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.