities than as individual fragments of a social totality. "Today," as image theorists Adrian MacKenzie and Anna Munster convincingly argue, "it is the "image *ensemble*—images not simply quantified, but labelled, formatted and made 'platform-ready'—that enables the emergence of a new mode of perception, and indeed a reformulation of visuality itself" (5). But how did we get from the early technologies of machine vision—like photography and film—to the post-cinematic and AI-powered forms of "platform seeing" that now define visual culture? Or, to put it in Benjaminian terms: what happened to the image form in the transition from the age of mechanical reproduction to the so-called age of AI?

A central focus that is addressed in a selection of the articles in this special issue deals directly with how the automation of vision has changed the image form, specifically concerning an historical relationship between the image and notions of truth and evidence. This is explored within the wider contemporary context of what has been termed a "post-truth" society. Indexicality, a concept describing the relationship between a sign and its referent and originating in the semiotic theory of Charles

Sanders Pierce, has historically grounded this relationship between truth and the image. The concept has been further theorized in photographic discourse framing the photographic image as evidence, in its record of a physical trace of an event. The contributions that take up this concept in their analyses hone in on the machinic image, its disruption of indexical relationships, and how its forms of representation can be understood in light of this concept. Indexicality is both reexamined in its historical development as well as reconstituted with relevance to the contemporary contexts of machine vision implementation, its production of an algorithmic image, and the challenges it presents to notions of an evidentiary aesthetic. The contemporary machinic image is approached as a site where historically contested ideas about vision and visuality are present. Rather than relying on a notion of an objective truth, the machinic image and its reconstitution of indexicality is understood as representing the *conditions* from which truth is constructed in contemporary society.

The following two contributions further complicate an understanding of indexicality in two distinct arenas of machine vision implementation: deepfake image production and the development of automated facial recognition technologies. Rebecca Uliasz, in her article titled "On the Truth Claims of Deepfakes: Indexing Images and Semantic Forensics," explores indexical relations in their production of deepfakes images, that is, digitally manipulated images and videos which falsely represent individuals and circulate widely online. She situates her inquiry into the wide scale production of deepfakes within a broader socio-political context of a contemporary struggle for truth and meaning in the visual. Her approach is framed by addressing the semiotic infrastructures of generative adversarial networks (GANs), a machine learning technique which aids in the production and implementation of deepfakes. The critical inquiry into GANs is grounded within contemporary sociopolitical discourse surrounding deepfakes, including a study into the status of deepfakes as evidence by DARPA's Semantic Forensics Program. Although representing a falsehood, deepfakes operate as evidence of the construction of truth in our contemporary visual regime. Through Uliasz's analysis on the phenomenon of deepfakes, she opens up an understanding of indexicality within the contemporary context of digital images, as traces of a relationship between technical and social registers including data sets, human

bias, and engineering decisions. In turn, Uliasz argues for an understanding of the ontology of deepfake images, which not only represent but generate ethicopolitical meaning within the wider contemporary context of post-truth media ecologies.

In Asko Lehmuskallio and Roland Meyer's article, titled "Experimental Indices: Situational Assemblages of Facial Recognition," the authors revisit the historical contexts upon which indexicality was premised both in the theory of Charles Sanders Pierce as well as through situated photographic practices to pluralize its understanding. Similar to Uliasz, Lehmuskallio and Meyer argue against a misguided understanding of indexicality in photography as incurring a transparent process and instead explore the processes of mediation, manipulation, and contingency of claims to a truth that continue in the contemporary context of automated facial recognition technologies. Lehmuskallio and Meyer focus on what they call the "experimental indices," that is, the varying situational assemblages of image apparatuses, including facial recognition technology, which form a material understanding of indexicality through its parameters and settings. They outline how mutable the notion of indexicality is and explore how different practices of contemporary facial recognition technologies produce their own set of experimental indices. Within this context, the authors address notions of a contemporary truth which is premised on a probability rather than certainty and problematize a long historied relationship of indexical relations which link bodies, images, and data.

If these contributions address the question of how machine vision technologies further unsettle the always-already unstable indexical relations between visual signs and their referents, the final two contributions to this special issue focus on the way that power inheres in imaging technologies, historically and presently. In the article "Automated Visions, Algorithmic Imageflows: The Technopolitics of Black Lives Matter Videos on YouTube," media theorist Jason Buel examines how proprietary algorithms intervene in the image flows of popular streaming platforms like YouTube. Far from representing a neutral patchwork of uploaded users' content, the video flows on YouTube are filtered through opaque recommendation algorithms that tailor content to specific viewer profiles while, paradoxically, appealing to an "idealized viewing position" free from ideological constraints. The

crucial question Buel raises is not what kind of ideological trickery may happen inside the "black box" but how forms of algorithmically curated representations actively *shape* and radicalize political subjectivities *by design*.

As Buel points out, YouTube, like many similar social media platforms, is built for "holding viewer attention (or, more accurately, collecting data assumed to stand in for viewer attention) and holding that attention as long as possible so as to commodify it and transform it into revenue." The result of this automating of visuality, Buel argues, is that a platform like YouTube tends to cater to spectacular and attention-grabbing content that too easily slides from, for instance, pro-Black Lives Matter videos toward more and more extreme forms of right-wing extremism. By algorithmically catering to viewers' engagement at any price, YouTube enacts a kind of late capitalist "subject formation" where it is ultimately, as Buel argues, "less about what any specific video has to say about the need for radical socio-political change and more about the viewing subject."

Buel's examination of Black Lives Matter videos in the context of YouTube's algorithmic curating of content challenges the idea that technological progress is inherently democratic, universal, and color-blind. As AI researcher Kate Crawford notes, long before most images ever reach the human eye, they have been bundled into "training sets," labelled, and sorted into categories that carry social norms and biases over into the visual field. What is usually referred to as the "black box," then, as sociologist Ruha Benjamin notes in her important book *Race After Technology*, often perpetuates forms of "routine anti-Blackness" and social inequity: "What I call the *anti-Black box* links the race-neutral technologies that encode inequity to the race-neutral laws and policies that serve as powerful tools for White supremacy" (35).

The idea that technology is never politically neutral but always reflects the social structure in which it was developed and employed is also a recurrent theme in this special issue's final contribution: an interview with the Egyptian-born contemporary artist Heba Y. Amin. Key to Amin's internationally acclaimed artistic practice is a critical exploration of the historical development of visual technologies in a context of colonial land-grabbing and exploitation, particularly in the Middle East and Africa. In

the conversation that took place over Zoom, Amin explains how early imaging technologies like panoramic photography were "at the core of visualizing the colonial project." Beginning from advanced imaging systems, such as drone and satellite vision technologies, Amin moves backwards in history to dissect some of the core assumptions that are built into and perpetuated by today's vision machines. Through a poetic lens, at once deeply historical and political, Amin's artistic practice and research thus offers a critical corrective to a "history of vision that has been, and continues to be, narrated through a Western 'universalist' perspective." Amin's contribution thus provides a timely and sobering view of the "future" of automation in the realm of vision as one necessarily haunted by a colonial past and fragmented along the broken, geopolitical lines and borders of capitalist modernity. What the "future"—itself an arguably time-worn Western trope—will *look* like is hard to tell. Most likely it will look different depending on where you look at it from. The effects of automation technologies, for better or worse, tend to be unevenly distributed.

Without further speculations, then, we happily hand over the word to the outstanding scholars and artists who have defied the pandemic condition of the past few years to generously share their research, time, and effort to help diagnose aspects of the "automation of visuality."

Works Cited

Andrejevic, Mark. "Automating Surveillance." *Surveillance & Society*, vol. 17, no. 1/2, 2019, pp. 7–13.

Benjamin, Ruha. *Race After Technology*: *Abolitionist Tools for the New Jim Code.* Polity Press, 2019.

Benjamin, Walter. "The Work of Art in the Age of Mechanical Reproduction." *Illuminations*, Schocken Books, 1968.

Browne, Simone. *Dark Matters: On the Surveillance of Blackness*. Duke UP, 2015.

Chamayou, Grégoire. "Patterns of Life: A Very Short History of Schematic Bodies." *The*

Funambulist Papers, no. 57, Dec. 2014, https://thefunambulist.net/editorials/the-funambulist-papers-57-schematic-bodies-notes-on-a-patterns-genealogy-by-gregoire-chamayou. Accessed 18 Mar. 2022.

Crawford, Kate. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale UP, 2021.

Farocki, Harun. "Phantom Images". *Public*, no. 29, Jan. 2004, https://public.journals.yorku.ca/index.php/public/article/view/30354. Accessed 18 Mar. 2022.

Hoel, Aud Sissel, and Frank Lindseth. "Differential Interventions: Images as Operative Tools." *Photomediations: A Reader*. Open Humanities Press, 2016.

Kaplan, Caren, Gabi Kirk, and Tess Lea. "Editors' Letter. Everyday Militarisms: Hidden in Plain Sight/Site." *Society & Space Online Forum*, 8 Mar. 2020, https://www.societyandspace.org/articles/editors-letter-everyday-militarisms-hidden-in-plain-sight-site. Accessed 18 Mar. 2022.

Kissinger, Henry A., Eric Schmidt, and Daniel Huttenlocher. *The Age of AI and Our Human Future*. Little Brown and Co., 2021.

Mackenzie, Adrian, and Anna Munster. "Platform Seeing: Image Ensembles and Their Invisualities." *Theory, Culture & Society*, vol. 36, no. 5, 2019, pp. 3–22.

Paglen, Trevor. "Invisible Images (Your Pictures Are Looking at You)." *New Inquiry,* Dec. 2016, https://thenewinquiry.com/invisible-images-your-pictures-are-looking-at-you/. Accessed 18 Mar. 2022.

Parikka, Jussi. *Operational Images*. University of Minnesota Press, Forthcoming 2023, https://jussiparikka.net/2022/02/16/operational-images-preface-in-the-forthcoming-book/. Accessed 18 Mar. 2022.

Parisi, Luciana. "Negative Optics in Vision Machines." *AI and Society*, vol. 36, no. 1281–1293, 2021, https://doi.org/10.1007/s00146-020-01096-7.

Steyerl, Hito. "A Sea of Data: Apophenia and Pattern (Mis)Recognition." *e-flux Journal*, no. 72, Apr. 2016, https://www.e-flux.com/journal/72/60480/a-sea-of-data-apophenia-and-pattern-mis-recognition/. Accessed 18 Mar. 2022.

Zuboff, Shoshana. *The Age of Surveillance Capitalism*. PublicAffairs, 2019.

Dominique Routhier is a post-doctoral researcher at the Department for the Study of Culture, University of Southern Denmark. Currently, he is affiliated with the research project Drone Imaginaries and Communities (Independent Research Fund Denmark, 2020-2024). His writings have appeared in *Rethinking Marxism*, *Nordic Journal of Aesthetics*, *Boundary 2 online*, and *LA Review of Books*, among other places. He is the author of a forthcoming book about the Situationist International (Verso Books, 2023). Email: dominique@sdu.dk.

Lila Lee-Morrison is currently a post-doctoral researcher at the University of Southern Denmark. Her research interests include cultures of machine vision and the intersection of contemporary art, critical theory and contemporary modes of perception, specifically by algorithmic processes. She is currently investigating the topic of machine vision in relation to the art historical genre of landscape imagery and at the theoretical intersection of the aesthetics of the anthropocene and posthumanism. Previous to this position she completed a PhD at the Dept of Art History and Visual Studies in Lund University with a published dissertation titled, *Portraits of Automated Facial Recognition: On Machinic Ways of Seeing the Face* (Transcript Verlag, 2019). She has written on the visual politics of drone warfare and of biometric systems as well as on media images of the European immigration crisis. She has published with MIT Press, Liverpool University Press and Brill Publishing.
Email: lile@sdu.dk.

Kathrin Maurer is a professor MSO of Humanities and Technology, the leader of the Center for Culture and Technology at the University of Southern Denmark, and also the leader for the research cluster Drone Imaginaries and Communities. She has a PhD in German Literature from Columbia University and Dr. Phil from SDU. Her research interests are surveillance technology, drones, German literature, visual culture, and aesthetic theory. Her current book *The Sensorium of the Drone and*

Communities is forthcoming with MIT Press. She has published *Visualizing History: The Power of the Image in German Historicism* (Walter de Gruyter, 2013), *Visualizing War: Images, Emotions, Communities* (Routledge, 2018) and published internal peer reviewed journals (*New German Critique*, *Media, War, and Conflict*; *Germanic Review*). She is also a Senior Fellow at the Danish Institute for Advanced Study at SDU and the recipient of a research grant from the Alexander von Humboldt Foundation. Email: kamau@sdu.dk.

The Journal of Media Art Study and Theory
Volume 3, Issue 1, 2022
Automating Visuality
www.mast-journal.org
MAST

The Haunting of the Automated Gaze

Karen Louise Grova Søilen
University of Copenhagen

Abstract

This article analyzes the artistic exploration of machine vision in the video installation *Modern Escape* (2018) by Danish artist duo Hanne Nielsen and Birgit Johnsen. The artwork recreates a modern Western home pervaded by surveillance technologies and the automated vision of a robotic vacuum cleaner. The main conceptual idea of the work is *the automated gaze*, and with few exceptions, all the scenes in the video are filmed without a human behind the camera. As a result, the installation evokes the haunted atmosphere of a home penetrated by a gaze that drifts incessantly, opening up randomly strange and awkward vantage points of the interior. Specifically, the article offers a reading of *Modern Escape* in the light of Avery Gordon's notion of "haunting" and argues that the modern Western home is saturated with traces of the military industrial complex through its commonplace technologies, which are haunted by a history of war, complicity, and masculine desires for control. Haunting, according to Gordon, is one way in which "abusive systems of power make themselves known and their impacts felt in everyday life" (*Ghostly Matters* xvi). The present article discusses how *Modern Escape* responds to and questions the subtle militarization of the everyday the technologies carry into the estranged home, "that quintessential space of the uncanny, the haunted house" (Gordon, *Ghostly Matters* 50).

Keywords

the automated gaze; contemporary art; haunting; machine vision; surveillance

> Only the Lighthouse beam entered the rooms for a moment, sent its sudden stare over bed and wall in the darkness of winter, looked with equanimity at the thistle and the swallow, the rat and the straw.
>
> Virginia Woolf, *To the Lighthouse*

Introduction

Darkness. White noise. The gaze moves along a floor; the reflection of a robotic vacuum cleaner passes by the shiny surface of a mirrored pedestal. It is the vantage point of the robot. The screen is overwritten by a grid map, the floor plan created by the smart vacuum cleaner as it is moving. Mapping. Searching. Suddenly, the lamp to my right turns on. Danish Design, just like the one inside the video living room. A woman is looking through the blinds of a modern home, glass windows as walls. Now the perspective is higher, slowly circling around. Flickering TV images and news, a second woman watching TV, a voice reporting on the building of a border wall. Images of barbed wire.

Shift to mid-height. Everything is tasteful. A woman, an expensive armchair, her knees, her skin, her upper body, her head cut off by the frame, a cat lounging on a footstool. I recognize it as a generic Scandinavian upper-middle class home. The automated gaze passes the woman in the chair; a voice on the news uttering "elsewhere." The TV comes into sight; images from a refugee tent camp, earth, white tents, greenery, and people. The gaze is making its round. Shiny surfaces, glass, steel. Those who are kept out.

Drone point of view from above: dark shiny pillars in an empty space. A review of Lighthouse *comes into view, AI smart surveillance for the home. The steady work of the vacuum*

robot, cleaning. A piling up of white plastic canisters, bags of sand or flour, and large stacks of tin cans inside the apartment; bright shiny silver-colored cans. What is going on? Who is behind this? Events and vantage points blend into a distorted, confusing plotline. Now the home is almost filled up with canisters, cans, bags, and thermal blankets. Barricaded.[1]

Danish artist duo Hanne Nielsen (b. 1959) and Birgit Johnsen's (b. 1958) immersive video installation *Modern Escape* (2018) recreates a modern Western home pervaded by surveillance technologies and the automated vision of a robotic vacuum cleaner. The main conceptual idea of *Modern Escape* is *the automated gaze*, and with few exceptions, all the scenes in the video are filmed without a human behind the camera. As a result, the installation evokes the haunted atmosphere of a home penetrated by an automated gaze, opening strange and awkward vantage points of the interior. Throughout *Modern Escape,* the automated gaze records the inside of the home from multiple angles and from four strictly determined levels, simultaneously orienting and disorienting the viewer. This artistic strategy creates seemingly random images, ranging from the low floor angle of the robotic vacuum cleaner and a mid-height angle capturing a woman at knee-height to a drone view from above. As spectators, we do not get the full view nor any sense of a cohesive narrative, resulting in a sense of discomfort inside the installation space. Whose gaze is this? What is being seen?

This article offers a reading of *Modern Escape* in the light of Avery Gordon's notion of "haunting," in order to unpack how the artwork addresses the home's saturation with traces of the military industrial complex. Specifically, I argue that *Modern Escape* foregrounds how the networked technologies present in the home are haunted by a history of war, complicity, and

[1] The opening vignette and the following analysis are based on my visits to the solo exhibit *REVISIT: HANNE NIELSEN & BIRGIT JOHNSEN* at Overgaden Institute for Contemporary Art, Copenhagen, Denmark.

masculine desires for control. The entanglement of surveillance, warfare, border control, and the networked technologies of consumers in the Global North is established through the work's formal and thematic structure of mapping and tracking practices. As an organizing principle, this artistic strategy effectively articulates a connection between consumer technologies such as the GPS-enabled smart vacuum cleaner equipped with intelligent sensors for charting the home, the facial and object recognition technology of home surveillance systems, and the state and military mapping and tracking practices of which these technologies simultaneously are heirs and further inform. In what follows, I first establish the conceptual framework of haunting, before turning to a reading of how *Modern Escape* responds to and questions the domestication of technologies of surveillance and machine vision and the subtle militarization of the everyday the technologies carry into the estranged home, "that quintessential space of the uncanny, the haunted house" (Gordon, *Ghostly Matters* 50).

Haunting

Drawing on the work of sociologist and feminist scholar Avery Gordon, the notion of haunting can be understood as material social forces that make themselves present affectively (*Ghostly Matters*; *Some Thoughts on Haunting*). Simply put, haunting is something sensed in our surroundings. It is a matter of shadowy manifestations of past or present wrongs. To be haunted, then, is "to be tied to historical and social effects" (Gordon, *Ghostly Matters* 190). Haunting emerges as affective encounters, uncanny experiences where unresolved or repressed social violence—the traces of "modernity's violence and wounds" (*Ghostly Matters* 25)—are felt in the present. According to Gordon, "[w]e are haunted by something we have been involved in" (*Ghostly Matters* 51), such as "all the different facets of the profound and elemental deprivations of modernity . . . slavery and racism, state authoritarianism, Enlightenment science, gendered

repression" (*Ghostly Matters* 197). Haunting, moreover, draws attention to injustices in which we are still involved. In this article, I employ the notion of haunting in order to cast new light on machine vision and surveillance. Generally, surveillance is not equally distributed; rather it reinforces inequality, racialization, and marginalization (Browne; Lyon; Monahan; Smith). The lens of haunting, I argue, activates the weight of the past in present surveillance practices, as well as the complicity and injustices we currently participate in, which are kept out of sight. As Annie Ring argues, the complicity on the part of consumers of networked technologies in providing vast amounts of data for capture and analysis can be said to be "at once forgetful and at least partially unconscious" (88). Likewise, our complicity includes the intimate entanglement and shared material histories between the military and consumer technologies of the networked era, and Ring observes that "[w]ar was the context in which the Internet was developed and where machine vision was first employed . . . War is where artificial intelligence and data analytics are most brutally applied" (90). Simultaneously, developments in consumer technologies also feed into the technologies of war. Particularly regarding machine vision and image recognition software, some of the most advanced systems are developed by private tech companies, which have access to vast amounts of users' data upon which to train their technologies (Michel). As my reading of *Modern Escape* will emphasize, the notion of haunting productively addresses the more or less subtly felt and consciously recognized power relations, violence, and complicity intrinsic to surveillance practices and technologies. The artwork effectively establishes the entanglement between privileged consumers and structures of domination and offers the spectator a bodily recognition of their own implication (Rothberg). Returning to Gordon, she proposes that haunting is "one way in which abusive systems of power make themselves known and their impacts felt in everyday life" (*Ghostly Matters* xvi). Thus, being attentive to the

haunting dimension of surveillance means listening to the workings of power, to the moments and spaces "when disturbed feelings cannot be put away" (xvi).

Modern Escape

Hanne Nielsen and Birgit Johnsen's video installation *Modern Escape* (2018) was created specifically for a solo exhibit at Overgaden Institute of Contemporary Art in Copenhagen. The artist duo has collaborated since 1993 and are among the leading exponents of video and media art in Denmark (Davidsen). Their oeuvre includes a long-term interest in the formal qualities of the video medium, politics, power, gender, and surveillance. Artistic explorations of the media, themes, and aesthetics of the automated gaze and surveillance have been present from an early stage of their collaborative practice, which can be seen in works such as *Grenåvej* (1995), *Installations in Urban Space* (2004), and *Replay* (2004). Lately, topics of war, migration, and security have moved to the center of Nielsen and Johnsen's artistic practice with the works *Drifting* (2014), *Defense Against the Unpredictable* (2014), *Camp Kitchen* (2014), *Modern Escape* (2018) and *Ved Hegnet* (2019) (Wolthers).

Modern Escape consists of a large screen with a one-channel color video with sound running in a loop, foregrounded by a podium with a robotic vacuum cleaner kept inside a grid structure of magnetic tape (see fig. 1). To the right of the screen stands a Verner Panton Panthella floor lamp (a popular classic of Danish Design), and to the left of the screen, an acrylic mobile shaped as the floor plan of the vacuum cleaner is hanging from the ceiling. The filmic installation itself is 21 minutes, organized in an array of fragmented scenes from inside a setting which alternates between resembling a home, an art space, and something in-between a laboratory and a storage space for cans, canisters, and thermal blankets, recalling humanitarian aid.

Fig. 1. Installation view, Nielsen and Johnsen, *Modern Escape*. Photograph by Anders Sune Berg. All images courtesy of the artists.

Overall, there is little action on the screen. The opening scene starts from a low angle by the floor, accompanied by the loud noise of a robotic vacuum cleaner. Next, the vacuum cleaner moves along on the floor. Soon it is overwritten by a "pacman grid," which is the floor plan created by the AI enabled vacuum cleaner. There is an electric fan. The interior on the screen is hard to place at first: there are large, mirrored pedestals, art, and the reflection of the robot (see fig. 2). We see a woman standing in front of a window looking out, another woman sitting in an armchair in front of the TV, a cat, and a dog. This sums up the cast and mode of action in the work overall. The artists themselves perform the two women in the work, embodying the

passivity of the privileged Western middle class, recorded by the automated gaze of the machines.

Fig. 2. Still from film, Nielsen and Johnsen, *Modern Escape*. All images courtesy of the artists.

Yet the story is not so clear-cut. Fragmented glimpses of uncanny parallel narratives interrupt the "main" story: a feeling of paranoia materializes in a scene showing a woman dressed in tin foil peering out across the street with a small circular mirror. Outside, a man in a white protective suit peers back out from his window. This anxious mood intensifies in a later glimpse of two figures inside the apartment also wearing protective suits (see fig. 3).[2] Indeed, *Modern Escape* evokes a haunted atmosphere, an eerie yet unlocalizable feeling that something

[2] While the global news audience would at the time be familiar with the white protective anti-virus suits from, for example, the 2014–16 Ebola outbreak in West Africa when the work was first exhibited in 2018, viewers of *Modern Escape* during and after the 2020–present COVID-19 pandemic are bound to bring further associations to the work, as social distancing, home quarantines, hoarding of food, and "prepping" for the worst became a mass activity.

is not quite right within the Western home. This is achieved through strategies which address the bodies and affective involvement of spectators inside the dark installation space. First, a sinister soundscape indicates the affective tone of the work. The cold white noise of the vacuum cleaner hits the ears, together with the steady mechanic rhythm of its expanding floor plan as the robot is mapping its surroundings, and an unnerving electric noise as the camera is rotating. The soundscape is further layered with a polyphony of voices from the constant stream of global news, speaking of migration and the building of border walls, national security, tent camps, surveillance technologies, and gender equality. We hear the TV speaking of "those who have little shelter in these flimsy tents" and "more suffering for those already displaced." The reported news stories ground the work in our contemporary time and its politics.

Fig. 3. Still from film, Nielsen and Johnsen, *Modern Escape*. All images courtesy of the artists.

Second, the spatial fluidity of the installation creates confusion in terms of where the work starts and ends. The installation is immersive. The lamp in front of the screen switches on and off, recalling domestic smart technologies. The robotic vacuum cleaner on the podium suddenly starts moving. Thus, the living room on the screen pours forth spatially into the gallery. The placement of "twin" objects in the installation space reflects the objects seen on the screen and creates a spatial extension of the surveillance home, which implicates visitors by addressing their embodied experience of the environment. This adds a borderlessness to the spectator's experience. Simultaneously, the flickering news stories from the TV inside the video living room take over the main installation screen, further dissolving the boundaries between the living room on the screen and the gallery space. Finally, the use of light addresses the viewer's sensual perception of the atmosphere inside the installation; the cold light of the video screen suggests a detached and isolated mood, in contrast to the warm light of the floor lamp in the installation (as seen in fig. 1). These artistic strategies effectively address the viewer as a literal presence in space, a participant and an accomplice.

A feeling of paranoia is reinforced in *Modern Escape* through the evolving barricade of the home and short glimpses of figures in protective suits and thermal blankets. Concurrently, the binary opposition between outside world and inside home, them and us, danger and safety, is obscured. This same artistic strategy was employed in Nielsen and Johnsen's closely affiliated work *Camp Kitchen* (2014). By combining war with domestic settings, both works nod to Martha Rosler's series of photo-collages, *House Beautiful: Bringing the War Home* (1967–72), but despite the similarities in Nielsen and Johnsen's two works, an important difference remains. While there is action in *Camp Kitchen*, the living room of *Modern Escape* is characterized by passivity and withdrawal from the world. The many borders and walls emerging globally as seen

on the news are paralleled by the increasing barricading taking place inside the home. But still the world seeps in, suggesting that there is no escape from our complicity in structures of inequality and violence. In this way, *Modern Escape* places the spectator in a position close to what Michael Rothberg has termed "the implicated subject"—i.e., the notion that while we may not be directly responsible for historical and contemporary injustices, we may nevertheless inhabit positions where we benefit from and indirectly contribute to structures and systems of injustice (*The Implicated Subject*).

The Automated Surveillance Gaze

Nielsen and Johnsen work conceptually, and the idea behind the work is usually their starting point. The main conceptual idea of *Modern Escape* is *the automated gaze*, and the self-imposed rule of the work was to film most of the scenes without a human behind the camera. This is achieved by filming the lowest and highest images with a GoPro Hero camera glued either to the robotic vacuum cleaner or to a turntable attached to the ceiling with double-sided tape, while the mid-range angles are filmed by a professional camera placed on turntables (Nielsen and Johnsen, email of May 2019). As such, the work can be inscribed into a range of artistic explorations of how machines "see" the world, in the tradition after Soviet filmmaker Dziga Vertov's *The Man with a Movie Camera* (1929), Canadian visual artist and filmmaker Michael Snow's *Wavelenght* (1967) and *La Région Centrale* (1971), and the contemporary Iranian-German artist Harun Farocki's *Eye/Machine* trilogy (2001–3).

Throughout *Modern Escape,* the automated gaze registers the inside of the home from multiple angles and from four strictly determined levels, ranging from floor level to a "drone view" from above. This artistic strategy creates random and awkward images such as the mid-height angle capturing the woman at knee-height (see fig. 4). Overall, the work exhibits a strong

tension between technologies and logics of mapping, security, borders, and surveillance on the one hand, and narrative labyrinths, blind spots, distortions, mirrors, and shiny impenetrable surfaces on the other. Accordingly, the desire to map, to see, to surveil, and to know is turned inside out and distorted.

Fig. 4. Still from film, Nielsen and Johnsen, *Modern Escape.* All images courtesy of the artists.

What is more, the artwork directly addresses the notion of the smart surveillance home and the link between networked consumer technologies and the military-industrial complex through appropriated imagery of an actual AI-enabled interactive surveillance system called *Lighthouse*. *Lighthouse* existed on the US market for approximately a year and a half, from its launch on May 11, 2017, to December 2018, and was at the forefront in introducing features such as computer vision and facial and object recognition technologies into the home (Lighthouse). The founders of the company came with backgrounds from the field of computer

vision, and the press release launching the surveillance system on the market highlighted the association between *Lighthouse* and the military industrial complex as a major selling point, noting how "Lighthouse uses deep learning and 3D sensing technology developed as part of the DARPA Grand Challenge to introduce an unprecedented level of awareness within the home while you're away." The DARPA—the U.S. Defense Advanced Research Projects Agency—Grand Challenge was first initiated in 2004 as a competition to "accelerate the development of autonomous vehicle technologies that could be applied to military requirements" (DARPA). The association between a home surveillance system and the U.S. military services' desires for new strategic opportunities recalls Cynthia Enloe's argument that "[t]hings start to become militarized when their legitimacy depends on their associations with military goals" (145). The subtle militarization of the home through computer vision developed for self-driving vehicles funded by DARPA puts a new perspective on the promise of providing "an unprecedented level of awareness within the home while you're away." Understood in this way, it attests to the coming together of the logics of warfare, paranoia, and the intimate sphere. In the artwork, the *Lighthouse* smart surveillance assistant enters through the flat screen TV in the living room, displaying images of an AI enabled camera sending a light out to circle and measure the space—a glare. The color scheme and sleek design of *Lighthouse* seamlessly integrate with the modern home of the artwork.

Overall, the image quality of the work is high and crisp, rather than pixelated and blurry—which sets it apart from the typical CCTV artworks of the 2000s and moves it into the drone era, thus updating the way we have come to understand surveillance imagery as a cultural trope. Nonetheless, the cold and faded color scheme of the work contributes to a detached, cold atmosphere, which is further established by the circumstance that there are no relations and no

eye contact between the women in the work and the spectator. We are looking in from the outside. The humans in the work are passive and withdrawn, captured by the steady mapping of the automatic gaze.

Fig. 5. Still from film, Nielsen and Johnsen, *Modern Escape*. All images courtesy of the artists.

The Machine's Mirror and Militarization

The vague sense of paranoia that *Modern Escape* conveys to the viewer can be ascribed to the formal exploration of machine vision as decoupled from human agency and the lack of control regarding what is "seen" and why. Considered as a critical artistic intervention, the work explores questions related to the automated surveillance gaze and perspective. Who is the creator of these images? Who is the recipient? The artwork depicts ambivalent and fleeting views, and the act of watching is an interpretative action. However, this is no longer a privileged human activity. As Hito Steyerl notes, "contemporary perception is machinic to a large degree. The

spectrum of human vision only covers a tiny part of it. Electric charges, radio waves, light pulses encoded by machines for machines are zipping by at slightly subluminal speed" (*Duty-Free Art* 47). How does this shape what is "seen"? The decoupling from human agency and human vision (Zylinska) contributes to another uncomfortable feeling created by *Modern Escape*, which is a vague sense of a loss of control and agency. Benjamin Bratton has referred to this as the "reverse uncanny valley," a discomfort and disillusion experienced "when we see ourselves through the 'eyes' of a machinic Other who does not and cannot have an affective sense of aesthetics . . . We are just stuff in the world for 'distributed machine cognition' to look at and make sense of." We are seen, recognized, or known by something which does not possess human capacities, such as emotional connection. This lack of emotional connection is emphasized by the overall lack of eye contact in *Modern Escape*. Bratton suggests that the uncomfortable recognition of ourselves in the machine's mirror reverses the "uncanny valley" originally coined by Masahiro Mori, which describe the feeling of revulsion which can arise in us when a robot becomes *too* human-like. In Bratton's reversed version, the uncanny arises when we see ourselves as un-human through the eyes of the creature and become aware of ourselves as "objects of perception from the position of the machines."

The automated gaze is a haunted gaze. Through it, *Modern Escape* suggests a deep entanglement of warfare, national security, surveillance, and the technologies of the Western middle-class home. Indeed, a range of the surveillance technologies embedded in everyday life originates in, or is funded by, the military sphere, as is the case with Global Positioning systems (GPS), close-circuit television (CCTV) cameras, locative media, the Internet, drones, and facial recognition technologies (Morrison). In the artwork, one expression of this entanglement is the deliberately inserted perspective from above, which alludes to the twenty-first century "eye in

the sky," the remote vision of the drone or satellite. On a formal level, this marks a shift from the older filmic trope of mimicking the angles from surveillance cameras mounted in the corner of a room. Nielsen and Johnsen explain that "surveillance cameras used to be installed in the corners of a room, and employing that filmic angle in a video or film will easily associate to surveillance. Today´s satellite perspectives are a direct angle from above, and we very deliberately installed the cameras to achieve these angles" (personal interview).

Fig. 6. Still from film, Nielsen and Johnsen, *Modern Escape*. All images courtesy of the artists.

Historically, the aerial perspective's association to warfare traces back to the prominent nineteenth-century photographer Nadar, who, when he took the first aerial photograph from a hot air balloon in Paris in 1855, "immediately grasped the future benefit of photography to warmakers" (Sontag 176). Practices of aerial photography and aerial surveillance greatly

advanced during World War I, and Jeanne Haffner notes how "views from above . . . allowed military officers to locate troop movements, camouflaged artillery, and the trenches at a glance" (8). Thus, aerial photography provided a new way of seeing, "a distanced, holistic outsider's perspective," which laid the ground for new techniques of observation (14). With the view from above, then, a synthesis between the surveillant and the military gaze takes place, or perhaps rather a militarization of the surveillance gaze.

In the current era, yet another new way of seeing is conditioned. It is the disembodied and remote-controlled gaze from above. Returning to Steyerl, she observes that this reorientation of the surveillance gaze is enabled by new technologies such as the satellite or drone and Google map views. Today, Steyerl argues, the linear perspective as we know it from traditional art history is increasingly replaced by the aerial perspective, as new technologies of surveillance, tracking, and targeting emphasize and normalize the god's-eye view from above ("In Free Fall"). This gaze is intimately linked to warfare. The shift in perspective noted by Steyerl replaces the stable and single point of view of the linear perspective, which was tied to a body, with a dehumanized "perspective of overview and surveillance for a distanced, superior spectator safely floating up in the air . . . which establishes a new visual normality—a new subjectivity safely folded into surveillance technology and screen-based distraction" ("In Free Fall" 8). Moreover, this emergent subjectivity should be seen in conjunction with the strong link between vision and domination, and with the argument that modern surveillance technologies operate on "the masculine and controlling end of the gender spectrum," primarily through disembodied control from a distance (Monahan 113). The subjectivity under construct, then, is the subjectivity of the age of remote warfare: a detached, observant gaze enabled by new technologies which is inherently militaristic.

Mapping, Searching, Recognizing

Returning to *Modern Escape*, the entanglement of warfare, surveillance, and the technologies of the Western middle-class home is further established through the work's formal and thematic structure of mapping and tracking practices. In fact, the association between the ordinary everyday object of the robotic vacuum cleaner and the military industrial complex is established right away, in the opening scene where the shape of the vacuum cleaner's path planner overlaying the image resembles a helicopter, a well-known trope of war in visual culture (see fig. 2). The path planner overwrites the video in three places, "eating" itself into the image, as the robot is searching, charting, and mapping the home, effectively blurring the boundaries between the commonplace vacuum cleaner and the technologies of reconnaissance, tracking, and tracing. Another blurring of boundaries appears when the TV shows a geomap image of a city, a satellite view which is soon overlaid by the grid of the vacuum cleaner. The scene alludes to aerial reconnaissance and computer vision technologies, and the green square further overlaying the geomap is familiar from object and facial recognition applications from, e.g., smart phones and social media apps (see fig. 7).

The mutually informative relationship between the military industrial complex and the private tech companies of Silicon Valley, and, by extension, the consumers of everyday technologies, evoked here leaves a sticky sense of complicity lingering in the networked environment. A recent example of this entanglement is Project Maven, a U.S. military initiative to integrate artificial intelligence into battlefield technology. In March 2018, it became known that Google was collaborating with the U.S. Department of Defense on the project with the objective of building an AI surveillance imagery analysis system (Fang, "Google is Quietly"). A key objective was to develop AI enabled image-recognition software for analyzing video feeds

captured by drones to improve drone strikes on the battlefield. While strong protest from employees at Google for staying out of the "business of war" led the company to—at least officially—withdraw from renewing the contract, Google's collaboration with DoD arguably casts a shadow on its AI image recognition products for the commercial market (D'Onfro).[3]

Fig. 7. Installation view, Nielsen and Johnsen, *Modern Escape*. Photograph by Anders Sune Berg. All images courtesy of the artists.

The promise of the home surveillance system *Google Nest*, for example, which features "familiar face alerts" to notify users' smart phones whether a camera detects a family member or a stranger, might gain a troubling undertone: "Your home, safe and sound. Google helps you look out for your family, day and night" (Google). Similarly, in the case of drones for civilian

[3] However, *The Intercept* reports that Google never fully renounced work with the DoD. Rather the collaboration continues through a venture capital arm, Gradient Ventures, which provides access to Google's AI training data (Fang, "Google Continues").

versus military use, Daniela Agostinho et al. makes it clear that "any differentiation between military and nonmilitary drones is contentious, if not impossible, insofar as the development and logic of each technology informs the other" (253–4).

As my reading of *Modern Escape* suggests, the artwork foregrounds how the violence of war haunts the surveillance technologies of everyday life. It is a matter of what sticks to technologies and systems of surveillance—to its haunting historical continuities and present complicity—no matter how smoothly and seemingly innocent they integrate into the home and our everyday lives. What mediates and co-creates the haunting in *Modern Escape*, I argue, is the automated vision of the surveillance gaze and the robotic vacuum cleaner's mapping and surveying technologies, both embedded in an inheritance of militarism, imperatives for control, and oppressive power relations.

The Haunted House

"We are haunted by something we have been involved in," Avery Gordon writes (*Ghostly Matters* 51), and as my reading of *Modern Escape* has shown, we are haunted by something we are still involved in. To be haunted is, as noted in the introduction, "to be tied to historical and social effects" (190). In *Modern Escape* the entanglements between everyday technologies and practices of surveillance and the military-industrial complex, national security, and the recast colonial binarism of post-9/11 politics (Hall) comes to the fore through the estranged home, "that quintessential space of the uncanny, the haunted house" (Gordon, *Ghostly Matters* 50). Sigmund Freud proposed that the category of sensations or feelings of the fearful and frightening belonging to the uncanny (*das Unheimliche*) is a matter of "that species of the frightening that goes back to what was once well known and had long been familiar" (124). Along the same lines, the etymological connection between haunting and home can be traced back to one origin

of the English word *haunt* which is likely to derive from the Old Norse *heimta*: "to bring home," or "to lead home" (*Online Etymology Dictionary*; *Merriam-Webster*; Heggstad, Hødnebø, and Simensen).[4]

Fig. 8. Still from film, Nielsen and Johnsen, *Modern Escape.* All images courtesy of the artists.

In *Modern Escape,* the warfare and suffering "elsewhere" are brought back to the affluent Western home as a troubling reminder of modernity's inheritance of "violence and wounds" as well as current politics of exploitation. In a critique of Freud, Gordon maintains that the unconscious is collective, rather than individual, and she argues that the uncanny is fundamentally a social matter—uncanny experiences are hauntings of worldly contacts, "in the world of common reality" (*Ghostly Matters* 54). Returning to the old Norse *heimta*, there is a

[4] This connection comes about via the Old French *hanter*. However, the etymological derivation from old Norse is a subject of ongoing academic discussion (*Merriam-Webster*; *Online Etymology Dictionary*).

second meaning of the word, which is "a claim" or "what one has to demand" (Heggstad, Hødnebø, and Simensen). This second meaning opens up a sense of inescapability that someone might come to collect their due share, and I suggest that this repressed knowledge, our lingering complicity in past and present social wrongs, contribute to the paranoia and sense of haunting in *Modern Escape*. As an affective experience, the artwork induces an atmosphere of surveillance haunted by the violence and power relations in which we are implicated, yet which are often kept out of sight. As Gordon further observes, haunting describes "those singular and yet repetitive instances when home becomes unfamiliar, when your bearings on the world lose direction, when the over-and-done-with comes alive, when what's been in your blind field comes into view" (*Ghostly Matters* 2). My closing argument in this reading is that *Modern Escape* achieves exactly this. The work calls forth the connection between the Western middle class and what is kept out—out of the home, out of the privileged Global North, and out of the collective Western consciousness. It displays the close entanglements between our everyday networked technologies and those of war and the increasing efforts needed to barricade the home as well as the Western world. In the work, a layer of suggested complicity is added as the spectator suddenly becomes aware of her own watching presence when the lamp turns on and the automatic vacuum cleaner starts moving close by. In this way, the work questions the notion of the technologically-connected surveillance home and the spatial tensions between inside and outside, home and world, work and spectator.

Works Cited

Agostinho, Daniela, Kathrin Maurer & Kristin Veel. "Introduction to The Sensorial Experience of the Drone." *The Senses and Society*, vol. 15, no.3, 2020, pp. 251–58, https://doi.org/10.1080/17458927.2020.1820195.

Bratton, Benjamin. "Machine Vision: Benjamin Bratton in conversation with Mike Pepi and Marvin Jordan." *DIS Magazine*, http://dismagazine.com/discussion/73272/benjamin-bratton-machine-vision/. Accessed 5 Aug. 2020.

Davidsen, Camilla. "Ugens kunstner: Hanne Nielsen & Birgit Johnsen." *Kunsten.nu*, 13 Nov. 2019, https://kunsten.nu/journal/ugens-kunstner-hanne-nielsen-birgit-johnsen/. Accessed 30 Jun. 2021.

D'Onfro, Jillian. "Google will not renew controversial Pentagon contract, cloud leader Diane Greene tells employees." *CNBC,* 1 Jun. 2018, https://www.cnbc.com/2018/06/01/google-will-not-renew-a-controversial-pentagon-contract.html. Accessed 28 Sep. 2021.

Enloe, Cynthia. *The Curious Feminist: Searching for Women in a New Age of Empire*. University of California Press, 2004.

"Haunt." *Merriam-Webster*, https://www.merriam-webster.com/dictionary/haunt. Accessed 6 Jul. 2020.

"Haunt." *Online Etymology Dictionary*, 2020, https://www.etymonline.com/word/haunt. Accessed 29 Apr. 2020.

Fang, Lee. "Google Continues Investments in Military and AI Technology Through Venture Capital Arms." *The Intercept*, 23 Jul. 2019, https://theintercept.com/2019/07/23/google-ai-gradient-ventures/. Accessed 23 Sep. 2021.

---. "Google is Quietly Providing AI Technology for Drone Strike Targeting Project." *The Intercept*, 6 Mar. 2018, https://theintercept.com/2018/03/06/google-is-quietly-providing-ai-technology-for-drone-strike-targeting-project/. Accessed 23 Sep. 2021.

Freud, Sigmund. "The Uncanny." 1919. Translated by David Mclintock, Penguin Books, 2003.

Google. "Google Nest Help: Familiar face detection and how to manage your library." 2020, https://support.google.com/googlenest/answer/9268625?co=GENIE.Platform%3DAndroid&hl=en. Accessed 10 Dec. 2020.

Gordon, Avery. *Ghostly Matters: Hauntings and the Sociological Imagination.* 1997. University of Minnesota Press, 2008.

---. "Some Thoughts on Haunting and Futurity." *Borderlands*, vol. 10, no. 2, 2011, pp. 1–21.

Hall, Rachel. "Terror and the Female Grotesque: Introducing Full-Body Scanners to U.S. Airports." *Feminist Surveillance Studies*, edited by Shoshana A. Magnet and Rachel E. Dubrofsky, Duke UP, 2015, pp. 127–49.

Haffner, Jeanne. *The View from Above: The Science of Social Space*. MIT Press, 2013.

Heggstad, Leiv, Finn Hødnebø, and Erik Simensen, eds. *Norrøn Ordbok*. Det Norske Samlaget, 1975.

Lighthouse. "Press release: Introducing the Lighthouse Interactive Assistant: *Company Brings Advances in AI to the Home.*" 11 May 2017, www.light.house. Web archive captured by author 18 Dec. 2018.

Lyon, David, ed. *Surveillance as Social Sorting: Privacy, Risk, and Digital Discrimination*. Routledge, 2003.

Michel, Arthur Holland. *Eyes in the Sky: The Secret Rise of Gorgon Stare and How It Will Watch Us All*. Houghton Mifflin Harcourt, 2019.

Monahan, Torin. *Surveillance in the Time of Insecurity*. Ruthers UP, 2010.

Mori, Masahiro. “The Uncanny Valley.” Translated by Karl F. MacDorman and Norri Kageki. *IEEE Robotics & Automation Magazine*, vol. 19, no. 2, 2012, pp. 98-100, doi:10.1109/MRA.2012.2192811.

Morrison. Elise. *Discipline and Desire: Surveillance Technologies in Performance*. University of Michigan Press, 2016.

Nielsen, Hanne, and Birgit Johnsen. *Modern Escape*. 2018. One channel color video installation, sound, podium, vacuum cleaner, objects. 21 min 12 sec. Exhibited in *REVISIT: HANNE NIELSEN & BIRGIT JOHNSEN*, Overgaden Institute for Contemporary Art, Copenhagen, Denmark, 10 Nov. 2018–6 Jan. 2019.

---. Personal communication via email. May 2019.

---. Personal interview. Aarhus, Denmark, 19 Oct. 2019.

Ring, Annie. “Complicity”. *Uncertain Archives: Critical Keywords for Big Data*. Edited by Nanna Bonde Thylstrup, Daniela Agostinho, Annie Ring, Catherine D'Ignazio, and Kristin Veel. MIT Press, 2021, pp. 87–97.

Rothberg, Michael. *The Implicated Subject: Beyond Victims and Perpetrators*. Stanford UP, 2019.

Smith, Andrea. “Not Seeing: State Surveillance, Settler Colonialism, and Gender Violence.” *Feminist Surveillance Studies*, edited by Shoshana A. Magnet and Rachel E. Dubrofsky, Duke UP, 2015, pp. 21–38.

Steyerl, Hito. “In Free Fall: A Thought Experiment on Vertical Perspective.” *e-flux Journal*, no. 24, April 2011, pp. 1–11, https://www.e-flux.com/journal/24/67860/in-free-fall-a-thought-experiment-on-vertical-perspective/. Accessed 19 Mar. 2022.

---. *Duty-Free Art: Art in the Age of Planetary Civil War.* Verso, 2017.

Sontag, Susan. *On Photography*. 1977. Penguin Books Limited, 1979.

Wolthers, Louise. "The Inappropriate." *Inklusion/Ekslusion Exclusion/Inclusion*, edited by Hanne Nielsen and Birgit Johnsen, Aros, 2014, pp. 169–97.

Woolf, Virginia. *To the Lighthouse.* 1927. Projekt Gutenberg Australia eBook, 2001, http://gutenberg.net.au/ebooks01/0100101h.html.

Zylinska, Joanna. *Nonhuman Photography*. MIT Press, 2017.

Karen Louise Grova Søilen recently received her PhD on the topic *Atmospheres of Surveillance* (University of Copenhagen 2021). Her research interests center on surveillance perceived as bodily, emotional, and multisensory experiences, surveillance in contemporary art and culture, nonrepresentational and feminist methodologies, cultural theory, and the changing notions of the home through the affordances of smart technologies. She is recipient of Surveillance Studies Network's 2020 *Early Career Researcher Award*, and her work is published in *Surveillance & Society* and *NTIK: Nordic Journal of Information Science and Cultural Mediation*.
Email: karenlouise.soilen@gmail.com.

The Journal of Media Art Study and Theory
Volume 3, Issue 1, 2022
Automating Visuality
www.mast-journal.org
MAST

Calm Surveillance in the Leaky Home: Living with a Robot Vacuum Cleaner

Kassandra Wellendorf, Karen Louise Grova Søilen, and Kristin Veel
University of Copenhagen

Abstract

Understanding the attachment owners can feel to their robot vacuums, which also map and collect data about their homes, is key to understanding the ambivalences involved in the integration of automated visualities in the home. Drawing on qualitative video interviews and observations of people interacting with their robot vacuums, this article identifies three key factors in understanding how cohabitation with a robot vacuum and its particular form of automated sensoria is experienced by its user: firstly, the robot assists with work that we would otherwise do ourselves with the aid of a broom or traditional vacuum cleaner; it is thus often regarded as an extension of ourselves, the equivalent of a cleaning assistant, or even a kind of pet with which you can interact. Secondly, its ability to move autonomously increases the inclination to anthropomorphize the robot as a being with some level of agency and intelligence. Thirdly, the robot vacuum cleaner is a very visible part of the intimate sphere. It has its charging station in the home; it cannot be hidden away in a cupboard like an ordinary vacuum cleaner; more often than not, furniture needs to be moved around for it to run smoothly. This article argues that these three factors are important for understanding people's difficulty in perceiving the robot as an entity that potentially participates in surveillance practices, and to understand the nature of this form of surveillance that emanates from the leaky home.

Keywords

robot vacuum; everyday culture; surveillance; leaky home; automated vision; smart home; cleaning habits

In the course of the twentieth century, the labor of looking involved in surveillance has become automated, and, in that process, the visuality of surveillance has been extended to encompass other sensoria. Surveillance today maps, calculates, predicts, and preempts—at the same time as more and more connected devices enter the home. By 2030 it is estimated that there will be 125 billion connected Internet of Things (IoT) devices worldwide (IHS Markit). Such devices include Internet-connected locks, interactive assistants with video and audio recording, surveillance systems, and sensors that chart humidity, heat, and light. They also include the newer generation of robot vacuum cleaners that are Wi-Fi enabled and can be controlled via one's smartphone or through speakers in the home such as Amazon Echo or Google Nest. They come with different forms of automated vision, linked to their navigation system, ranging from cameras to lidar technology. Such automation renders the home "leaky" in the sense that they permeate it and open it up to remote control. Connected devices developed to assist with mundane everyday chores, such as the robot vacuum, thus walk a fine line between being helpful assistants and being surveyors. Indeed, some brands market their products explicitly on their surveillance capabilities. For instance, Trifo's Lucy is described as "Super sensing. I'm always aware of my surroundings. I scan and detect the smallest obstacles to avoid. I quickly identify people, pets, furniture, and even items down to an inch!" (Trifo). However, even those models that do not emphasize surveillance features in their marketing may be repurposed for surveillance. In 2020, researchers from Singapore University and the University of Maryland demonstrated that they could make a Xiaomi Roborock vacuum cleaning robot eavesdrop on conversations taking place

in the room where the robot was located (Sami et al.). Although brands such as the MIT-based iRobot, with its iconic Roomba, have been consistently vocal about the precautions they take to secure users' data (iRobot, "Data Security"; Astor), most models today map your home in order to be able to do the vacuuming. As an advanced documentation of an automated visuality that transgresses the visual faculty, this map may hold valuable information about you as a consumer, ripe for exploitation by marketing. Indeed, a large survey carried out in Australia, Canada, France, Japan, the UK, and the US found that "63% of people surveyed find connected devices 'creepy' in the way they collect data about people and their behaviours" (Consumers International and Internet Society 2) and "28% of people who do not own and do not intend to purchase a connected device make this decision because of lack of trust in security and privacy" (8).

Drawing on qualitative video interviews and observations of people interacting with robot vacuum cleaners, this article explores the robot's dual movement: on one hand, trotting habitually across the floorboards as a trusted companion that is deemed most benign when it fails to "see" obstacles in its way; on the other, the robot's ability to perform mapping of its own route that is able to be "seen" by the robot itself and thus potentially also by others. This dual perspective is at the heart of the paradox of the machine: part-benign pet whose sensory capabilities are regarded as inferior in comparison with that of its owner, part-surveillant intruder that sees more than its keeper. To unfold this paradox, we need to start by looking at the properties of the scene where these interactions take place: the home.

The Leaky Home

In this article we argue that the robot vacuum cleaner can function as a vehicle to address the wider implications of what we propose to call "leaky homes," i.e., automated homes that leak

through the connected devices and sensors that collect, transmit, receive, and share data (Sofia). These homes leak beyond their traditional architectural and spatial boundaries in so far as they are accessible from afar, with mobile phones often functioning as remote controls for the technologies in the home. Scholarship in surveillance, digital media, and cultural studies has addressed how the home's concrete walls are superseded by "incorporeal informational barriers that continuously monitor and document" (Rapoport 326) and how the connected home's smart sensors, interactive assistants, and networked machines are sources of extractable information and data. Moreover, as Tanja Wiehn aptly remarks with regard to digital infrastructures more broadly: "Digital infrastructures we encounter today and maintain our interpersonal relations with are not bound to binary categories of public and private; On the contrary, they thrive and expand on leakage, but nevertheless invoke the impression of containment" (62). This sense of containment and privacy can make users even more vulnerable because they believe that their information is protected in ways that often do not correspond with reality (Agostinho and Thylstrup 763–4). In the home, this sense of security and containment is particularly prevalent because the home traditionally embodies connotations of containment and shelter. In Western liberal traditions, the home is usually conceptualized as a shelter, an enclosed space of privacy and retreat. Classic phenomenological writings posit the home as a site for "dwelling" (Heidegger 145–61) and the house as a shelter for daydreaming that "protects the dreamer" (Bachelard 6), defined by qualities of safekeeping. Significantly, it is the trope of the home-as-shelter that the marketing of IoT technologies in the home often stresses. However, as feminist critique has long noted, conceptions of the home as a shelter contrast starkly with the experiences of women, for whom the home through the ages has been a site of labor—of tending to children, household chores, and family members—not to speak of those that are victims of domestic

abuse, who need shelter *from* the home (Dobash and Dobash). The experience of the home as a place of safety is thus both illusory and exclusionary, and we need to pay attention to this longer cultural history when conceptualizing the properties of the leaky home today. In this article, we contend that the conception of the home as a contained shelter secured by technology needs to be qualified if we are to understand the calm surveillance dynamics involved in the automation of home devices that "see" the home in different ways than humans because (as our small-scale empirical study indicates) technologies are not necessarily at their most invasive when they are invisible. Rather, their noisiness, blatant clumsiness, and the fact that we need to move obstacles for them may cause us to see them as benign when perhaps we should not.

"Well, There is Nothing Exciting Here to Record"

Influenced by ethnographic methods found in work on human-computer interaction (HCI), human-robot collaboration (HRC), and design anthropology, we conducted video interviews and observations of eight people in five different households.[1] The people that opened their homes to us all live in the greater Copenhagen area in Denmark. Stella (26) is a recent graduate in comparative literature; her partner Rune (36) is a system developer and administrator; they live in an apartment in central Copenhagen. Pernille (44) works in publishing, and Jesper (45) works for the postal service; they both live with Pernille's twin daughters in a detached house in a residential area 25 km north of Copenhagen. Mia (47) works with children with brain impairments and lives with her three-year-old son in an apartment north of Copenhagen. Marianne (72) is a retired nursery school head; she lives in an apartment in the north-western part of Copenhagen during the winter, and in a small house in her allotment garden over the

[1] When we make use of observations and interviews we use only first names. At the end of the reference list, there is a list of video interviews and observations with specific dates.

summer. Moreover, our study springs from auto-ethnographic reflections, and the informants therefore include co-author Kristin (40) and her engineer husband Rasmus (37), who live with their two-year-old daughter in a house in a forested area north of Copenhagen.

Our qualitative interviews are in line with other ethnographic studies that focus on how people project intelligence onto machines (Fink et al.; Taylor) and how robot vacuum cleaners are gendered (Sung et al., "Housewives or Technofiles?" 133–4; Strengers and Kennedy 23–48), as well as a large group of ethnographic HCI and HRC studies intended to enhance the design of robot vacuum cleaners (Sung et al., "Domestic Robot Ecology"; Vaussard et al.; Forlizzi; Forlizzi and DiSalvo; Hendriks et al.). We draw on these studies to support the observations we have made in our own small-scale fieldwork in order to shed light on the ways in which users experience the particular form of automated sensoria that the robot vacuum represents from a cultural studies perspective. However, our approach is inspired by sensory ethnography and design research (Pink; Pink et al. 99–105) and involves a combination of semi-structured qualitative video interviews and "home video tours" where our informants re-enact cleaning routines with their robot vacuum cleaners while they talk us through what they are doing and how they imagine their robot vacuum's capacity to move, see, and navigate. These kinds of home video tours, Sarah Pink et al. argue, can help to raise reflections on habitual everyday actions and allow researchers to access those reflections (108–12). Moreover, we expand on this methodology by focusing on the microanalysis of informants' body language when they speak about and show us the functions of their robot vacuums. We draw here on a filmic methodology inspired by Harold Garfinkel's ethnomethodology, which is developed by Kassandra Wellendorf, where camera recordings allow us to observe in detail the small nuances in informants' gestures, body language, choice of words, etc. This method allows us to gain insight into the everyday

negotiations and interpretations that these people perform when interacting with the automated sensorium of the robot vacuum.

The main purpose of our interviews was not to map people's sentiments about the surveillance implications of their robot vacuum cleaners. However, this issue came up in every interview. Mia does not consider her robot vacuum a potential source of surveillance, and this was not a concern for her when she was buying the robot. When we talk to her, she is unsure whether the model she bought has a camera, but her stance is anyhow that her life is far too boring to be of any interest for surveillance: "Well, there is nothing exciting here to record. I would rather log into something else, the royal family or something more exciting than my dust bunnies." She thereby displays a classic response to surveillance of "nothing to hide and nothing to fear," which has been particularly prevalent with regard to CCTV cameras and other kinds of overt surveillance (Solove). What is at stake here is the conception that surveillance is only surveillance if it is individualized; systematic data-gathering on a metadata level to identify consumer groups of "people-like-you" (which is an intrinsic part of the automated visuality of surveillance that we are interested in here) is seen as less problematic (van Dijck 200).

Marianne reveals a similar stance. When prompted, she explicitly states that she has nothing to hide, although she adds that she is aware that her phone might be listening in on her and that what she Googles has an effect on the advertisements that appear on her computer. It seems, however, that these concerns do not affect her everyday use of or sentiments about the robot vacuum. Her enthusiasm for the robot and its ability to make her life easier is predominant: "I think that it is creepy that they can . . . listen in, I actually think that it is really creepy, but . . . I also forget it. It is only when you remind me that I think about it, but otherwise I don't really think about it every day."

The youngest and most tech-savvy couple, Stella and Rune, have the strongest concerns. Rune works as a developer and systems administrator at a small company that specializes in networks for IoT devices. He has set up a main network with several subnetworks in their apartment, and he is keen to isolate the TV and robot vacuum and to constrain their access to anything beyond the absolutely necessary, in order to avoid hijacking and potential infection by malware. Stella's arguments focus less on the technical aspects, and she displays a more affective relationship to the potential risk. She explains that she would feel uncomfortable if the robot had an embedded camera. Stella regards the robot they acquired as "less invasive" than she had expected and links this with what she perceives as its disappointing navigation skills, which she accepts in order to avoid feeling under surveillance.

These statements on the surveillance potential of the robot vacuum point to a constellation of issues that we have observed in the material, to do with: 1) the robot vacuum's integration into the household, 2) its navigation skills and perceived intelligence due to imagined sensory capacity, and 3) the attention the robot vacuum calls to itself as a smart technology, especially when it does not function correctly and gets stuck on carpets or under chairs. These factors all contribute to the robot owners paying less attention to the potential for data leakage once the machine enters their home. It is these three factors and what they have to say about automated visuality that we shall now turn to.

An Automated Member of the Household

Our perceptions of robot vacuum cleaners' intelligence and agency is key to how we relate to them. In *The New Breed: What Our History with Animals Reveals about Our Future with Robots*, robot ethics scholar Kate Darling writes about humans' relationships with different kinds of robots, arguing: "Even very simple robots, when they move around with 'purpose,' elicit an

inordinate amount of projection from the humans they encounter" (100). According to Darling, animal detection capabilities that are biologically hardwired into our bodies are activated when objects move on their own, and we perceive such objects as agents (101). This may explain why so many owners of robot vacuums give them names and talk about them as if they were household members, projecting human sensory capabilities onto the devices.

This corresponds well with our findings. Mia compares the robot vacuum to the cleaner she employed before she bought it, but she also differentiates between what she calls "*my* vacuum cleaner"—her traditional vacuum, for which she uses a possessive pronoun signifying a sense of ownership—and the robot vacuum, which she has named Roberta as if it were something or someone in its own right. Moreover, her three-year-old son treats it like a pet, following it around and stroking it.

Others create more elaborate narratives around their robots. When asked to describe their robot vacuum, Pernille and Jesper explain that it is called Preben and is gendered with the pronoun "he." They say that he is "quite human" and that they talk about him in an appreciative manner when he has completed a task. They also volunteer the information that he is having an unhappy love affair with their robot lawnmower, who is called Bertha: "They drive around looking at each other, but—naturally—they never meet." When prompted on how the devices "see," they remark that evidently the robots cannot see each other, as they do not have eyes. Nonetheless, they continue describing the robot vacuum as if it is able to see.

While Pernille and Jesper thus seem to have embraced the projection of agency and happily tell the story of Preben and Bertha who keep an eye on each other through the window, Marianne affirms (when asked) that she does not regard her robot vacuum as intelligent and does not have any kind of affective relationship with it. She describes it as a tool and compares it to a

food processor or dishwasher that makes everyday life easier; she is careful to assign intelligence not to the machine but to the people who designed it. Nonetheless, although she mostly refers to it with the gender-neutral "it," she has named the robot Robert, after a cleaner she employed before she acquired it. She also describes the process of acquiring and setting up the robot as having been akin to a battle of wills which she reenacts with her feet when telling us about it: "It constantly wanted to go over the doorsills, and I wanted to determine where it should vacuum. We talked a bit about that, Robert and me. But I won, because I just shut the door, and then it just does one room at a time." For Marianne there seems to be a sense of resistance to the anthropomorphization of the device. She tells us that she does not want to come across as a "crazy old person" who does not relate to it as a pure object. Yet during the interview, it repeatedly becomes clear that this perception of the robot is difficult to maintain, especially when it moves around.

Moreover, for Marianne, her own and the robot vacuum's sensory capabilities intermingle when it comes to detecting dirt. Her own way of checking if it is time to vacuum is when she feels crumbs underneath her toes. After she starts the robot vacuum she repeatedly empties its dust chamber to check how much dirt it finds. Only when she sees that it is completely empty does she feel convinced she can stop the cleaning process.[2] We may argue that she uses the robot as an automated extension of her eyes and feet while at the same time she monitors it with her own senses by peeking into the dust chamber to determine when it is done.

[2] The primary trick of traditional vacuum sellers has often been to show how their product is able to detect and eliminate more dust than the average vacuum cleaner (Scott 222). As part of a demonstration, it was common to bring a pile of dust to pour out on the carpet, after which the housewife was asked to hoover with her own vacuum cleaner. When the carpet looks clean, the salesman would repeat the hoovering with his model. He then shows a full dust chamber as proof of how this model is able to suck up debris that is not visible to the human eye.

For Marianne, Mia, her three-year-old son, Pernille, and Jesper, in different ways, their relations with their robot vacuums involve the projection of agency and sensorial capabilities onto the machine. It is another being onto which character traits can be projected or with whom conversations can be had. The automation of tasks we might do ourselves here results in anthropomorphizing the device. With Rasmus we find another way of engaging with the machine in so far as he uses his own body to explain the robot's functionality, the obstacles it encounters, and its comparative shortcomings. The vocabulary he uses also anthropomorphizes the machine—he talks about its "mouth" and "legs," and he uses body language to exemplify it with his own body, thereby displaying a form of identification with the machine. Rune wriggles his bottom to demonstrate how the robot has problems entering the docking station, and he makes choking, guttural sounds when he talks about how it gets stuck on the fringes of the carpet. He is keenly aware that he is anthropomorphizing the machine and repeatedly comments on the fact, in this way maintaining an intellectual distance to the way in which his body language keeps displaying identification.

What is particularly striking in these examples is that both the anthropomorphizing and the identification are almost always connected to the robot vacuum's movements and capacity of seeing or sensing its surroundings, which is predicated on its navigation system. This points to a need to further understand the perceived correlation between intelligence and navigation.

What the Robot Vacuum "Sees"

Several of our respondents remark that sensing dirt underneath their bare feet works as a repulsive indicator of uncleanness. For instance, Rasmus distinguishes between "visible dirt" that blocks the view of a surface and "invisible dirt" that is not seen but which can be sensed underneath your feet. But how does the robot vacuum help us detect dirt?

The iconic Roomba, launched in 2002, did not map its surroundings until its seventh generation. Instead, it operated through randomized navigation, which meant that it moved in a random direction whenever it bumped into a wall or an object, thereby eventually covering the whole floor. At the time this was more efficient than its competitor, the now-discontinued Electrolux Trilobite, which used ultrasonic sensors to map obstacles. Today most models use different kinds of sensors, cameras, or laser navigation to create a floor plan, which is stored and can often be accessed remotely on a mobile phone, showing the robot's route, mapping its trail. Moreover, the Roomba's patented "dirt detect technology," implemented in some models in 2004, means that if its sensors detect more particles in a given area, it will pass over that spot several times (iRobot, "What is Dirt Detect™ Technology?"). However, observing a robotic vacuum cleaner for longer stretches of time almost inevitably affords a sense of puzzlement. What patterns, invisible to the human eye, does it detect on the floor? What rationales determine its choreography? This may prompt the eerie feeling of an unintelligible logic at play, or alternatively the outright dismissal of its intelligence—both of which, in different ways, point to how automated machinic vision and the intelligence behind it is imagined.

When we ask Mia if the robot vacuum is intelligent, she responds: "I suppose so, because it can find its way around." But then she goes on to describe the seemingly illogical routes it takes: "It does not take the smartest or most intelligent route, but I suppose there is a purpose to it." She has confidence that "some smart people have calculated that this is the best way and have programmed it for that." Nonetheless, for her, the robot's apparently irrational navigation is a sign that it lacks intelligence in the human sense of the word, which she also finds reassuring: "It cannot think for itself. I don't want it to take over my home." She would not want it to be more intelligent than it is: "It is me that pushes a button and says it should start. I don't want it to

be able to register anything itself. It doesn't need to be able to do that. I prefer to be the one in control." This aligns with a 2011 user design study that investigated how people preferred their robot vacuums to behave, which concluded: "People prefer a calm, polite, and cooperative robot vacuum cleaner that works efficiently, systematically and likes routines" (Hendriks et al. 194). Moreover, an empirical study from 2014 contended that people wished to understand how the robot navigated because this gave them the sensation of being in control: "Most households were skeptical about the Roomba's random path, and one mother expressed her disappointment: 'How does it decide where it goes? It is stupid, it does not see where the dirt is, it always moves away from it!'" (Vaussard et al. 386). This resonates heavily with our findings. For most of our informants, the robot vacuum's ability to navigate and employ an automated sensorium seems to be a measure of its intelligence. Jesper and Pernille find it strange that Preben is unable to find the shortest route home to his docking station. They are able to see the docking station right behind him, but since he does not orient himself in the room in the same way as they do, he often takes a detour. Automating visuality here results in a sense of superiority in relation to the robot. The robot's navigation is also a mystery to Marianne:

> A robot vacuum has its own life and moves as it pleases. And I cannot figure it out. It must have something coded so it runs diagonally and then straight and then diagonally, and if I think that it needs to clean more thoroughly in a place, then I can block its route.

The robot's navigation is thus a sign both of the machine's enigmatic nature—it "has its own life"—and of Marianne's ability to outsmart it and take control. However, those that anthropomorphize the robot to a greater extent, such as Pernille and Jesper, seem to have a greater tolerance for behaviors that appear irrational to the human sensorium. This aligns with what social psychologist Sherry Turkle refers to as the ELIZA effect, i.e., "that desire to cover

for a robot in order to make it seem more competent than it actually is" (131). Although the robot vacuum is probably less of a black box to Rune, who is more tech-savvy than our other informants, he also displays tolerance for its inadequacy and finds it amusing to watch the robot "puttering around" the home. He explains that when it enters a room, it stops to orient itself, looking a bit confused, before resuming its cleaning job, which he finds hilarious. Yet Stella and Rune are also among the most concerned about privacy: they consciously chose a model that is intelligent enough to do the job but not so intelligent that it makes them feel surveilled. This may be another reason why they accept its shortcomings, although its tendency to keep getting stuck under the same section of pipe is a cause of annoyance. The robot vacuum's flawed ability to "see" is thus at the same time a cause of concern and what renders it a benign pet that can be regarded as part of the family.

Calm Surveillance

Despite the wish for "calm, polite, and cooperative" robots expressed in the 2011 study mentioned above (Hendriks et al. 194), and despite what many of our informants tell us about their puzzlement at the robot vacuum's movement patterns, it is apparent from our small-scale study that living with a robot vacuum cleaner is anything but a calm experience. It often requires a considerable reorganization of the home. Kristin and Rasmus had a big debate when they acquired the robot vacuum about whether their long Moroccan felt carpets would be compatible with the robot or would have to be removed; they were eventually put into storage because it was too much work to move them every time the robot vacuum was started. Pernille and Jesper usually start the robot vacuum when they leave the house because it is noisy. In an interview we conducted with Stella before she and Rune acquired their robot vacuum cleaner, she explained that she would have no patience with remodeling her home to accommodate the machine: "If I

have to pay thousands of kroner for an intelligent tool that can clean for me, then I won't bother with having to move everything around." However, when we visit her later, after they have had the robot vacuum for a couple of months, they have set up pieces of cardboard to stop it from getting stuck under a pipe, and they have arranged their shelves so that they do not touch the floor, to allow the robot to move beneath them. They have also fixed some loose cords so the robot will not get tangled up, and Rune has installed "no-go lines" around their carpets so that it will not get stuck on them. Stella explains (while laughing) that the robot needed two weeks of "house-training" before they could put it to work on its own because it would get stuck on their carpets.

Living with a robot vacuum cleaner thus requires its own set of work on the cohabitants' part to clear the way for the robot to be able to move as smoothly and unobtrusively as possible. Its presence is anything but "calm" in the way computer scientist Mark Weiser imagined that technologies in the home would become. In his seminal 1996 text "Open House," he proclaimed:

> Over the next twenty years computers will inhabit the most trivial things: clothes labels (to track washing), coffee cups (to alert cleaning staff to moldy cups), light switches (to save energy if no one is in the room), and pencils (to digitize everything we draw). In such a world, we must dwell with computers, not just interact with them.

For Weiser, dwelling means a comfortable cohabitation that is unobtrusive, rendering automation invisible. Our interviews suggest that the robot vacuum cleaner is rarely an invisible device that performs its function without calling attention to itself. Rather, it takes on a life of its own. Paradoxically, this noisy presence is part of the reason why our informants pay less attention to the machine's surveillance potential: its ability to move and its clumsiness contribute to the impetus not only to anthropomorphize it but also to regard it as less intelligent and less of a

threat. The capabilities of its navigation system and how the owners imagine its sensorium are central to how intelligent it is perceived to be, and the robot vacuum's helplessness often also reassures its cohabitants that it does not have the ability to surveil them. This means that even if the presence of the robot vacuum is anything but calm, paradoxically the form of surveillance it represents *is* calm, albeit in a different understanding of the word than what Weiser sets out above. It is a form of surveillance that we may fail to notice precisely *because* the robot vacuum does not blend into the surroundings. In our data-saturated everyday lives, we expect surveillance either to be overt and explicit, as with CCTV cameras or baby monitors, or to take the form of silent, ubiquitous data-tracking (Steiner and Veel). The disturbing presence of a robot vacuum that we regard as a member of the household somehow falls in between these categories.

Going back to the notion of the leaky home that we set out in the beginning of this article, we now see how the example of cohabitation with a robot vacuum that we have explored shows how the cultural trope of the home-as-shelter is challenged in new ways by the technology-saturated leaky home. Here the home-as-shelter trope breaks down not so much from the tedium of housework or the threat of domestic abuse, as feminist critique has historically argued, but rather through the seemingly benign calm surveillance of the robot's automated gaze. The automated home is perceived not only by us but also by the connected devices that inhabit the same space as us, in this way harboring, on one hand, an increased level of surveillance potential and, on the other hand, an affective engagement with non-human agents which causes us to pay less attention to the former. Most IoT and artificial intelligence devices are designed to make people integrate and accept them rather than raising privacy or security concerns. For that very reason, a wealth of insight is available from looking at people's body language, vocabulary, and concrete interactions, as we have done in our interviews. Further research may expand our

findings to consider the implications of other automated devices in the home, such as Alexa, Google Nest, smart locks, or thermostats. However, as our small-scale fieldwork shows, the robot vacuum cleaner is a particularly interesting case study because of the autonomy it displays as it moves around the house—often causing its owners to personify it and treat it as a household member in a way that downplays its surveillant implications but also highlights how automated visuality is perceived by those living with it on an everyday basis.

Works Cited

Agostinho, Daniela, and Nanna Bonde Thylstrup. "If Truth Was a Woman: Leaky Infrastructures and the Gender Politics of Truth-telling." *Ephemera: Theory & Politics in Organization*, vol. 19, no. 4, 2019, pp. 745–75.

Astor, Maggie. "Your Roomba May Be Mapping Your Home, Collecting Data That Could Be Shared." *New York Times,* 25 Jul. 2017, https://www.nytimes.com/2017/07/25/technology/roomba-irobot-data-privacy.html. Accessed 8 Oct. 2021.

Bachelard, Gaston. *The Poetics of Space*. Beacon Press, 1964.

Consumers International and Internet Society. *The Trust Opportunity: Exploring Consumers' Attitudes to the Internet of Things*. Reston: Internet Society 2019, https://www.internetsociety.org/wp-content/uploads/2019/05/CI_IS_Joint_Report-EN.pdf. Accessed 8 Oct. 2021.

Darling, Kate. *The New Breed: What Our History with Animals Reveals about Our Future with Robots*. Henry Holt and Co., 2021.

Dobash, R. Emerson, and Russell Dobash. *Violence Against Wives: A Case Against the Patriarchy*. Free Press, 1979.

Fink, Julia, Valérie Bauwens, Frédéric Kaplan, and Pierre Dillenbourg. "Living with a Vacuum Cleaning Robot." *International Journal of Social Robotics*, vol. 5, 2013, pp. 389–408.

Forlizzi, Jodi. "How Robotic Products Become Social Products: An Ethnographic Study of Cleaning in the Home." *HRI 2007 - Proceedings of the 2007 ACM/IEEE Conference on Human-Robot Interaction - Robot as Team Member*, pp. 129–36, https://doi.org/10.1145/1228716.1228734.

Forlizzi, Jodi, and Carl DiSalvo. "Assistive Robots and Domestic Environments: A Study of the Roomba Vacuum in the Home." *Proceedings of HRI06*, 2006, pp. 258–65, https://doi.org/10.1145/1121241.1121286.

Garfinkel, Harold. *Studies in Ethnomethodology*. Polity Press, 1989.

Heidegger, Martin. *Poetry, Language, Thought*. Harper Colophon, 1971.

Hendriks, Bram, Bernt Meerbeek, Stella Boess, Steffen Pauws, and Marieke Sonneveld. "Robot Vacuum Cleaner Personality and Behavior." *International Journal of Social Robotics*, vol. 3, 2011, pp. 187–95.

IHS Markit. *The Internet of Things: A Movement, Not a Market*. IHS Markit, 2017, https://cdn.ihs.com/www/pdf/IoT_ebook.pdf. Accessed 8 Oct. 2021.

iRobot. "Data Security." https://www.irobot.dk/Legal/Data-Security. Accessed 12 Oct. 2021.

---. "What Is Dirt Detect™ Technology?" https://homesupport.irobot.com/s/article/68. Accessed 12 Oct. 2021.

Pink, Sarah, Kerstin Leder Mackley, Roxanna Morosanu, Val Mitchell, and Tracy Bhamra. *Making Homes: Ethnography and Design*. Bloomsbury Academic, 2017.

Pink, Sarah. *Doing Sensory Ethnography*. 2nd ed. Sage, 2015.

Rapoport, Michele. "The Home Under Surveillance: A Tripartite Assemblage." *Surveillance & Society*, vol. 10, no. 3–4, 2012, pp. 320–33.

Sami, Sriram, Yimin Dai, Nirupam Roy, Jun Han, and Sean Xiang. "Spying with Your Robot Vacuum Cleaner: Eavesdropping via Lidar Sensors." *SenSys '20: Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, Association for Computing Machinery, 2020, pp. 354–67, https://doi.org/10.1145/3384419.3430781. Accessed 12 Oct. 2021.

Scott, Peter. *The Market Makers: Creating Mass Markets for Consumer Durables in Inter-war Britain*. Oxford UP, 2017.

Sofia, Zoë. "Container Technologies." *Hypatia: A Journal of Feminist Philosophy*, vol. 15, no. 2, 2000, pp. 181–201.

Solove, Daniel J. "I've Got Nothing to Hide and Other Misunderstandings of Privacy." *San Diego Law Review*, vol. 44, no. 4, 2007, pp. 745–72.

Steiner, Henriette, and Kristin Veel. "A Window to the Black Box: Stories from a Calm Surveillance Society." *The Window: Motif and Topos in Austrian, German and Swiss Art and Literature*, edited by Heide Kunzelmann and Anne Simon, IUDICIUM Verlag, 2016, pp. 149–63.

Strengers, Yolande, and Jenny Kennedy. *The Smart Wife: Why Siri, Alexa, and Other Smart Home Devices Need a Feminist Reboot.* MIT Press, 2020.

Sung, Ja-Young, Rebecca E. Grinter, and Henrik I. Christensen. "Domestic Robot Ecology: An Initial Framework to Unpack Long-Term Acceptance of Robots at Home." *International Journal of Social Robotics*, vol. 2, no. 4, 2010, pp. 417–29.

Sung, Ja-Young, Rebecca E. Grinter, Henrik I. Christensen, and Lan Guo. "Housewives or Technophiles? Understanding Domestic Robot Owners." *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*, 2008, pp. 129–36.

Taylor, Alex C. "Machine Intelligence." *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, 2009, pp. 2109–18, https://dl.acm.org/doi/10.1145/1518701.1519022. Accessed 19 Feb. 2022.

Trifo. "Lucy: With Great Power, Comes Great Incomparability." https://trifo.com/. Accessed 28 Dec. 2021.

Turkle, Sherry. *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books, 2011.

van Dijck, José. "Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology." *Surveillance & Society*, vol. 12, no. 2, 2014, pp. 197–208.

Vaussard, Florian, Julia Fink, Valérie Bauwens, Philippe Rétornaz, David Hamel, Pierre Dillenbourg, and Francesco Mondada. "Lessons Learned from Robotic Vacuum Cleaners Entering the Home Ecosystem." *Robotics and Autonomous Systems*, vol. 62, 2014, pp. 376–91.

Weiser, Mark. "Open House." March 1996, https://calmtech.com/papers/open-house.html. Accessed 12 Oct. 2021.

Wellendorf, Kassandra. "Elastic Looking and Negotiating of Invisibility in Public Spaces." *Transvisuality: The Cultural Dimension of Visuality, Vol. II: Visual Organizations*, edited by Anders Michelsen, Frauke Wiegand, Tore Kristensen, Liverpool UP, 2015, pp. 215–28.

Wiehn, Tanja Anna. "Algorithmic Intimacies: A Cultural Analysis of Ubiquitous Proximities in Data." PhD thesis, University of Copenhagen, 2021, https://static-curis.ku.dk/portal/files/272242205/Ph.d._afhandling_2021_Wiehn.pdf. Accessed 12 October 2021.

Video Recordings of Interviews and Observations

Kristin. Personal video interview. 8 Aug. 2019.

Rasmus. Personal video interview. 10 Aug. 2019.

Pernille and Jesper. Personal video interview. 13 Nov. 2019.

Rune and Stella. Personal video interview on Zoom. 14 Sep. 2020.

Mia. Personal video interview. 21 Sep. 2020.

Marianne. Personal video interview. 22 Sep. 2020.

Rasmus. Personal video interview. 1 Oct. 2020.

Rune and Stella. Personal video interview. 1 Jul. 2021.

Kassandra Wellendorf is a film director and a media artist besides her employment as Teaching Associate Professor in visual culture, in Department of Arts and Cultural Studies, University of Copenhagen. She has produced a number of internationally recognized films and interactive installations on the fleeting transitions between public and private space (Close 2002; Invisible 2003; Sleeping Obstacles 2010; Insideout2400 2012). Her work draws on ethnographic, ethnomethodological, and filmic methods. Her research interests center on practice-based research, participatory art forms, development of new filmic methods within sensory ethnography and the role of digital technologies in our everyday life. Email: kassdorf@hum.ku.dk.

Karen Louise Grova Søilen completed her PhD on the topic *Atmospheres of Surveillance* (University of Copenhagen 2021). Her research interests center on surveillance perceived as bodily, emotional, and multisensory experiences, surveillance in contemporary art and culture, nonrepresentational and feminist methodologies, cultural theory, and the changing notions of the home through the affordances of smart technologies. She is the recipient of Surveillance Studies Network's 2020 *Early Career Researcher Award*, and her work is published in *Surveillance & Society* and *NTIK: Nordic Journal of Information Science and Cultural Mediation*. Email: karenlouise.soilen@gmail.com.

Kristin Veel is Associate Professor at the Department of Arts and Cultural Studies, University of Copenhagen. Her research focuses on the impact of digital technology and datafication on the contemporary cultural imagination, with a particular interest in invisibilities, uncertainties, and overload, and the ways in which they are negotiated in film, art, and literature. Among her publications are *Tower to Tower: Gigantism in Architecture and Digital Culture* (MIT Press, 2020), and *Touch in the Time of Corona: Reflections on Love, Care and Vulnerability in the Pandemic* (De Gruyter, 2021) both co-authored with Henriette Steiner, and she is co-editor of ten collected volumes and special journal issues, including *Uncertain Archives: Critical Keywords for Big Data* (MIT Press, 2021). Email: kristinv@hum.ku.dk.

The Journal of Media Art Study and Theory
Volume 3, Issue 1, 2022
Automating Visuality
www.mast-journal.org
MAST

On the Truth Claims of Deepfakes: Indexing Images and Semantic Forensics

Rebecca Uliasz
Duke University

Abstract

When news media shared a video of outgoing president Donald Trump acknowledging the victory of president-elect Joe Biden, some social media users conspired that it was a deepfake, a synthetic image made with machine learning (ML) algorithms, despite evidence to the contrary. Employing this example in the following, I focus on how images generate veracity through the interrelated actions of human and machine learning (ML) algorithms. I argue that ML presents an opportunity to revisit the semiotic infrastructures of images as an approach towards asking how photorealistic images produce truth claims in ways that exceed the purely visual. Drawing from photographic theories of the image index and diagrammatic understandings of ML, I argue that *meaning*, described here as what images do in the world, is a product of negotiation between multiple technological processes and social registers, spanning data sets, engineering decisions, and human biases. Focusing on Generative Adversarial Networks (GANs), I analyze sociopolitical and scientific discourses around deepfakes to understand the ways in which ML affords hegemonic ways of seeing. I conclude that ML *operationalizes* the evidentiary power of images, generating new thresholds of visibility to manage uncertainty. My aim is to critically challenge post-truth paranoias by analyzing how ML algorithms come to have ethicopolitical agency in visual culture, with implications for how images are made to *matter* in post-truth media ecologies.

Keywords

deepfake; machine learning; generative adversarial network; truth claims; indexicality; diagram; post-truth

Introduction

> Whatever the truth of images in media might be, we will have to reckon with their radically contradictory reputation as "everything" and "nothing," the most valuable and powerful elements of the messages transmitted by media, or the most trivial, degraded, and worthless. (Mitchell 35)

When news media shared a video on January 7, 2021, of outgoing president Donald Trump denouncing Capitol protestors and acknowledging the victory of then president-elect Joe Biden, some social media users conspired that it was faked. Users across Facebook, Twitter, and Parler claimed Trump's "head doesn't seem to move . . . properly," citing "his cadence" and "the obvious face movements they're attempting to pass as real" as evidence that the video was a deepfake, despite that it was released by the White House and posted to Trump's own social media account (@ddale8; "Fact Check").

Deepfakes belong to the category of "synthetic media" which includes digital video, image, text, and audio created autonomously by or with the assistance of machine learning (ML) (Babu; Paris and Donovan).[1] Because they are believable in a visual sense, they contribute to public paranoia concerning our ability to perceive and understand digital images culturally (Chesney and Citron; Mirsky and Lee). As one Reddit user pointed out, in contemporary media

[1] They are typically traced to November 2017, when a Reddit user called "deepfakes" made a post containing machine-generated pornography of actresses Gal Gadot and Scarlett Johansson (Paris and Donovan).

ecologies "video evidence can no longer be taken as gospel; anything can be discounted as fake" (@Eric_Fapton). That this claim was made of the video of Trump's address highlights existing tensions within media theories of the image, paradoxes that earn them the contradictory reputations as "everything" and "nothing." Images appear to our minds and eyes as resemblances of things while at the same time only ever appearing in mediation, circulating through networks that shape their productions, distributions, and receptions. With the increasing role that ML plays in influencing human perception and knowledge, the visual accuracy of an image, understood as its resemblance to a thing in the world, may not testify to its connection to its referent in the traditional sense that a photorealistic image might. At stake with this example are questions of how ML participates in the production of the ambiguities that typify post-truth media ecologies, and how its participation factors into the shifting onto-epistemological status of images in visual culture today.

While the relationship of visual media to truth takes many discursive, abstract, and historical forms, this article narrows in on digital images that aim towards photorealism, mobilizing deepfakes to track the shifting function of photorealism in networked culture. One part of this inquiry is asking how imagistic indexicality is negotiated by ML. The pre-digital photographic index was understood by some, following Charles Sanders Peirce's definition developed through his second trichotomy of sign types, to function by way of an existential connection between a sign and its object, a trace of the real that confirms its referentiality. Some have claimed that digitization severs the photograph's indexical links because image information is converted into digital code, rendering the image "programmable" (Manovich). Others point out that the rhetoric of crisis is often evoked in regards to the digital, where the threat of the "death" of the index would usher in an era in which any photorealistic image can be forged, and

thus doubted (Paulsen). It seems that further algorithmic developments in the field of image synthesis have only intensified such anxieties around imagistic veracity.

My aim is to critically challenge such anxieties through analyzing the ways in which ML *operationalizes* the generation of visual truth claims. The operational notion of machinic images is defined in non-representational discourses of the image (Farocki). The meaning of an operational image would be derived from the action it prompts in a technical system rather than a result of its "reality effects" (Barthes). Today, our visual paradigm would seem linked to the "algorithmic turn" (Uricchio), a condition in which images are produced through processes that perform statistical calculations on the vast quantities of image information contained in online data sets. Our opening example suggests that the algorithmic turn is linked to a qualitative shift in the way machinic operations intervene in very human parameters of visibility, with consequences for visual cultures and their epistemologies (Ananny and Crawford; Pasquinelli and Joler). The "truth" of the image, if it is to be found, is in this instance a product of negotiation between various technological infrastructures, social relations, and epistemic practices.

To locate the "epistemic agency" (Werner) of deepfakes, the *how* of how deepfakes make meaning, requires we move between the multiple semiotic registers through which they operate, which span technical and social scales. Deepfakes are complex epistemic things, linked to infrastructures composed of data, models, and algorithms that pattern machine behaviors and human knowledge practices. The issue at stake is not just one of deciphering what a deepfake represents but also what it does, which is related to how ML is used to claim meaning and the types of meaning it makes possible. The following inquiry into how deepfakes come to matter in the context of post-truth discourses is one that asks, in N. Katherine Hayles's words, how

human-computer interactions "form semiotic relationships that exceed the limits of biological cognition alone" and to what end (Hayles 41). Its aims are twofold: first, to describe the ontology of synthetic images through the non-sensuous semiotic processes of ML, and second, to understand the significance of these imaging operations for the politics and epistemologies of "truthful" visual media today.

Indexing Images and Truth Claims

Deepfakes fill a certain cultural role as harbingers of post-truth politics. Their advent, argues Yves Citton, "strikes a fatal blow to the trust one could (in most cases) put in what looked like (and was presented as) indexical representations of reality" (47). In this case, the mere existence of deepfakes, and synthetically generated media more broadly, forces one into a state of trustlessness, where one suspects that behind every imagistic truth claim there lurks the possibility of a lie, and therefore the possibility of deceit.

Concerning photorealism, some have argued that the concept of the "death" of the index stems from a misdiagnosis of the image index as primarily a "trace" that would mark its material connection to the real. Such accounts point out that images are by virtue subject to processes of manipulation (Doane; Paulsen). Both digital images and traditional chemical photographs alike are "subject to elaborate procedures before a picture will result" (Gunning 40). Photographic images, rather than containing a direct "imprint" of a real object, are realized through lenses, film stock, and other material constraints. Instead of the material trace, some have argued that we understand digital images with regard to the *deixis,* the "pointing" function of the index (Paulsen). This particular type of indexicality focuses on the way a sign is mediated, how it "reaches out" to point to something that may not yet be known but is called to attention, such as how a pointing finger indexes something yet to be seen. If material traces are what gave optical

photographs a trustworthy existential link, the pointing function of the index in general is to communicate something across time, "to reference a real without realism" (Doane 4). The index can be understood as the very *possibility* of communication through mediation.

As such, the stratum of the index as deixis is always somewhat ambiguous, open to interpretation and doubt. Still, others remark that despite this indeterminacy inherent to the deixis, images still somehow possess the power to "tell the truth." Rather than presenting contradiction, Tom Gunning notes that the ambiguity of the index is precisely what allows for things to be said about an image. For Gunning, digitization doesn't undermine the image's indexical qualities but shows us how truth claims are technically and socially produced by way of establishing iconic resemblances. For instance, Photoshop tools might give users unprecedented freedom to manipulate an image of a face. But for a portrait to function as an accurate image of its subject, it can only be transformed in a way that maintains this iconic relationship (Gunning 41). This quality of "visual accuracy" compels our belief in the image's referent while also giving it a mutable characteristic. In this sense, a fake image must maintain visual accuracy, a relation to a notion of "likeness" that is configured through technical or artistic manipulations. Put differently, a preformatted and agreed-upon sense of iconic similarity is required to produce accurate (described as "believable" in the case of deepfakes) fake images. As Gunning argues, "indexicality intertwines with iconicity" in our assessment of photorealistic images—the former establishes the possibility of interpretation while the latter grounds meaning through qualitative resemblance to a known object (Gunning 41).

Considered from the standpoint of photography theories of the image, deepfakes testify to the ongoing socio-technical value we place on visual accuracy which manifests in our continued investment in imagistic realism as truthful. As Gunning argues, the truth claim of photorealistic

images can even be said to *depend* on the ways in which we value the visual domain to provide a "continuing sense of the relation between the photograph and pre-existing reality" (45). Nonetheless, there are technical differences between the icon-index function of photographs and the operations of ML. Deepfakes do away with the photographic image's previous relationship with an optical lens. Instead, they are given their form by statistically weighted models trained on data sets of pre-existing images. As we will explore, their indexicality is operationalized by algorithms, with consequences for how ML systems generate their own truth claims, producing images that take on prescriptive rather than representative qualities.

GANs and Diagrams

It is part of the "trick" of ML that we treat their outputs iconically, i.e., as pictures that communicate meaning through their qualitative resemblance to some object. As Leif Weatherby and Brian Justie explain, we fall in the trap of "naïve iconic interpretation" when we construe machine behavior as intelligence because a system expresses behavior we take to be *like* intelligence (383). In such instances, we draw upon metaphysically ungrounded assumptions about *what* intelligent behavior looks like.[2] That we socially and culturally ascribe verisimilitude to certain outputs produced by ML speaks to the bias inherent to iconic interpretation, but also raises questions about how the technical operations specific to ML factor into qualitative interpretations. Framed as a question, when ML intersects with the visual field, how does it change cultural and social assumptions of likeness? How is meaning, understood in the pragmatic sense as the action of a sign, intermediated between cognitive registers which span the technical and social? In Peirce's tripartite, the icon is the sign type "from which information may

[2] With concerns related to my own in this article, Weatherby and Justie attend to the ways in which deep neural networks treat icons (as images) indexically. They argue that the indexical pointing function *can* result in an iconic (imagistic) output, although the operations of the net are "indexical all the way through" (390).

be derived" by a cognitive interpreter (309). Arguing that the icon sign can function non-trivially, or by means other than visual similitude, Frederik Stjernfelt argues that the icon works in Peirce's semiotics to describe non-sensuous semiotic operations.[3] As Stjernfelt describes, a particular type of icon, which Peirce calls a diagram, is at stake whenever the "rational relations" between parts, which mirror the relations between the parts of the object they correspond to, can be manipulated to gain new information about the object. Diagrammatic manipulation involves the operationalization of similarity to yield "implicit possibilities involved in the icon" (72). In this sense, ML can be described as a particular *type* of icon, a diagram, which captures likenesses as the relationships between heterogeneous parts in vector space rather than by visual similitude (Mackenzie).

The interplay between the iconic and indexical elements of a GAN makes verisimilitude possible. A GAN (see fig. 1) contains a diagrammatic relation of nodes, layers, and thresholds in which indexical relationships are algorithmically calibrated to achieve optimal icon outputs, and icons are converted to indices to provide information about how this optimization process should occur. A GAN takes a defined training set of images (icons), recasts them as vectorial representations (indexes), and learns to output images (icons) that look similar enough to this training set to a human observer such that the human cannot decipher between the two—a sort of visual Turing Test (Mirsky and Lee). This means the output images, although synthetically produced, "pass" as real, in a conventional sense.

[3] Stjernfelt gives an *operational* concept of Peirce's account of diagrams as non-trivial icons as "moving pictures of thought." This concept works to separate the icon from psychologism: "The decisive test for its iconicity rests in whether it is possible to manipulate the sign so that new information as to its object appears" (90).

To describe this process in some technical detail, a GAN is comprised of two competing neural networks (NNs).[4] A generator (G) network takes N-dimensional random variables (noise) as input. G captures a probability distribution function P(x) of samples drawn from the data set and aims to output an image with the same distribution. A discriminator (D) takes an input and determines whether this input was drawn from the training set or, rather, from the generator. G is trained to maximize the error of D ("fool" it into taking fake images as part of the training set) while D aims to minimize this error—a zero-sum game. In this game, or training process, indexical pathways between nodes in hidden "deep" layers are optimized over time to find the ideal routes to maximize the possibility of producing the desired types of outputs. The result is that the output image tends toward the "likeness" of the input image set, even while it remains uniquely singular.

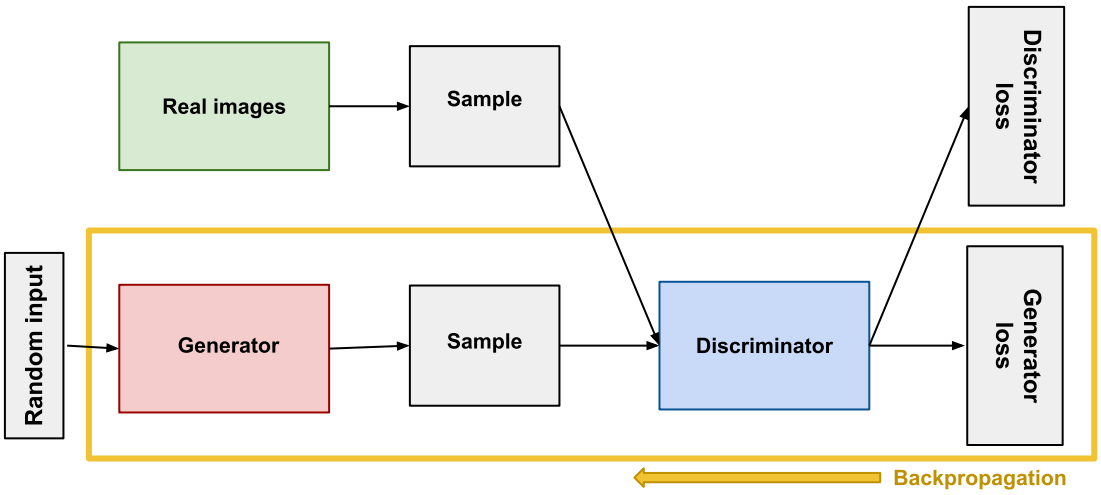


Fig. 1. "Backpropagation in generator training." Google Developers, https://developers.google.com/machine-learning/gan/generator. Accessed 1 Aug. 2021.

[4] This method was designed by MIT Deep Learning textbook author Ian Goodfellow and colleges in 2014. For surveys of further developments in deepfake methods, see Nguyen et al.; Mirsky and Lee.

Although the technological novelty of a GAN is that it can generate completely new categories instead of depending on a priori rules, we should see this as less of a departure from human cognition than an intensification of iconic reasoning. In a certain sense, the possibilities of establishing likenesses are expanded by a GAN, for it incorporates semiotic processes that operate at speeds and granular scales outside of human perceptive registers into its diagrammatical operations, opening the possibilities for deepfake generation beyond the purview of what is humanly sensible. At the same time, the GAN is optimized such that it may automate claims to imagistic verisimilitude following specific prefigured criteria. It is because of this prefiguration or biasing of the GAN in a certain way, that GANs preserve the truth function of images. Like photorealistic images before, deepfakes preserve iconic continuities. This has to do with the way *meaning* is articulated as a problem to be *solved* through the optimization of ML algorithms. Providing an example of how meaning is problematized by ML, Jonathan Roberge and Michael Castelle describe a deep learning object recognition model trained on the ImageNet database, a popular dataset containing over 14 million categorized images. Instead of using the taxonomies of ImageNet's object categories, "ML/DL's conquest of iconicity—its ability to calculate the likeness between a picture of a tiger and an arbitrary value or category denoted to as 'tiger'—is performed through a layered, directional (and thus indexical) flow of linear and nonlinear transformations . . . by producing categorial outputs, ML/DL necessarily morphs into something prescriptive" (9). Problem solving for ML is a matter of discovering the right rules to calculate resemblance to some input rather than classifying an existent object. In the case of a GAN, the optimization of the model is the problem to which the solution has already been decided upon, in the sense that one has already specified what sort of output one would like to get.

It would seem that the indexical operations of a GAN are what gives them their epistemic agency, as they carve out the semiotic paths through which meaning can flow. In doing so, they both produce and constrain the possibilities of their outputs. But the iconic aspects of ML operations should also be examined, not because they perform some sort of ideological trick but because such operations are entwined with epistemology at the level of social reality. The reality imagined by engineers through ML is one where diagrammatic substitutions can be applied methodologically to find the ruleset to achieve the types of outputs one would like to see in the world. Part of the value in reading ML diagrammatically is to understand how it imposes certain regularities on an epistemology based on the type of problem it is trained to solve. For instance, a GAN trained on images selected from the tiger category of the ImageNet database would learn to expand a random noise distribution into an image statistically similar to those tiger images and visually accurate to a human observer. "Tigerness" is the problem space the GAN operates within. Despite variations, "tigerness" will always be maintained because it is the solution given in advance of the formalization of a procedure, which is discovered over time through the optimization of interactions between training algorithms, vectorized images, and human decisions.

Synthetic Images and SemaFor

I have argued that the diagrammatic logic of ML facilitates and intensifies the possibilities of imagistic truth claims but what remains is the question of how ML techniques are used to maintain *specific* ways of seeing and knowing the world. Here, I use the term "regime of recognition" following Louise Amoore, who describes the way that ML algorithms don't just recognize already existing categories in the world, in the sense that statistical classification algorithms might make use of a priori classifications in a dataset such as faces, threats, vehicles

and so on (67). Instead "they actively generate recognizability as such" by deciding "what or who is recognizable as a target of interest in an occluded landscape" (69). Regimes of recognition describe a field of power in which the politics of ML techniques affect the parameters of what matters and what is brought to attention (Jacobsen).

Before recent developments in ML, digital forensics research was premised on the idea that a target of interest bears some accountable indexical link to its technical infrastructures. For instance, if an image were tampered with or manipulated there would be some form of recognizable digital fingerprint. Computer-generated artifacts added to a digital photograph could be discerned through their use of "idealized models of geometry, lighting, camera optics and sensors," which could be classified using statistics-based analysis (Farid 72). Three key technological developments in the early 2000s posed novel challenges to digital forensics— the invention of the high-powered NVIDIA GPU, the creation of large data sets for computer vision research, and the 2012 creation of AlexNet (DARPAtv, "Artificial Intelligence Colloquium"). AlexNet is used today as the architecture for most self-supervised learning. It is a deep convolutional neural network (CNN) that was designed with the notion that objects in the "real world" are imperfectly represented in images—they may be occluded or ambiguous in some way (Krizhevsky et al.). To overcome the problem of indeterminacy, the creators of AlexNet trained it on millions of images labeled by Amazon Mechanical Turk workers, capturing similarities as statistical distributions of weightings and probabilities rather than fixed categories. In turn, AlexNet would be able to recognize all future instances of an object of interest, even instances that varied in ways that were not previously recognizable.

Whereas digital forensics before these developments depended on access to "white box" statistical models which could link fake images to their manipulation techniques, ML is by virtue

black-boxed (Pasquinelli and Joler), which means its learning models evade previous detection techniques. As the director of DARPA Semantic Forensics (SemaFor) Program explains, "detecting GAN generated media requires detailed training data from those specific GANs" and detailed knowledge of those generators (DARPAtv, "Semafor"). Where this knowledge is not available, SemaFor aims to develop "semantic algorithms that automatically detect, attribute, and characterize falsified multi-media to defend against large scale automated disinformation attacks" (DARPAtv, "Semafor").

These three specified goals—detection, attribution, and characterization—each lay a specific claim to meaning. They are each given problems to be solved in regulating maliciously manipulated images. Take SemaFor's example of the application of this multi-modal approach toward assessing an article describing a protest in front of the US Capitol. *Detection* would aim to solve the problem of classifying semantic inconsistencies and manipulation errors at the level of a single output image. For instance, do the person in this image's earrings match? *Attribution* would solve the problem of source or authorship e.g., is there formatting constancy in the images in the source in which they appear? After screening the protest images to determine if they are fake, ML might be used to perform a textual analysis of a body of journalism by one purported author to recognize inconsistencies within the vocabulary and style. Once their inauthenticity is established, it could screen these articles for terms associated with a high likelihood of malintent. Finally, *characterization* would attempt to solve the problem of moral categorization, deciding on the intent behind media manipulation and prioritizing the fake content for human review. At the level of characterization, the protest article, for instance, could be flagged for posing a threat to influence voter decisions (DARPAtv, "Semafor"). Semantic forensics moves us beyond the scientific detection of data patterns that already exist within the image. It is instead figured as a

site from which one can speculate on the possible harmful social effects of images. In each case, semantic meaning "moves beyond low-level statistical fingerprints" to include high-level information such as "words, phrases, signs, visual elements, audio elements; relationships between these elements; and relationships between such elements and the real world" (SemaFor).

We can no longer say, in the case of semantic forensics, that verifying likeness is a matter of common-sense knowledge. Rather, the scope of semantic knowledge sets the regime of visibility by determining what types of information are of interest for establishing visual authenticity or, alternatively, falsifying it. Put another way, solving the problems of detection, attribution, and characterization necessitates that an ideal solution is given in advance for each—a formal assumption of what malicious relationships or patterns of artifacts would look like. In the example DARPA provides of the fake protest article, building an algorithm to interpret semantic meaning at each scale involves inbuilt decisions about what a consistent author style would look like, decisions that correlate terms or image features to social harm, and finally, presuppositions that these decisions can be modeled, schematized, and used to screen future media objects. To automate these solutions at scale requires that interpretation sneaks in the backdoor of data analytics to offer preformed answers to the question of what makes a pattern meaningful and to what this pattern should be determined to mean. Pattern finding thus requires decisions, in the political sense, over what types of information count as meaningful and which are spatially proximate or alike enough to offer a form of correlative proof (Apprich et al.).

For DARPA, which serves as a research and development branch of the US Department of Defense, deepfakes pose novel threats to national security because they undermine consensus reality, which could lead audiences to be duped by malign foreign actors. But it would seem that the problems of detection, attribution, and characterization don't deal with establishing an

understanding of existent sociotechnical phenomena but instead deal with countering "expected threats" (SemaFor). Taking stock of the synthetic imaging technologies of the present, SemaFor imagines vectors of future development, such as Identity Attack as a Service (IAaaS), new forms of synthetic generation that would have ripple effects up organizational and political scales. For DARPA, threat modeling becomes a form of threat anticipation, a means to "make all actions imaginable in advance" through arraying possible future threats (Amoore 79). SemaFor creates the conditions for acting in the present through operationalizing image interpretation, pre-emptively developing ML models to counter threats that always remain on a horizon of knowability.

Politics of Precomputation

To analyze how algorithms accrue agential power, Amoore adapts the concept of "precomputation," used in ML to describe how an algorithm is predisposed to be able to recognize the attributes of something in advance of its run time (78). As she explains, to "precompute" is also to imply what a proper procedure should look like and thus comes hardwired with normative social or political assumptions. This point is significant because as ML increasingly becomes part of everyday media ecologies, our ways of being, knowing, and acting in the world are delimited and afforded accordingly. There is a particular logic inherent to ML that is not just relegated to the world of tech but has become part of social life more broadly. This logic operationalizes the information at hand and adjusts the indexical relationships between features to discover the pattern one would need to confirm the precomputed knowledge one has provided as input. Finding the desired output is a matter, as Amoore explains, of establishing the "level of detection that is useful to you," which establishes a threshold of recognizability (68).

This action calls for decisions on questions such as: what types of information matter in a particular problem space? What ought to be screened out or discarded?

By deciding on precomputational aspects in advance (what type of algorithm or model is being deployed to solve for an unknown x?), one is deciding on "the register of what kinds of political claims can be made in the world, who or what can appear on the horizon, and who or what can count ethicopolitically" (Amoore 69). Precomputing is, plainly put, a way to take action in a condition of unknowability. It permits one to compress reality through the construction of a fragmented and limited simulation where any inputs can be mapped to the desired outputs. In a technical and social reality where data is in surplus, meaning isn't derived from causality but rather can be constructed through the infinitely possible patterning of data points (Halpern et al.). Precomputing establishes a register of perceptibility to describe a complex and unknowable reality through data. The conspiratorial claims around Trump's concession video can be understood correspondingly as a form of paranoid positivism bound to the fragmented (anti-)social realm of a filter bubble, bent on finding the desired outputs through reading patterns into the past to make actionable claims in the present. In this case, visual truth is subject to the rhetorical force of collective belief. For post-truth conspiracy-ridden epistemologies, a truth claim is made on the image to generate, rather than represent, some version of reality in a grasp for control of an unknowable future.

As we saw with GANs, their operations transform icons through diagraming vectorial relations and indexical substitutions. The diagram provides the precomputed elements that make manipulation possible, yielding image icons that are always propositionally accurate within the semantic boundaries determined by the GAN. Despite what post-truth discourses would suggest, fake images make truth claims. What ML shows us in this instance is that the thresholds of

regimes of recognition are porous and mutable, positioning photorealism as a site from which one can make claims to reality as a function of power or to its contestation.

Conclusion

When truth claims are made from various networked states, produced through collectives inclusive of technical and social fragments that do not map onto a consensual public sphere, it is important to ask what beliefs and desires govern our networks and how this in turn sets limitations on visibility and knowability. In security discourses, such as DARPA's, forensics isn't about discovering the causal truth behind an image, as if there were a repository of state knowledge through which all effects could be traced back to their origins. Instead, forensics is mobilized to create evidence of futures not-yet-come to prevent technological surprise and thus manage risk as a matter of national security. We could argue that paranoias around control and the loss of control are the animus of our networked reality, where power is maintained through the rejuvenation of its verifying logics. This is today's circular conspiratorial logic—meaning can be found precisely where it is sought out because the logics of verification are subject to any number of technical operations with more or less authority.

If the production of visuality has always been tied to questions of knowledge and power, we would do well to take imagistic claims to realism seriously, not just in terms of what images represent but in terms of the conditions of possibility they outline and the semiotic regularities they impose. As we have seen, ML is a technique for abstracting human decisions and biases and mobilizing them to arbitrate decisions outside of the scope of human epistemology. This should not be taken to mean that the claims of ML are somehow more objective, accurate, or truthful than human logic, nor that it tricks us into taking its claims as truths, but rather that it acts in the space of what is unknowable. ML systems allow us to create new indexical relations from

within. “The condition of indexicality,” claims Paulsen, is “the openness of the sign to interpretation and doubt” (98). ML expands the field of visibility beyond human sensory ratios, generating new possibilities for interpretation, but it also limits this field through precomputing those possibilities. Its operations are shot through with presumptions about what matters and how it should matter. As a result, we end up foreclosing alternatives, manifesting new worlds that are just superficial variations of the old one. Looking at the links between ML and photorealism lets us recognize how truth claims are syntactically conditioned, semantically bounded, and collectively constructed, and can therefore provide an opportunity to rethink the ontology of images not as representational objects but as ethicopolitical agents enmeshed in active sociotechnical negotiations around what matters and why.

Works Cited

@Eric__Fapton. "What Was /R/Deepfakes and How Did It Influence the Recent Reddit Rule Changes?" *Reddit*, 2017. https://www.reddit.com/r/OutOfTheLoop/comments/7vydoi/what_was_rdeepfakes_and_how_did_it_influence_the/. Accessed 15 Sep. 2021.

Amoore, Louise. *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*. Duke UP, 2020.

Ananny, Mike and Kate Crawford. "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability." *New Media & Society*, vol. 20, no. 3, 2018, pp. 973–89.

Apprich, Cramer, et al. *Pattern Discrimination*. Meson Press, 2018.

Babu, Sudharshan Chandra. "A 2020 Guide to Synthetic Media." *Paperspace*, blog.paperspace.com/2020-guide-to-synthetic-media/. Accessed Aug. 2021.

Barthes, Roland. "The Reality Effect." *The Novel: An Anthology of Criticism and Theory, 1900–2000*, Wiley, 1968, pp. 229–34.

Chesney, Robert, and Danielle Citron. *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*. U of Texas Law 107 California Law Review 1753, 2019, https://scholarship.law.bu.edu/faculty_scholarship/640. Accessed 20 Mar. 2022.

Citton, Yves. "Could Deep Fakes Uncover the Deeper Truth of an Ontology of the Networked Images?" *The Nordic Journal of Aesthetics*, vol. 30, no. 61–2, 2021, pp. 46–64.

Crary, Jonathan. *Techniques of the Observer: On Vision and Modernity in the Nineteenth Century*. MIT Press, 1992.

Dale, Daniel [@ddale8]. "Meanwhile on Parler, there are dozens and dozens of posts in which Trump supporters say or wonder if his concession video - which he tweeted out - is a “deepfake.”" *Twitter*, 8 Jan. 2021. https://twitter.com/ddale8/status/1347570822437863424?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1347570822437863424%7Ctwgr%5E%7Ctwcon%5Es1_&ref_url=https%3A%2F%2Fwww.reuters.com%2Farticle%2Fuk-factcheck-trump-consession-video-deep-idUSKBN29G2NL. Accessed 15 Sep. 2021.

DARPAtv. "Artificial Intelligence Colloquium: Media Forensics." *YouTube*, 2019, https://www.youtube.com/watch?v=Crfm3vGoBsM&t=2s&ab_channel=DARPAtv. Accessed 20 Mar. 2022.

---. "Semafor Proposers Day." *YouTube*, 2019, https://www.youtube.com/playlist?list=PL6wMum5UsYvZGmhx5NYRXf0OumdP0H8sN. Accessed 1 Jan. 2021.

Doane, Mary Ann. "Indexicality: Trace and Sign: Introduction." *Differences*, vol. 18, 2007, pp. 1–6.

"Fact Check: Donald Trump Concession Video Not a 'Confirmed Deepfake.'" Reuters, 2021. https://www.reuters.com/article/uk-factcheck-trump-consession-video-deep/fact-check-donald-trump-concession-video-not-a-confirmed-deepfake-idUSKBN29G2NL. Accessed Sep. 2021.

Farid, Hany. "Digital Image Forensics." *Scientific American*, vol. 298, no. 6, 2008, pp. 66–71.

Farocki, Harun. "Phantom Images." *Public*, no. 29, 2004, https://public.journals.yorku.ca/index.php/public/article/view/30354. Accessed 20 Mar. 2022.

Goodfellow, Ian, et al. "Generative Adversarial Nets." *Proceedings of the 27th International Conference on Neural Information Processing Systems*, vol. 2, MIT Press, 2014, pp. 2672–80.

Gunning, Tom. "What's the Point of an Index? Or, Faking Photographs." *Still Moving: Between Cinema and Photography*, edited by Karen Beckman and Jean Ma, Duke UP, 2008. pp. 39–49.

Halpern, Orit, et al. "Surplus Data: An Introduction." *Critical Inquiry*, vol. 48, no. 2, 2022, pp. 197–210.

Hayles, N. Katherine. "Can Computers Create Meanings? A Cyber/Bio/Semiotic Perspective." *Critical Inquiry*, vol. 46, no. 1, 2019, pp. 32–55.

Jacobsen, Benjamin N. "Regimes of Recognition on Algorithmic Media." *New Media & Society*, Oct. 2021. https://doi.org/10.1177/14614448211053555. Accessed 20 March 2022.

Krizhevsky, Alex, et al. "Imagenet Classification with Deep Convolutional Neural Networks." *Advances in Neural Information Processing Systems*, vol. 25, 2012. https://dl.acm.org/doi/10.1145/3065386. Accessed 25 Sep. 2021.

Mackenzie, Adrian. *Machine Learners: Archaeology of a Data Practice*. MIT Press, 2017.

Manovich, Lev. *The Language of New Media*. MIT Press, 2002.

Mirsky, Yisroel and Wenke Lee. "The Creation and Detection of Deepfakes: A Survey." *ACM Comput. Surv*., vol. 54, no. 1, 2021, arXiv:2004.11138. Accessed 20 Sep. 2021.

Mitchell, W.J.T, ed. *Critical Terms for Media Studies*. Co-edited by M.B.N. Hansen, University of Chicago Press, 2010.

Nguyen, Thanh Thi, et al. "Deep Learning for Deepfakes Creation and Detection." *arXiv preprint*, 2019, arXiv:1909.11573. Accessed 20 Sep. 2021.

Paris, Britt, and Joan Donovan. *Deepfakes and Cheap Fakes*. Data & Society Research Institute, 2019.

Pasquinelli, Matteo and Vladan Joler. "The Nooscope Manifested: AI as Instrument of Knowledge Extractivism." *AI & Society*, vol. 36, 2020, pp. 1263–80.

Paulsen, Kris. "The Index and the Interface." *Representations*, vol. 122, no. 1, 2013, pp. 83–109.

Peirce, Charles Sanders. “A Second Trichotomy of Signs.” *Elements of Logic*, edited by Charles Hartshorne and Paul Weiss, Harvard UP, 1931, pp. 247–50.

Roberge, Jonathan, and Michael Castelle. "Toward an End-to-End Sociology of 21st-Century Machine Learning." *The Cultural Life of Machine Learning: An Incursion into Critical AI Studies*, Springer, 2020, pp. 1–29.

SemaFor. "Frequently Asked Questions." DARPA, https://www.darpa.mil/attachments/FAQ_SemaFor%202019-09-03_FINALa.pdf. Accessed Dec. 2021.

Stjernfelt, Frederik. "Moving Pictures of Thought." *Diagrammatology: An Investigation on the Borderlines of Phenomenology, Ontology, and Semiotics*, edited by Frederik Stjernfelt, Springer Netherlands, 2007, pp. 89–116.

Uricchio, William. "The Algorithmic Turn: Photosynth, Augmented Reality and the Changing Implications of the Image." *Visual Studies*, vol. 26, no. 1, 2011, pp. 25–35.

Weatherby, Leif, and Brian Justie. "Indexical AI." *Critical Inquiry*, vol. 48, no. 2, 2022, pp. 381–415.

Werner, Gabriele. "Discourses About Pictures: Considerations on the Particular Challenges Natural-Scientific Pictures Pose for the Theory of the Picture." *The Technical Image: A History of Styles in Scientific Imagery*, edited by Bredekamp Horst et al., University of Chicago Press, 2021, pp. 8–13.

Rebecca Uliasz is an artist and PhD Candidate in the department of Computational Media, Arts and Cultures at Duke University. Her research interests include emerging issues in digital culture, philosophies of computational media, and new media art. She is one half of the theory-noise performance collective GOVERNANCE, alongside collaborator Quran Karriem. Her work has appeared in *Springer AI & Society: Special Issue, Ways of Machine Seeing*, *Review of Communication, Optimization: Towards a Critical Concept, A Peer Reviewed Journal About*, and transmediale. Email: rebecca.uliasz@duke.edu.

The Journal of Media Art Study and Theory
Volume 3, Issue 1, 2022
Automating Visuality
www.mast-journal.org
MAST

Experimental Indices: Situational Assemblages of Facial Recognition

Asko Lehmuskallio
Tampere University

Roland Meyer
Brandenburg University of Technology Cottbus-Senftenberg

Abstract

Facial recognition technologies are increasingly used outside of constricted, laboratory-like settings. While supporters of the technologies contend that they help in identifying threats by linking specific bodies to hard evidence, we argue that the indexical relations they exhibit are best described as experimental, pointing to specific situational constellations within which they were initially created. By revisiting key moments in the development of (semi-)automated facial recognition technologies from the late 1960s to the present, we identify varying situational assemblages of facial recognition that depend on different understandings of indexicality. These experimental indices rely on historical dynamics, including significant government interest in the development of facial recognition technology, expansion in the scale of experimental settings, and dissolution of the formerly strict boundaries between the social spheres of private image-sharing, commercial image distribution, and institutional image forensics for identification. In coupling experimental indices with the development of facial recognition technologies, we hope to show a way forward to comparing the histories of other evidential technical images too.

Keywords

face recognition; photography; surveillance; indexicality; automation; technical images; passports

Introduction

In January 2020, a hitherto unknown IT company in the US made worldwide headlines. Clearview AI had gathered billions of facial images from social networks without the knowledge or consent of those represented in those images, had collected them into a massive database, and was now promising its customers that it could identify nearly any individual in mere seconds by comparing facial images with its database. As this example made shockingly clear, companies in recent decades have capitalized on the ubiquity of cameras, social networks, and facial recognition to turn every face shot we share online into an operative portrait (Meyer, *Operative Porträts*): an image that might be algorithmically processed, compared to other images, and assembled with other portraits to create personal profiles, with the goal of establishing links between bodies, images, and data.

Related photo-theoretical work has long understood the camera as a more or less stable photographic device for producing evidence, with photography woven into a broader discourse about reliable traces. With this paper, by problematizing the notion of the index, we argue that photographic indexicality cannot be separated from the specific constellations in which cameras and the images they produce are used. In our view, the camera should be understood not as a mere recording device but, in accordance with its etymological roots in a chamber for deliberative, judicial, or legislative bodies, as a meeting place for decision-making (Lehmuskallio, "The Camera as a Meeting Place for Decision Making").

The truth claims often still associated with photography, rather than being simply the outcome of technical recording processes, in fact rely, as John Tagg has argued, on complex social assemblages of people and devices, social practices, and technical as well as legal norms. Indexical relations are not produced by technology alone but rather within specific modes of experimentation that need to be agreed upon to provide meaningful accounts. In other words, the ways in which images are supposed to provide evidence are the outcome of contingent practices and processes that have been historically stabilized for the specific purpose of generating truth claims. Following practice-theoretical work underscoring the importance of assemblages, i.e., the role of situationally ordered heterogeneous entities organized to make decisions (Müller), we focus in the following on how facial recognition has been employed to make claims about indexical relationships between physical bodies, photographic images, and the digital data used to identify individual faces.

Based on our findings, we argue that the digitally captured and algorithmically processed face remains an unstable and unreliable anchor. This is because facial recognition is limited: it can only determine probabilities, not produce certainties. Neither does it treat all faces equally, as ongoing debate on racial bias and social sorting attest well. Rather than having their foundations in a stable indexical relationship linking bodies, images, and data, facial recognition technologies produce what we call experimental indices that always leave some doubt as to their real-world applicability.

Our discussion begins with a brief introduction to the concept of the index and the varieties of photographic understandings that philosopher Charles Sanders Peirce worked with when developing the term. After this, we turn to key moments in the development of (semi-)automated facial recognition, highlighting the various ways in which digital images have been

used to analyze faces and establish indexical links between bodies and identities, links which in our view are inherently unstable without the contextual work needed to stabilize them. While, in its nascent stages in the 1960s and 1970s, automated facial recognition concentrated on a single face and its supposedly unique anatomical features, techniques of recognition later turned to statistical methods focusing on properties of large sets of visual data, leading toward the first wave of commercialization of the technology in the late 1990s. With the development of electronic biometric passports after 9/11, new questions arose, mainly concerning the indexical relationship between mobile bodies and machine-readable documents, calling for ongoing work on international agreements and standards of interoperability. In recent years, however, issues of norms and standardization have been overshadowed by new methods of facial recognition driven by machine learning coupled with the availability of vast quantities of images of faces online, which have not been captured in a preformatted manner. By focusing on these different assemblages of facial recognition, we aim to draw attention to the variance in the experimental indices on which they rely, underscoring the unruly nature of the relation behind these technologies that are used to decide on fundamental aspects of our day-to-day life. Rather than seeing their experimental nature as something unique to digital images and algorithmic images, we argue that the contingency of claims of indexicality has haunted photography and its various uses from its beginning—this being a reason why a look back into the nineteenth and early-twentieth centuries still provides theoretical clues for understanding image practices in the present.

Experimental Indices

Ever since the philosopher Charles Sanders Peirce discussed the indexical sign using the example of photography, the index has been a staple of photo-theoretical treatises. For many

scholars, it has served to guarantee a specific physical relation between a photograph and that which has been photographed (see Krauss, "Notes on the Index," Parts 1 and 2)—a claim also prominently disputed, but proven influential nevertheless. Since at least the 1970s, photography has been conceived in terms of either the index as a physical trace—a kind of imprint of reflected or refracted light, an "emanation of past reality" (Barthes 88)—or, often in polemical opposition to this earlier notion, an ideological construction entirely dependent on its context (Sekula). These accounts, however, fail to do justice to the complexity of Peirce's notion of the index and the specificity of photographic images he had in mind. The index, in Peirce's work, is not meant as a guarantee of evidence, nor can it be adequately described in terms of a critique of ideology. With this in mind, it is worthwhile to revisit Peirce's understanding of the index and pay closer attention to the role that photography played for it.

Importantly, Peirce's discussion of photography did not extend to many kinds of photos we might consider today, something that surprisingly tends not to be discussed in work focusing on the role of the index in photography. Even those photographs taken with a standard 35 mm film camera did not exist at the time of Peirce's writing. Instead, as Alexander Robins has suggested, it is likely that Peirce's understanding of the relation between photography and the index was an outgrowth of his own work with photographic technologies (Robins 2). Peirce's work on signs, including the index, was informed by a specific scientific setting within which photography was of particular importance: his first published book, *Photometric Researches*, was based on his efforts at the Harvard College Observatory, where he meticulously recorded variations in star luminance by photographic means (Robins 5-7). The recordings were based on Zöllner photometry, a photographic procedure designed for study of variations in light intensity, which does not yield the kinds of representations usually considered "photographic."

The index as conceived by Peirce should be considered in this light, and hence it is not automatically applicable to other photographic procedures that have emerged since. He discussed a tripartite system of signs consisting of icon, symbol, and index, with the first two of these components being regarded as important for numerous means of communication, such as drawing, painting, music and various forms of literature (Jappy). The index again, at least for many photo theoreticians, is essential to photography, setting it apart from various other modes of recording and communication. From a historical reading of Peirce's understanding of photography, we follow Robins, who has worked out that Peirce actually referred to three photographic techniques in his discussion of the index: Zöllner photometry (used in his visual stellar photometry), chronophotography (as had been used and developed by Étienne-Jules Marey and Eadweard Muybridge), and composite photography (a technique Francis Galton utilized for his photographic experiments). Each of these photographic technologies constitutes a distinct constellation of photographic devices, processes, people, and results which were explicitly used in scientific and quasi-scientific settings to make truth claims (see also Josh Ellenbogen for a comparison of these image logics). These truth claims are central to later applications to facial recognition technologies, too, as they always rely on particular experimental settings that include photographic technologies deemed to guarantee indexical relations.

Although physical relations exist between events and the photographs that represent them, the various photographic settings within which Marey, Muybridge, Galton, Bertillon, and Peirce used these technologies to produce scientific and quasi-scientific evidence manifested divergences: their respective camera devices differed significantly, and so did the photography involved. Hence the kinds of indexical relations they made visible varied, according to their

distinct experimental settings. The cameras they used should in this light be understood not only as particular photographic devices but also, in line with the word's etymology, as specific chambers within which judicial or legislative deliberation and decisions come into being, which again have real-world implications (see Asko Lehmuskallio, "The Camera as a Meeting Place for Decision Making," for a fuller discussion of this dual role of the camera). In short, the specific situated assemblages for "doing photography" relied on particular experimental settings to ensure an indexical relation between photographic inscriptions and what was photographed. We suggest, therefore, that when photographic technologies are connected to claims of evidence, such as in the realm of facial recognition, one can speak usefully of experimental indices to distinguish analytically among various claims to photographic indexicality. In this understanding, indexicality must be achieved by assembling a range of practices, devices, bodies, and modes of operation in ways that require infrastructural maintenance work (Kaltenbrunner).

In the following pages, we discuss developments in facial recognition technologies to illustrate how claims to indexicality in various fields have been made: from the first semi-automated, digital facial recognition technologies; through the widespread deployment of identification techniques in border control in the early 2000s; to today's seemingly ubiquitous use of facial recognition technologies "in the wild." This examination reveals that, in each case, an overall claim to indexicality exists when depicted faces can be determined to identify particular kinds of people, although the means to reach such claims differ significantly in historical comparison. While today astonishingly low error rates have been achieved for very particular kinds of conditions and settings, on the other hand when some conditions of photo capture, processing, analysis, and archiving are not ideal, the figures are much less impressive. That said, relatively high error rates have seldom precluded these technologies' use. Historically,

their implementation and societal implementation has often had less to do with the accuracy of experimental indices than with a political desire to surveil and control (Gates; Introna and Nissenbaum).

The Beginnings of Automated Facial Recognition

Automated facial recognition is not a new technology (for a historical overview, see Kelly Gates 25–63). Its beginnings date back to the 1960s, when AI pioneer Woody Bledsoe undertook moderately successful experiments with computer-assisted comparison of mugshots under the sponsorship of the CIA (Boyer and Boyer). Bledsoe's work was directly inspired by that of Alphonse Bertillon (Bledsoe 25). In fact, his method, which involved human operators marking distinct feature points on standardized facial images, could be seen as semi-automation of nineteenth-century-style anthropometry (Raviv). The image here functioned as a direct representation of anatomical features, which could be measured, compared, and put toward identification. The indexical character of photographic images in relation to the bodies depicted was seen as a given, and the transformation of visual representations into computable data was not problematized as having an effect on the indexical relations claimed.

While these earliest experiments were largely kept secret, it did not take long for the technology to reach public spotlight. Already in 1970, Nippon Electric Company was able to present a form of automated facial analysis at the Expo Osaka (Gates 25–6). In this case, the experimental setting involved the voluntary participation of members of the audience. As part of the "computer physiognomy" attraction, audience members were invited to have their facial features recorded by a video camera, which were then digitized, analyzed, and automatically measured by a computer program, which was supposed to compare them to a set of seven celebrity faces. Each of the celebrities, from Winston Churchill to Marilyn Monroe, purportedly

represented a certain "type" with which the faces of the visitors were matched, but the visitors never learned anything about the criteria employed to determine their "celebrity type." Rather, the results were presented like the wisdom of an oracle. Whatever happened inside the black box of the computer was hidden from the public, and only subsequent data analysis showed that the program was highly susceptible to error and its output largely random (Wayman 266). The claimed indexical relation between recorded visitor faces and celebrity faces was hence both black-boxed and largely random, not quite unlike some of the more recent commercial applications of facial recognition and analysis technologies in use today.

Its shortcomings notwithstanding, the Osaka attraction laid the groundwork for further developments in automated facial recognition. On the basis of the data sets from the event, young computer scientist Takeo Kanade developed one of the first fully automated systems for facial recognition, which relied on capturing facial feature points and measuring distances between them (Kanade 33–4). For this, a process of graphic reduction was put in place, which abstracted the facial features to make them machine-readable. In the first step, the frontally oriented and artificially isolated face was transformed into a black-and-white pixel matrix, through which Kanade's program then sought to localize predetermined features such as eyes, mouth, and nose (Kanade 12–3). The electronically rendered and then digitized, filtered, and reduced image acts as an immutable mobile, to use Latour's words, a visual matrix that allows for the recording, transformation, and transportation of empirical data across locations and technical settings. The earliest computer-based facial recognition system was composed of a partially automated process of measuring facial features on the basis of images captured from immobile bodies—while the majority of the visual information available and captured was filtered out. The indexical relationship between bodies, images, and data, rather than being

merely the effect of the technical recording apparatus, was the product of a series of prescribed formattings, abstracting and transforming facial features into machine-readable and analyzable form—hence, an outcome of an experimental assemblage.

In this respect, computer physiognomy stands in a long tradition of interpreting bodies and faces by standardizing, quantifying, and measuring technical images, which stretches back at least to its namesake, Johann Caspar Lavater, a Swiss pastor who in the 1770s made physiognomic character interpretation fashionable throughout Europe, famously demonstrating a preference for the proto-photographic silhouette over the artistic portrait. The silhouette was a "poor image," showing only the outline of the facial profile, a mere shadow of the living being, but allowed for precise measurements and, thereby, in Lavater's view, exact interpretations (90). Just as they would for Kanade two centuries later, reduction and quantification went hand in hand for Lavater, who even designed a special technical apparatus, the silhouette machine, for mechanically recording facial features. This was not unlike the photographic apparatus devised by the French Bertillon a hundred years later for production of standardized and precisely measurable mugshots (Sekula). One might characterize the earliest automated facial recognition process as one example in a two-hundred-year history of experiments in extracting exactly quantifiable data from technical images of isolated bodies—a history deeply entangled with European racism and colonialism (Gray). At its heart was an experimental setting aimed at immobilizing bodies and standardizing conditions of capture for purposes of guaranteeing the indexical relation between facial features and image properties: whatever distance could be measured on the surface of the image should correspond exactly to a measurement taken directly on the body.

While the earliest systems focused on plotting feature points on an individual face and comparing these to a universal biometric face model, a new approach was developed in the late 1980s and early 1990s. This was based on statistical analysis of facial image data sets (Sirovich and Kirby; Turk and Pentland). Rather than rely on a predetermined model of the face, the *eigenface* approach used hundreds of standardized images to learn what a face is in the first place. Several significant deviations from the average were determined for each face in this "training set," but rather than focus on geometric distances, here distributed brightness differences over the entire image matrix are of interest. This process of analytical image decomposition yields a set of so-called *eigenfaces*, each representing specific differences from the average. Unlike measurable features, such as eye distances, these *eigenfaces* cannot be extracted directly from living bodies: they constitute not so much anatomical features of individual faces as statistical properties of digital image archives. The indexical relation is thus calculated as a statistical probability in relation to an image data set. It may not be entirely coincidence that the ghost-like *eigenfaces* bear some similarity to the statistician and eugenicist Galton's "composite portraiture" from the 1880s. As with Galton's ultimately futile experiments, the *eigenface* representations are meant to visualize statistical variations within large collections of images (Ellenbogen 164–9).

With the *eigenface* approach, automated facial recognition became a technique of digital image analysis: the technical images now used no longer functioned as mere photographic indexes of preexisting anatomical features; rather, visual properties of digital images themselves, namely their two-dimensional brightness distribution, could now be quantified at the level of discrete pixel-level data—a quantification made possible by digitization beyond what Galton could have dreamed. The indexical relationship between bodies, images, and data thus became

even more precarious and entirely dependent on experimental processes of algorithmic transformations, as the data now used to identify faces were no longer abstracted from single images of isolated bodies but the statistical result of algorithmically comparing large sets of images.

This new approach triggered a veritable research boom in the 1990s—largely promoted by the US-government-funded Facial Recognition Technology (FERET) program between 1994 and 1996, in which the most promising algorithms from university computer labs competed against each other in standardized tests (Crawford 104–5). To guarantee comparability of the results, an image database was set up, for which hundreds of army employees were digitally photographed under standardized conditions. The standardization brought by these and subsequent databases was one not of method but performance: how these algorithms recognized faces was by no means the central concern—the program was meant to ensure that various algorithms' performance could be tested against each other (Phillips, Rauss, and Der). A new experimental setting emerged that was aligned well with the neoliberal agenda of contemporary economic reforms: in this context, the standardization of images was not focused on guaranteeing stable indexical relationships between facial features and quantifiable data; rather, it was seen as a precondition for deregulated competition. Eventually, this government-sponsored contest laid foundations for the further commercialization of the technology, as some of the researchers involved used the test results to raise venture capital for their newly founded private companies (Gates 47–51).

Machine-readable Biometric Passports Normalize Facial Recognition

It is often claimed that the further percolation of automated facial recognition technology into a host of domains of society was closely connected to the aftermath of the 2001 terrorist attacks on

attacks on the World Trade Center, the Pentagon, and a Pennsylvania field. A common thread of public discussions in both the US and Europe articulated musings about how better and, especially, more accurately employed computer-vision technologies could have been decisive in identifying and stopping the perpetrators before the fact. Facial recognition technology offered reassurance for a post-9/11 future, as a means of stopping terrorist threats. Senator Dianne Feinstein voiced a common concern at the time: "In the case of at least two of the hijackers, authorities had pictures of them as suspects prior to the attack, and airport cameras actually photographed them. But because these cameras didn't use facial biometric systems, security was not alerted and the hijackers remained free to carry out their bloody plans" (qtd. in Gates 2). This concern, which in hindsight proved to be inaccurate (Kean et al.), was coupled with three arguments commonly taken up when changes are called for in surveillance technologies: 1) visibility is a useful trap, with the visibility just having to be organized correctly; 2) suspects can be identified *ab initio*, through profiling in advance; and 3) the face has a pivotal role, as a particularly information-rich surface to work with.

Whereas earlier implementations of facial recognition technologies had targeted only specific populations (such as "criminals") or locations (e.g., downtown Tampa), this changed after 9/11. With the introduction of electronic machine-readable passports, including the possibility to read biometric information in digital form, automated facial recognition became a concern for anyone wanting to travel to the US and, later, the European Union. The US mandated the collection and exchange of electronic biometric information in 2002 with the USA PATRIOT Act, with this leading to the implementation of EU biometric passports featuring machine-readable biometric details, including so-called face prints by 26 October 2006 and fingerprints by 28 June 2009 (Torpey; Hausken).

In the eyes of politicians and legislators, the passport, as part of an assemblage of control exercised mainly at borders, had failed to provide enough information on possibly suspect travelers to aid in identifying future perpetrators. Focus was placed on one-to-one correspondence: the machine-readable electronic chip, containing biometric identifiers, had to maintain an indexical relationship with only one specific body, which could be identified as a distinct person whose information is stored in a relevant register. Hence, collective passports, such as group passports, became prohibited, and so were passports covering both minors and their guardians. The passport as a device for assuring indexical relations did so now with partially different means than earlier. A strictly regulated facial photograph was stored digitally on a single chip, so as to be machine-readable, and needed to point at only one embodied person, without allowing confusion with others. Hence the passport, which bore particular indexical connections to people and databases, was significantly changed in response to the novel machine-reading technologies implemented and moreover to the limitations in who could be inscribed into each passport. Of special note with regard to the novel EU legislation was a concern related to both illegal migration and terrorism. Worries about the two were bundled together in arguments that novel biometric technologies had to be implemented for travel documentation. The reasoning was similar to what was used to justify earlier developments in passport technologies: public concerns about security became instruments for further implementation of security technologies (Robertson).

While passports had contained biometric information ever since the nineteenth century (Caplan) in the form of signalments, signatures, and, especially after World War I, photographs, the novel aspect of machine-readable electronic biometric passports was greater interoperability among nation-states and other entities following specifications and standards from bodies such as

the International Organization for Standardization and the International Civil Aviation Organization (ICAO). Legislators believed that focusing on the face would provide useful information both for security personnel at airports, who could compare a traveler's physical face with the one depicted in the passport photograph, and for machine vision systems, which could compare the images they capture *in situ* and *in actu* at airports with the ones stored on the passport's chip. The face, thus, could be compared by human and machine agents with the one stored on the passport, either as a visible portrait or as an invisible latent image held on the chip. The earlier assumed indexical relation between facial photographs and human bodies was thus complemented by an additional computational layer, which used machine vision for verification (Lehmuskallio and Haara). The computational layer introduced was intended not so much to provide a final guarantee of a tight indexical relation between body and document but instead to facilitate interoperability among service providers so that nation-states and local authorities were not pressed to adopt any single automated facial recognition technology. The ICAO knew that facial recognition was less reliable than other biometric technologies but pushed for its introduction into passports on the assumption that it would be more likely to be accepted by citizens and politicians. As Liv Hausken has argued, a major reason for this was the assumed familiarity of having one's facial photo taken, as well as the fact that faces are visible in public spaces anyway (Hausken 167-8, 172). Part of the argument relied specifically on the assumed indexical connection between photography and the depictions created, although the automated facial recognition was performed with probabilistic models which always bear a possibility of providing false positives.

Belying claims that the events of 9/11 were central in developing electronic machine-readable passports, the ICAO had established its Panel on Passport Cards in 1968, tasking this

group with providing recommendations for standardized machine-readable passport cards. In 1984 and 1998, the agency established further groups to work on machine-readable travel documents, including means of biometric identification and data storage. As the ICAO notes in its own documentation, "[t]he bulk of the work had been completed by the time the events of 11 September 2001 caused States to attach greater importance to the security of a travel document and the identification of its holder" (1).

What 11 September 2001 did provide was impetus for a turn in public debate. Desire for a sense of security for air travel allowed momentum for implementing facial recognition technologies on a grand scale (Kember). Techniques originally developed with electronic means since the 1960s and 1970s by the likes of Bledsoe and Kanade could now be targeted toward all travelers. While early developments showed low accuracy rates in other than extremely controlled settings, trust in facial recognition technologies was so high at the turn of the century that various companies went unchallenged when overpromising what they could actually achieve. The claimed indexical relation between computer models and actual physical faces was early on an experimental index, relying not only on technical feasibility but also, to a large extent, on the ways in which visibility was structured using standards, international organizations, legislation, and temporality. This visibility encompassed all the elements necessary for creating and using "correct" photographic depictions for identification, including how photos had to be taken, processed, stored, read, and made interoperable with various databases. Because the specifications and standards needed to fit several elements of national, as well as local, infrastructure, they demanded only minimal overlap in the indexing between physical faces and their models in digital form. Whether the actually implemented models relied

on feature points, statistical probabilities in image data sets, or other forms of creating indexical relations was up to different service providers.

Image Ecologies beyond Standardization

Since the first steps in automated facial recognition technology, from its deployment in restricted local settings to the huge boost of mass-scale deployment for shared travel bureaucracy, facial recognition has seen increasing implementation, appearing often now in the wild. For example, in April 2021, 31-year-old Stephen Chase Randolph was arrested for several crimes committed during the attack on the US Capitol on 6 January. What was notable was how federal agents identified the suspect: "Capitol riot suspect arrested after FBI use face recognition on girlfriend's Instagram" read one headline (Hall). The story behind that headline can be traced in the Statement of Facts published online by the US Department of Justice. A key role in the identification and eventual arrest of Randolph was played by the Twitter account @SeditionHunters, which isolates individual still images of recognizable faces from the vast mass of digital footage captured during the January insurrection and publicly calls for their subjects' identification. In the eyes of the FBI, one of these anonymous faces, quickly and for obvious reasons denoted as "Grey Carhartt Hat," belongs to a person seen, in another video circulating online, climbing a barricade and assaulting a police officer. In this video, however, his face is not clearly visible, so the FBI took one of the still images posted by @SeditionHunters and, using an open-source facial recognition algorithm, searched the Web for other images of the same face. Eventually, they came across the Instagram page of the suspect's alleged girlfriend, and the same young woman's Facebook profile led them to Stephen Chase Randolph.

As this recent example illustrates, automated facial recognition today operates beyond the more regulated domain of governmental travel biometrics in an increasingly vast, deeply

saturated, and poorly regulated ecology of networked digital images. While early facial recognition was limited primarily to selected images created specifically for surveillance and identification purposes, the coupling of smartphone photography and social media has opened a whole new field of investigation to law-enforcement agencies. Masses of digital faces circulating via social-media accounts can now be searched for recurring patterns that link dispersed images across platforms. Although not every image eventually leads to identification and facial recognition, the expectation alone is radically changing how we treat the images we take of ourselves and share online (Meyer, *Gesichtserkennung*).

How deeply facial recognition and social media are already intertwined became obvious when, in January 2020, Clearview AI made global headlines for secretly scraping billions of facial images from social networks and fueling huge facial recognition databases with them (Hill, "The Secretive Company That Might End Privacy as We Know It"). The convergence of social media, smartphone photography, and facial recognition has not only compromised personal privacy in the digital sphere but also threatened to make anonymity in public places a thing of the past. It has also stimulated the development of a new generation of facial recognition algorithms that work very differently from their predecessors. Today's "deep learning" algorithms do not use predefined models of measurable facial features or follow predefined rules for abstracting statistical features from isolated, standardized images; rather, they are trained to recognize patterns in series of images showing the same face in multiple situations, poses, lighting conditions, etc. (Kelleher).

Training databases such as "Labeled Faces in the Wild" (Huang et al.) contain tens of thousands of images taken from the Web. Fed with these huge masses of images, artificial neural networks "learn" how to identify faces; that is, they autonomously develop the criteria by which

they recognize recurring facial patterns in different images without humans being able to control—or even understand—how they come to their conclusions. These new, deep-learning-based approaches have been highly effective, but one might ask what exactly they are effective for. After all, today's facial recognition is best at what it is trained for: comparing digital images in vast quantities and finding recurring patterns, as done by the FBI in the case of "Grey Carhartt Hat." Facial recognition, in this respect, is less a technology of "biometric" identification than an automated form of image comparison based on pattern recognition and the calculation of probabilities. The indexical relationship between the depicted and the depictee is assessed here very differently compared to the claims made about photography's indexical relationship to a here and now in classic photo theory, as perhaps most famously articulated by Roland Barthes.

In one sense, image comparison has been important from the very outset—only by translating living human bodies into stable two-dimensional images could discrete data be captured for comparison and identification of facial features. But whereas the older techniques of classification and identification used by Lavater, Bertillon, and Bledsoe treated the single, standardized image as a source of measurable features, today's deep-learning algorithms compare recurring visual patterns in vast, unstandardized ecologies of networked images and calculate their probability of matching. In a way, their operation resembles Galton's approach to pictorial statistics, but where Galton tried to mechanically synthesize the standardized portraits of multiple people into one image in order to reveal collective types, artificial neural networks extract statistical patterns from series of very different images of one and the same face (Alpaydin 23–5). In both cases, however, pictorial values and visual patterns such as brightness distributions play a more important role than do otherwise visible and measurable features. When Facebook was sued in the USA in 2015 for storing and processing "biometric" data of its users,

the company argued that it "did not collect anyone's biometric identifiers because its technology does not rely on 'human-notable' facial features" (Settlement Motion, 5).

Therefore, one could say that, rather than striving to establish an indexical truth about a single body via a single image, these algorithms use recurring patterns to establish probable links between different pictorial representations of one single body. Experimental indices nowadays are not limited to standardized forms of creating images, to isolated settings, or purpose-built databases. Rather, our networked image ecology as a whole has become the site of production of experimental indexical relations between recurring patterns, using methods validated only on statistical grounds. Facial recognition, it could be said, is now a large-scale technical, cultural, and social experiment linking dispersed images from previously unconnected situations to gain control over living bodies—an experiment that involves both images and metadata, police databases and social-media platforms, human and nonhuman actors, commercial companies and state institutions. However, as recent cases of wrongful arrests based on automatic facial recognition have demonstrated (Hill, "Wrongfully Accused by an Algorithm"), facial recognition is massively flawed, notoriously racially biased, and misidentifies especially people of color significantly more often than white people (Benjamin 112–3; Buolamwini and Gebru). In consequence, this experiment acts mostly to the detriment of those least able to defend themselves.

Unruly Relations between Physical Faces and Their Digital Counterparts

The history of automated facial recognition is much more than a history of technology. It is, as we have attempted to show in this essay, a history of shifting experimental settings wherein various technologies have been implemented for decision-making, a history of varying situational assemblages of bodies and cameras, images and data, formats and processes, material

infrastructure, social practices, and algorithmic operations. Each of these experimental settings was established to assert evidential truth claims and stabilize indexical connections that link faces, images, and identities. And each of those settings was, at the same time, a product of contingent circumstances and multiple factors, among them political and commercial interests, discursive conjunctions, transformations in visual culture, and technological developments.

Without reducing this complex history to an overly linear narrative, the experimental indices in automated facial recognition have been developed not least because of the following:

1. significant government interest in the use of facial recognition technology, which early on led to attempts to standardize the conditions for image capture and comparison to ensure interoperability, as visible in the work of Bertillon, the US FERET program, and the implementation of internationally standardized machine-readable biometric passports;
2. a tendency in the algorithmic models and procedures to use more and more abstract features for identification, from measurable distances to statistical patterns that are independent from reliance on a comparison with human vision;
3. an expansion of experimental settings in terms of scale, from very limited experiments with isolated physical bodies and single images to algorithmic processing of millions of images circulating online; and
4. a more fundamental transformation that reflects a general trend in today's networked image ecologies: growing dissolution of the formerly strict boundaries between the social spheres of private image-sharing, commercial distribution of images, and institutional use of images for identification purposes.

The different indexical relations that have become important throughout facial recognition's history—whether relying on an archival system and a way to measure standard deviations, as done by Bertillon, or on mechanical objectivity as was the case in the early days of using photography in passports, or on comparison of image patterns in today's searches for perpetrators in vast image ecologies—all show that a search for a stable indexical relationship based on technology alone is less useful than a focus on the kinds of assemblages needed to make indexical claims in the first place. We suggest examining the ways in which images are always enmeshed in networks of material infrastructure, technical operations, discursive attributions, and cultural practices from the outset. To understand indexical relationships within automated facial recognition, the attached claims to truth, and these images' social role and epistemic function, one has to detach the gaze from the individual image and begin to focus on the role of situational assemblages, which we understand as situationally ordered heterogeneous entities used to make claims and decisions. Only then can we properly critique the unruly relations that are claimed to exist between bodies and the technical inscriptions made of them.

Works Cited

Alpaydin, Ethem, *Machine Learning*. MIT Press, 2016.

Barthes, Roland. *Camera Lucida: Reflections on Photography*. Farrar, Straus and Giroux, 1982.

Batchen, Geoffrey. *Burning with Desire: The Conception of Photography*. MIT Press, 1999.

Benjamin, Ruha. *Race after Technology*: *Abolitionist Tools for the New Jim Code*. Polity, 2019.

Bledsoe, Woodrow Wilson. "A Facial Recognition Project Report," Panoramic Research, 1963, https://archive.org/details/firstfacialrecognitionresearch.

Boyer, Anne Olivia, and Robert S. Boyer. "A Biographical Sketch of W. W. Bledsoe," *Automated Reasoning: Essays in Honor of Woody Bledsoe*, edited by Anne Olivia Boyer and Robert S. Boyer, Springer, 1991, pp. 1–30.

Buolamwini, Joy, and Timnit Gebru. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." *Conference on Fairness, Accountability and Transparency*. PMLR, 2018.

Caplan, Jane. "'This or That Particular Person': Protocols of Identification in Nineteenth-Century Europe." *Documenting Individual Identity: The Development of State Practices in the Modern World*, edited by J. Caplan and J. C. Torpey. Princeton UP, 2001, pp. 49–66.

Crawford, Kate. *Atlas of AI: Power, Politics and the Planetary Costs of Artificial Intelligence*. Yale UP, 2021.

Daston, Lorraine, and Peter Galison. "The Image of Objectivity." *Representations*, vol. 40, 1992, pp. 81–128.

Ellenbogen, Josh. *Reasoned and Unreasoned Images: The Photography of Bertillon, Galton and Marey*. Pennsylvania State UP, 2012.

Gallagher, Sean. "Why Facial Recognition Tech Failed in the Boston Bombing Manhunt." *ArsTechnica*, 7 May 2013, https://arstechnica.com/information-technology/2013/05/why-facial-recognition-tech-failed-in-the-boston-bombing-manhunt/. Accessed 19 Mar. 2022.

Gates, Kelly A. *Our Biometric Future: Facial Recognition and the Culture of Surveillance*. New York UP, 2011.

Gray, Richard T. *About Face: German Physiognomic Thought from Lavater to Auschwitz*. Wayne State UP, 2004.

Grossenbacher, Timo, and Felix Michel. “Automatische Gesichtserkennung—So einfach ist es, eine Überwachungsmaschine zu bauen.” *SRF Data*, 7 Feb. 2021, https://www.srf.ch/news/schweiz/automatische-gesichtserkennung-so-einfach-ist-es-eine-ueberwachungsmaschine-zu-bauen. Accessed 19 Mar. 2022.

Hall, Louise. “Capitol Riot Suspect Arrested after FBI Use Face Recognition on Girlfriend’s Instagram.” *The Independent*, 2 Apr. 2021, https://www.independent.co.uk/news/world/americas/capitol-riot-stephen-randolph-instagram-b1835304.html. Accessed 19 Mar. 2022.

Harvey, Adam, and Jules LaPlace. Microsoft Celeb. *Exposing AI*, 2021, https://exposing.ai/msceleb/. Accessed 30 Jun. 2021.

Hausken, Liv. "The Face in the Biometric Passport." *The Art of Identification. Forensics, Surveillance, Identity*, edited by Rex Ferguson, Melissa M. Littlefield, and James Purdon, Penn State UP, 2021, pp. 159–181.

Hill, Kashmir. “The Secretive Company That Might End Privacy as We Know It.” *The New York Times*, 18 Jan. 2020, https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html. Accessed 19 Mar. 2022.

---. “Wrongfully Accused by an Algorithm.” *The New York Times*, 24 Jun. 2020, https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html. Accessed 19 Mar. 2022.

Huang, Gary B., Manu Ramesh, Tamara Berg, and Erik Learned-Miller. “Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments.”

University of Massachusetts at Amherst, Technical Report 07-49, Oct. 2007, http://vis-www.cs.umass.edu/lfw/lfw.pdf.

International Civil Aviation Organization (ICAO). *Doc 9303. Machine Readable Travel Documents*. Part 1: Introduction, 7th ed., 2015.

Introna, Lucas, and Helen Nissenbaum. *Facial Recognition Technology: A Survey of Policy and Implementation Issues*. New York University, 2017.

Jappy, Tony. *Introduction to Peircean Visual Semiotics*. Bloomsbury, 2013.

Kaltenbrunner, Wolfgang. 2018. Situated Knowledge Production, International Impact: Changing Publishing Practices in a German Engineering Department. *Minerva*, vol. 56, no. 3, pp. 283–303.

Kanade, Takeo. *Computer Recognition of Human Faces*. Birkhäuser, 1977.

Kean, Thomas H., et al. *The 9/11 Commission Report*. National Commission on Terrorist Attacks upon the United States, 2004.

Kelleher, John D., *Deep Learning*, MIT Press, 2019

Kember, Sarah. 2014. Face Recognition and the Emergence of Smart Photography. *Journal of Visual Culture*, vol. 13, no. 2, pp. 182–199.

Krauss, Rosalind. "Notes on the Index: Seventies Art in America." *October*, vol. 3, Spring 1977, pp. 68–81.

---. "Notes on the Index: Seventies Art in America." Part 2. *October*, vol. 4, Autumn 1977, pp. 58–67.

Latour, Bruno. "Visualisation and Cognition: Drawing Things Together." *Knowledge and Society: Studies in the Sociology of Culture Past and Present*, edited by H. Kuklick, Jai Press, 1986, vol. 6, pp. 1–40.

Lavater, Johann Caspar. *Physignomische Fragmente, zur Beförderung der Menschenkenntnis und Menschenliebe. Zweyter Versuch. Mit vielen Kupfertafeln*. Weidmanns Erben und Reich, 1776.

Lehmuskallio, Asko. "The Camera as a Meeting Place for Decision Making." *The Camera as Actor: Photography and the Embodiment of Technology*, edited by Amy Cox Hall, Routledge, 2020, pp. 17–28.

Lehmuskallio, Asko, and Haara P. "The Passport as a Medium of Movement." *Varieties of Cooperation. Mutually Making the Conditions of Mutual Making*, edited by Clemens Eisenmann, Kathrin Englert, Cornelius Schubert, and Ehler Voss, Wiesbaden, Springer VS, 2022, in press.

Lee-Morrison, Lila. *Portraits of Automated Facial Recognition: On Machinic Ways of Seeing the Face*. Transcript, 2019.

Meyer, Roland. *Gesichtserkennung* (*Digitale Bildkulturen*). Wagenbach, 2021.

---. *Operative Porträts. Eine Bildgeschichte der Identifizierbarkeit von Lavater bis Facebook*. Konstanz UP, 2019.

Müller, Martin. "Assemblages and Actor-networks: Rethinking Socio-material Power, Politics and Space." *Geography Compass*, vol. 9, no. 1, 2015, pp. 27–41.

Nelson, Lisa S. *America Identified: Biometric Technology and Society*. MIT Press, 2011.

Norval, Aletta, and Elpida Prasopoulou. "Public Faces? A Critical Exploration of the Diffusion of Face Recognition Technologies in Online Social Networks." *New Media & Society*, vol. 19, no. 4, 2017, pp. 637–54.

Phillips, P. Jonathon, Patrick J. Rauss, and Sandor Z. Der, *FERET (Face Recognition Technology). Recognition Algorithm Development and Test Results*, Technical Report ARL-TR #995), 1996.

Randolph Statement of Facts. United States Department of Justice, 16 Apr. 2021, https://www.justice.gov/usao-dc/case-multi-defendant/file/1388841/download.

Raviv, Shaun. "The Secret History of Facial Recognition." *Wired*, 22 Jan. 2020, https://www.wired.com/story/secret-history-facial-recognition/. Accessed 19 Mar. 2022.

Robertson, Craig. *The Passport in America: The History of a Document.* Oxford UP, 2010.

Robins, Alexander. "Peirce and Photography: Art, Semiotics, and Science." *The Journal of Speculative Philosophy*, vol. 28, no. 1, 2014, pp. 1–16.

Sekula, Allan. "The Body and the Archive." *October*, vol. 39, 1986, pp. 3–64.

Settlement Motion in Case 3:15-cv-03747-JD, United States District Court for the Northern District of California, San Francisco Division (document 445), filed 8 May 2020, https://www.courthousenews.com/wp-content/uploads/2020/06/Facebook-FacialDataSettlementMotion.pdf.

Sirovich, Lawrence, and Michael Kirby. "Low-Dimensional Procedure for the Characterization of Human Faces." *Journal of the Optical Society of America*, vol. 4, no. 3, 1987, pp. 519–24.

Stanley, Jay, and Barry Steinhardt. *Drawing a Blank: The Failure of Facial Recognition Technology in Tampa*, *Florida*. ACLU Special Report, January 2002.

Tagg, John. *The Burden of Representation: Essays on Photographies and Histories*. University of Minnesota Press, 1993.

Toister, Yanai. *Photography from the Turin Shroud to the Turing Machine*. Intellect Books, 2020.

Torpey, John C. *The Invention of the Passport*. Cambridge UP, 2018.

Turk, Matthew, and Alex Pentland. "Eigenfaces for Recognition." *Journal of Cognitive Neuroscience*, vol. 3, no. 1, 1991, pp. 71–86.

Wayman, James L. "The Scientific Development of Biometrics over the Last 40 Years." *The History of Information Security: A Comprehensive Handbook*, edited by Karl de Leeuw and Jan Bergstra, Elsevier, 2007, pp. 263–74.

Asko Lehmuskallio is Professor of Visual Studies in the Communication Sciences Unit at Tampere University. His work focuses on the intersection between visual studies and media studies, with a special interest in media anthropology and digital cultures. Among his publications are the coedited works *Digital Photography and Everyday Life* (2016) and *Mobile Digital Practices* (2017) and the Finnish-language *Entanglements between Bodies and Documents: A History of the Finnish Passport* (2020), co-authored with Paula Haara. In addition, he has published in a variety of journals and edited books, including in *Visual Communication*, *International Journal of Communication*, *IEEE Pervasive Computing*, and *Journal of Aesthetics and Culture*. Email: asko.lehmuskallio@tuni.fi.

Roland Meyer is a postdoctoral researcher and lecturer in art history at the Brandenburg University of Technology Cottbus-Senftenberg. His research focuses on the history and theory of technical images, with a special interest in the history of facial recognition, identification techniques, and image forensics. Recent publications include *Operative Porträts. Eine Bildgeschichte der Identifizierbarkeit von Lavater bis Facebook* (Konstanz 2019); "Formatting Faces: Standards of Production, Networks of Circulation, and the Operationalization of the Photographic Portrait," in *Format Matters: Standards, Practices, and Politics in Media Cultures* (Lüneburg 2020); and *Gesichtserkennung*, Series *Digitale Bildkulturen* (Berlin 2021). Email: roland.meyer@b-tu.de.

The Journal of Media Art Study and Theory
Volume 3, Issue 1, 2022
Automating Visuality
www.mast-journal.org
MAST

Automated Visions, Algorithmic Imageflows: The Technopolitics of Black Lives Matter Videos on YouTube

Jason W. Buel
North Carolina Wesleyan College

Abstract

This essay considers how mechanisms of machine vision intervene as forms of "social sorting" and subject formation in the context of YouTube's algorithmic flows of images. Too often, algorithms are treated as neutral, unbiased processes. In reality, many algorithms reinscribe and reinforce human biases. This essay focuses on the power of YouTube's algorithms to shape viewers' understandings of the Black Lives Matter movement, focusing specifically on what Chris Ingraham calls the micro-rhetorical tier of algorithmic processing. The essay employs critical cultural studies methods to rigorously contextualize and compare case studies of algorithmically-suggested content connected to pro-Black Lives Matter videos. In this context, I argue that these automated flows of images become less about what any specific video shows about the need for radical socio-political change and more about the articulation of an idealized viewing position and idealized viewing subjects.

Keywords

algorithms; YouTube; Black Lives Matter; social movements; digital media

In the early hours of New Year's Day 2009, a transit cop shot and killed Oscar Grant on a subway platform in Oakland, California. Grant was unarmed. There is no evidence that he was

posing any threat to the officer. At least five cell phone videos of the incident taken by other subway passengers show that Grant was lying face down with his hands behind his back and another officer's knee holding his head against the ground when he was shot. Without video evidence of this event, it might be hard for the portions of the general public who do not experience police terror on a regular basis to believe that it happened at all. With video evidence, though, it became undeniable. Still, even with an abundance of publicly available audiovisual evidence, the officer who killed Grant served less than two years in jail on a charge of involuntary manslaughter. Grant's death and the incongruity of the crime's severity with the killer's sentence sparked three founders of Black Lives Matter, Alicia Garza, Patrisse Cullors, and Opal Tometi, to begin organizing around racial injustices in policing during an era when many pointed to the election of Barack Obama as evidence that the US had become a "post-racial" nation that had moved beyond the need to critically examine issues of race (Segalov).

Such videos can help spread awareness and engage viewers who might not otherwise be moved to act (Malkowski). Bay Area broadcast station KTVU-TV posted cell phone videos it received to its website, where they were viewed nearly half a million times in just a few days after Grant was killed by the police (Stannard and Bulwa). An annotated version of one of the videos was posted to YouTube, where it received more than one thousand views per hour on average for the first week after Grant's death. As Jennifer Malkowski argues, YouTube played a primary role in making videos of Grant's death not just available but spreadable.

To consider YouTube videos as self-contained texts whose power exists exclusively in their ability to represent events, though, would be to overlook the platform's role in doing what its affordances and economic conditions dictate it must do: produce audiovisual flows capable of holding viewer attention, then hold that attention as long as possible in order to commodify it

and transform it into revenue. Part of the way YouTube does this is through its default autoplay function combined with its recommendation algorithms. Through these, the platform produces audiovisual flows. The representational components within any given video become secondary to the flow's capacity to modulate affect. The social imaginaries produced by these audiovisual flows, then, are also not reducible to the representational production of meaning within any given video. The platform produces a social imaginary of its own. What are the political horizons of this imaginary? To what extent is it able to redirect energy and attention away from the political horizons envisioned in any given video?

This essay seeks answers to these questions by interrogating the automated audiovisual flows that the platform produces in response to the input "Black Lives Matter." I examine the micro-rhetorical effects of the platform through the lens of a materialist approach to media in order to investigate representational effects as one subset of material effects produced by and through the platform. Specifically, I focus on the boundaries between representational and non- or super-representational effects as well as the crossing and, in some cases, erosion of these boundaries. I consider how mechanisms of machine vision intervene as forms of "social sorting" and subject formation in the context of YouTube's algorithmic flows of images and how these flows express the power of YouTube's recommendation algorithm to shape viewers' understandings and affective predispositions toward the Black Lives Matter movement. In this context, I argue that these automated flows of images become less about what any specific video has to say about the need for radical socio-political change and more about the viewing subject: the articulation of an idealized viewing position—a seemingly omniscient position that exists as an assemblage distributed across the "machine" of internet infrastructures—as well as the articulation of an idealized viewer that interpellates any actual viewers. This process automates

the production of a capitalist realism that fundamentally limits the potential of social movement discourses to bring about change. I examine this phenomenon through what Félix Guattari calls a "non-reductive pragmatic analysis" aimed toward better understanding the "economy of desires in the social field" produced on YouTube in part through its use of recommendation algorithms and endless flows of images stitched together by its autoplay function (*Lines of Flight* 9).

Beyond Representation: Data, Flows, Affect

Too often, algorithms are treated as neutral or objective processes that are somehow free of bias. In recent years, there has been a wave of scholarship that is working to debunk that misconception, but the misconception remains—especially in popular discourses (Eubanks; O'Neil; Gillespie; Vaidyanathan). Safiya Umoja Noble's work has been particularly important in not just drawing attention to the biases of algorithmic processes but also to their power to reinscribe racist narratives and ways of thinking. To understand how such processes play out within YouTube's algorithmic ecology, the next section will focus specifically on what Chris Ingraham calls the micro-rhetorical tier of algorithmic processing by providing close readings not just of texts themselves but of the automatic connections between videos that the platform makes as well as the way that the platform blends audiovisual materials together to manufacture automated flows.

Joining together videos through automated processes creates juxtapositions based not on the content of the videos but on metadata created from tracking past viewers' engagement with each video. That metadata suggests some filiation between videos based on past viewers' habits taken in aggregate. This process, then, creates an ideal viewing subject that is an amalgamation of all past viewers. This articulation of an imagined ideal viewer is particularly powerful when it comes to political nonfiction videos. YouTube's interface constructs a passive viewer who

observes rather than acts. The platform's logics imagine the ideal viewer to be someone who makes no decisions apart from selecting an initial video to watch and allows the algorithms to determine their unending flow of content for them.

There are, of course, key differences between the sort of flow we experience on YouTube and televisual flow as Raymond Williams first considered it. The two most significant differences are the automatability of YouTube's flows as well as the disconnect between video creators and the programmers who write the code that automates the flow. In television, the imagemakers and flow-constructors would be working within the same institutions, aware of one another and capable of interacting directly. In the case of YouTube, the person who creates the image, the person who uploads it, and the many different people who write the code all exist in their own places and times. They may have no knowledge of one another, much less any ability to communicate directly with one another in order to exert some limited agency over how their work fits into the larger flow of images. In other contexts, algorithmic editing is used to interrogate the performative aspects of software (Enns). The default sort of algorithmic "editing" that YouTube generates, however, is not developed by artists or activists but by software engineers. They are not incentivized to explore or critique but to maximize the ability for the platform to capture the attention of users and ultimately monetize that attention. It is this overarching reduction to capitalistic capture that forms the basis of Alex Juhasz's critique of YouTube: even in the best of circumstances, any activist video on YouTube is being mobilized to serve the interests of private industry at least to some extent. This criticism is similar to what Jodi Dean suggests about all forms of user-generated content on the web: no matter the aims of a particular video, tweet, blog post, or picture, the affects produced by their transmission are inevitably captured by consumer capitalism and further reinforce its power.

Taking any given video that YouTube returns in a search for "Black Lives Matter" as a starting point, the automated flow of videos quickly moves away from—or many times directly against—the Black Lives Matter movement. In the hundreds of times I began with a pro-Black Lives Matter video and tracked the flows that followed from it, these algorithms almost never directed me to more than one other overtly pro-Black Lives Matter video. Instead, the flows tended to move quickly toward content that explicitly opposed the movement. Often, these flows became saturated with remediated clips from twenty-four hour cable news programs, especially FOX News. This is consistent with existing findings that videos from accounts associated with mainstream media outlets, which consist almost entirely of television clips, have twenty-four times as many videos as accounts associated with the more overtly reactionary "YouTube Right" (Munger and Phillips). If I followed a flow for long enough, it often eventually moved on to other political issues that had been put on the public agenda by these traditional media giants and/or to topics that have nothing to do with the Black Lives Matter movement and that do not appear to be overtly political at all: shark attacks, discussions of whether or not a particular musician is overrated, and discussions of the Kardashians, for example.

Notably, my first foray through these algorithmic flows provided, in addition to cable news critiques of Black Lives Matter, critiques of the movement's tactics from people who presented themselves as generally sympathetic to the movement as well as critiques from the far right. One particularly troubling example, which has since been removed, came from a company called Florida Gun Supply. The video presents Black Lives Matter activists as "a gang." The host asserts that "we" need to stand up against this gang by arming ourselves and "carrying daily," presumably with guns purchased from Florida Gun Supply itself. Not all of the critiques were from the far right, though. There were also critiques from the far left in the case of one evening-

news-style segment from Maoist Rebel News. This critique comes to the conclusion that Black Lives Matter's tactics were fundamentally flawed because the movement lacks a strong central authority figure.

The overwhelming tendency within these algorithmic flows is to move toward de-politicized or counter-politicized messages. Doing so explicitly reframes the videos seeking to objectively or favorably represent the movement as outliers that are isolated and unworthy of further consideration. The platform does not facilitate discussion here. There is no genuine exchange of ideas between those who support and those who oppose the movement. Instead, the discourse that this movement is nothing but "thugs" and entitled radicals persists through what appears to be a preponderance of "evidence" provided through the flow itself. Even if a particular video crafts a perfectly flawless argument with substantial visible evidence, its capacity to persuade is limited when it is stitched together into an automated flow. In the most extreme instances, such flows contain one person after another reinforcing the already-dominant discourses around the movement. Even in the best cases, the flows present little to nothing that would enable the viewer to further explore the movement from the perspective of its participants.

At the most, the one video immediately following my chosen entry-point video contained a similar political formation of the truth that the movement seeks to address, though often even the very first algorithmically selected video departed significantly in terms of its framing of the movement or even its subject matter generally. There also appears to be a heavy bias driving the flow toward established media companies, which is also consistent with existing studies (Munger and Phillips). Often, these are video clips that remediate TV news (most often FOX News in my experiences immersing myself in this flow, but also frequently CNN, MSNBC, and CBS). Sometimes they are clips from other streaming media outlets like The Young Turks or

self-contained videos (not clips) produced by professional media companies like Pitchfork. In the cases of FOX News, The Young Turks, and Pitchfork, entering the flow of that particular media company provided no escape without human intervention. That is, once the algorithms drove the flow into one remediated FOX News clip, it seemed there was no choice but to swirl around in other FOX News clips indefinitely until the viewer stopped it. It appears, then, that there is a tipping point within these flows: they continue to present media that frames itself as nonfiction, but they move quickly and irreversibly toward anti-Black Lives Matter media, eventually to unrelated clips typically from professional media companies. Once each of those thresholds is crossed, it appears to rarely if ever go back. Of course, because YouTube's algorithms are proprietary, there is no way to know exactly how they function to process video and user data in developing these flows; we are instead left to infer their functionality from the traces of their processes that we receive as viewers. As the algorithms connect these videos by a logic based on the established "old media" institutions, it becomes clear that this process of flow is the antithesis of the movement itself. The movement is, after all, based on logics of decentralization, polyvocality, nonhierarchical organization, heterogeneity, and disruption.

There is risk in this flow of images, and that risk is magnified by the amateur status of so many creators of documentary work on YouTube. An activist's video in support of any aspect of Black Lives Matter risks having itself joined into a series of moving images in which it functions as an entry point into a world of self-regenerating consumer capitalism at best. More often, it is an entry point into a world of explicit hate and oppression. YouTube's algorithms are not solely responsible for these outcomes nor do they function in complete isolation: viewers must be complicit and accept the platform's enticement to passively accept the content it strings together. The algorithms capture, abstract, and distill datafied traces of past viewers' collective habits and

behaviors, then use such aggregates to automate the experiences of individual viewers in the present. In this way, they automate and inscribe hegemony at a scale that is removed from individual human sense perception. They inscribe their logics onto us as a collective, amorphous machine-imagined "community" of isolated users.

Automation and Algorithmic Control

Under Integrated World Capitalism, humans have ceded a great deal of control over our lives to automated decision-making processes (Guattari, *The Three Ecologies*). For example, algorithms increasingly manage the workplace by scheduling workers' shifts, surveilling workers on the job, quantifying attention in the form of a worker's "time on task" and flagging those who fall short for disciplinary action, and even making hiring and firing decisions (Todolí-Signes; Adler-Bell and Miller; Aloisi and Gramano; Rosenblat). While YouTube's recommendation algorithms are not making such decisions, this context nevertheless illustrates the power of algorithms not just to shape representations but to enact material changes that impinge on people's lives. With YouTube's joining of algorithmic recommendations with an autoplay system structured to create flows, the platform develops for itself a great deal of power to shape meaning as well as viewer experience beyond representation. As Stefania Milan put it, "No longer mere platforms, social media have become actors in their own right, intervening in the meaning-making process of social actors by means of their algorithmic power" (887). Social media recommendation algorithms, including YouTube's, have been connected with a rise in white nationalism and far-right politics (Daniels, "The Algorithmic Rise"). For example, white supremacist murder Dylann Roof's own manifesto describes his experience of searching for "black on white crime" and being directed toward white supremacist content as being an integral part of his developing a violent, reactionary political worldview (Daniels, "The Algorithmic Rise" 62). As Noble points

out, this is not an aberration but part of a larger tendency of internet algorithms to perpetuate racial bias and inequality ("Google Search").

Of course, algorithms are not the only aspects of digital media that are driving people toward white supremacist content (Daniels, "Race and Racism"; DeCook; Nagle). Celebrities within the alt-right media ecosystem play a major role in this trend by cultivating parasocial relationships with their audiences, then driving audiences toward other far-right figures (Lewis, "This Is What the News"). Becca Lewis's work mapping alternative influence networks on YouTube found that anti-feminist sentiment is a driver of radicalization across various political ideologies: for example, self-described "classical liberal" Dave Rubin's YouTube channel is only two degrees of separation away from white nationalist Colin Robertson's channel because they have both featured the explicitly anti-feminist Sargon of Akkad on their channels, thereby introducing the devoted followers of one channel to the other through their shared anti-feminism ("Alternative Influence" 11). The affective power of homophobic and anti-feminist sentiments appears to drive men toward white nationalism (Bjork-James). White nationalist videos on YouTube frequently attempt to cultivate fear around Muslims, immigrants, and feminists by explicitly foregrounding the perceived threats these groups pose to the status of white men (Hawkins and Saleem). Such connections have everything to do with the affective moods generated by the content of the videos as well as the circulation of such videos as they are shared on other social media platforms. Algorithmic recommendation, then, is but one vector in a larger cluster of quasi-causes driving this mainstreaming of white nationalist content.

Algorithmic filtering and recommendation are still important, though, particularly in the way that they carry the past forward into the future by their very nature. Their predictions about what will hold any given viewer's attention in this moment and into the next are necessarily

based on data collected from past viewers. This allows any social biases, values, or attitudes to be reflected in those suggestions. The algorithm's flows prescribe normative values: they convey a sense that this is what a normal viewer ought to want. This enacts a conditioning of desire that, while it may not be effective for most or even many viewers, is nevertheless present. Such an effect is particularly powerful when, in the example of searching for videos on "Black Lives Matter," one is presented in sequence with an endless flow of other concerns that, from the apparent perspective of the platform, the ideal viewer ought to care about instead. Of course, such an imagined subject will never exist, but it is the platform's reaching for it—the *becoming* inherent in this subjectivation process—that matters. This is an enactment at the level of the platform of what Ingraham calls "concerned gestures" that "beckon toward some potential that they seldom see actualized except through the realization of reaching for it" (2). The platform gestures toward a subject who is too thoroughly immersed in the endless flows of both images and capital to ever imagine an alternative to either.

YouTube's Automation of Hegemony

Attempting to use YouTube to learn about the Black Lives Matter movement has major political limitations. The most pervasive limitation comes from the platform's automated enacting of capitalist realism: "the widespread sense that not only is capitalism the only viable political and economic system, but also that it is now impossible even to *imagine* a coherent alternative to it" (Fisher 2). While Mark Fisher discusses this as a socio-cultural phenomenon, YouTube enacts it at the level of the machine, thereby further insulating it from direct human agency. If we accept Becky Kazansky and Stefania Milan's claims that technology constitutes the "digital backbone" of social imagination and that social movements are engines driving alternative imaginaries

(366), then what happens when social movement discourse is itself bounded and redirected toward different visions by technological systems like YouTube?

This bounding and redirection of discourse within a platform that otherwise presents itself as being a radically open and democratic space is even more powerful given that it appears as a seemingly unbiased technical decision being made by an objective machine that is free from direct human influence or bias. The exact workings of the algorithm are proprietary and, thus, opaque. This creates a "black box" effect whereby we can observe the functioning only from the outside and cannot look "beneath the hood" of the algorithm to see exactly what is happening (Gillespie; O'Neil). This black box effect is something that does not serve the public interest but does serve YouTube's ability to generate revenue. What is hidden is not only the precise technical functioning of the algorithms but also the human labor and human values that go into shaping the platform. As Nick Seaver puts it: "The 'black box' is full of people who design, build, and maintain it; algorithmic systems can extend and scale up their all-too-human biases and worldviews" ("Seeing Like an Infrastructure" 773). This ability to not only preserve but also scale up biases is precisely what makes the political role of algorithms worthy of special attention. The algorithm also exerts influence in the short term in the way that it interprets a viewer's failure to disrupt the flow of images as the viewer's desire to continue seeing similar content in the future. For example, Anthony Burton et al. found that after just one day of allowing right-wing recommendations to play, YouTube's main page "was swamped with predominantly right-wing content."

YouTube's algorithms exert control within a mixed semiotic system that, like other forms of surveillance, "produces a social body, rather than straightforwardly reflect[ing]" one (Nakamura 150). Guattari developed mixed semiotics as a theoretical framework that considers

how linguistic signifiers, non-linguistic signs like gestures, and codes that are not directly accessible to humans and instead operate directly on material flows all interact and shape one another (Lazzarato; Guattari, "The Place of"). In digital communication, such an understanding is key precisely because of the ways in which the machines structuring these communications mask their internal processes and because information, no matter how ephemeral it may appear in practice, is always embodied (Chun). The production of meaning through semiotic representation, then, is only one aspect of communication and not necessarily the most important one. If we understand rhetoric as the discovery of all available means of persuasion, then there is no reason to limit such means to the realm of semiotic signification. Moreover, such semiotic signification is often subservient to other a-semiotic encodings of asignifying systems. This is especially true in digitally mediated environments. As Guattari puts it, "a-signifying machines remain based on signifying semiotics, but no longer use them as anything but a tool" (*Molecular Revolution* 75). Asignifying semiotics "connect an organ, a system of perception, an intellectual activity, and so on, directly to a machine, procedures and signs, bypassing the representations of a subject" (Lazzarato 40). This has particularly important consequences in the context of late capitalism because "what matters to capitalism in controlling the asignifying semiotics apparatuses . . . through which it aims to depoliticize and depersonalize power relations" (Lazzarato 41).

The articulation of an ideal viewing subject further entrenches capitalist realism by making it more difficult for collective enunciations of alternative social imaginaries to emerge. Social imaginaries are the outcomes of "collective sense-making activities" that produce shared ideas including "fears, hopes, and expectations" (Kazansky and Milan 364). One need not trace the audiovisual flows from an initial search of the phrase "Black Lives Matter" for long before

observing a habitual drift away from a politics of hope and toward a politics of fear. Such a shift is useful to the platform: a politics of hope might require logging off, but a politics of fear often requires staying put. What emerges is the expression of a power relation articulated through collecting and sequencing audiovisual fields. The expressive impact of these audiovisual fields is not reducible to the meaning produced representationally within any individual video. As Guattari puts it, "The organisation of contents, the constitution of a homogeneous field of representation, always corresponds to the crystallisation of a power formation" (*Lines of Flight* 137). In the case of flows produced by YouTube in response to searching "Black Lives Matter," the power of the platform is crystallized through its habitual production and modulation of presence-effects. "Presence-effects" describe a phenomenological layer of encounters that have impacts beyond the extrapolating "meaning" to make sense of an encounter through interpretation (Gumbrecht 79). Presence-effects "are especially important in a time when we are so inundated with information that now nearly anything can be signal and anything noise, depending on whom you ask and which algorithms are doing the sorting" (Ingraham, *Gestures* 4). It is because of the power of presence-effects that Ingraham considers meaning an "epiphenomenon" of communication (*Gestures* 4). In the case of YouTube-produced flows following from searching "Black Lives Matter," the viewer's experience of the flow state itself may be considered a presence-effect. More specifically, any videos containing footage of street demonstrations and embodied direct-action protests are, through these flows, deterritorialized from the streets and from any coherent left-wing political projects. They are reterritorialized, in many instances, into flows of far-right paranoia. The affective moods produced through the confluence of endless flows of images and the temporary audiovisual fields they call into being out of theoretically endless possible combinations matter at least as much as the actual meaning

of the words spoken and actions represented. They gesture away from the movement and the alternative social imaginaries it seeks to create. In gesturing away, the destination matters less than the presence-effect of the gesture: always away, always to elsewhere.

Conclusion

Automated decisions shape our audiovisual environments, and their power to do so continues to grow. Even media giants like Warner Bros. are beginning to use artificial intelligence software to automate decisions that would traditionally be made by humans, like assessing a star's worth to a particular project (Siegel). One can easily see how such decisions, once they are offloaded to digital automation, could be used to frame racist and sexist hiring practices as accidental "bugs" of complex technical systems. Automation provides the businesses that operate and use these platforms with a way of further diffusing responsibility across an assemblage that includes both human and nonhuman components.

YouTube is a sociotechnical assemblage. Its algorithmic flows pull together never-ending streams of images. What associations and patterns appear in its stream of not-quite consciousness? What political imaginaries emerge through this assemblage? We must consider the ways that both the social and the technological recursively shape one another. To do so, we must consider this assemblage not only in terms of the more or less likely outcomes the algorithms tend to produce but also as accidental experiments in the production of automated imaginaries. Such imaginaries regularly come into being no matter how unlikely any one particular flow of images may be. Examining these, we can attend "to how things are actually working out relationally betwixt manifestation and possibility" (Genosko 10). In YouTube's case, the platform's automated production of audiovisual flows enacts a drifting away. This gesture of drifting away recontextualizes the meaning of any given video that appears at the

same time that it produces a drifting away at the level of presence-effects. It conditions the body to remain engaged in the flow rather than to heed any calls to action that may be expressed within any given video, no matter how compelling such calls may be on their own.

Works Cited

Adler-Bell, Sam, and Michelle Miller. "The Datafication of Employment." The Century Foundation, 19 Dec. 2018, tcf.org/content/report/datafication-employment-surveillance-capitalism-shaping-workers-futures-without-knowledge/. Accessed 1 May 2021.

Aloisi, Antonio, and Elena Gramano. "Artificial Intelligence is Watching You at Work: Digital Surveillance, Employee Monitoring, and Regulatory Issues in the EU Context." *Comparative Labor Law & Policy Journal*, vol. 41, 2019, pp. 101–127.

Bjork-James, Sophie. "Racializing Misogyny: Sexuality and Gender in the New Online White Nationalism." *Feminist Anthropology*, vol. 1, no. 2, 2020, pp. 176–83.

Burton, Anthony, et al. "Streams of the Deep Web: Rebel Media, YouTube, and the Algorithmic Shaping of Media Ecosystems." Digital Methods Initiative, 2019, https://wiki.digitalmethods.net/Dmi/SummerSchool2019StreamsoftheDeepWeb. Accessed 3 Apr. 2021.

"Captured by 6 different cameras BART police shoot and kill unarmed Oscar Grant." YouTube, uploaded by StreetTV, 4 Jul. 2010, www.youtube.com/watch?v=rSN5WF9qD3g. Accessed 20 Mar. 2022.

Carey, James. *Communication as Culture*. Routledge, 1994.

Chun, Wendy. *Programmed Visions: Software and Memory.* MIT Press, 2011.

Daniels, Jessie. “Race and Racism in Internet Studies: A Review and Critique.” *New Media & Society*, vol. 15, no. 5, 2013, pp. 695–719.

---. “The Algorithmic Rise of the ‘Alt-Right.’” *Contexts*, vol. 17, no. 1, 2018, pp. 60–5.

Dean, Jodi. “Affective Networks.” *MediaTropes*, vol. 2, no. 2, 2010, pp. 19–44.

DeCook, Julia R. “Memes and Symbolic Violence: #proudboys and the Use of Memes for Propaganda and the Construction of Collective Identity.” *Learning, Media and Technology*, vol. 43, no. 4, 2018, pp. 485-504.

Enns, Clint. “Navigating Algorithmic Editing: Algorithmic Editing as an Alternative Approach to Database Cinema.” *Millennium Film Journal*, 2012, pp. 66–72.

Eubanks, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* St. Martin's Press, 2018.

Fisher, Mark. *Capitalist Realism: Is There No Alternative?* Zero Books, 2009.

Florida Gun Supply. “America is under attack by Black Lives Matter: Unite Against BLM!” 4 Sep. 2015, www.youtube.com/watch?v=afE8mymaE1o. Accessed 14 Mar. 2016.

Genosko, Gary. “Felix Guattari: A Critical Introduction.” Pluto Press, 2009.

Gillespie, Tarleton. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale UP, 2018.

Guattari, Félix. *Lines of Flight: For Another World of Possibilities*. Translated by Andrew Goffey, Bloomsbury Publishing, 2015.

---. *Molecular Revolution: Psychiatry and Politics*. Translated by Rosemary Sheed, Penguin, 1984.

---. *The Three Ecologies.* Continuum, 2008.

---. "The Place of the Signifier in the Institution." *The Guattari Reader*, edited by Gary Genosko, Blackwell, 1996.

Hawkins, Ian, and Muniba Saleem. "Rise UP!: A Content Analytic Study of How Collective Action Is Discussed within White Nationalist Videos on YouTube." *New Media & Society*, 2021, doi:10.1177/14614448211040520.

Ingraham, Chris. *Gestures of Concern.* Duke UP, 2020.

---. "Toward an Algorithmic Rhetoric." *Digital Rhetoric and Global Literacies: Communication Modes and Digital Practices in the Networked World*, edited by Gustav Verhulsdonck and Marohang Limbu, IGI Global, 2014, pp. 62–79.

Juhasz, Alexandra. *Learning from YouTube.* MIT Press, 2011, vectors.usc.edu/projects/learningfromyoutube/texteo.php?composite=14. Accessed 20 Mar. 2022.

Kazansky, Becky, and Stefania Milan. "'Bodies Not Templates': Contesting Dominant Algorithmic Imaginaries." *New Media & Society* vol. 23, no. 2, 2021, pp. 363–81.

Lazzarato, Maurizio. *Signs and Machines*. Translated by Joshua David Jordan, Semiotext(e), 2014.

Lewis, Rebecca. "Alternative Influence: Broadcasting the Reactionary Right on YouTube." Data & Society, 2018, datasociety.net/wp-content/uploads/2018/09/DS_Alternative_Influence.pdf. Accessed 2 Mar. 2021.

---. "'This Is What the News Won't Show You': YouTube Creators and the Reactionary Politics of Micro-Celebrity." *Television & New Media*, vol. 21, no. 2, 2020, pp. 201–17.

Malkowski, Jennifer. "Streaming Death: The Politics of Dying on YouTube." *Jump Cut,* vol. 54, 2012, https://www.ejumpcut.org/archive/jc54.2012/malkowskiYoutubeDeaths/text.html. Accessed 22 Mar. 2022.

Milan, Stefania. "From Social Movements to Cloud Protesting: The Evolution of Collective Identity." *Information, Communication & Society*, vol. 18, no. 8, 2015, pp. 887–900.

Munger, Kevin, and Joseph Phillips. "Right-Wing YouTube: A Supply and Demand Perspective." *The International Journal of Press/Politics*, 2020, doi:10.1177/1940161220964767.

Nagle, Angela. *Kill All Normies: Online Culture Wars from 4chan and Tumblr to Trump and the Alt-Right,* Zero Books, 2017.

Nakamura, Lisa. "The Socioalgorithmics of Race: Sorting it out in Jihad Worlds." *The New Media of Surveillance*, edited by Shoshana Magnet and Kelly Gates, Routledge, 2009, pp. 149–62.

Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press, 2018.

---. "Google Search: Hyper-Visibility as a Means of Rendering Black Women and Girls Invisible." *InVisible Culture*, vol. 19, 2013, pp. 1–23.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Books, 2016.

Rosenblat, Alex. *Uberland: How Algorithms are Rewriting the Rules of Work*. University of California Press, 2018.

Seaver, Nick. "Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems." *Big Data & Society,* vol. 4, no. 2, 2017, pp. 1–12.

---. "Seeing Like an Infrastructure: Avidity and Difference in Algorithmic Recommendation." *Cultural Studies*, online preprint, 2021, pp. 771–91, doi: 10.1080/09502386.2021.1895248

Segalov, Michael. "We Spoke to the Activist behind #BlackLivesMatter about Racism in Britain and America." *Vice*, 2 Feb. 2015, www.vice.com/read/patrisse-cullors-interview-michael-segalov-188. Accessed 5 May 2015.

Siegel, Tatiana. "Warner Bros. Signs Deal for AI-Driven Film Management System." *The Hollywood Reporter,* 8 Jan. 2020. www.hollywoodreporter.com/business/business-news/warner-bros-signs-deal-ai-driven-film-management-system-1268036/. Accessed 1 May 2021.

Stannard, Matthew B., and Demian Bulwa. "BART Shooting Captured on Video." *SFGate.com,* 7 Jan. 2009, www.sfgate.com/bayarea/article/BART-shooting-captured-on-video-3177850.php#photo-2302188. Accessed 4 May 2015.

Todolí-Signes, Adrián. "Making Algorithms Safe for Workers: Occupational Risks Associated with Work Managed by Artificial Intelligence." *Transfer: European Review of Labour and Research*, Aug. 2021, doi: 10.1177/10242589211035040.

Vaidhyanathan, Siva. *Antisocial Media: How Facebook Disconnects Us and Undermines Democracy*. Oxford UP, 2018.

Williams, Raymond. *Television: Technology and Cultural Form.* Fontana, 1974.

Jason W. Buel is an assistant professor in the Communication and Cinema & Media Arts programs at North Carolina Wesleyan College. He earned his Ph.D. in Communication, Rhetoric, and Digital Media from North Carolina State University. Dr. Buel studies the intersections between media, technology, and politics focusing on emerging documentary

practices, especially within the media ecologies of contemporary social movements. His article "Assembling the Living Archive: A Media-Archaeological Excavation of Occupy Wall Street" appears in *Public Culture* 30.2. His research on activist video in the Idle No More movement is available in the edited collection *InsUrgent Media from the Front*. Email: jwbuel@gmail.com.

The Journal of Media Art Study and Theory
Volume 3, Issue 1, 2022
Automating Visuality
www.mast-journal.org
MAST

Bringing Colonialism into the Frame: A Conversation with Heba Y. Amin

Dominique Routhier
University of Southern Denmark

In 2013, Egyptian authorities detained a stork on suspicions of espionage. The stork was equipped with an electronic tracking device that was mistakenly assumed to be a piece of hostile surveillance equipment. To the entertainment of a then recently booming social media public, the story of the overly paranoid Egyptian authorities soon went viral. To the Egyptian artist Heba Y. Amin, whose art project The General's Stork *(2016–ongoing) critically mines this incident, the story of the captive bird reveals something less amusing about how visual technologies leading up to automated drone surveillance have been developed and deployed in a colonial context. Amin's art thus prompts us to fundamentally reconsider the "automation of visuality" by bringing colonialism into the frame.*

DOMINIQUE ROUTHIER: Beginning from your art project *The General's Stork*, could you elaborate on the context of this work and what is it essentially about?

HEBA Y. AMIN: In general, my art and research centers on questions of power in relation to technology and its role in visual representation. I've been investigating the correlation between the development of technology and the political implications of image production in the Middle East and Africa in the last 150 years or so. Broadly speaking, this is the thread that ties my body

of work together, which matured in a meaningful way during the 2011 Egyptian revolution and the years that followed. This was a turning point, when social media and other technological platforms played a big role in allowing us to contextualize the uprisings and politics within a broader history, especially with the massive production of visual material.

In 2013, I started investigating a viral media story about a stork that was captured in southern Egypt and accused of espionage. The story was broadcast by multiple Western media outlets in a tone that, presumably, had something to do with their disapproval of Egyptian politics and the turn of events that had unfolded at that time. The so-called democratically elected leader of the Muslim brotherhood had been overthrown by a second wave of uprisings, and Egypt's revolution was exposing itself as a farce. The spy stork was seemingly instrumentalized to mock the deteriorating political situation.

However, suspicion to the extent of detaining a bird is indicative of something more deeply dysfunctional, and I was motivated to find its source. What I did not know is that it would lead me to uncover the complicated history of aerial imaging technologies and how Middle Eastern landscapes have served as the backdrop for the development of drone warfare since the late nineteenth century. If one considers the whole arc of this narrative and the context of how these technologies were developed, it is not so strange that a bird with an electronic device attached to its body would cause concern. *The General's Stork* is a project that brings these findings together and tells the story of how the bird's-eye view actively transformed and shaped the geopolitics of the Middle East in the last century.

ROUTHIER: There's a related aspect of *The General's Stork*, which is poetically addressed in your video work *As Birds Flying* (2016), namely the question of the "naturalization" of drone warfare (see fig. 1). The stork, obviously, functions discursively as a cipher for the paranoid age

of drone warfare. But in your work, there is a sense in which the saying that "birds of a feather flock together" acquires an uncanny *material* truth, as for instance when you draw attention to the convergence of military and civilian drone engineering trends that model drones on actual, living birds. As a researcher working on these same topics, it can sometimes feel difficult to find a form that can adequately address these historical convergences. Is there a method to your work as an artist?

Fig. 1. 7'11" video still from Heba Y. Amin, As *Birds Flying*, 2016. Image courtesy of the artist and Zilberman Gallery.

AMIN: Drone warfare, colonialism, occupation, and violence are not poetic; they are horrific things. But we need to find ways to relay our humanity and contextualize our contemporary conditions, and art has the power to reveal injustice in affective ways. My research as an artist is driven by visual content which, in many ways, allows me to confront otherwise difficult narratives. We need other ways of engaging with how we want our futures to look and spaces to

nurture discourse on what we consider to be *technological development* and *human progress*, for example. Who wants to live in a world where machines of death simulate beautiful creatures to the point of confusion?

The short allegorical film *As Birds Flying* (2016) speaks to these issues of concern. The stork, of course, becomes the vehicle through which the paranoid age of drone warfare is confronted, but the film also seduces you with beautiful drone footage of what turns out to be the occupied Palestinian landscapes that storks migrate through. It takes its title from a Biblical prophecy from the Book of Isaiah. I made the connection after finding a particularly peculiar portrait of Lord Allenby with his pet marabou stork in his villa in Cairo (see fig. 2). Allenby was the British Commander-in-Chief of the Egyptian Expeditionary Force in Egypt in the early twentieth century and is attributed with his successful capture of Jerusalem from the Ottoman Turks in 1917. In keeping with British interests to expand political and economic control over the region as well as to "restore" Palestine to an imagined biblical landscape, Allenby was inspired by the birds in this passage to justify an air assault and colonial take-over of Jerusalem.

Perhaps it is in the nuance of these absurd details that allows one to see history in other ways. Perhaps it will make you laugh before you dwell on your anger, but ultimately it attempts to shake what we understand to be the *truth* through assumed historiographies.

ROUTHIER: Your narrative arc follows the historical metamorphoses from the bird into the drone (and vice versa), and through the example of Lord Allenby and British colonialism you highlight how new technologies are always embedded in broader cultural narratives. In this particular case, part of the justification for the imperialist settler-mentality grew out of Christian religious narratives about divine and vengeful "birds." So, in a sense, you are not merely describing "cultural narratives" but exposing actual, historical frames of war?

Fig. 2. Heba Y. Amin, *The General's Stork I*, 2020: 100 x 80.86 cm, archival color print. Image courtesy of the artist.

AMIN: We know that Lord Allenby was deeply religious and much of his actions were inspired by his Christian convictions. We also know that British colonialism was very much driven by religious ideology: in Palestine, a move towards "British Restorationism" was supported by the idea that the second coming of Christ was only achievable if the British played a role in ensuring that the Holy Land belonged to Jews. There is no doubt that the imperialist settler-mentality grew out of religious narratives, among other motivations. More importantly, how do we expose these historical frameworks to better understand our contemporary context?

The amount of digital content being produced and the relatively new access to digitized historical archives allow us to revisit and reframe history. What I am particularly concerned with is the role that imaging technologies play in shaping ideology, particularly in relation to colonial warfare. The colonial context has, in many ways, been written out of the story of Western technological progress, and I think it is important to bring occupation back into that narrative.

ROUTHIER: Part of your exhibited archive footage includes the first aerial photographs of Palestine and serves as a reminder, among other things, that the bird's-eye view is inextricable from the history of colonization. But more specifically, do you see present-day drone surveillance and new forms of "machinic vision" as structurally embedded in Western universalist perceptions of space?

AMIN: Imaging technologies from the nineteenth century were already being used to justify colonial land grab. Panoramic photography which emerged shortly after the invention of the daguerreotype, for example, was utilized as a tool to visualize the vast scale of territory "available" for occupation on the African continent. Depicting the land as a vast open territory was intended to act upon the desire for the openness of "primitive" African landscapes where a new aesthetic of fantasy geographies was at the core of visualizing the colonial project. German missionary and photographer Carl Hugo Hahn developed a technological device, a camera with a revolving panoramic lens, capable of photographing 180-degree landscapes for the purpose of visualizing this expansive environment. This was intended as an invitation for Europeans to occupy "empty" land.

The view of the world from above, however, introduced a new imagination of territory drawn from the fantasies of colonial ideology. The early twentieth century saw British military interests in aerial technologies which presented opportunities for land expansion in Europe's

territories in Africa and the Middle East. Aerial images, in particular, played a significant role in the desire and fulfillment of Europe's vision of the modern nation-state.

ROUTHIER: How so? By portraying the landscape in which ways?

AMIN: In the context of Palestine, aerial images were used to frame the landscape as empty, human-less, and free for the taking. In this case, it was not only about relaying the primitiveness of landscape and the possibility of superimposing European fantasies on virgin territory. Palestine was of particular interest for its Biblical history, not a land with an existing modern people and society, and therefore depicted as such. The aerial photograph became an extension of utopian thinking, a future world mirroring European (Christian) ideology. Aerial photography was framed as a tool of scientific research, for surveying and imperial cataloguing; "machinic vision" therefore came to represent progress as the bird's-eye view became a symbol of modernity. The narrative of religion combined with the "evidentiary" nature of technology became almost impossible to question.

ROUTHIER: So, do you see this colonial prehistory as embedded in the technical apparatus itself, or in the "universalist" gaze that is constructed through these machines? I'm asking this because one thing I noticed in your works, and really found intriguing, was that you seem to be investigating "vision" in this more fundamental, historically anchored sense. I mean, obviously, many contemporary artists work with visual representations and the question of what is seen and what is *not* seen—I'm thinking more specifically about artists such as Harun Farocki, Trevor Paglen, and Hito Steyerl here—but I get the feeling from your artworks like *Vision is one of the Senses* (2016) through to *The General's Stork* (2016–ongoing) that you are perhaps investigating the question of visuality at a deeper level (see fig. 3)? Is that an objective of yours?

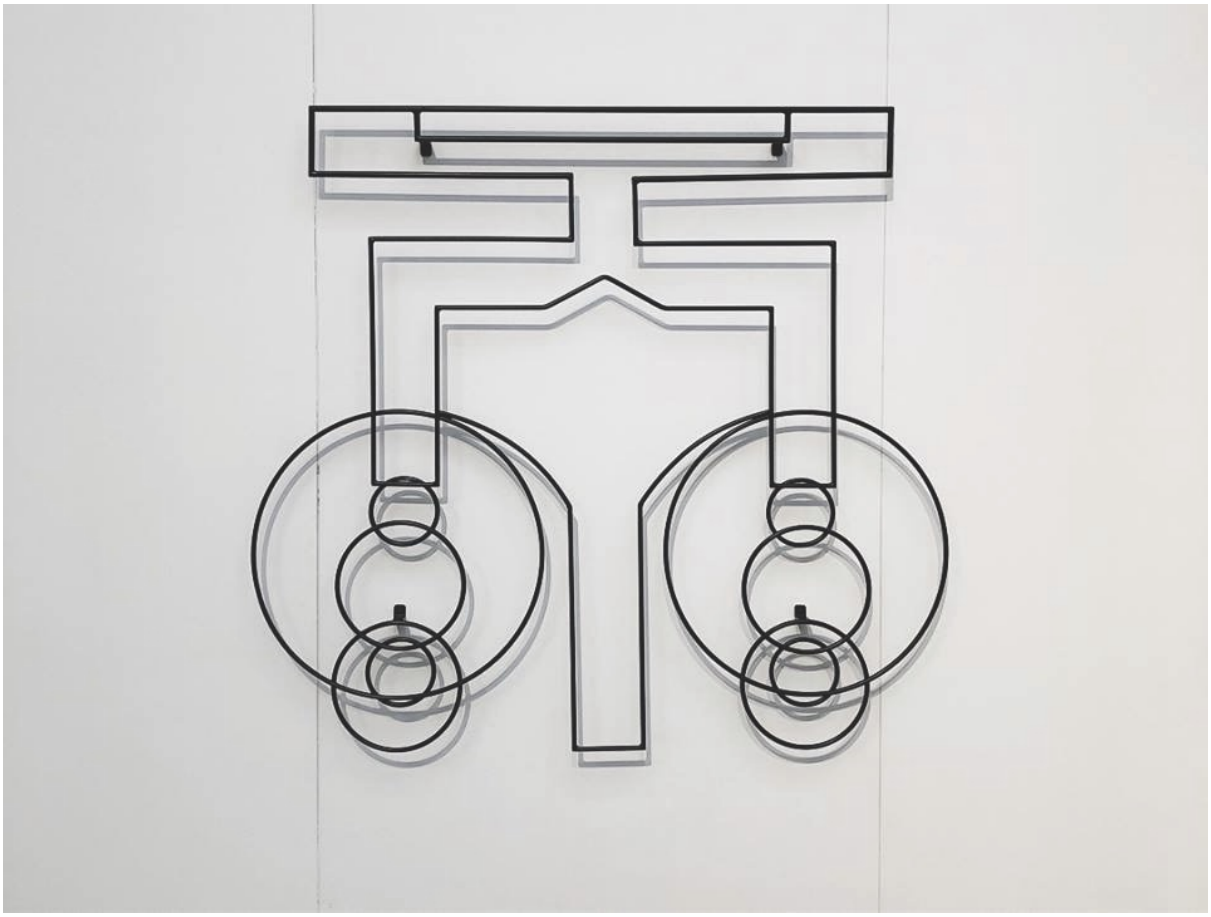


Fig. 3. Heba Y. Amin, *Vision is One of the Senses*, 2016: iron, powder coated, 110 x 120 x 6 cm. Image courtesy of the artist and Zilberman Gallery.

AMIN: The colonial history is absolutely vital in contextualizing the role of the technical apparatus. I look at how the earliest photographs of the region were extrapolated from a long tradition of the orientalist gaze. For example, only three months after the invention of the daguerreotype was gifted to the world, French artists rushed to North Africa to capture erotic images of North African women. Due to the inaccessibility of the North African female subject, European artists invented an idea of romance by photographing the native woman to fit their vision, their fabricated delusion. Their image, constructed by the dreamscapes of orientalist painting, was a tool of political propaganda for the colonial project. Sexualized representations

of women came to represent the domination of territory; their exploited bodies merged with the idea of claiming land. Indeed, the technical apparatus emerged from this predatory context of seeing, except now it was validated through the technological lens.

Indeed, I am very much influenced by the artistic practice of Harun Farocki—as well as Trevor Paglen and Hito Steyerl—for their groundbreaking scholarship on "machinic vision." I found that it was also important for me to fundamentally understand how the history of vision has been, and continues to be, narrated through a Western "universalist" perspective.

ROUTHIER: So, by going back in time, you are essentially contesting the historical construction of Western visuality?

AMIN: Precisely. My research on early photography takes a closer look at scientific and philosophical developments on optics and vision dating back to tenth- and eleventh-century Arabic manuscripts. Until recently, these important treatises on optics were essentially written out of history. I discovered that many of the ideas that were attributed to the Renaissance era far preceded it, and that the knowledge that had already been produced and published by that time was usually not acknowledged. Until recently, a scholar like me was often not granted permissions to explore archives in European institutions, and now, with the efforts in digitizing archives, many historical manuscripts are more easily available. While one still has to have the right affiliations, with the right institutions and the right credentials in order to access archives, scholars from the Global South are suddenly privy to parts of their own histories that they've never had access to before.

As a result, the scholarship of the tenth-century Arab thinker and scientist Ibn al-Haytham, for example, is finding its way back into contemporary scholarly discourse. Ibn al-Haytham's groundbreaking manuscript, *The Book of Optics (Kitab al-Manazir)*, made significant scientific

observations about the mechanics of vision and the philosophy of perception. His work was the first to explain vision as a function of the brain; he demonstrated vision by intromission of light rays to the eye rather than rays being emitted from the eye. Furthermore, his book contemplates the manipulative potential of perception at length. For me, this became a potent source for cultural critique.

ROUTHIER: Tellingly, I wasn't familiar with al-Haytham's theories other than by your mentions of him as a source for your optical sculptures. It reminds me of how in the scholarly discourse sometimes referred to as visual studies, the term "visuality" points to a "social fact" (Hal Foster) rather than to the physical processes involved in seeing. An essential mediating component of this compound "social fact" is, of course, technology.

In Dziga Vertov's groundbreaking film, *Man with a Movie Camera* (1929), the technical apparatus, or the mechanical eye, appears as a world-making apparatus, "freely" distributing its visual coordinates and creating a new world. Since Vertov, much of the Western avant-garde filmmaking tradition has reveled in the artistic affordances of new forms of machinic vision. Today, with the drone, Vertov's "mechanical eye" literally reaches new heights and appears as something like the fulfilment of the avant-garde dream of technical progress. Given what we know, should we be more careful in assessing the "mechanical eye's" so-called objective visual constructs, its alleged capacity for world-building and community-making?

AMIN: I would argue that there is no such thing as the "mechanical eye's" objective visual construct. Even Farocki's notion of "operational images"—where machines speak to machines through "images"—does not function on an objective level in so far as the logic of these technologies and the systems on which they are built stem from racially driven colonial constructs. The history of aerial photography is inherently linked to political cartography and a

vertical power hierarchy that was, and continues to be, strategically enforced through the structuring of space and the policing of bodies from above. In that sense, it is very difficult to look at the development of these technologies and consider them in terms of world-building without recognizing the military and corporate frameworks they serve. Automating visuality, or machine-led forms of image production, become an extension of a system that is already racially deterministic in nature. But it's not just about visuality and image-making, it's also about how the systematic weaponization of imaging technologies has become increasingly normalized through the algorithmic apparatus, especially as we remove ourselves from the responsibility and complicity of image-making.

ROUTHIER: This kind of "genealogical" investigation of yours into vision and visuality seems to me a kind of corrective not just to contemporary art but also, more broadly, to the trajectory of European modern art as such. I mean, if we consider the lineage of avant-garde filmmakers from Vertov through to mid-century high modernists such as Lazlo Moholy-Nagy and other forerunners of Op-art, we find that, for all the intricate aesthetic explorations of the mechanics of vision, there is little to no interest in inquiring into the historicity of the Western image-making tradition and the material construction of the "universal" gaze. Your "optical sculptures" share an aesthetic affinity with these kinds of mid-century sculptures, but I assume that they are intended as a different kind of artistic intervention?

AMIN: My optical sculptures portray the original scientific diagrams from Ibn al-Haytham's *Book of Optics*; they speak to that moment in history when we accurately understood the mechanics of vision. I would say that these sculptures are a critique of the material construction of the "universal" gaze, and a general questioning of perception and knowledge. Whose framework of knowledge? Whose history?

ROUTHIER: This reminds me of your series of self-portraits called *Portraits of Woman with Theodolite (I-III)* which show you posing with an intricate technical apparatus called a "theodolite," which as far as I can tell was originally used for land surveying purposes (see fig. 4)? But visually and aesthetically, these portraits come across as a *détournement* of some iconic portraits of avant-garde filmmakers posing with their camera. Could you tell me a little bit about the context of these self-portraits—and if and how they relate to questions of art historical "representation"?

Fig. 4. Heba Y. Amin, *Portrait of Woman with Theodolite Series*, 2019. Photograph by Markus Rack. Image courtesy of the artist and Zilberman Gallery.

AMIN: These images were made in the spirit of early studio portraiture where the display of technological objects as a representation of human progress was a common photographic theme. The triptych portrays a woman with a technological device (*Portrait of Woman with Theodolite I - II*, 2019) and a woman with the technological device in use (*Survey of German Landscapes by Night (New Morgenthau Plan) I*, 2019). I began to experiment with ways to articulate the power gap between observer and observed, and to make the embedded politics of tools of observation and image-making as transparent as possible.

In 2014, I embarked on a five-month road trip along contemporary migration routes from Africa to Europe with a surveying tool known as a theodolite. By flipping the power dynamics and positioning myself as a *voyeur*, I attempted to magnify the predatory frameworks of African landscapes that have been inscribed in dominant histories. Through the surveying and surveillance of contested territories, the dominant political powers embedded in the landscapes become visible through visual documentation. While the non-human subjects of my images are clearly contemporary, their aesthetics resemble early photographic techniques that regurgitate a colonial logic.

The project is again about the gaze: who is doing the gazing? What does it mean to be the voyeur, to adopt the male gaze? The European gaze? I took a performative approach in embodying the technologies, embodying the archives as a way to better understand the logic of the technical apparatus. By addressing the fundamental act of seeing, my work regards a confrontation of scopic regimes as imperative to a thorough reading of images.

Artist **Heba Y. Amin** engages with political themes and archival history, using mixed-media including film, photography, lecture performance and installation. Her artistic research takes a speculative, often satirical, approach to challenging narratives of conquest and control. Amin is a

Professor of Digital and Time-Based art at ABK-Stuttgart, the co-founder of the Black Athena Collective, curator of visual art for the MIZNA journal, and currently sits on the editorial board of the Journal of Digital War. She was awarded the 2020 Sussmann Artist Award for artists committed to the ideals of democracy and antifascism, and was selected as a Field of Vision Fellow, NYC (2019). Her latest publication, *Heba Y. Amin: The General's Stork* (ed. Anthony Downey) was recently published by Sternberg Press (2020) and her works and interventions have been covered by *The New York Times, The Guardian*, *the Intercept*, and *BBC* among others. Email: hebaamin@yahoo.com.

Dominique Routhier is a Post-doctoral Researcher at the Department for the Study of Culture, University of Southern Denmark. Currently, he is affiliated with the research project Drone Imaginaries and Communities (Independent Research Fund Denmark, 2020-2024). His writings have appeared in *Rethinking Marxism*, *Nordic Journal of Aesthetics*, *Boundary 2 online*, and *LA Review of Books*, among other places. He is the author of a forthcoming book about the Situationist International (Verso Books, 2023). Email: dominique@sdu.dk.