Introduction

Experimenting with the Face

> Of all the dramatic images to emerge in the hours and days following the September 11 attacks, one of the most haunting was a frame from a surveillance-camera video capturing the face of suspected hijacker Mohamed Atta as he passed through an airport metal detector in Portland, ME. Even more chilling to many security experts is the fact that, had the right technology been in place, an image like that might have helped avert the attacks. According to experts, face recognition technology that's already commercially available could have instantly checked the image against photos of suspected terrorists on file with the FBI and other authorities. If a match had been made, the system could have sounded the alarm before the suspect boarded his flight.
>
> —Alexandra Stikeman, "Recognizing the Enemy," *Technology Review*, December 2001

The September 11 terrorist attacks generated an enormous flood of imagery, and among the deluge was a grainy shot of two of the alleged attackers taken early that morning at a security checkpoint in the Portland, Maine, airport. The recorded video image, which appears to show Mohammad Atta and Abdulaziz Alomari passing through airport security, is a familiar part of 9/11 iconography. Although difficult to discern the men's faces in the image, it is virtually impossible to reference it without also invoking the claim that facial recognition technology could have identified the men as wanted terrorist suspects. Already existing commercially available systems, according to this regretful yet strangely hopeful assertion, "could have instantly checked the image against photos of suspected terrorists" and alerted airport security.[1]

The suggestion that an automated facial recognition system may have helped avert the September 11 terrorist attacks was perhaps the most ambi-

tious claim circulating about biometric identification technologies in the aftermath of the catastrophe. The precise origin of the claim is hard to identify; it seemed to spring forth simultaneously from multiple sources. If it first came from someone in the security industry, it was quickly embraced and repeated by other public figures who felt sure it was true. This hopeful, regretful possibility was the basis for hearings held on Capitol Hill following September 11. On November 14, 2001, a Senate subcommittee on Technology, Terrorism, and Government Information held a hearing on "Biometric Identifiers and the Modern Face of Terror: New Technologies in the Global War on Terrorism." In her opening remarks, Senator Dianne Feinstein (D-CA) asked, "How could a large group of coordinated terrorists operate for more than a year in the United States without being detected, and then get on four different airliners in a single morning without being stopped?" The answer, she noted, "is that we could not identify them." Voicing again the assertion that had become part of the repertoire of public responses to the 9/11 events, she asserted, "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."[2]

The idea that the events of 9/11 could have been prevented with the sophisticated technological products of modernity was laden with what the cinema studies scholar Pat Gill has called "technostalgia"—the desire to revise the past to redetermine the present by harnessing technology toward human ends, all the while recognizing the impossibility of the endeavor (a common theme in science fiction films, like *The Terminator* series).[3] The claim might be said to embody a collective psychological need to believe that the nation was not as vulnerable as it appeared, that U.S. technological sophistication remained intact and in fact would have stopped the men had it been in place. This technostalgic longing to revise the past—a nostalgic sensibility wrapped up with an almost gleeful futurism—provides a paradoxical sort of origin myth for facial recognition technology. In the post-9/11 context, the technology emerged as an already existing, reliable, and high-tech solution to the newest, most pressing problem facing the nation. This move effectively elided the history of this technology, even as it inserted it fully formed into the past. In reality, well before 9/11 a whole set of social actors was already engaged in ongoing struggles and negotiations over the development and use of automated facial recognition technology. While it was *not* already fully formed and ready to identify the nation's post–Cold War enemy Other, it

was already "embedded in and shaped by a rich web of cultural practices and ideas."[4]

This book explores the web of cultural practices and ideas, along with the policy programs and institutional priorities, in which automated face perception technologies are embedded. I investigate the effort, underway since the 1960s and gathering momentum since the 1990s, to teach computers to "see" the human face—to develop automated systems for identifying human faces and distinguishing them from one another, and for recognizing human facial expressions. This effort is best understood not as a unified program but as an interdisciplinary field of research and set of technological experiments. It is part of the broader effort to automate vision—to create machines that not only can generate images, but also analyze the content of those images. Computer scientists are interested in developing automated face perception technologies primarily as a means of creating more intelligent machines and more sophisticated forms of human-computer interaction.[5] Other social actors—especially military, state security, and law enforcement agencies—have viewed these technologies as uniquely suited to the development of "smart" surveillance, monitoring systems that perform the labor of surveillance with less human input and less need to rely on the perceptual capacities and analytical skills of human beings, with the possibility of creating new divisions of perceptual labor between humans and computers. New forms of human-machine integration promise to make surveillance systems function more effectively and extend their reach over time and space. But whether these experimental technologies can or should be made to accomplish these goals remains open to debate, one that often plays out in press and policy discussions as a trade-off between "security" and "privacy."

The purpose of this book is to move beyond the security versus privacy debate and make more adequate sense of the combination of technological projects that aim to create machines for identifying human faces and facial expressions, to understand the constellation of social forces that are shaping these technologies, and to consider what interest in their development and uses tells us about the society in which they are embedded. Why the perceived need to automate the human capacity for facial recognition and expression interpretation at this historical juncture? How are particular social problems conceptualized such that these technologies are posited as potential solutions? What forms are these technologies taking, and what is at stake in their development and institutionalization? In framing the main questions of the book in this way, I take my cue from communication scholars like Raymond Williams, James Carey, Susan Douglas, Carolyn Marvin,

and Jonathan Sterne, who have made the unassailable case that technologies are thoroughly cultural forms from the outset, embodying the hopes, dreams, desires, and especially the power relations and ideological conflicts of the societies that produce them. My main focus is the United States, but it would be a mistake to view automated face perception technologies as uniquely "American." Research and development has occurred in other countries as well, and the technologies are finding applications beyond U.S. borders. But the spaces along and within the U.S. borders have become special sites of experimentation for automating the deceptively complex processes of facial identification and expression recognition. U.S. borders, airports, cities, suburbs, shopping malls, bars, casinos, banks, schools, workplaces, homes—and the bodies and behaviors of the inhabitants of these spaces—are special areas of experimentation for new surveillance technologies, in no small part because contemporary U.S. culture exhibits an intense preoccupation with the combined priorities of security and technology. Automated face perception technologies promise to provide high-tech security for the spaces and bodies of value in late capitalist societies, and belief in their ability to fulfill that promise plays a fundamental role in their institutionalization.[6]

Scholars writing in the field of surveillance studies have offered diagnoses about the preoccupation with security in late modern societies. The social theorist David Lyon is perhaps the strongest advocate of the view that these are essentially "surveillance societies," characterized by pervasive forms of social sorting and orchestration, and that surveillance should be taken seriously as an ethical, moral, and political concern.[7] A central political concern in debates about surveillance is whether the obsession with security and spread of new monitoring technologies are ushering in oppressive totalitarian societies akin to George Orwell's dystopic vision in *1984*. In his foundational study of five mass surveillance systems in Britain and the United States in the late 1960s and early 1970s (both state and commercial systems), the sociologist James Rule examined the extent to which each system approximated the "total surveillance society" depicted in Orwell's novel. He found that they shared "many of the sociological qualities . . . though of course they are much less powerful and they do not necessarily pursue the same malevolent purposes."[8]

Rule conducted his seminal study on the cusp of computerization, providing an invaluable snapshot of the moment and identifying many of the shortcomings in mass surveillance systems—in terms of size, degree of centralization, speed of information flows, points of contact with clientele, and

capacity or effectiveness—that computer databases and networks promised to overcome. Subsequent studies have examined the influence of computerization on monitoring practices, including David Burnham's work on the rise of the "computer state," Oscar Gandy's examination of the computer-assisted market research machine (the "panoptic sort"), and Mark Poster's analysis of the intensified panoptic functions of the database form.[9] Much of this scholarship suggests that the novelty of these new computerized forms of surveillance derives from the extent to which computers have enabled distinct surveillance systems to converge into a larger "information infrastructure" or "surveillant assemblage."[10] Networked systems allow data to flow from one site to another, back and forth between state and private-sector organizations, enabling ever more sophisticated and effective forms of social control. These scholars maintain that, while late capitalist societies may not precisely mirror Orwell's vision, computerization is nevertheless enabling significant advancements in institutionalized forms of surveillance.

These arguments clearly have a basis in real developments. But a problem with theoretical claims about the convergence of surveillance systems afforded by computerization is a tendency to gloss over the amount of effort that goes into developing and integrating new technologies and systems.[11] The introduction of new surveillance technologies and the convergence of surveillance systems do not happen seamlessly or automatically, instead presenting major logistical, technical, and political challenges involving conflicts and negotiations among various vested social actors. In the case of automated facial recognition systems, a common perception is that these technologies are either already deployed in a variety of settings, or that their deployment is happening at a rapid pace. One encounters these assumptions repeatedly not only in industry discourse and press accounts but also in the surveillance studies literature, and even those who recognize that the technologies are not widely diffused tend to see them as a largely inevitable part of the not-too-distant future.[12] But while experimental systems have already been deployed in a variety of contexts, the widespread use of these technologies has never been a foregone conclusion. For reasons that I explore in this book, whether in fact automated facial recognition and expression analysis systems can accomplish what their proponents aim to accomplish remains an open question. Although developers are making incremental improvements in algorithms and other dimensions of software and hardware development, so far these technologies do not work very well outside constrained settings. Computerized face perception is proving to be an incredibly difficult technology to engineer.

The problem with assumptions about the rapid diffusion and integration of new surveillance technologies is not only that the process is more complicated than it seems. Understanding the experimental status of the technology is critical, because the prevalent myth of inevitability surrounding this and other new forms of surveillance itself performs an important role in their institutionalization, and in the broader effort to shape the future toward certain ends. Making authoritative predictions about increasingly ubiquitous and intrusive surveillance techniques encourages public acquiescence, while suppressing alternative, less technocratic ways to address complex social problems and envision a better future. Assumptions about the rapid development and convergence of surveillance systems support what William Bogard calls "the imaginary of surveillant control"—a hyper-real vision of perfectly functioning and totalizing surveillance that is more real than the real in the Baudrillardian sense.[13] Creating the illusion creates the reality, as technologies of simulation supersede the material forms and effects of actual monitoring systems. In short, people come to believe in the power and sophistication of surveillance systems, and this belief itself has important effects on social organization and practice. For this reason, according to Bogard, it is important to understand not so much "what surveillance is and does"—what a surveillance system is actually capable of—but instead how the image of totalizing surveillance itself informs the logic of system development and functions as a form of social control.[14]

While Bogard is right that the surveillant imaginary is itself a problem in need of critical analysis, this does not mean that social theories of surveillance can abandon consideration of the actual capacities of surveillance systems (leaving it to legal and policy analysts, as Bogard suggests). For social theory no less than legal and policy analysis, it remains crucial to understand "what surveillance is and does," to leave the reality principle intact rather than assuming that what really matters is the surveillant imaginary. Instead of making assumptions about the convergence of surveillance forms or the effects of a totalizing image of surveillance, it is important to investigate precisely how and to what extent system convergence is happening, and how beliefs about the power of surveillance help to forward the aims of developing more sophisticated and ubiquitous surveillance systems. While there is clearly a need to make theoretical speculations about the convergence of surveillance systems, there is also a need to understand precisely how surveillance technologies are developed, how convergence happens, and whose interests are served in the process. Why are certain types of technologies developed and certain systems integrated in the first place (to serve what

specific needs or priorities)? What policies and procedures are put in place to make system integration possible? When and why does convergence fail, and what are productive effects of those failures? When are new technologies and efforts at system convergence effective and how is that effectiveness determined? "Technologies are not 'effective' by virtue of some straightforward appeal to science," write Haggerty and Ericson. "Instead, a determination of effectiveness is the net outcome of often contentious political struggles."[15] Although proponents of facial recognition and expression analysis systems often define their effectiveness in narrowly technical terms, in reality, determinations of their effectiveness involve contentious political struggles and intense persuasive efforts.[16]

By focusing on the politics of developing and deploying specific new technologies, this book aims to demonstrate not the inevitability of a particular technological future, but its profound contingency and contestability. In order to understand the social implications and historical significance of new computerized forms of surveillance, it is especially important to avoid making determinist assumptions about the inevitability of new technologies, and to understand the process by which proponents attempt to establish their authority, legitimacy, and necessity. I investigate what proponents want to achieve with automated face perception technologies and what would be required to put that vision into place. In the process of examining what would be necessary to bridge the gap between the current state of the technology and the vision, I identify the tensions and contradictions that would have to be designed into these systems—in other words, the ambiguity and error that the technologies would have to absorb in order to function effectively in real-world contexts. These inherent tensions and contradictions call into question claims about the accuracy and authority of new automated forms of face perception, and also raise questions about their effectiveness as technologies of security.

Drawing insights from communication scholarship, this book examines how the interrelated priorities of institutional actors are shaping the effort to program computers to identify human faces and facial expressions, including business entities, law enforcement, and state security agencies. Work on the history of media and communication technologies has much to offer the study of new surveillance technologies, in no small part because these technologies are themselves forms of media that both derive from and help to shape the evolution of technologies like video, television, telephony, and computing. The drive to produce more effective forms of surveillance has

had a consistent role to play in the development of communication technologies, including visual media technologies.[17] As Tom Gunning has shown, a specific and direct historical relationship exists between the physiognomic analysis of the face and the development of photography and motion pictures. "The desire to know the face in its most transitory and bizarre manifestations was stimulated by the use of photography," writes Gunning, "but that desire, in turn, also stimulated the development of photography itself, spurring it to increasing technical mastery over time and motion, prodding it toward the actual invention of motion pictures."[18] A cultural analysis of automated facial recognition and expression analysis technologies provides evidence that the drive to "know the face" continues to be stimulated by new photographic technologies, while at the same time pushing the development of these technologies in particular directions.

As technical projects, automated facial recognition and automated facial expression analysis pose distinct problems, and they represent related but somewhat distinct research and development efforts. Strictly speaking, facial recognition technology treats the face as an index of identity, disregarding its expressive capacity and communicative role in social interaction. This is nowhere more apparent than in new rules prohibiting drivers from smiling for their driver's license photos, in order to improve the accuracy of computer matching.[19] The aim is to use the iconicity of facial images as a means of establishing their indexicality, their definitive connection to real, embodied persons. Automated facial expression analysis, on the other hand, targets precisely what facial recognition technology attempts to control for—the diverse meanings that an individual face can convey—and in this way promises to accomplish what facial recognition technology fails to do, using the surface of the face to see inside the person. Where automated identification of individual faces disregards their affective qualities, automated facial expression analysis treats those affective dimensions as objects for precise measurement and computation.

However distinct these projects are in a technical sense, the face is at once both a marker of identity and a site of affect, and the differentiation of the face along these lines serves the purposes of scientific investigation and technical engineering more than it does a theoretically rich understanding of human communication. The diverse range of ways human beings use their faces and interpret the faces of others does not reduce easily to a set of technical processes. In a basic sense, the face is never a static object, and different facial expressions can change the appearance of the face considerably. Not surprisingly, there is also overlap in technical efforts to develop

automated forms of facial identification and expression analysis. The initial steps of automated facial recognition—"face detection" and "feature extraction," or locating a face in an image and extracting relevant features—are also the necessary first steps in automated facial expression analysis. In addition, computer scientists have investigated individual differences in facial expressiveness as a means of augmenting the accuracy of facial identification algorithms.[20] More significantly, the possibility of robust computer vision systems depends on the fusion of technical systems designed to simulate these interrelated forms of visual perception. While there may be distinct applications for identification versus expression analysis, at least in the short term, the predominant view in computer science is that more advanced forms of computer vision require designing systems that can adequately perform both of these functions, and especially to do so as well as, or preferably *better than,* humans.

The question of whether computers can be made to identify faces and facial expressions as well as or better than humans raises philosophical questions concerning the nature of sight and visual perception that have animated theoretical debates about both artificial intelligence and visual media technologies. The question at the heart of the Turing test, about whether a machine can "think" or exhibit humanlike intelligence, focused on the ability of a computer program to convincingly manipulate natural language, but a similar question likewise pertains to whether a machine can "see." The answer of course depends on what it means to see, and while "seeing" is obviously a physiological process, it is also a cultural practice, shaped by social and historical forces. The art historian Martin Jay uses the term "scopic regime" to describe an ideal typical combination of visual theories and practices that together create more or less unified "ways of seeing."[21] While there may be multiple, competing scopic regimes operating in any particular context, most theorists of visual culture agree that the dominant scopic regime of modernity is what Jay calls "Cartesian perspectivalism," the privileging of a rational observing subject capable of surveying the world objectively and from a distance, in its totality. Kevin Robins has extended this concept to an analysis of new visual imaging technologies, arguing that they embody the characteristic Cartesian desire for visual sovereignty, signaling the progressive rationalization of vision in order to establish mastery and control over a chaotic world.[22] As these arguments suggest, sight and vision should not be understood as essential qualities or strictly physiological processes with universal, transhistorical meanings or functions. There is no such thing as natural or "true" vision. Instead, "vision is always a question of the *power to*

see," and "struggles over what will count as rational accounts of the world are struggles over *how* to see," as Donna Haraway has argued.[23] Because seeing is as much a cultural practice as a physiological process, Suzannah Biernoff notes, "it cannot provide a historical common ground or the basis of a shared aesthetic experience."[24]

It follows that *if* machines can see, they must necessarily embody particular ways of seeing, rather than possessing a universal, disembodied, objective form of vision, outside of any particular vantage point or subject position. This is as true of automated face perception as it has been of other photographic technologies. The *digitizing impulse* behind the development of computer vision techniques is in one sense a re-articulation of the *mechanizing impulse* that motivated the adoption of photography by the positivist sciences in the nineteenth century, as Daston and Galison have documented—another way of standardizing images and eliminating individual human judgment in their interpretation.[25] Surveying the world from afar in a detached, authoritative way is the goal that underpins most machine vision projects.[26] In discussions about the merits of automated face perception technologies, one repeatedly encounters claims to their authority and technical neutrality. Simply by nature of being computerized, facial recognition systems are deemed more accurate and objective and less subject to the prejudices and apparent inadequacies of human perception. Automated facial expression analysis similarly promises to introduce a level of precision to the interpretation of facial expressions, detached from and exceeding human perceptual capabilities. The claim is that these technologies will create accurate, precision-guided, objective, all-seeing machines that function much better, more efficiently, and more powerfully than human perception alone.

This book unpacks such claims by examining what is involved in the effort to invest computers with the capacity to recognize human faces and facial expressions. Quite a bit of effort goes into this project: the painstaking development of algorithms to digitize the analog world of faces, the amassing of vast storehouses of facial images to serve as the basis of computerized visual memory, the development of efficient image retrieval methods, the tagging of billions of images with metadata to make them more searchable, the wiring of physical spaces with networks of cameras and computers and other hardware, the development of software interfaces that make sense to human users of these systems, the training of people to use the technologies and continuous retraining to help them keep pace with software and hardware upgrades, and much more. Peering inside the construction of this massive machinery

of vision reveals that computers "see" only in a metaphorical sense, only in highly constrained ways, and only with a significant investment of human effort. Computer vision systems are very much constrained by the purposes of their design, and suggesting that a computational model of vision represents an objective, detached form of vision elides the intentions and labor behind the design, deployment, and uses of these technologies. The desire to overcome the ambiguities and interpretive flexibility inherent in human perception and social interaction is driving the development of incredibly complex machines that can do what humans do everyday but in a totally imperfect way. These new techniques of observation will not lead to the fulfillment of the rationalist program, as Kevin Robins suggests. The rationalist program will never be complete, but the longing to fulfill it, and the vain belief that it can be fulfilled—that total, perfect knowledge of the world is possible—is one of the major motivating forces behind the pursuit of new technologies of vision and new human-machine visual assemblages.

Just as there is no standard or universal way of seeing, there is no universal way of seeing the face. A significant contingent of psychologists and other researchers who study the face suggest otherwise, but I take the position in this book that there are many ways of seeing the face, and that the meaning and experience of face-to-face interaction is a historically and culturally variable set of practices. We find evidence of this variability, for example, in research suggesting the existence of an "other race effect" or an "own-race bias" in people's ability to recognize faces. People tend to have more difficulty recognizing the faces of people outside their primary social group, meaning that individuals learn to see some faces better than others. Humans' capacity for facial recognition is a specialized area of research in fields like psychology and neuroscience, and much of this research suggests that there may indeed be something physiologically special about the way humans see and recognize faces that distinguishes this practice from the way they see and identify other types of objects.[27] But even the obvious existence of a physiological dimension to face perception does not mean that there exists a universally human way of seeing the face. While face perception may in fact be "the most developed visual perceptual skill in humans,"[28] the critical role that it plays in human social interaction—and the wide array of technologies and practices designed and implemented to facilitate facial representation and interpretation—means that face perception cannot help but vary widely in different contexts. The ways humans see the faces of others change, of necessity, along with changing cultural practices, social conventions, and forms of social and technological organization. In short, face perception does not

reduce to a universal physiological process that can be manifested in a standardized computational system; instead, there are a wide variety of "scopic regimes" of the face, a wide range of ways in which people use their faces and interpret the faces of others. Computation is itself a culturally and historically specific way of analyzing faces and modeling visual perception.

Facial recognition technology is more advanced in its development and applications than automated facial expression analysis, and for this reason receives considerably more attention in this book. Computer scientists began developing algorithms for detecting faces in images and distinguishing faces from one another in the 1960s, as part of the range of problems being addressed in the areas of automated reasoning and pattern recognition. Not surprisingly, the Department of Defense funded much of the research in the United States, marrying the automation of facial recognition to military priorities since its inception. After several decades of sporadic development, limited to some extent by the available computing technology but also by the lack of well-formulated social uses, prototype systems began to take shape. New companies with names like Visionics, Viisage, and Miros Inc. started marketing commercial facial recognition systems in the 1990s. The early period of commercialization—still ongoing—has involved the search for markets for unproven products, with proponents working to move the technology beyond military applications. Vendors have directed most of the effort toward building a customer base among institutional users—military as well as civilian government agencies, police departments, and business enterprises, including manufacturers of computers, automated teller machines, and other electronic equipment—working with these potential customers to define compelling social needs for more high-tech forms of surveillance and identification in the form of expensive, large-scale, database-driven facial recognition systems.

In its applications for biometric identification, facial recognition technology is one of array of technologies being developed to address a fundamental concern of modern societies: the problem of "disembodied identities," or the existence of visual and textual representations of individuals that circulate independent of their physical bodies. Long before the development of audiovisual media and electronic databases, the circulation of visual and textual representations created the conditions whereby certain classes of human identities became unmoored from their bodily existence. But it was the new communication technologies of the nineteenth century, like telegraphy, photography, telephony, and the phonograph, that gave these representations

new mediated forms and amplified the uncanny phenomenon of ubiquitous incorporeal replicas moving through society disarticulated from their embodied human counterparts. In 1886, Frederic Myers, a member of the British Society for Psychical Research, coined the phrase "phantasms of the living," to refer to the proliferation of these humanoid replicas.[29] "What men and women in the late nineteenth century faced with alarm," writes John Durham Peters, "is something we have had over a century to get used to: a superabundance of phantasms of the living appearing in various media."[30] While we are now quite accustomed to the existence of these "phantasms of the living" and the representational roles they play in our lives, the period of computerization has seen a renewed explosion in their quantity, forms, and uses, intensifying the problem of how to connect them back to embodied persons. One of the main reasons why digital biometric technologies are taking institutionalized forms at this historical juncture is because they promise to resolve one of the central problems of communication in large-scale societies: bodies missing in action from mediated communicative contexts.[31]

Insofar as they promise to re-embody disembodied identities, new biometric forms of identification supplement, and even replace in some cases, what Craig Robertson calls a "documentary regime of verification."[32] This system of standardized documents, archives, and administrative procedures for the management of individual identities itself displaced the more personal and informal forms of trust and recognition characteristic of smaller-scale forms of social organization. The aim of a documentary regime of verification was to assign each individual an official identity that could be verified in repeated transactions with the state and other institutions. These official forms of bureaucratic identification cobbled together a set of existing and already mediated markers of identity—such as names, addresses, signatures, and photographs—to create a more stable and standardized form of identity that could be verified via the very bureaucratic apparatus that constitutes that identity. In short, our seemingly self-evident "official identities" are in reality a product of bureaucratization and a relatively recent historical construction, and considerable effort has gone into designing systems that can produce and reproduce these identities.[33]

Much like the drive to produce more advanced and precise forms of visual perception, states and other institutional users of identification systems are in constant pursuit of the perfect means of identification. Each new innovation in identification systems (fingerprinting, standardized documents, ID photos, computerized record keeping, and machine-readable documents, to name a few) has been designed to resolve some of the tensions inherent in

the mediated process of identification, and especially to lay claim to making an absolute, immediate connection between bodies and their official identities. The aim of biometric identification technologies—like optical fingerprinting, iris scanning, and voice recognition—is to bind identity to the body using digital representations of unique body parts, or, in the case of voice printing, by capturing, digitizing, and analyzing the sounds that the body produces. Digitization—translating images and other analog texts into binary code—is the latest in a long line of techniques posited as a definitive, accurate, and objective means of binding identity to the body. The claim is that by digitizing visual representations of the body, the body itself will be laid bare and tied directly into information networks.

Today, DNA in particular is viewed as the ultimate identifier, the most precise and scientific means of linking bodies to identities. Our genes are thought to be quite literally the code that determines who we are, to represent our identities in an absolute sense. But genes themselves are not codes that identify bodies; they must be translated into coded form using a specialized technique of inscription, a process of mediation that involves layers of technical integration. A genetic code ostensibly makes a direct link to physical bodies when in fact it does no such thing. Not only does it require a process of technical mediation to generate, but a genetic profile establishes only a probability. It cannot point definitively to a specific body, and it cannot serve as an index of identity unless it is connected to other identifying information about the person it represents. Belief that a genetic code is literally extracted from the physical body (and thereby connected to it in an unadulterated, absolute sense) is a form of what Donna Haraway calls "gene fetishism." Drawing on Alfred Whitehead's notion of "the fallacy of misplaced concreteness," Haraway explains that gene fetishism mistakes the abstraction of the gene for a concrete entity: "A gene is not a thing, much less a 'master molecule,' or a self-contained code. Instead, the term *gene* signifies a node of durable action where many actors, human and nonhuman, meet."[34]

The claim to authority of biometric identification rests in part of on a similar form of fetishism, a disavowal of the mediated relationship between bodies and identities and the overdetermined process by which a "biometric" is produced. In laying claim to a direct link to bodies, biometric technologies promise to stabilize the messy ambiguity of identity, to automatically read a stable, individual identity directly off the body. To be effective, the connection that biometric technologies establish between identity and the body must appear natural and self-evident: *of course* our true identities can be ascertained by scanning and digitizing our faces, eyes, voices, hands, and fin-

gertips, the logical next steps in identification systems. But the assumption that biometrics are derived from and link directly to physical bodies conceals a complex technological process of mediation, as well as a whole set of historical relationships that create the very conditions of possibility for biometric identification.

In her work on protocols of identification in nineteenth-century Europe, Jane Caplan explains that any standardized identification system depends for its effectiveness on the stability or replicability of its operations, and their relative invisibility (i.e., their naturalization).[35] Stable, standardized identification systems were a crucial development for societies increasing in size, mobility, and anonymity. But "in spite of every effort at stabilization," writes Caplan, "the culture of identification was—is—essentially unruly, not only because of its vast scope and variety, but because even in its most controlling and technologized forms it is based on a concept that is itself difficult to stabilize and control"—the concept of *identity*.[36] The term "identity" signifies both what is unique about an individual and what associates her with others like her, a dual meaning that keeps the concept "constantly in motion against itself, even before one addresses any of the mechanical problems of operating an efficient system of identification in practice."[37] The inherent instability and mediated characteristics of identity make building identification systems a process fraught with difficulty, and these challenges have in turn led to perpetual efforts at standardization. The turn to digital biometric identification represents the latest in a long line of efforts to stabilize and standardize identification systems, and to push those standardized categories of identity back out onto individual bodies.

In its application to identification systems, the dual aim of facial recognition technology is to automate the mediated process of connecting faces to identities, and to enable the distribution of those identities across computer networks in order to institutionalize a more effective regime of mass individuation than afforded by documentary identification alone. By *mass individuation* I mean the process of extending across an entire population technologies and procedures for treating each individual as a specific case, a process greatly facilitated by computerization and the development of networked databases.[38] Mass individuation supports other social processes and political-economic priorities, such as *mass customization*, or the mass production of customized goods and services, and the *individualization of labor,* the process of identifying and individuating workers in order to measure their precise contribution to production.[39] Mass individuation is also a modern governmental strategy for security provision and population manage-

ment, a social regulatory model that involves knowing in precise detail the identity of each member of the population in order to differentiate individuals according to variable levels of access, privilege, and risk. In other words, the possibility of digital biometric identification should not be understood in narrow terms as the natural outgrowth of technical advancements in identification systems. Instead, these technologies are being envisioned and designed to fulfill certain perceived social necessities and political-economic demands of large-scale, late capitalist societies—societies characterized by a predominance of mediated forms of social organization and vastly asymmetrical distributions of wealth. The expansion of computer networks has created new problems of communication-without-bodies, necessitating new techniques for identifying people, verifying their legitimate identities, and otherwise gaining knowledge about who they are.

But while facial recognition and other biometric technologies hold out the promise, they can never completely resolve the inherent problems with building stable identification systems for the mass individuation of large-scale societies. Digitizing faces, storing infinitely reproducible images of faces in databases, networking those databases, and designing more sophisticated image retrieval techniques are technical practices that can increase the scale of identification systems. Automation facilitates the standardization of identification systems and the expansion of their reach over time and space. Facial recognition technology in particular promises to provide a means for identification "at a distance," in terms of both the measurable distance between cameras and people in particular local spaces, and the more extended reach of automated identification over networks covering distant locations. But the automation of facial identification cannot definitively stabilize identity. No matter how sophisticated the algorithms for matching facial images or how comprehensive the image databases, the tension remains between identity "as the *self-same*, in an individualizing, subjective sense, and 'identity' as *sameness with another*, in a classifying, objective sense."[40] What Stuart Hall has argued about "cultural identity" is in fact also true of the official, bureaucratic form of identity: it is not a fixed and stable object or "an already accomplished fact," but a "'production,' which is never complete, always in process, and always constituted within, not outside, representation."[41]

This does not mean that the automation of facial recognition is a meaningless endeavor, bound to fail as a means of establishing consistently identifiable bodies over time and space. There will be technological failures, to be sure, but the technology is in many ways already effective, because the very pursuit of computerized, biometric forms of identification suggests that we

are witnessing a reconfiguration of identity, the attachment of new meaning and new practices to what identity is and to how it works. Identity remains that which both differentiates us and associates us with others. But we now have new ways of identifying ourselves and being identified that did not exist before, and there is an increasing quantity of instances where we are required to establish our identities definitively so that our status or level of access can be determined, and information about those transactions can be recorded, with all that data in turn becoming part of our identities.[42] Identity is now understood as a disembodied aggregate of data, a digital representation of the person constructed over time and space based on the perpetual collection of more data. Although difficult to measure empirically, it is hard to deny that people living in modern societies today engage in more transactions on a daily basis that require them to interface with identification systems than people did in the past. Rather than stabilizing identity, the proliferation of these transactions is making it messier than ever, in turn leading to the perpetual pursuit of new efforts at stabilization. It is difficult to know where this seemingly self-fulfilling process is going, but it would certainly not be insightful to say that the practices of identity remain the same as they ever were.

Facial recognition technology should not be conflated with other types of biometrics. It is unique relative to other forms of biometric identification because the content of the medium is the image of the human face. This makes it an especially challenging technical problem and puts it at a disadvantage relative to other biometrics in terms of its level of development, ease of adoption and use, and general viability. Optical fingerprinting is in many ways a more accurate and reliable means of binding identities to bodies, for example. As two MIT researchers put it, "Developing a computational model of face recognition is quite difficult, because faces are complex, multidimensional, and meaningful visual stimuli."[43] An individual face changes considerably not only with its surface movements, but also with aging, trauma, surgery, makeup, and lighting. Faces themselves change over time, and images captured of a face can be of highly variable quality. The variability of faces across populations, as well the dynamic states of the individual and the range of images that can be rendered of a particular person, make automated facial recognition a very challenging technical problem.

In short, like identity, the face is a difficult object to stabilize. Computer scientists have developed a variety of different techniques designed to translate an image of the face into a "facial template," a smaller amount of data that can be compared against existing images stored in a comparison database.[44]

And the digitization of facial images is only one small part of the design of facial recognition systems. Faces must be detected in images, extracted from the background, and "normalized" so that they conform to a standard format. The matching process typically results in not one but a range of possible matches, depending on a set "matching threshold." High matching thresholds increase the chances of missing a positive match, while low matching thresholds can produce a large number of "false positives." At the level of their applications, automated facial recognition systems are divided into two general types: those that use static images of the face and those that analyze dynamic images of faces from video.[45] Applications can also be differentiated according to whether the aim is to *verify* the identities of individuals (to determine whether people are whom they claimed to be, for example, at a border crossing station or when engaging in a financial transaction), or to *identify* people whose identities are unknown (in urban or crowd surveillance scenarios, for example). The first problem requires a one-to-one facial image comparison, while the second problem involves a more technically challenging and information-intensive process of comparing an image captured of an unknown person's face against a database of facial images (a form of content-based image retrieval, or CBIR).

While facial recognition technology presents considerable technical challenges, it also possesses certain advantages over other biometrics. One set of advantages involves the practical improvements it portends for identification systems; for example, "it poses fewer demands on subjects and may be conducted at a distance without their knowledge or consent."[46] Along with the practical upgrades it offers for surveillance and identification systems, the technology's use of the face as an object of identification invests it with certain cultural and ideological capital. Facial recognition technology combines an image of high-tech identification with a set of enduring cultural assumptions about the meaning of the face, its unique connection to individuality and identity (in its multiple, conflicting senses), and its distinctive place in human interaction and communication. It is for these reasons, as much as its practical advantages, that facial recognition technology has received special attention.

Long before computer scientists began developing techniques for automated face perception, visual media technologies were being developed and used to analyze, classify, and identify human faces. The use of the face as an object for the social and biological classification of people has a long and sordid history, indelibly tied to the use of the human sciences to justify social inequality. The science of physiognomy, widely accepted in the West until

the mid to late nineteenth century, held that people's faces bore the signs of their essential qualities and could be visually analyzed as a means of measuring moral worth. Photography was invented at the height of this "physiognomic culture" and was put to use for facial analysis and classification. Most famously, the psychiatrist Hugh Welch Diamond (1809–1886) used photography to analyze and document the alleged facial indicators of insanity, and the eugenicist Francis Galton (1822–1911) developed a form of composite photography that he used to claim the existence of criminal facial types. As historians have shown, physiognomic classifications exhibited a consistent tendency to conflate ostensible facial signs of pathology with racial and class differences.[47] Galton's work in particular was "the culmination of all the nineteenth-century attempts to objectify, classify, and typify humans through portraiture," and Galton's ideas went on to form the basis of many of the racist, classist, and biologically determinist theories about human difference promulgated in the twentieth century.[48]

In what ways does the automation of face perception connect to this history? For its part, facial recognition technology appears to diverge radically from the business of using the face as an object for the social or biological classification of people. In practical terms, the aim of facial recognition systems is to identify individuals, to use the face like a fingerprint—as an index or recorded visual trace of a specific person. Computer scientists have taken a variety of approaches to developing algorithms for automated facial recognition, and the specific techniques devised for digitizing facial images are not necessarily based on assumptions about facial typologies. In technical terms, the development of algorithms for translating images of faces into digital "facial templates" is more or less divorced from the explicit social and biological classification of faces in the conventional sense. Developers of facial recognition algorithms have attempted to operationalize a more individualizing form of facial identification. They are not primarily interested in measuring differences in facial features between people of different racial or ethnic identities, for example. This does not mean that things like skin color and other facial signs of racial, ethnic, or gender difference are irrelevant to the development of these techniques, but the classification of faces according to race, ethnicity, or gender is not, by and large, the problem that computer scientists are trying to solve in their efforts to design automated forms of facial recognition.

The individualizing logic that underpins the development of facial recognition systems gives them more obvious genealogical ties to the system of anthropometry developed in the late nineteenth century by Francis Galton's contemporary, the Paris police official Alphonse Bertillon (1853–1914), than

to Galton's photographic composites. As Allan Sekula has shown, Bertillon's system of anthropometry disregarded the possibility of generic categories and instead concerned the more practical, administrative aim of identifying individual criminals (much like fingerprinting, though significantly more laborious, since anthropometry involved not only taking photographs but also recording a series of bodily measurements and coded descriptions of each subject's body). Like anthropometric measurement, the algorithms developed to digitize faces and link those "facial templates" to identities are not designed to reveal anything about the essence of the individual. Unlike physiognomic analysis, facial recognition algorithms are not meant to derive knowledge of the interior of the person from the surface of the face. The individualizing logic that informs the design of facial recognition technology is part of the basis for claims about its technical neutrality—the aim is to identify individual faces rather than facial types.

But there are a number of problems with claims about the individualizing logic and technical neutrality of automated facial recognition. First, as I have already discussed, "identity" itself fundamentally embodies an individualizing and classifying logic, an inescapable tension that manifests at the level of system design. Even if the explicit classification of identities does not occur at the level of algorithm development, it does happen at the level of database construction and in the specific applications that give facial recognition technology a functioning form.[49] Again, facial recognition algorithms, or techniques for digitizing the face, represent only one part of the operation of facial recognition systems; most of these systems are designed to make use of an archive of facial images that define the parameters for the class of individuals that the system will identify. And forms of *social* classification, if not outright *biological* classification, inevitably happen at the level of database development, whether the database is a terrorist watchlist or an A-list of preferred customers. This was likewise true of the archival practices developed in the early application of photography to criminal identification. In Bertillon's system of anthropometric identification, Sekula explains, the camera alone was a limited technology. The contribution of photography to criminal identification came with its integration into a larger ensemble: "a bureaucratic-clerical-statistical system of 'intelligence."[50] A similar bureaucratic-statistical apparatus is required to make facial recognition algorithms effective for large-scale identification systems. "The act of recognition relies on comparison," writes Gunning, "and knowledge resides not in the single photograph, but within a vast photographic archives, cross-indexed by systems of classification."[51]

A related problem with claims about the technical neutrality of facial recognition technology and its divergence from the problematic assumptions of physiognomy has to do with its symbolic associations with antiquated notions of facial typologies. Facial recognition technology is inescapably tied to cultural assumptions about the relationship between the face and identity, including enduring beliefs about faces as carriers of signs that reveal the essential qualities of their bearers. As we will see in chapter 3, the technology's metaphoric connection to archaic ideas about physiognomic facial types gave it a special edge in the effort to define it as a security solution in the post-9/11 moment, especially through the trope of the "face of terror." The repeated use of this trope in press and policy discussions and in industry discourse about facial recognition technology did important ideological work, implying that embodied evil could be read off the faces of "terrorists," even if those faces had to be identified one by one, with the help of a new, methodically individualizing form of computerized facial identification. The racist implications of the "face of terror" trope underscored the underlying logic of social classification that informed the strategy of deploying new identification technologies for "homeland security." Although facial recognition algorithms were *not* designed to classify faces according to particular identity typologies, the deployment of large-scale, database-driven identification systems very much depended on, and promised to facilitate, a set of "biopolitical" security strategies that aimed to differentiate the population according to racially inflected criteria for determining who belonged and who did not, who was entitled to security and who posed a threat to that security.[52] Far from incorporating a neutral, all-seeing mode of visual perception, facial recognition systems promise to facilitate the diffusion of particular institutionalized ways of seeing, ones that rely on and seek to standardize essentialized identity categories under the guise of what appears to be a radically individualizing form of identification.

Both automated facial recognition and automated facial expression analysis gain metaphoric leverage in their historical connection to what Gunning refers to as "the semantically loaded and unceasingly ambiguous representation of the human face."[53] Both of these technological projects have genealogical ties to the history of using the face as an object of scientific investigation and social differentiation. But while facial recognition algorithms are not designed to use the surface of the face to reveal something about the interior of the person, automated facial expression analysis is another matter. The aim of automated facial expression analysis, or AFEA, is to peer inside

the person, using the dimensions and intensities of facial movements as a means of determining what people are feeling and thinking. Where facial recognition technology treats the face as a "blank somatic surface" to be differentiated from other faces, AFEA treats the dynamic surface of the face as the site of differentiation—not a blank somatic surface but a field of classifiable information about the individual.[54]

The facial expression classification scheme devised by the psychologist Paul Ekman and his colleagues in particular promises to play centrally in the development of AFEA. Ekman is best known for his work on deception detection, the influence of which has extended beyond the academic field of psychology to the development of police and military interrogation techniques.[55] In the 1970s, Ekman and his colleague Wallace Friesen undertook an eight-year-long study of facial expressions, creating a scheme of forty-four discrete facial "action units"—individual muscle movements combinable to form many different facial displays. This "Facial Action Coding System," or FACS, is now the "gold standard" system for analyzing facial expressions in psychology, and as we will see in chapter 5, computer scientists see it as a promising approach for the automation of facial expression analysis. In turn, a host of social actors see computerized facial expression recognition as holding special potential to facilitate automated deception detection and other forms of affect analysis in a wide range of settings.

The stakes can be high in applying a system of classification and a computational logic to the human world of emotional or affective relations. Something inevitably happens to affect—to our understanding of what affect is and how it works—when it is treated as data to be processed by computers. One of the primary aims of FACS and FACS-based AFEA is to make human affective behaviors more *calculable*, to open them up to precise measurement and classification, thereby making them more amenable to forms of intervention, manipulation, and control. In this way, the technology promises to function like other forms of psychological assessment—as a technique of subjectification, a means of applying normalizing judgments to individual behaviors in order to shape or direct those behaviors in particular ways. Like earlier applications of photography to the analysis of the facial expressions, automated facial expression analysis is intimately connected to the social regulation of affect, and to the corresponding project of "making up people,"[56] marking out the parameters of normalcy and establishing a set of prescriptions for conduct, including the range of appropriate emotional responses to the world and the ways those responses are enacted through the face.

The face, like the body, does not exist somehow outside of history, and the very possibility of automating the human capacity for face perception gives us cause to consider the cultural reconstruction of the face in these times. What exactly do we mean when we refer to "the face"? The face is often conceived as the site of our identity and subjectivity, the source of our speech, the location of much of our sensory experience. But thinking about the face in these terms requires disarticulating it from the rest of the body as well as from its social, historical, and material context, slicing it off from the physical and discursive networks that allow it to take in air, food, and water; to express, blink, and sniffle; to acquire and produce signals; and to otherwise engage in meaningful exchange with the world. Faces are assemblages of skin, muscle, bone, cognition, emotion, and more. But certainly we have faces, however inseparable they are from our bodies and the worlds we inhabit. We know what faces are and what they do, and we learn to use and interpret them in particular ways. We are now perfectly comfortable metaphorically cutting them off from our bodies by photographing them and treating those images as objects-in-themselves. Our photographed faces do not diminish our subjectivities, our identities, or our relations with others. Rather, photography is now used as a means of constructing, facilitating, and enhancing these dimensions of what it means to human. Of course it is also used to classify and individuate us in ways over which we have no control, and these different impulses of photography—its honorific and repressive tendencies, to borrow Allan Sekula's terminology[57]—continue to intersect in complex ways along with the development of new photographic techniques and practices.

Photographic and archival practices have helped create a false unity of the face—the very idea of the face as a singular, unified object, detached from the body and from the world. The idea that we might think of "the face" as an object-in-itself begins to fall apart when we see faces as assemblages, and when we consider the vast range of differences among faces, what they look like, what they do, and especially what they mean, within cultures, across cultures, and over time. Deleuze and Guattari claim that "all faces envelope an unknown, unexplored landscape," by which they surely mean something profound.[58] In "Year Zero: Faciality," they argue that concrete faces do not come ready made but are produced by an "abstract machine of faciality," a process of "facialization" whereby bodies and their surroundings are reduced to the face, a "white wall, black hole system."[59] The face that is produced by the abstract machine of faciality is

an *inhuman* face, and facialization can be a violent process, transforming the radical and deeply relational potential of becoming into a dichotomous relationship between signifier and signified.[60] The paradigmatic instance of facialization is the close-up of the face in cinema, which destroys what is recognizable, social, and relational about the face and "turns the face into a phantom."[61]

Although precisely what Deleuze and Guattari mean by the abstract machine of faciality is open to interpretation, one imagines that the translation of an embodied face into a digital "facial template," the circulation of millions of facial templates over networks, or the meticulous classification of facial movements according to a standardized coding system would fuel the engine of such a machine. If an abstract machine of faciality actually exists, then the automation of facial recognition and expression analysis would have to count as part of that machine's "technological trajectory." In his classic study of nuclear missile guidance systems, Donald MacKenzie uses the term "technological trajectory" to refer to an institutionalized form of technological change, a course of technical development that appears natural and autonomous from a distance because it has a relatively stable organizational framework, because resources are channeled to support the activities of that framework, and because the prediction that the technology can be made to work is perceived as credible.[62] In other words, a "technological trajectory" is a sort of "self-fulfilling prophecy."[63]

This book documents efforts to build a stable framework for the development of automated face perception technologies and to channel resources to support that framework. It also examines efforts on the part of proponents to make a credible case for the viability and desirability of these technologies. Functioning face perception technologies depend, now and in the future, on the formulation of compelling social uses and on a measure of faith in the capacity of the technology to do what it purports to do, regardless of whether it can do so with complete accuracy. The ability of computer scientists and other social actors to design machines that can "see" human faces depends on whether people believe that it is both possible and useful. Automated facial recognition and facial expression analysis do not have to work perfectly to be effective, and our belief in the inevitability of these technologies has important effects on their development and their uses. To take shape as functioning technologies, automated facial recognition and expression analysis must be viewed as obvious, necessary, and inevitable next steps in the technological trajectory of facial representation, identification, and interpretation—in the abstract machine of faciality.