journal of visual culture

Face Recognition and the Emergence of Smart Photography

Sarah Kember

Abstract

This article examines face recognition as a key instance of the emergence of smart photography. Smart photography, drawing on Artificial Intelligence (AI) and Ambient Intelligence (AmI) manifests a 'habit of mind' (Karen Barad), or a way of thinking that is humanist in so far as it is predicated on human and machine autonomy, and representationalist in its quest for unmediated objects-in-themselves. Faces are among the objects that smart photography seeks (autonomously) to represent. By examining two of the principal algorithms of face recognition technology, the article will show how ways of seeing allied to ways of thinking that are also, ultimately, discriminatory and essentialist, materialise through software. Finally, if the 'smart' in smart photography means learning to discriminate between classes of faces that are fixed, essentialised and ultimately elusive (the stereotypical face of terror is both gendered and racialised) then how could smart be made smarter? This is a question of politics rather than progress.

Keywords
algorithms • ambient • face recognition technology • intelligence • photography • representationalism

Introduction: The Emergence of Smart Photography

How is it possible to speak of the emergence of smart photography? What may seem to be, and to a very significant extent still is, a futuristic scenario – involving a medium and technology that is responsive, adaptive and able to learn – is in fact a consequence of the convergence[1] of imaging, information

and biotechnologies. As Donna Haraway (2000: 25) remarked in her interview with Thyrza Nichols Goodeve, 'we live intimately "as" and "in" a biological world'. From the health and entertainment industries to the food industry, intellectual property law, environmentalism, management 'and so on' there is, she says, 'almost nothing you can do these days that does not require literacy in biology' (p. 26). Her biology is that which is 'woven in and through information technologies and systems', and although Haraway has little to say about the role of media in and as those technologies and systems, one aim of this article is to advance an argument I have already made about media in general (see Kember, 2006) and show how photography today promises to make biotechnologists of us all (Haraway, 2000: 26).

Of course, the promise – or threat, depending on your point of view – of biotechnology, like all forms of techno-futurism, is unlikely to be fulfilled as stated. It is only in science-fiction and fantasy films that photographs (let's assume they are digital) literally come to life, offering unmediated access to events and individuals who engage with each other and interpellate the viewer in ways that Louis Althusser[2] almost certainly didn't see coming. Nevertheless, critical techno-futurists like Haraway and social scientists like Lucy Suchman are invariably willing to acknowledge the performativity of the promise that we have always, somehow, heard before – that is not new, but generative of some degree of newness (Haraway, 1991; Suchman 2007). Smart photography will not shout your name because its existence is not a given. It is not a facet of technology alone, or indeed an *inevitable* consequence of the convergence of imaging, information and biotechnologies. This article will explore the extent to which smart photography is coming about and the contexts and conditions that make it possible. It will also consider the technological and extra-technological forces that enable and constrain the evolution of photography after photography.

In *After Photography*, Fred Ritchin (1991) – who once declared that digitization would bring about 'the end of photography as we have known it' – speculates about the uncertain status and future of an expanded medium whose identity is increasingly bound, not only to other media, but to cybernetic systems, biological forms and physical processes (Ritchin, 2009). Wielding somewhat mixed metaphors of the quantum, cyborgian and genetic image that moves like a particle, functions like a machine and is distressingly easy to manipulate, Ritchin (2009: 25) reaches through a nihilistic Baudrillardian scenario in which he foresees the dissolution of photography in life and of life in photography[3] in order to outline more productive (if not progressive) ways in which the world and its inhabitants are 'becoming image'. Pixellated virtual realities and identities, Ritchin suggests, can be and are being augmented and automated. Every image, conceived as a 'map of squares', can be modified seamlessly and 'serve as a pathway elsewhere' (p. 141). Like Charlie Brooker's (2011) drama in which the protagonist's eye doubled as both a projection screen and a constant recording device, Ritchin's photography machine 'will do it for us, using face recognition to remind us with whom we are talking at a party, or recording

what we missed when inebriated' (p. 163). The once self-contained image can now be overlaid or augmented with information tailored to the viewer's location and specification (see the Layar app, for example), and the digital camera itself 'will be further absorbed into other devices' including not only mobile phones but 'refrigerators, walls, tables, jewelry, and ultimately our skin' (p. 143). Here, Ritchin touches on a sense of the vernacular, the everyday in post-photography that Geoffrey Batchen significantly misses. Concerned primarily with what he considers to be the marginal status of popular practices in photography history and theory, Batchen (2002: 59) makes a case for the vernacular as 'the organizing principle of photography's history in general'. What he misses - noting that he is writing almost a decade before Ritchin - is precisely the fact that this history of photography as key to the quotidian and a central ritual as well as visual practice of everyday life, is currently being written by the technoscience[4] industries. The vernacular is very much contested ground, and planted firmly within it are camera-enabled objects and spaces that are so taken-for-granted as to be effectively invisible.

Face Recognition as Smart Photography

The taken-for-granted environments of the home and city constitute the terrain of smart photography, emerging as part of a new technoscience called Ambient Intelligence (AmI). AmI is a development of Artificial Intelligence (AI) and a re-branding of ubiquitous computing. It emphasizes the environmental and social aspects of networked, distributed and intelligent forms of computing. AmI systems that include photography – and face recognition in particular – are conceived of as being fully embedded within the environment and capable of adapting and responding to its inhabitants. An 'intelligent' environment is one that can 'recognize the people that live in it', adapt to them, 'learn from their behavior, and possibly show emotion' (Aarts et al., 2004: 6). In this industry-driven vision of an ambient intelligent world, 'people will be surrounded by electronic systems that consist of networked intelligent devices that are integrated into their surrounding [sic] and that provide them with information, communication, services, and entertainment wherever they are and whenever they want.' Counteracting a discourse of business with one of leisure, the stress on environments and sociality incorporates but ultimately exceeds the confines of the smart home so that, as AmI researcher Stefano Marzano (2003: 8) says, 'we may find that any non-interactive objects or systems around us have been replaced by almost invisible, intelligent interactive systems – an "AI" that could soon form a natural part of our everyday lives.'

Applicable to home environments; mobile spaces like cars, buses, trains and planes; public spaces like shops and airports; and even private spaces such as clothing, AmI systems normalize and naturalize a culture in which the joint operation of marketing and surveillance is becoming dominant. Home-based as well as mobile networks 'bring about the possibility of surveillant systems fully integrated within consumer landscapes' (Allon, 2003: 264).

Within these systems, 'every act of consumption and transaction can be recorded and processed, entailing the accumulation of an unprecedented degree of personalized information'. This 'ontology of everyday control' – as Fiona Allon describes it in her work on smart homes – is not new, but rather newly naturalized.

Face recognition technology is becoming central to this naturalized – embedded and invisible – ontology of everyday control. As a marketing and surveillance-based biometric as well as photographic technology, one of its main advantages over other biometrics such as finger printing or iris-scanning is that it operates at a distance and does not require consent or participation. Face recognition is a default setting on social networking sites such as Facebook, offering automatic tagging suggestions as the user uploads photographs of friends and family. It is becoming ubiquitous in international airports and other social environments where security and/ or commerce are at stake. The objective of face recognition is to be able to pick out a face from a crowd and identify the target by comparing it with a database. Where this objective is, as I will show, hard to achieve, another goal entails learning to discriminate between classes of faces based on gender, race and age. This is easier in so far as it relates to groups rather than individuals and appeals to biological differences. Face recognition systems are both autonomous and ambiently intelligent according to a set of industry claims, but what is the basis of such claims and how smart is face recognition really?

The methods and modes of critique associated with feminist science and technology studies are arguably among the best placed to address this question. Suchman (2007), for example, engages ethnographic methods in order to challenge assumptions about machine autonomy. Whether she is analysing software or hardware, a performance artist's AI-enabled, speech-based avatar or a seemingly stand alone robot that emotes, she engages the artefact in situ and details the operational failures by means of which its autonomy is unmasked. The affects associated with such artefacts are, for Suchman, 'effects of multiple agencies' including the hidden labour of designers, trainers and programmers (p. 246). Machine autonomy simply does not exist, or rather exists as performance and as a projection of qualities – such as intelligence and emotion – associated with the human. Within feminist science and technology studies, humanism is regarded as a discriminatory way of thinking, predicated on a set of Cartesian dualisms, false divisions that are gendered and hierarchical. For Haraway (1991), whose methods embrace ethnography as one form of social and scientific story-telling, Cartesianism constitutes a matrix of domination, a story about nature and culture that can and must be retold. For the physicist and philosopher Karen Barad, Cartesianism is a habit of mind and of sight, a way of thinking allied to a way of seeing that produces both particles and people as apparently autonomous entities (Barad, 2007). Her theory of entanglement challenges the false divisions of humanism and representationalism which, for her, is the 'belief in the ontological distinction between representations and that

which they purport to represent' (p. 46). People and particles, humans and machines are entangled in so far as they lack 'an independent, self-contained existence' (p. ix). Ontological and epistemological entanglements undermine the tenets of representationalism that allow humans to represent machines as if there was an essential distinction or 'gap' between them (p. 47). Suchman (2007) is quick to point out that this particular division is highly expedient from an industry point of view since it legitimizes the technological innovations that seek to overcome it, for example, by making machines and machinic systems more autonomous and intelligent.

Face recognition systems seek to overcome the division between human and machine vision or, specifically between human and machinic capacities for appearance-based face recognition and identification. Questions of system accuracy and performance come to the fore because the comparison remains unfavourable. Subsequent performance anxiety serves to legitimize a range of technological innovations designed to close the gap, and among them is the use of AI: 'AI approaches utilize tools such as neural networks and machine learning techniques to recognize faces' (Jafri and Arabnia, 2009: 48). How do machines learn? The issue has been widely debated but in this context it is clear that in addition to techniques of pattern recognition and sorting that will be discussed in the following sections, the principal mechanism of machine learning is reductionism. Matthew Turk and Alex Pentland (1991) have made a significant contribution to the development of face recognition. For them, 'developing a computational model of face recognition is quite difficult, because faces are complex, multidimensional, and meaningful visual stimuli' (p. 71). Face recognition systems substitute the meaning of faces for a mathematics of faces, reducing their complexity and multidimensionality to measurable, predictable criteria. Moreover, face recognition technology requires a reduction in the variation of face images and environments and must ultimately replace faces with vectors (principal components of faces) or with standardized templates in order to learn anything at all. System accuracy and performance depend on 'constrained environments such as an office or a household' (p. 71). The face image presented to the system for recognition must be centred, 'the same size as the training images' and fully frontal or in profile, so reproducing – as input – the mug shot photograph generated by 19th-century ways of seeing (p. 76). Following Suchman, Kelly Gates (2011) has identified the elision of labour that secures the illusion of autonomy in face recognition technology. There is also an inventory of technological failures (see Introna and Nissenbaum, 2009) that, combined with reductionism, delimit the claim to smartness in this instance of smart photography. Can we, therefore, dismiss the claim altogether? I will argue that we cannot dismiss it in so far as it is manifest in the very architecture of the system. Here, smartness materializes in pattern-recognizing and sorting algorithms that are learning to identify faces by discriminating among them, generating ontological and epistemological divisions – between male and female, black and white, old and young – that in this case must remain un-reconciled, reduced to a set

of essentialised categories that guarantee system performance by ensuring that input (a recognizable face) is equivalent to output (a recognized face).

Sorting Algorithms

The aim of a facial recognition system is to either verify or identify someone from a still or video image. Following the acquisition of this 'probe' image, the system must first of all detect the face or distinguish between the face and its surroundings. To do this it selects certain landmark features in order to compare them with the database. Either that or it generates what are called standard feature templates – averages or types. Once detected, the face is normalised or rather, the image is standardised with respect to established photographic codes such as lighting, format, pose and resolution. Again, this aids comparison with the database. However, the normalisation algorithm is only capable of compensating for slight variations, and so the probe image must already be as close as possible to a standardized portrait. In order to facilitate face recognition, the already standardised image is translated and transformed into a simplified mathematical representation called a biometric template. The trick, in this process of reductive computation, is to retain enough information to distinguish one template from another and thereby reduce the risk of creating 'biometric doubles' (see Introna and Nissenbaum, 2009).

According to the manufacturers and promoters of face recognition systems, the complex sequence of technical operations and transformations performed on the face image does nothing to undermine the objectivity of the process. This is partly because the underlying principle of the system is photographic, and historically, the authority of photography derives not only from its strong claim to indexicality, but from its development and use in the very institutions in which it continues to be deployed. The history of photography as an imaging technology that is inseparable from the disciplinary institutions of the 19th century is very well documented (Sekula, 1986; Tagg, 1988). In the context of the industrial revolution in Western Europe, there was a perceived need to cater for and control the newly urbanised masses, to combat and reform the spread of poverty, disease and crime and to render productive an unproductive population of the sick, the mad and the bad. A new police force emerged – alongside hospitals, schools, asylums and workhouses – and it was here that Alphonse Bertillon developed the first criminal identification system using photography in conjunction with anthropometric and statistical methods. Anthropometrics are broadly equivalent to contemporary biometrics, and Bertillon took 11 measurements of each individual criminal's body, recording them on index cards alongside frontal and profile portraits and a brief verbal description of distinguishing features. As his archive of the portrait *parlé* built up, he needed to organise it, which he did by incorporating the concept of the average man. Statistically, the average or mean could be expressed using the bell-shaped curve, but as Allan Sekula (1986) points out, it was also

conflated with normality and the social ideal while difference from the mean was similarly conflated with deviance. Moreover, this slippage from a purely statistical to a discriminatory social law of averages was backed up by a climate of belief in the quasi-sciences of physiognomy and phrenology. While phrenology was concerned with correspondences between skull topography and localised mental faculties, physiognomy assigned essential characteristics to the arrangement of facial features so that a narrow forehead, for example, was taken as a sign of low intelligence. More generally, in his chapter on criminal physiognomy, Havelock Ellis (1901: 87) wrote that beautiful faces 'are rarely found among criminals. The prejudice against the ugly and also against the deformed is not without sound foundation.'

Sekula (1986) is clear about the central if problematic role of photography in generating this huge and ignoble archive of the 19th century. Photography operated as an effective mechanism of surveillance, recording, normalisation and social control but could not alone secure the identities of criminals, requiring instead the addition of verbal, anthropometric and statistical measures. The authority of Bertillon's criminal identification system rested not on the camera but on a 'bureaucratic–clerical–statistical system of "intelligence"' (p. 16). This system has continued to be updated ever since and, despite the so-called shift from analogue to digital imaging, was manifest in the computerised criminal identification systems of the late 20th century. These systems, in operation internationally, drew on a database of analogue and digital photographs and were organised and operated according to recognisable principles. Jacques Penry, the inventor of PhotoFIT (which became the basis of E-FIT and CD-FIT) based his identification system on the establishment of facial norms and on measuring deviance from that norm. He argued that faces are either angular, rounded or a combination of the two (Penry, 1971: 27). Since facial features vary, he suggested that it would be helpful 'to have one set of standards or face-measures by which to judge whether a feature is "large" or "small" (or "long" or "short") – that is, in relation to the area of the face it appears on'. Penry provided a blueprint 'of normal proportions' (Figure 1) against which all faces could be measured. As a physiognomist, he also linked facial features with personality, arguing that, for example, a mouth with upturned corners would have a perpetually cheerful owner while 'a full, fleshy-lipped or loosely moulded mouth in itself suggests a basic general lack of control over emotional urges' (p. 42). While a traditional photographic album might still be used for recognition or identification purposes (where the suspect was not known to the witness but might be known to the police), computerised systems were used to assist the witness in facial recall. These systems depended on two different forms of coding and on the use of specialist 'cognitive' interview techniques designed to help eyewitnesses recreate the context of the crime and search their memories in a systematic way (Kember, 1998: 50). Geometric coding involved measuring facial features from images and is still the basis of what happens now in face recognition technology (FRT). Syntactic coding used descriptions of faces rather than measurements, and was perhaps more closely related to the sort of observations witnesses make, such as 'he had a long nose' or 'she had protruding ears'. Extracting a face and thereby
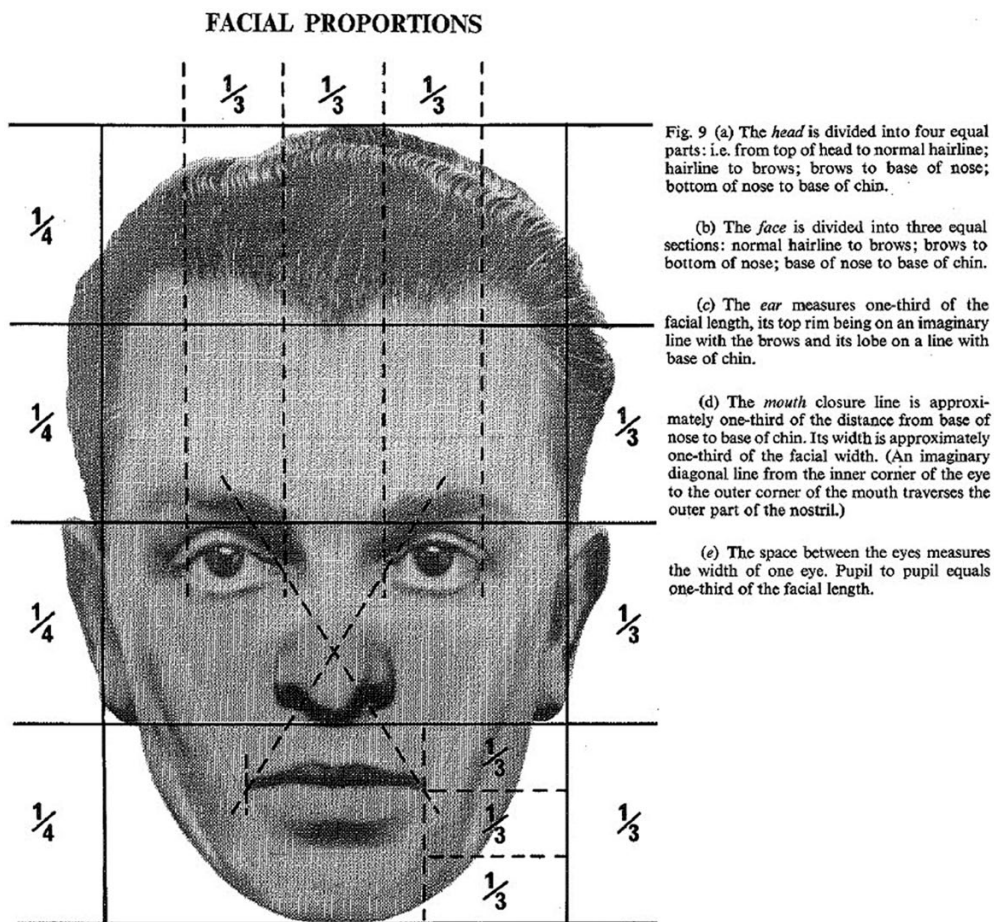
FACIAL PROPORTIONS

Fig. 9 (a) The *head* is divided into four equal parts: i.e. from top of head to normal hairline; hairline to brows; brows to base of nose; bottom of nose to base of chin.

(b) The *face* is divided into three equal sections: normal hairline to brows; brows to bottom of nose; base of nose to base of chin.

(c) The *ear* measures one-third of the facial length, its top rim being on an imaginary line with the brows and its lobe on a line with base of chin.

(d) The *mouth* closure line is approximately one-third of the distance from base of nose to base of chin. Its width is approximately one-third of the facial width. (An imaginary diagonal line from the inner corner of the eye to the outer corner of the mouth traverses the outer part of the nostril.)

(e) The space between the eyes measures the width of one eye. Pupil to pupil equals one-third of the facial length.

Figure 1 Normalised facial proportions, from Jacques Penry *Looking at Faces and remembering them*, Elek Books, 1971.

an identity from the witness's memory has never been an easy thing to do. The image of the face was built up in quite a painstaking way, feature by feature and the process was long and arduous, more often than not resulting in witness fatigue – and in failure. The efficacy of these then new computerised systems was questioned at the time, and yet they continued to be produced, pushed through not by technical as much as market forces (Shepherd and Ellis, 1993).

Twenty-first century facial recognition technologies are embedded with this legacy of technically limited, pseudo-scientific and politically problematic ways of seeing, regardless of any claim to neutrality and improved efficiency. Nevertheless, after the events of 9/11, the demands on these inadequate bureaucratic–clerical–statistical systems of intelligence have increased exponentially. Increasingly, they are required to act retrospectively and proactively, not only capturing a face, and thereby an identity from an image but securing us from terrifying, terrorising events that have already

taken place – and could therefore take place again. Associated with the development of CCTV cameras installed, arguably, to protect property rather than people, these somewhat frustrated and angry calls for a technology that has let us down to time travel on our behalf – to undo bad things that have happened, to make good and prevent bad things from ever happening again – are not new. Characteristic of public responses to the grainy 'security' camera images that failed to prevent the abduction and murder of James Bulger in 1993, such calls were heard again, repeated with a vengeance in the wake of 9/11. Then, the problem of witness fatigue and failure that had marred earlier systems was writ large, so large that the role of the eyewitness has subsequently been marginalised and slowly eliminated from increasingly automated systems. From a policing and intelligence point of view, the terrorist attacks of 9/11 were marked by a major failure, on behalf of airport security staff, to identify Mohammed Atta and his associates who had already been 'seen' by the security cameras and who, in two cases, were already 'captured' in the US intelligence database. Always the weakest link, from this point onwards the effectiveness of facial recognition has been premised on the elimination of the eyewitness as the main bug in the machine.

It is important to emphasise that automated face recognition systems are not post 9/11 technologies. Their history, as I've suggested, is much longer, their remit is marketing as much if not more than surveillance and they are intrinsic to the new technoscience and its attempted conquest of the everyday. In the near future, cameras linked to face recognition technology may well be installed in shopping malls and higher end shops. Their task will be to isolate individual customers, and tailor marketing and displays to suit them. Ivor Tossell (2011) reports that Intel, the computer chip manufacturer, is working on facial recognition systems that profile customers: 'a camera mounted on a large LCD screen watches for faces that come within four to six metres. The screen can switch ads depending on *the kind of face* that walks by' (emphasis in original). Although it was based on measurements of the individual criminal, Bertillon's system related the individual to the group by establishing both norms and deviants. Contemporary face recognition makes the same moves whether the context is institutional or commercial, classifying and segregating individuals into groups and types depending on their appearance as an indicator of behaviour, and evincing a form of biopolitical[5] control that is perhaps more effective, or at least more insidious, for being at a distance.

The face recognition system as a whole is comprised of technologies and users, images, infrastructure, investment, labour, expectation and belief. In the context of social networking and elsewhere, privacy concerns are real, but they are arguably overshadowed by the image and infrastructure of total surveillance and marketing which significantly out-perform the actual technology. While this continues to be extremely limited – by everything from poor lighting, viewing angles that are not standard frontal or profile, obstacles like hair and glasses, low resolution and expressions in excess of the average mug shot – we should still consider what the system as a whole produces. It

produces faces as quasi-objects, at once detached from, and conflated with bodies that are, in turn, detached from and conflated with identities. These faces are re-coded as static, standardised photographic images that were perhaps falsely divided by Sekula into honorific and repressive categories. It was John Tagg (1988) who showed how each category leached into the other. Sekula (1986) was nevertheless right to argue that the photographic codes established in the 19th century would persist.

One of the algorithms used in face recognition is principal component analysis (PCA). It produces images akin to Francis Galton's eugenicist photographic composites by removing extraneous information, including the outline of the face itself (Turk and Pentland, 1991). PCA reduces faces to their vectors and refigures them as eigenfaces. In 'Eigenfaces for recognition', Turk and Pentland explain that the system functions 'by projecting face images onto a feature space that spans the significant variations among known face images' (p. 71). Significant features are referred to as eigenfaces 'because they are the eigenvectors (principal components) of the set of faces'. They may correspond to familiar features like eyes and noses whose geometric relation is then measured and computed – along the lines set out by Penry. Each input, or individual face image is 'a weighted sum of the eigenface features, and so to recognize a particular face it is necessary only to compare these weights to those of known individuals' (p. 71). Turk and Pentland acknowledge that an eigenface is 'an extremely compact representation' not only of the face but of the original face image (p. 73). It is a practical rather than elegant solution to the problem of face recognition (p. 84).

Linear Discriminant Analysis (LDA) is another key algorithm. It creates classes of faces, much like Havelock Ellis did in his physiognomy of criminals. In their survey of face recognition techniques, Jafri and Arabia (2009: 47) explain that LDA 'maximises the ratio of the between-class scatter' and is better at classifying and discriminating between classes of faces than PCA. This may be partly because this approach starts by selecting faces that are already distinctive. As LDA researchers Kamran Etemad and Rama Chellappa (1997: 1726) state: 'First, we need a training set composed of a relatively large group of subjects with diverse facial characteristics. The appropriate selection of the training set directly determines the validity of the final result.' Sorting algorithms discriminate between classes and types of faces. Both LDA and, increasingly, PCA are being used to discriminate on the basis of gender.

Difference and Dissent

Contemporary face recognition systems differ from earlier analogue and digital systems in that they are exclusively oriented towards recognition rather than recall. They are designed according to the surveillance and marketing imperatives of targeting, tracking and location. However, picking out one face in a crowd is harder and more prone to error than identifying one class of faces from another, especially when that class appeals to the

biological categories that inform gender, race and age. These categories are naturalised through geometric coding techniques (where syntactic coding is reserved for face recall) and the default subject of these techniques is still the young white male.

Penry's PhotoFIT pack came into use in the 1970s and consisted of photographic images of five features (hair and forehead, eyes, nose, mouth and chin) mounted on card (Kember, 1998). Penry included a male and female database but established what he claimed was a universal – genderless – facial topography. This was actually derived from a norm, a young white male that face recognition systems continue to use (Figure 2) but with the aim, for example, of 'restricting access to certain areas based on gender' or 'collecting valuable demographics' such as 'the number of women entering a retail store on a given day' (Moghaddam and Yang, 2002: 1). The segue from disciplinary to biopower is, for Foucault (2008), contingent on the increasing use of demographics and statistics that orient governance more towards the populace than the individual. Face recognition systems demonstrate both forms of power and perhaps even the shift from one to the other. This becomes clearer as we track back from the biopolitical uses and applications of face recognition technology to the disciplinary design and architecture of the technology itself.

Koray Balci and Volkan Atalay (2002) present two algorithms for 'gender estimation'. They point out that the same algorithms can be used 'for different face specific tasks' such as race or age estimation, 'without any modification' (p. 364). In the first algorithm, the training face images are normalised and the eigenfaces are established using PCA.[6] PCA is described here as a

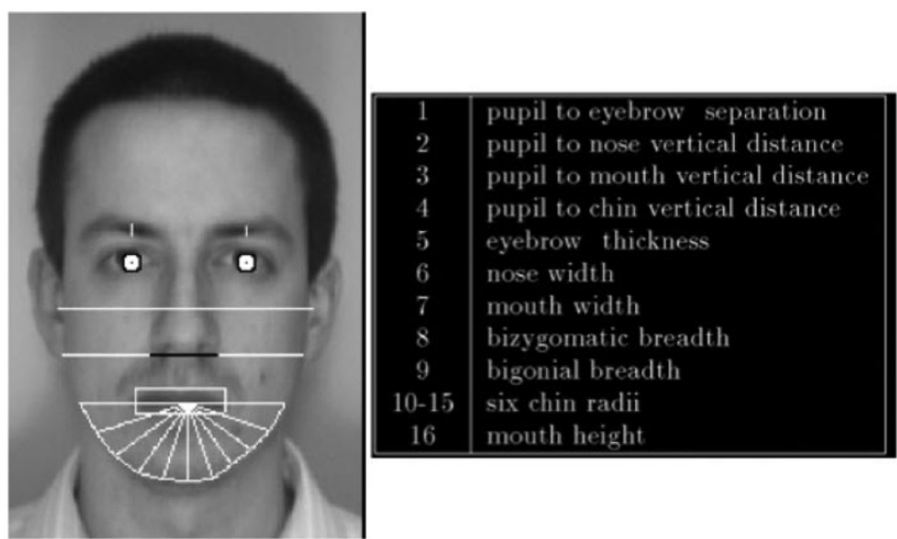| 1 | pupil to eyebrow separation |
2	pupil to nose vertical distance
3	pupil to mouth vertical distance
4	pupil to chin vertical distance
5	eyebrow thickness
6	nose width
7	mouth width
8	bizygomatic breadth
9	bigonial breadth
10-15	six chin radii
16	mouth height

Figure 2 Geometrical features (white) used in contemporary face recognition, from R. Brunelli and T. Poggio 'HyperBF Networks for Gender Classification' Proceedings DARPA image understanding workshop, August 1995, 311–314.

statistical technique for 'dimensionality reduction and feature extraction' (p. 364). The performance of the system is improved by the subsequent use of a 'pruning' algorithm which identifies statistical connections extraneous to gender (race or age) estimation and deletes them. 'After deletion, the system is re-trained' and the pruning is repeated until 'all the connections are deleted' (p. 365). A performance table is produced, showing the relation between each iteration of pruning, the percentage of deleted connections and the accuracy of the system. The accuracy of gender estimation in Balci and Volkan's experiment actually diminishes after the eighth iteration, albeit by only a few percentage points, allowing them to claim that the system is stable. They maintain that pruning or the deletion of statistical connections improves gender estimation not in a linear or absolute sense but by enhancing the process of classification itself.

For Geoffrey Bowker and Susan Leigh Star (2002), classification is a largely invisible, increasingly technological and fundamentally infrastructural means of 'sorting things out'. It is an instrument of power-knowledge that is productive of the things it sorts, things such as faces that are by no means 'unambiguous entities' that precede their sorting (p. 320). The existence of a pruning algorithm that renders faces less ambiguous testifies to their elusiveness, or their inherent resistance to classification as one mode of representationalism. It would, perhaps, be going too far to suggest that there is a crisis of representationalism in appearance-based face recognition systems. However, their designers and engineers are clearly aware that faces are things that 'resist depiction' (Elkins, 2008: xv) because they are 'complex and multidimensional' (Turk and Pentland, 1991: 71) and not 'unique, rigid' objects (Jafri and Arabnia, 2009: 42). From the perspective of visual as well as science and technology studies, Aud Sissel Hoel is concerned with the difference that computational forms of visualisation – such as face recognition – make. She raises the important question of 'whether the new generation of computational picturing in all its forms gives the last push to representationalism in favour of more dynamic and relational approaches' that might be better at tackling the entanglements of technology and knowledge, power and perception (Hoel et al., 2012: 2). For me, the advantage of a more dynamic and relational approach to the production of faces in face recognition technology would include recognising representationalism, and indeed humanism as a claim, a defensive manoeuvre in the face of faces' non-essential ontology and dynamic co-evolution with technological systems. Still, this defensive manoeuvre matters in a double sense: it is both meaningful and material, reproducing norms, for example of gender in a machine that is learning to classify, sort and discriminate among the population, better than it could before. If this is a last push to representationalism, it is one that reinforces it rather than shows it the door. Face recognition technology upholds a belief in the existence of ontological gaps between representations and that which they represent. It re-produces the norms of 19th-century disciplinary photography even as photography becomes allied to the security-based biopolitics of computational vision and smart algorithmic sorting. In this

sense, Gates (2011) is right to argue that new vantage points can underscore old visions as well as old claims to unmediated visuality. Like her, I've questioned the autonomy of face recognition systems without denying that, in conjunction with human input of various kinds, they enact what Barad (2007) calls 'agential realism', generating both categories and entities by cutting and sorting male from female, black from white, old from young. In a context in which security systems are fully integrated with those of marketing, these particular epistem-ontologies intersect, in predictable ways with the category of criminal/citizen–consumer (Lyon, 2008). Since the events of 9/11, the stereotypical face of terror (gendered, racialised) has been perhaps the most represented and most elusive of all. If the problem, from a system point of view, is that the categories leak and the classification structure does not hold, the solution is to reinforce it by pruning it. This process of agential cutting and sorting strengthens statistical groups by deleting connections between them and is precisely the point of a possible intervention, the means by which the biopolitics and ethics of computational vision can be intercepted in order to make a difference.

Conclusion: Connecting Algorithms, Re-Cognising Faces and Re-Writing the Smart in Photography

For Barad (2007) and Haraway (2008), ethics is figured as a process of becoming-with constitutive others. Specifically, it is about recognising and taking responsibility for the 'lively relationalities' that both connect and differentiate humans and machines, subjects and objects, selves and others (Barad, 2007). Ethics here is predicated on a non-essentialist approach to ontology and an active disregard for the categories that humanism upholds – our 'Cartesian habits of mind' (p. 49). Haraway (1991) in particular is quite literal in her drive to make a difference, advocating a hands-on approach to technology that recognises it as a tool of political intervention. She is sympathetic to Derrida's deconstructions while seeking the opportunity to retell the stories that technoscience tells and to engage in some 'serious play'.

What are the opportunities for intervention and revision in face recognition technology? I've already signalled the operational failures and technological limitations by means of which the system deconstructs. Face recognition fails in uncontrolled environments. It cannot cope with poor lighting or resolution, struggles with facial hair and glasses and can only sort six basic types of expression which it must produce by reducing variation and automating expression analysis (Shan and Braspenning, 2010). The aim is to 'recognize a set of prototypic emotional facial expressions' from a non-exhaustive list: 'anger, disgust, fear, joy, surprise, and sadness' (p. 480). To complete what Foucault (2007: 21) recognized as the inherent circularity of biopower,[7] researchers in this area of face recognition confide that 'most of the existing works have been carried out on expression data that were collected by asking subjects to deliberately pose facial expressions' (p. 489).

Elsewhere, I have questioned whether commercially available face distortion software such as Apple's Photo Booth offers a way out of, or through such circularity (Kember, 2011/2012). It is hard to argue that it does when the software merely replaces one set of prototypes for another. Instead of anger, disgust and so on we are offered a rather safe set menu of face distortion that includes the twisted face, the squeezed face and the kaleidoscopic face.[8] Commercial applications contain face distortion and in this sense complement rather than contradict security applications, offering, ultimately, a means to extend them. Nevertheless, face recognition technology remains over-reliance on inputting the frontal flat-lit mug shot – further trimmed to remove hair and face outline – in order to generate gendered stereotypes and generic differences. Categories of male/female, black/white, old/young are pruned at the boundary and connections are deleted. But what if they weren't? What if a sorting algorithm became a connecting algorithm by means of the substitution of a few basic instructions:

> **for all** connections **do not** compute error gradient
> **end for**
> compute threshold for connection
> add connections according to threshold
> **until** all connections are completed[9]

Software writing could not, by itself, retell the story of face recognition, but it might, as Anderson and Pold (2011) suggest, be a good place to start. The opportunity is clearly presented by the fact that the system struggles with ambiguity, including, especially, gender ambiguity. For one research lab, the line between male and female is neither straight nor certain. Moghaddam and Yang (2002) draw it as a curvy, snaking, incomplete trajectory with faces on either side but very close to the boundary (Figure 3). 'It is interesting', they write, 'to note not only the visual similarity of a given pair but also their androgynous appearance' (p. 710). Indeed it is, especially when furnished with the additional insight that there are 'higher error rates in classifying females' which is 'most likely due to the general lack of prominent and distinct facial features in these faces' (p. 710). If, as Brunelli and Poggio (1992) report, 'eyebrow information' is considered to be amongst the most helpful in discriminating male from female, then, depending on which side of the line they find themselves, those who wish to embrace their androgyny know whether or not to pluck.

A connecting algorithm would re-cognize, re-think faces as female–male–black–white–old–young. These faces would constitute feminist, anti-racist, anti-ageist figurations, performative images and political imaginaries akin to the cyborg. They would make manifest a non-discriminatory politics and ethics predicated on entanglement and relationality if not – or not yet – symmetry. While it is not a solution to the problem of asymmetric power relations, relationality is a means of acknowledging, a good start in taking
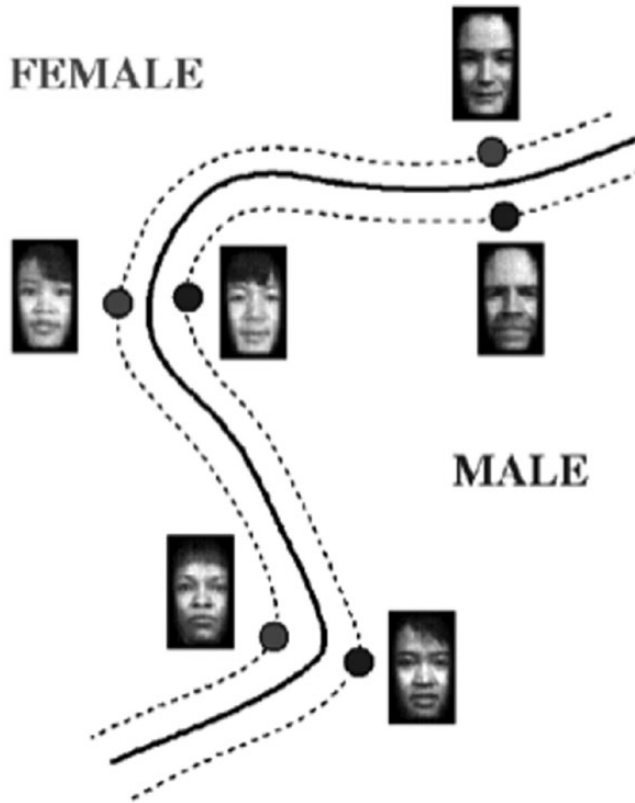
Figure 3 Faces at the boundary. From Moghaddam and Yang 'Learning Gender with Support Faces', *IEEE* Vol. 24, Issue 5, 2002.

responsibility for the fact that 'what is on the other side of the agential cut is not separate from us' (Barad, 2007: 393). A connecting algorithm would take a leaky boundary and play with it in order to envision the world of faces with more potential for ambiguity. This potential is not limitless – the algorithm completes its connections – because classificatory cutting and sorting is 'human', it is what humans do. As Bowker and Star (2002: 1) suggest, 'not all classifications take formal shape or are standardized in commercial and bureaucratic products'. Rather, 'we all spend large parts of our days doing classification work, often tacitly, and we make up and use a range of ad hoc classifications to do so' (p. 2). All of the categories that emerge have 'material force' but the pertinent questions remain: 'What *are* these categories? Who makes them, and who may change them? When and why do they become visible?' (p. 3).

In face recognition technology, gender categories are being remade in a broader socio-political context in which they appear to matter less. They are a key instance of increasingly covert, invisible, infrastructural containment, a paradoxical closing down of the possibilities continually opened up in systems of biotechnology.

The face in face recognition systems is re-naturalised, retro-fitted as either/or within an environment – one that is too hot to handle and must therefore be contained in the office or home – in which 'either/or' is increasingly 'both-and'. But identity, including facial identity, is not an accumulation strategy and cuts have to be made, categories have always to be negotiated. A stable identity is not a fixed and finished one. Faces are forever changing and forever being cut and sorted by habits of mind and sight that are hard to break. The question is not whether to classify, but who classifies, how, when and why?

Face recognition is a technology of everyday control that thoroughly integrates the interests of security with those of marketization. Faces are instrumentalised at a distance by being cut and sorted into categories that re-naturalise them. De-naturalisation is a matter of making the boundary-work visible. This happens internally, within the system itself, which thereby opens up a space for intervention. If cutting and sorting are inevitable, the question is not whether to do that but also, I would suggest, how to do that better, smarter, where smart is not an indicator of technological (or human) autonomy and progress but rather of a human–machine photographic system able to see faces as objects 'both-and' as ambiguous becomings-at-the-boundary.

Notes

1. A contested term, convergence describes cultural and economic as much, if not more than, technological processes and for Bolter and Grusin (2002) is another word for remediation or the mutual refashioning of old and new media.
2. Althusser's (2006) notion of interpellation refers to the way in which subjects are produced in and through ideology.
3. Baudrillard (1983: 555) posits a reciprocal end-game in the relation between events and media: 'the 'dissolution of TV into life, the dissolution of life into TV'.
4. Technoscience is a term used by Haraway (1997) to insist on the connection between science and technology, theory and practice. Biotechnologies are exemplary forms of technoscience.
5. Biopolitics is Foucault's term for the way that power operates at the level of individual and social bodies as well as the state. Crucially, for him, power is not a one-way, top-down, state to subject process caught up exclusively with technologies of domination, but it is also negotiated by means of technologies of the self that are both restrictive and enabling (see Foucault, 2008).
6. With Multi-Layer Perceptron (MLP) gender classifier (Balci and Atalay, 2002).
7. Biopower operates within a milieu in which 'a circular link is produced between effects and causes' (Foucault, 2007: 21).
8. See Apple's Photo Booth: https://www.apple.com/osx/apps/#photo-booth
9. This is a rewriting of Balci and Atalay's (2002: 365) pruning algorithm.

References

Aarts E, Korst J and Verhaegh WFJ (2004) Algorithms in ambient intelligence. In: Verhaegh WFJ et al. (eds) *Algorithms in Ambient Intelligence*. Boston, MA: Kluwer Academic.

Allon F (2003) An ontology of everyday control: Space, media flows and 'smart' living in the absolute present. In: Couldry N and McCarthy A (eds) *Media Space: Place, Scale and Culture in a Media Age*. London: Routledge.

Althusser L (2006) *Lenin and Philosophy and Other Essays*. Delhi: Aakar Books.

Balc K and Atalay V (2002) PCA for gender estimation: Which eigenvectors contribute? *IEEE* 3: 363–366.

Barad K (2007) *Meeting the Universe Halfway*. Durham, NC: Duke University Press.

Batchen J (2002) *Each Wild Idea: Writing, Photography, History*. Cambridge, MA: The MIT Press.

Baudrillard J (1983) *Simulations*, trans. P Foss et al. Los Angeles: Semiotext(e).

Bolter JD and Grusin R (2002) *Remediation: Understanding New Media*. Cambridge, MA: The MIT Press.

Bowker GC and Star SL (2002) *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: The MIT Press.

Brooker C (2011) The entire history of you. *Black Mirror*, Sunday 18 December, Channel 4.

Brunelli R and Poggio T (1992) HyberBF networks for gender classification. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.54.2814 (accessed 28 December 2011).

Elkins J (2008) *Six Stories from the End of Representation*. Stanford, CA: Stanford University Press.

Ellis H (1901) *The Criminal*. London: Walter Scott.

Etemad K and Chellappa R (1997) Discriminant analysis for recognition of human face images. *Journal of the Optical Society of America* 14(8): 1724–1733.

Foucault M (2007) *Security, Territory, Population: Lectures at the Collège de France*, 1977–78. New York: Palgrave Macmillan.

Foucault M (2008) *The Birth of Biopolitics: Lectures at the Collège de France, 1978–79*. New York: Palgrave Macmillan.

Gates K (2011) *Our Biometric Future: Facial Recognition Technology and the Culture of Surveillance*. New York: New York University Press.

Haraway DJ (1991) *Simians, Cyborgs and Women: The Reinvention of Nature*. London: Free Association Books.

Haraway DJ (1997) Modest_Witness@Second_Millenium. London: Routledge.

Haraway DJ (2000) *How Like a Leaf: An interview with Thyrza Nichols Goodeve*. New York: Routledge.

Haraway DJ (2008) *When Species Meet*. Minneapolis: University of Minnesota Press.

Hoel AS, Carusi A and Webmoor T (2012) Editorial. *Interdisciplinary Science Reviews* 37(1): 1–3.

Introna LD and Nissenbaum H (2009) *Facial Recognition Technology: A Survey of Policy and Implementation Issues*. New York: New York University, Centre for Catastrophe Preparedness and Response. Available at: http://www.nyu.edu/projects/nissenbaum/papers/facialrecognitionreport.pdf (accessed 28 December 2011).

Jafri R and Arabnia H (2009) A survey of face recognition techniques. *Journal of Information Processing Systems* 5(2): 41–67.

Kember S (1998) *Virtual Anxiety: Photography, New Technologies and Subjectivity*. Manchester: Manchester University Press.

Kember S (2006) Doing technoscience as (new) media. In: Curran J and Morley D (eds) *Media and Cultural Theory*. London: Routledge.

Kember S (2011/2012) Face re-cognition. *Photoworks*, Autumn/Winter, November–April: 50–56.

Lyon D (2008) *Surveillance after September 11*. Cambridge: Polity.

Marzano S (2003) Cultural issues in AI. In: Marzano S and Aarts E (eds) *The New Everyday: Views on Ambient Intelligence*. Rotterdam: 010 Publishers.

Moghadddam B and Yang MH (2002) Learning gender with support faces. *IEEE* 24: 707–711.

Penry J (1971) *Looking at Faces and Remembering Them: A Guide to Facial Identification*. London: Elek Books.

Ritchin F (1991) The end of photography as we have known it. In: Wombell P (ed.) *Photovideo: Photography in the Age of the Computer*. London: Rivers Oram Press, 8–16.

Ritchin F (2009) *After Photography*. New York: W.W. Norton & Company.

Sekula A (1986) The body and the archive. *October* 39, Winter.

Shan C and Braspenning R (2010) Recognizing facial expressions automatically from video. In: Nakashima H et al. (eds) *Handbook of Ambient Intelligence and Smart Environments*. Berlin: Springer.

Shepherd J and Ellis H (1993) Face recognition and recall using computer-interactive methods with eye witnesses. In: Bruce V and Burton M (eds) *Processing Images of Faces*. Norwood, NJ: Ablex, 129–148.

Suchman L (2007) *Human–Machine Reconfigurations*. Cambridge: Cambridge University Press.

Tagg J (1988) *The Burden of Representation: Essays on Photographies and Histories*. London: Macmillan Education.

Tossell I (2011) Facial-recognition technology needs limits, privacy advocates warn. *The Globe and Mail*, Sunday, 24 July. Available at: http://www.theglobeandmail.com/news/national/time-to-lead/time-to-lead-archives/facial-recognition-technology-needs-limits-privacy-advocates-warn/article/2108118/ (accessed 28 December 2011).

Turk M and Pentland A (1991) Eigenfaces for recognition. *Journal of Cognitive Neuroscience* 3(1): 71–86.

Sarah Kember is a writer and academic. She is Professor of New Technologies of Communication at Goldsmiths, University of London. Recent publications include a novel – *The Optical Effects of Lightning* (Wild Wolf Publishing, 2011) – and a monograph – *Life after New Media: Mediation as a Vital Process* (The MIT Press, 2012). She co-edits the journals of *photographies* and *Feminist Theory*. Current research includes a feminist critique of smart media (*iMedia: The Gendering of Objects, Environments and Smart Materials*, Palgrave, forthcoming) and an affiliated novel, provisionally entitled *A Day in the Life of Janet Smart*. Sarah is also co-PI of an RCUK funded project on digital publishing and part of the Centre for Creativity, Regulation, Enterprise and Technology (CREATe).

Address: New Technologies of Communications, Goldsmiths University of London, New Cross, London SE14 6NW, UK. [email: s.kember@gold.ac.uk]