**Chapter 17**

**Biometrics**

**And the Gileadites took the passages of Jordan before the**

**Ephraimites: and it was so, that when those Ephraimites which**

**were escaped said, Let me go over; that the men of Gilead said unto**  
 **him, Art thou an Ephraimite? If he said, Nay; Then said they unto**

**him, Say now Shibboleth: and he said Sibboleth: for he could not**  
 **frame to pronounce it right. Then they took him, and slew him at**

**the passages of the Jordan: and there fell at that time of the**

**Ephraimites forty and two thousand.**

**– Judges 12:5–6**

**17.1** **Introduction**

The above quotation may be the ﬁrst recorded military use of a security protocol  
 in which the authentication relies on a property of the human being – in this  
 case his accent. (There had been less formal uses before this, as when Isaac  
 tried to identify Esau by his bodily hair but got deceived by Jacob, or indeed  
 when people recognized each other by their faces – which I’ll discuss later.)

Biometrics identify people by measuring some aspect of individual anatomy

or physiology (such as your hand geometry or ﬁngerprint), some deeply ingrained  
 skill or behavior (such as your handwritten signature), or some combination of  
 the two (such as your voice).

In the 21st century the market has really taken off, with three major changes

since the second edition of this book in 2008.

1. There are many large-scale programs by states to identify citizens using

biometrics, of which the biggest single programme may be India’s Aadhaar  
 project, which has recorded the iris codes and ﬁngerprints of over a billion  
 people. International travel has been speeded up by international standard  
 biometric travel documents, the US-VISIT program which ﬁngerprints  
 visitors to the USA, and face-recognition passport booths at the borders  
 of the European Union.

2. There has been a massive improvement in face recognition technology,

brought about by the revolution in deep neural networks since 2012. This  
 has made passport booths steadily faster and more reliable, made mass  
 surveillance easier, and led to concerns about privacy and human rights –  
 particularly given its deployment in China.

3. Automatic ﬁngerprint readers are no longer a niche product for bank vaults

and welfare offices, but are deployed on hundreds of millions of mobile  
 phones. Now that people keep their entire lives in their phones, or on web  
 services for which their phones have the credentials, they are relied on to  
 stop a lost or stolen phone turning from annoyance into disaster.

The biometric systems market has taken off like a rocket, growing from $50m

in 1998 to over $1.5bn in 2005 [997] and $33bn in 2019 [2038].

I’ll start off by describing the biometric techniques that predate the computer

age – handwritten signatures, facial features and ﬁngerprints – then describe  
 how they have been automated, and then go on to explore some more modern  
 techniques.

**17.2** **Handwritten signatures**

Handwritten signatures had been used in classical China, but carved personal  
 seals came to be considered higher status; they are still used for serious trans-  
 actions in China, Japan and Korea. Europe was the other way round: seals  
 had been used in medieval times, but as writing spread after the Renaissance  
 people increasingly just wrote their names to assent to documents. Over time,  
 the signature became the standard. Every day, billions of dollars’ worth of con-  
 tracts are still concluded by handwritten signatures; how these will be replaced  
 by electronic mechanisms remains a live policy and technology issue.

Handwritten signatures are a weak authentication mechanism in that they’re

easy to forge, but they worked well enough for centuries because of the context  
 of their use. An important factor is the liability for forgery. Britain’s Bills of  
 Exchange Act 1882 provides that a forged handwritten signature is null and  
 void, and this has survived in the laws of many countries that were part of the  
 British Empire at the time, such as Canada and Australia. In these countries,  
 manuscript signatures are better for the customer, as the bank carries most of  
 the risk, but PINs and electronic tokens can be better for the bank – and so  
 have largely replaced them. Europe also went for electronic signatures following  
 lobbying by the French and German smartcard industries. In the USA, the

law makes banks liable for the electronic systems they deploy, so US banks  
 generally stuck with chip and signature cards rather than going for chip and  
 PIN. Courier companies also collect handwritten signatures as proof of receipt  
 as they’re the only thing that works for all recipients. So the veriﬁcation of  
 handwritten signatures continues to matter.

Now the probability that a forged signature will be accepted as genuine

mainly depends on the amount of care taken when examining it. Many bank  
 card transactions in stores are accepted without even a glance at the specimen

|  |  |  |
| --- | --- | --- |
|  |  |  |

signature on the card – so much so that many Americans don’t even bother to  
 sign their credit cards1. But even diligent signature checking doesn’t reduce the  
 risk to zero. An experiment showed that 105 professional document examin-  
 ers, who each did 144 pairwise comparisons, misattributed 6.5% of documents.  
 Meanwhile, a control group of 34 untrained people of the same educational level  
 got it wrong 38.3% of the time [1010], and the nonprofessionals’ performance  
 couldn’t be improved by giving them monetary incentives [1011]. Errors made  
 by professionals are a subject of continuing discussion in the industry but are  
 thought to reﬂect the examiner’s preconceptions [198] and context [587]. As the  
 participants in these tests were given reasonable handwriting samples rather  
 than just a signature, it seems fair to assume that the results for verifying sig-  
 natures on checks or delivery receipts would be even worse.

In most of the English-speaking world, most documents do not need to be

authenticated by special measures. The essence of a signature is the intent of  
 the signer, so an illiterate’s ‘*X*’ on a document is perfectly valid. A plaintext  
 name at the bottom of an email message therefore has full legal force [2042],  
 except where there are speciﬁc regulations to the contrary.

The exceptions come from conventions and special rules that vary from one

country to another. For example, to buy a house in England using money bor-  
 rowed from a bank of which you’re not an established customer, the procedure  
 is to go to a lawyer’s office with a document such as a passport, sign the prop-  
 erty transfer and loan contracts, and get them countersigned by the lawyer.  
 The requirement for government-issued photo ID was originally imposed by the  
 lender’s insurers, and became a ‘know-your-customer’ (KYC) provision of anti-  
 money-laundering regulations; the requirement that a real-estate purchase be in  
 writing was imposed centuries ago in order to collect tax on property transac-  
 tions.

Other types of document (such as expert testimony) may have to be no-

tarized in particular ways. Many of the anomalies go back to the nineteenth  
 century, and the invention of the typewriter. Some countries require that ma-  
 chine written contracts be initialled on each page, while some don’t; clashes in  
 conventions still cause serious problems.

It’s rare for signatures to be disputed in court cases, as the context mostly

makes it clear who did what. So this weak biometric mechanism actually works  
 fairly well in practice – the real problems come from a thicket of procedural  
 rules that vary by country and by application. Lawmakers have made various  
 attempts to sort out the mess, and impose uniform rules for electronic docu-  
 ments.

In section 26.5.2 I discuss the Electronic Signatures in Global and National

Commerce (‘ESIGN’) Act of 2000, which legitimised contracts made by clicking  
 on buttons in web pages, and the much more heavyweight European eIDAS  
 Regulation (910/2014) which requires all Member States to accept electronic  
 signatures made using approved products. This was originally designed to help  
 the smartcard industry, but as many people and ﬁrms need to sign things occa-

1Indeed it’s not in the cardholder’s interest to give a specimen signature to a thief – if the

thief makes a random signature on a voucher, it’s easier for the real cardholder to disown it.  
 Signing the card is in the bank’s interest but not the customer’s.

|  |  |  |
| --- | --- | --- |
|  |  |  |

sionally and don’t want to buy special hardware, the latest regulation now allows  
 online signature service ﬁrms to generate signatures in their cloud service that  
 are considered legally binding, even though the security of the customer’s phone  
 or laptop may leave a lot to be desired. Signature services typically generate an  
 electronic document with a machine-written signature that we’re supposed to  
 pretend was handwritten; there’s also an electronic signature whose veriﬁcation  
 by the service provider we’re supposed to trust.

A separate topic is the automatic recognition of handwritten signatures,

such as on checks. This became one of the earliest topics of serious biometric  
 research in the 1980s by ﬁrms selling check-processing equipment to banks. In  
 early systems, an operator was presented on screen with the check image and  
 the customer’s reference signature, and took the decision. For cost reasons this  
 was only done for amounts over a few thousand dollars; smaller checks just went  
 straight through and it was up to the account holder to dispute them. From  
 the early 1990s there were signature tablets which record not just the shape of  
 the curve but also its dynamics (the velocity of the hand, where the pen was  
 lifted off the paper, and so on). These are used by delivery drivers to collect  
 receipts for goods and also for credit card transactions. Since the early 1990s  
 the better products can compare captured signatures against specimens enrolled  
 previously.

Like alarms, most biometric systems have a trade-off between false accept

and false reject rates, often referred to in the banking industry as the *fraud* and  
 *insult* rates and in the biometric literature as *type 1* and *type 2* errors. Many  
 systems can be tuned to favor one over the other. The trade-off is known as  
 the *receiver operating characteristic*, a term ﬁrst used by radar operators; if  
 you turn up the gain on your radar set too high, you can’t see the target for  
 clutter, while if it’s too low you can’t see it at all. It’s up to the operator to  
 select a suitable point on the curve. The *equal error rate* is when the system  
 is tuned so that the probabilities of false accept and false reject are equal. For  
 tablet-based signature recognition systems, the equal error rate is at best 1%; for  
 purely optical comparison it’s several percent. This is not fatal in an operation  
 such as a check processing center, as the automated comparison is used as a  
 ﬁlter to select checks for human scrutiny. However, it is a show-stopper in

a customer-facing application such as a retail store. If one transaction in a

hundred fails, the aggravation to customers would be unacceptable. So back  
 in the 1990s, UK banks set a target for biometrics of a fraud rate of 1% and  
 an insult rate of 0.01%, which was beyond the state of the art in signature  
 veriﬁcation and ﬁngerprint scanning – and indeed still is [719]. In fact, even  
 the 1% equal error rate for tablets was achieved by excluding *goats* – a term  
 used by the biometric community for people whose templates don’t classify well.  
 Vendors typically exclude people without eyes from statistics on iris scanners  
 and manual workers with worn ﬁngertips from ﬁngerprint statistics. This can  
 lead to deceptive performance claims and hide issues of social exclusion.

In general, biometric mechanisms tend to be more robust in attended oper-

ation where they assist a guard rather than replacing them.

|  |  |  |
| --- | --- | --- |
|  |  |  |

**17.3** **Face Recognition**

Recognizing people by their facial features is the oldest identiﬁcation mechanism  
 of all, going back at least to our early primate ancestors. Biologists believe that  
 a signiﬁcant part of our cognitive function evolved to provide efficient ways of  
 recognizing other people’s facial features and expressions [1604]. For example,  
 we are very good at detecting whether another person is looking at us or not.

The human ability to recognize faces is an important baseline for many rea-

sons, of which one is the reliance placed on photo ID. Drivers’ licenses, passports  
 and other kinds of identity card are not only used to control entry to computer  
 rooms directly, but also to bootstrap most other systems. The issue of a pass-  
 word, or a smartcard, for access to a system is often the end point of a process  
 that was started by that person presenting photo ID when applying for a job or  
 opening a bank account.

So how good are we at identifying strangers by photo ID, as opposed to

identifying friends in the ﬂesh?

The simple answer is that we’re not. Psychologists at the University of

Westminster conducted a fascinating experiment with the help of a supermarket  
 chain and a bank [1035]. They recruited 44 students and issued each of them  
 with four credit cards each with a different photograph on it:

*•* one of the photos was a ‘good, good’ one. It was genuine and recent;

*•* the second was a ‘bad, good one’. It was genuine but a bit old, and the  
 it was typical of the photo that most people have on their photo ID;

*•* the third was a ‘good, bad one’. From a pile of a hundred or so random  
 looked like the subject. In other words, it was typical of the match that  
 criminals could get with a stack of stolen cards;

*•* the fourth was a ‘bad, bad’ one. It was chosen at random except that it  
 of the match that really lazy, careless criminals would get.

The experiment was conducted in a supermarket after normal business hours,

but with experienced cashiers on duty, and aware of the purpose of the experi-  
 ment. Each student made several trips past the checkout using different cards.  
 It transpired that none of the checkout staff could tell the difference between  
 ‘good, bad’ photos and ‘bad, good’ photos. In fact, some of them could not even  
 tell the difference between ‘good, good’ and ‘bad, bad’. Now this experiment  
 was done under optimum conditions, with experienced staff, plenty of time, and  
 no threat of embarrassment or violence if a card was declined. Real life perfor-  
 mance can be expected to be worse. In fact, many stores do not pass on to their  
 checkout staff the reward offered by credit card companies for capturing stolen  
 cards. So even the most basic incentive is absent. Yet at least two banks that  
 had experimented with photos on credit cards had experienced a substantial

|  |  |  |
| --- | --- | --- |
|  |  |  |

drop in fraud [154]. The conclusion was that the beneﬁt to be had from photo  
 ID at the time was basically its deterrent effect [689].

So maybe people won’t use their facial-recognition skills effectively in iden-

tiﬁcation contexts, or maybe the information we use to identify people in social  
 contexts is stored differently in our brains from information we get by looking at  
 a single photo. Recognising passing strangers is in any case much harder than  
 recognising people you know. It’s reckoned that misidentiﬁcations are the main  
 cause of false imprisonment, with 20% of witnesses making mistakes in identity  
 parades [2044] – not as bad as the near-random outcomes when comparing faces  
 with photos, but still not good.

Since photo-ID doesn’t work well with human guards, many people have

tried to automate the process. Attempts go back to the nineteenth century,  
 when Francis Galton devised a series of spring-loaded “mechanical selectors” for  
 facial measurements [738]. But automated face recognition actually subsumes  
 a number of separate problems, in most of which we don’t have the luxury of  
 taking careful 3-d measurements of the subject. Automated passport control  
 booths may be the easiest: the subject looks straight at the camera under  
 controlled lighting conditions, and their face is compared with the one on ﬁle.  
 In forensics, we may be trying to establish whether a suspect’s face ﬁts a low-  
 quality recording on a security video. The hardest of all is surveillance, where  
 we may want to scan a moving crowd of people at an airport and try to pick  
 out anyone who is on a list of thousands of known suspects.

Early applications of face recognition were often just security theater. In

1998, the London borough of Newham placed video cameras prominently in  
 the high street and ran a PR campaign about how their new computer sys-  
 tem constantly scanned the faces in the crowd for several hundred known local  
 criminals. They got a signiﬁcant reduction in reported burglary, shoplifting and  
 street crime, but later admitted that they only had 20 or 25 villains’ faces on  
 the system, and it never recognised any of them [1282]. After 9/11, a num-  
 ber of places tried this. In Tampa, Florida, a similar system was abandoned  
 in 2003 after an ACLU freedom-of-information request discovered that it had  
 recognised no villains [1597]. Face recognition was also tried at Boston’s Lo-  
 gan airport; passengers passing through security screening were observed and  
 matched. The system was found to be impractical, with no useful balance be-  
 tween false matches and false alarms [316]. The Illinois Department of Motor  
 Vehicles adopted face recognition in 2003 to detect people applying for extra  
 drivers’ licenses in false names [663]. In such an application, it may be worth-  
 while to try to detect wrongdoers even if you only catch a quarter of them.

As a baseline, tests done in 2001 by the UK National Physical Laboratory

(NPL) of a number of biometric technologies found that face recognition was  
 almost the worst; its single-attempt equal-error rate was almost 10% [1217].  
 A UK Passport Office trial in 2005, that was a better approximation to ﬁeld  
 conditions, found it recognised only 69% of users (and only 48% of disabled par-  
 ticipants) [1920]. Face recognition was still adopted by the ICAO as a standard  
 for passports and ID cards with embedded chips; iris codes and ﬁngerprints were  
 optional extras. The typical installation has a row of booths relaying both live  
 and ﬁle photos to a human operator who is alerted to suspected mismatches.

|  |  |  |
| --- | --- | --- |
|  |  |  |

However, since the neural network revolution began in 2012, the performance

of facial recognition has improved remarkably, with error rates falling by an or-  
 der of magnitude. Getting through a passport booth is often a lot quicker now  
 than in 2010, and you don’t always have to take off your glasses. But what  
 about data? The best are probably from NIST’s Face Recognition Vendor Test  
 (FRVT) which tests products against millions of law-enforcement mugshots,  
 prison webcam images and wild photos for 1:1 veriﬁcation, one-to-many iden-  
 tiﬁcation, face morph detection and face image quality assessment. According  
 to the 2018 report, massive gains in accuracy were achieved in 2013–2018, and  
 largely due to the adoption of convolutional neural networks (CNNs). The most  
 accurate algorithms will ﬁnd matching entries when present, in galleries contain-  
 ing 12 million individuals, with a miss rate approaching 0.1%; but in about 5%  
 of images the identiﬁcation succeeds with low conﬁdence and human adjudica-  
 tion is necessary. A few algorithms correctly match side-view photos to galleries  
 of frontal photos; such *pose invariance* has been a long-sought milestone in face  
 recognition research.

There is measurable racial bias. U.S.-developed algorithms had signiﬁcantly

higher rates of false positives in one-to-one matching of Asians, African Amer-  
 icans and American Indians, while for one-to-many matching the highest false  
 positive rates were for African American females. Algorithms developed in Asia  
 did equally well for Asians and whites. The remaining errors are in large part  
 due to long-run ageing, facial injury, poor image quality or a second face in shot,  
 such as a face printed on a t-shirt [828].

A 2018 study pitted face recognition algorithms against professional forensic

face examiners, untrained superrecognisers (highly talented individuals), and a  
 control group of random people. It found that both types of human expert

were signiﬁcantly better than the control group, and that four deep CNNs,  
 developed between 2015 and 2017, identiﬁed faces within the range of human  
 experts, with the most recent scoring above the median of the forensic experts.  
 However, the best results could be achieved if algorithms and human experts  
 worked together [1522].

As for what’s under the hood, a 2019 survey paper by Guodong Guo and

Na Zhang explores the use of deep learning in face image analysis and recogni-  
 tion, and discusses how systems handle variations in pose, age, illumination and  
 expression [834]. Most systems are CNNs but with a range of adaptations, e.g.  
 with multiple CNNs looking for different types of feature in different regions  
 of two candidate faces simultaneously and an autoencoder looking for common  
 latent features to give pose robustness; there are then various kinds of fusion,  
 aggregation and ﬁltering. There may also be mechanisms to correct for makeup  
 and for facial expressions. There are complex trade-offs in algorithm choice,  
 with the best algorithm in ROC terms taking time linear in the gallery size,  
 meaning half a second to match against 10m other faces; accuracy can double  
 if three or more mugshots are available, as this enables the CNN to allow for  
 ageing. But blur in video images is still a signiﬁcant problem, as is matching  
 still images to video and visible-light images to near-infrared.

The face-recognition revolution is continuing apace, with NIST reporting

that some algorithms doubled in accuracy during 2018 alone. It is also becoming  
 controversial. Do we face a dystopian future where every other lamp post has

|  |  |  |
| --- | --- | --- |
|  |  |  |

|  |
| --- |
| a 5g base station with an embedded CCTV that recognises all passers-by? Allof a sudden, CCTV changes from a tool for crime-scene forensics to one thatdoes real-time person recognition and tracking. This appears to be the Chinesevision; ﬁrms there are training cameras not just to recognise individuals but alsogroups, with classiﬁers that alert if the subject appears to be an ethnic Uighuror Tibetan. This has been interrupted by mandatory face masks during thecoronavirus pandemic, but it will no doubt resume afterwards. Russia has beenusing its cameras to spot people breaking coronavirus quarantine orders, andclaims to have deployed 178,000 of them [1907]. Even in the West, do we face afuture in which the police get a feed not just from the automatic number-platerecognition systems that already track road vehicles, but a system that trackspedestrians too? Cynics would say that mobile phone location history alreadyworks ﬁne even if you’re wearing shades or a mask, so what’s the fuss about?But there are now companies with much larger collections of faces than lawenforcement, as they don’t face the legal restrictions, and whose services help lawenforcement solve crimes committed by people with no mugshots on ﬁle. Theseﬁrms appear set to o↵er services more widely; they could potentially enable usersof augmented-reality glasses to identify most of the people they see – whether anattractive stranger on a subway, or a protester at a demonstration. You couldﬁnd out their names, where they live and what they do online. The company’sbacker remarks “Laws have to determine whatˆa˘A´Zs legal, but you canˆa˘A´Zt bantechnology. Sure, that might lead to a dystopian future or something, but youcanˆa˘A´Zt ban it.” [897]. |

The political and legal pushback has started. A family in Evanston, Illinois

found that photos of their kids that they’d uploaded into Flickr in 2005 had  
 ended up in a database called MegaFace, used to train many of the new recogni-  
 tion systems. This is against Illinois law, and there are now several class actions  
 in progress. As a result, some face-tagging features on social media don’t work  
 in Illinois (or Texas for that matter) [898]. In 2018, Google decided not to make  
 face-recognition APIs available in its cloud platforms until their use was regu-  
 lated. If you train a system on criminal mugshots, it can look at any passer by  
 and say ‘This robber is the closest match’. Where the police are trigger-happy,  
 that can kill. In May 2019, San Francisco banned the use of face recognition  
 by its agencies including the transport authority and law enforcement. In June  
 2020, following worldwide protests over racism and biased policing, Amazon an-  
 nounced a one-year pause in making its Rekognition face-recognition software  
 available to law enforcement; their technology had been criticised for misidenti-  
 fying people of colour. The ACLU had shown that Amazon’s system generated  
 false matches of 28 members of Congress against mugshots of people who had  
 been arrested. IBM and Microsoft also announced that they would stop selling  
 face-recognition products [2004]. As the technology is now a commodity, the  
 self-restraint of the big four doesn’t stop second-tier ﬁrms selling it. So the big  
 four are now pushing for face-recognition products to be regulated. Courts are  
 already engaged: in August 2020, the Court of Appeal in London found that  
 the use of facial recognition by South Wales police breached privacy rights, data  
 protection laws and equality laws [1592].

Finally, facial recognition can be enhanced with special hardware. In 2017,

Apple introduced it on the iPhone X, in which a dot projector paints your  
 face with tens of thousands of dots and a camera reads them. This deals with

|  |  |  |
| --- | --- | --- |
|  |  |  |

makeup, some sunglasses and facial hair, and was claimed to have a false accep-  
 tance rate of one in a million – as opposed to one in 50,000 for the ﬁngerprint  
 reader that previous iPhones used. However my eldest granddaughter’s iPhone  
 can be unlocked by both of her younger siblings, and this is a general problem  
 for families [526].

**17.4** **Fingerprints**

Automatic ﬁngerprint identiﬁcation systems (AFIS) have been around for years.  
 In 1998, they accounted for 78% of the $50m sales of biometric technology; this  
 had fallen to 43.5% of $1,539m by 20052. AFIS products look at the friction  
 ridges that cover the ﬁngertips and classify patterns of *minutiae* such as branches  
 and end points of the ridges. Some also look at the pores in the skin of the  
 ridges [1213].

The use of ﬁngerprints to identify people was discovered independently a

number of times. Mark Twain mentions thumbprints in 1883 in *Life on the*  
 *Mississippi* where he claims to have learned about them from an old Frenchman  
 who had been a prison-keeper; his 1894 novel *Pudd’nhead Wilson* made the idea  
 popular in the States. Long before that, ﬁngerprints were accepted in a seventh  
 century Chinese legal code as an alternative to a seal or a signature, and required  
 by an eighth century Japanese code when an illiterate man wished to divorce  
 his wife. They were also used in India centuries ago. Following the invention of  
 the microscope, they were mentioned by the English botanist Nathaniel Grew  
 in 1684, and by Marcello Malpighi in Italy in 1686; in 1691, 225 citizens of  
 Londonderry in Ireland used their ﬁngerprints to sign a petition asking King  
 William for reparations following the siege of the city.

The ﬁrst modern systematic use was in India from 1858, by William Her-

schel, grandson of the astronomer and a colonial magistrate. He introduced

handprints and then ﬁngerprints to sign contracts, stop impersonation of pen-  
 sioners who had died, and prevent rich criminals paying poor people to serve  
 their jail sentences for them. Henry Faulds, a medical missionary in Japan,

discovered them independently in the 1870s, and came up with the idea of using  
 latent prints from crime scenes to identify criminals. Faulds brought ﬁngerprints  
 to the attention of Charles Darwin, who in turn motivated Francis Galton to  
 study them. Galton wrote an article in *Nature* [738]; this got him in touch with  
 the retired Herschel, whose data convinced Galton that ﬁngerprints persisted  
 throughout a person’s life. Galton went on to collect many more prints and  
 devise a scheme for classifying their patterns [739]. The Indian history is told  
 by Chandak Sengoopta, whose book also makes the point that ﬁngerprinting  
 saved two somewhat questionable Imperial institutions, namely the indentured  
 labor system and the opium trade [1701].

The practical introduction of the technique owes a lot to Sir Edward Henry,

who had been a policeman in Bengal. He wrote a book in 1900 describing a  
 simpler and more robust classiﬁcation, of *loops*, *whorls*, *arches* and *tents*, that  
 he had developed with his assistants Azizul Haque and Hem Chandra Bose,

2I don’t have comparable ﬁgures for 2019 as ﬁngerprint tech is now bundled with phones

or with other biometrics in systems such as Aadhaar.

|  |  |  |
| --- | --- | --- |
|  |  |  |

and that is still in use today. In the same year he became Commissioner of  
 the Metropolitan Police in London from where the technique spread round the  
 world3. Henry’s real scientiﬁc contribution was to develop Galton’s classiﬁcation  
 into an indexing system. By assigning one bit to whether or not each of a

suspect’s ten ﬁngers had a whorl – a type of circular pattern – he divided the  
 ﬁngerprint ﬁles into 1024 bins. In this way, it was possible to reduce the number  
 of records that had to be searched by orders of magnitude. Meanwhile, as Britain  
 had stopped sending convicted felons to Australia, there was a perceived need  
 to identify previous offenders, so that they could be given longer jail sentences.

Fingerprints are used by the world’s police forces for essentially two different

purposes: identifying people (their main use in the USA), and crime scene  
 forensics (their main use in Europe).

**17.4.1** **Verifying positive or negative identity claims**

In America nowadays – as in nineteenth-century England – quite a few criminals  
 change their names and move somewhere new on release from prison. This is ﬁne  
 when offenders go straight, but what about fugitives and recidivists? American  
 police forces have historically used ﬁngerprints to identify arrested suspects to  
 determine whether they’re currently wanted by other agencies, whether they  
 have criminal records and whether they’ve previously come to attention under  
 other names. The FBI maintains the *Next Generation Identiﬁcation* (NGI)

service system for this purpose; it identiﬁes about eight thousand fugitives a  
 month [1809]. Anyone wanting a US government clearance at Secret or above  
 must have an FBI ﬁngerprint check, and checks are also run on some people  
 applying to work with children or the elderly. Up to 100,000 checks are made a  
 day, and about a million federal, local and state officers have access. There’s a  
 ‘rap-back’ service to alert the employer of anyone with a clearance who gets into  
 trouble with the law [1378]; it’s also used to track reoffending by probationers,  
 parolees and sex offenders. The Department of Homeland Security’s IDENT  
 system holds ﬁngerprints on 200 million aliens who have arrived at US ports; it  
 matches them against a watch list of bad guys, compiled with the help of police  
 forces and intelligence services worldwide.

These are examples of one type of identity veriﬁcation – checking against a

blacklist. The other type is where the system checks a claim to identity, with the  
 main US applications being building entry control and welfare payment [588].  
 Banks have used them for years to identify customers in countries such as In-  
 dia and Saudi Arabia, where the use of ink ﬁngerprints was already common  
 thanks to high levels of illiteracy. India now has a national system, Aadhaar,  
 with ﬁngerprints and iris codes of most residents, designed initially to support  
 welfare payments and ensure that nobody can claim twice. Its use has become  
 mandatory for many other transactions too.

3In the Spanish version of history, they were ﬁrst used in Argentina where they secured

a murder conviction in 1892; while Cuba, which set up its ﬁngerprint bureau in 1907, beat  
 the USA whose ﬁrst conviction was in Illinois in 1911. The Croation version notes that the  
 Argentinian system was developed by one Juan Vucetich, who had emigrated from Dalmatia.  
 The German version refers to Professor Purkinje of Breslau, who wrote about ﬁngerprints in  
 1828. Success truly has many fathers!

|  |  |  |
| --- | --- | --- |
|  |  |  |

Fingerprints have never taken off for authenticating bank customers in North

America or Europe, though a few US banks do ask for ﬁngerprints if you cash  
 a check there and are not a customer. They ﬁnd this cuts check fraud by about  
 a half. Some have gone as far as ﬁngerprinting new customers, and found that  
 customer resistance is less than expected, especially if they use scanners rather  
 than ink and paper [716]. These applications are not authentication, but rather  
 an attempt to identify and even deter customers who later turn out to be bad  
 – another example being the large British van-hire company that demands a  
 thumbprint when you rent a van. If the vehicle isn’t returned, or if it’s used in a  
 crime, the thumbprint is given to the police. They’re thus really a crime-scene  
 forensics application, which I’ll discuss in the following section.

So how good are automatic ﬁngerprint identiﬁcation systems? A good rule

of thumb (if one might call it that) is that to verify a claim to identity, it may  
 be enough to scan a single ﬁnger, while to check someone against a blacklist of  
 millions of felons, you had better scan all ten. After the US DHS program set  
 out to scan the two index ﬁngers of each arriving visitor, it was overwhelmed  
 by false matches. With 6,000,000 bad guys on the database, the false match  
 rate in 2004 was 0.31% and the missed match rate 4% [2027]. The program  
 moved to ‘10-prints’, where each visitor must present the four ﬁngers of each  
 hand, and then both thumbs, in three successive scans. The European Union  
 will be adopting a combination of 4-prints plus facial recognition from 2020;  
 nonresidents will need both to get in, and either to get out.

This is all about the trade-off between false negatives and false positives – the

receiver operating characteristic, described in the previous section. The better  
 systems have an equal error rate of slightly below 1% per ﬁnger. False accepts  
 happen because of features incorporated to reduce the false reject rate – such  
 as allowance for distortion and ﬂexibility in feature selection [1610]. Spotting  
 returning fugitives with high enough probability to deter them and high enough  
 certainty to detain them (which means keeping false alarms at manageable lev-  
 els) requires several ﬁngers to be matched – perhaps eight out of ten. This does  
 cause delays; a UK Passport Office study found that about 20% of participants  
 failed to register properly when taking a 10-print, and that 10-print veriﬁcation  
 took over a minute [1920]. This is approximately my experience while ﬂying  
 in and out of the USA during the 2010s. The cost of ﬁngerprinting everybody  
 is that a US airport getting a planeload of 300 international arrivals every 15  
 minutes needs an extra 10 working immigration lanes. The extra building and  
 staffing costs swamp anything spent on hardware and software. (For more on  
 algorithms and systems, see [973, 1211, 1213].)

Errors are not uniformly distributed. A number of people such as man-

ual workers and pipe smokers damage their ﬁngerprints frequently, and both  
 the young and the old have faint prints [392]. Automated systems also have  
 problems with amputees, people with birth defects such as extra ﬁngers, and  
 the (rare) people born without conventional ﬁngerprint patterns at all [1120].  
 When I was a kid, I slashed my left middle ﬁnger while cutting an apple, and  
 this left a scar about half an inch long. When I presented this ﬁnger to the  
 system used in 1989 by the FBI for building entry control, my scar crashed the  
 scanner. (It worked OK when I tried again ten years later.)

Fingerprint identiﬁcation systems can be attacked in many ways. An old

|  |  |  |
| --- | --- | --- |
|  |  |  |

trick was for a crook to distract (or bribe) the officer ﬁngerprinting him, so that  
 the officer takes the ﬁngers in the wrong order and instead of the hand being  
 indexed under the Henry system as ‘01101’ it becomes perhaps ‘01011’, so his  
 record isn’t found and he gets the lighter sentence due a ﬁrst offender [1120].

The ﬁrst high-proﬁle technical attack was in 2002, when Tsutomu Mat-

sumoto and colleagues showed that ﬁngerprints could be molded and cloned  
 quickly and cheaply using cooking gelatin [1246]. He tested eleven commer-

cially available ﬁngerprint readers and easily fooled all of them. This prompted  
 the German computer magazine C’T to test a number of biometric devices that  
 were offered for sale at the CeBIT electronic fair in Hanover – nine ﬁngerprint  
 readers, one face-recognition system and one iris scanner. They were all easy to  
 fool – the low-cost capacitive sensors fell to such simple tricks as breathing on  
 a ﬁnger scanner to reactivate a latent print left by a previous user [1877]. La-  
 tent ﬁngerprints can also be reactivated – or transferred – using adhesive tape.  
 The more expensive thermal scanners could still be defeated by rubber molded  
 ﬁngers.

In 2013, Apple introduced a ﬁngerprint scanner on the iPhone 5S and other

phone makers raced to follow suit. Hackers duly demonstrated attacks, with  
 a 2014 CCC presentation of a model of the German defence minister’s ﬁnger,  
 created from a photograph [313]. Scanners on phones typically store 8–12 partial  
 prints on registration and will unlock against any of them, which makes the  
 scanner more usable but also more vulnerable. In 2016, Aditi Roy and colleagues  
 invented the ‘masterprint’: a fake ﬁngerprint that can be worn on your ﬁngertip  
 and that’s designed to match at least one of the partial prints derived from a  
 typical ﬁnger; it works against 6% of users’ prints [1625]. In 2017, Apple moved  
 from ﬁngerprints to face recognition, as I discussed above, but most Android  
 OEMs still use ﬁngerprints. In 2019, it turned out that a new ultrasonic scanner  
 on the Samsung S10 enrolled the screen protector instead of the customer’s  
 ﬁnger, leading to the phone being blocked from running a number of banks’  
 apps [466].

There are other angles too. For example, the San Bernardino shooter used

an iPhone 5C, the last made without a scanner; if he’d used a later version, the  
 FBI could have unlocked it by taking it to the morgue and pressing it against  
 his ﬁnger, or by making a ﬁngertip mould from his ﬁle print. And as govern-  
 ment agencies collect more and more prints, they will be less and less private.  
 (The Chinese already got all U.S. federal employees’ prints via the OPM hack  
 I discussed in section 2.2.2.) Fingerprint systems have also expanded rapidly  
 into low-assurance applications, from entry into golf club car parks to auto-  
 matic book borrowing in school libraries. (Most European countries’ privacy  
 authorities have banned ﬁngerprint scanners in schools; Britain allows them,  
 which causes pushback from privacy-conscious parents [190].) And the latest  
 twist comes from a Mitre project that developed software to harvest people’s  
 ﬁngerprints from photos they post on social media; these often show ﬁngers in  
 enough detail to get matches against FBI databases [321].

One ﬁnal reason for the success of ﬁngerprint identiﬁcation systems is their

deterrent effect, which is particularly pronounced in welfare payments. Even  
 though the cheap ﬁngerprint readers used to authenticate welfare claimants have  
 an error rate as much as 5% [383], they turned out to be such an effective way

|  |  |  |
| --- | --- | --- |
|  |  |  |

of reducing the welfare rolls that they were adopted in one place after another  
 during the nineties [1315].

**17.4.2** **Crime scene forensics**

The second use of ﬁngerprint recognition is in crime scene forensics – the main  
 application in Europe. Prints found at a crime scene are matched against

database records, and any that match to more than a certain level are taken  
 as evidence that a suspect visited the crime scene. They are often enough to  
 secure a conviction on their own. In many countries, ﬁngerprints are required  
 from all citizens and all resident foreigners.

The forensic error rate has become extremely controversial in recent years,

the critical limitation being the size and quality of the image taken from the  
 crime scene. The quality and procedure rules vary from one country to another.  
 The UK used to require that ﬁngerprints match in sixteen *points* (corresponding  
 minutiae), and a UK police expert claimed that this will only happen by chance  
 somewhere between one in four billion and one in ten billion matches [1120].  
 Greece accepts 10, Turkey 8, while the USA has no set limit (it certiﬁes exam-  
 iners instead). This means that in the USA, matches can be found with poorer  
 quality prints but they can be open to challenge in court.

In the UK, ﬁngerprint evidence went for almost a century without a success-

ful challenge; a 16-point ﬁngerprint match was considered hanging evidence.  
 The courts’ conﬁdence was shattered by the McKie case [1273]. Shirley McKie,  
 a Scottish policewoman, was prosecuted on the basis of a ﬁngerprint match on  
 the required sixteen points, veriﬁed by four examiners of the Scottish Crimi-  
 nal Records Office. She denied that it was her ﬁngerprint, and found that she  
 could not get an independent expert in Britain to support her; the profession  
 closed ranks. She called two American examiners who presented testimony that  
 it is not an identiﬁcation. The crime scene and ﬁle prints are side-by-side at  
 Figure 17.1.

|  |  |
| --- | --- |
|  |  |
| (a) Crime scene print | (b) Inked print |

Figure 17.1: The prints in the McKie case

|  |  |  |
| --- | --- | --- |
|  |  |  |

She was acquitted, which led to a political drama that ran on for years [1272].

The ﬁrst problem was the nature of the case against her [1273]. A number of  
 senior police officers had tried to persuade her to make a false statement in order  
 to explain the presence, at the scene of a gruesome murder, of the misidentiﬁed  
 print. Her refusal to do so led to her being prosecuted for perjury, as a means of  
 discrediting her. Her acquittal cast doubt on the reliability of police testimony,  
 not just in her speciﬁc case but more generally. The man convicted of the murder  
 was acquitted on appeal and sued the police for compensation. The government  
 panicked at the prospect of dozens more appeals in other cases, and prosecuted  
 its four ﬁngerprint experts for perjury. That didn’t get anywhere either. The  
 issue went back to the Scottish parliament again and again. The police refused  
 to reinstate Shirley McKie, the officers involved got promoted, and the row got  
 ever more acrimonious. Eventually she won £750,000 compensation from the  
 government [189].

The case led to wide discussion among experts of the value of ﬁngerprint iden-

tiﬁcation, and to ﬁngerprint evidence being successfully challenged in a number  
 of other countries [760]. Two high-proﬁle cases in the USA were Stephan Cow-  
 ans and Brandon Mayﬁeld. Stephen Cowans had been convicted of shooting  
 a police officer in 1997 following a robbery, but was acquitted on appeal six  
 years later after he argued that his print was a misidentiﬁcation and saved up  
 enough money to have the evidence tested for DNA. The DNA didn’t match,  
 which got the Boston and State police to reanalyze the ﬁngerprint, whereupon  
 they realised it was not a match after all. Brandon Mayﬁeld was an Oregon  
 lawyer who was mistakenly identiﬁed by the FBI as one of the perpetrators of  
 the Madrid bombing, and held for two weeks until the Madrid police arrested  
 another man whose ﬁngerprint was a better match. The FBI, which had called  
 their match ‘absolutely incontrovertible’, agreed to pay Mayﬁeld $2m in 2006.

In a subsequent study, psychologist Itiel Dror showed ﬁve ﬁngerprint exam-

iners a pair of prints, told them they were from the Mayﬁeld case, and asked  
 them where the FBI had gone wrong. Three of the examiners decided that the  
 prints did not match and pointed out why; one was unsure; and one maintained  
 that they did match. He alone was right. The prints weren’t the Mayﬁeld set,  
 but were in each case a pair that the examiner himself had matched in a recent  
 criminal case [586]. Dror repeated this with six experts who each looked at eight  
 prints, all of which they had examined for real in the previous few years. Only  
 two of the experts remained consistent; the other four made six inconsistent  
 decisions between them. The prints had a range of difficulty, and in only half  
 of the cases was misleading contextual information supplied [587].

Prosecutors and police still insist to juries that forensic results are error-

free, when FBI proﬁciency exams have long had an error rate of about one  
 percent [205], and misleading contextual information can push this up to ten  
 percent or in some cases over ﬁfty percent.

Four comments are in order.

*•* As Figure 17.1 should make clear, ﬁngerprint impressions are often very  
 (and prejudices) of the examiner enter into the equation in a much bigger  
 way than was accepted until the McKie case, the Mayﬁeld case, and the

|  |  |  |
| --- | --- | --- |
|  |  |  |

general uproar that they have caused. Dror’s work conﬁrmed that the  
 cases in which misidentiﬁcations occur tend to be the difficult ones [587].  
 Yet the forensic culture was such that only certainty was acceptable; the  
 International Association for Identiﬁcation, the largest forensic group, held  
 that testifying about “possible, probable or likely identiﬁcation shall be  
 deemed ... conduct unbecoming.” [205]

*•* Even if the probability of a false match on sixteen points were one in ten  
 pared against each other, probability theory starts to bite. A system that  
 worked ﬁne in the old days as a crime scene print would be compared man-  
 ually with the records of a hundred and ﬁfty-seven known local burglars,  
 breaks down once thousands of prints are compared every year with an  
 online database of millions. It was inevitable that sooner or later, enough  
 matches would have been done to ﬁnd a 16-point mismatch. Indeed, as  
 most people on the ﬁngerprint database are petty criminals who will not  
 be able to muster the resolute defence that Shirley McKie did, I would  
 be surprised if there hadn’t been other wrongful convictions already. And  
 things may get worse, because European police forces now link up their  
 biometric databases (both ﬁngerprints and DNA) so that police forces  
 can search for matches across all EU member states [1905]. They may  
 eventually need more robust ways of handling false positives.

*•* The belief that any security mechanism is infallible creates the compla-  
 eration appears to have been given to increasing the number of points  
 required from sixteen to (say) twenty with the introduction of computer  
 matching. Sixteen was tradition, and nobody wanted either to challenge  
 the system or make public funds available for defendants’ experts. In the  
 UK, all the experts were policemen or former policemen, so there were  
 no independents available for hire anyway. Even so, it would have been  
 possible to use randomised matching with multiple experts; but if the ﬁn-  
 gerprint bureau had had to tell the defence in the perhaps 5–10% of cases  
 when (say) one of four experts disagreed, then more defendants would  
 have been acquitted.

*•* A belief of infallibility ensures that the consequences of the eventual failure  
 helped torpedo claims about cash machine security, an assumption that  
 a security mechanism is infallible causes procedures, cultural assumptions  
 and even laws to spring up to ensure that its eventual failure will be  
 denied for as long as possible, and will thus have real impact when it can  
 no longer be postponed. In the Scottish case, there appears to have arisen  
 a hierarchical risk-averse culture in which examiners were predisposed to  
 conﬁrm identiﬁcations made by colleagues (especially senior colleagues).  
 This risk aversion backﬁred when four of them were tried for perjury.

However, even when we do have a correct match its implications are not

always entirely obvious. Fingerprints can be transferred using adhesive tape,  
 and moulds can be made, using techniques originally devised for police use.  
 So it’s possible that the suspect whose print is found at the crime scene was

|  |  |  |
| --- | --- | --- |
|  |  |  |

framed by another criminal (or by the police – most fabrication cases involve  
 law-enforcement personnel rather than other suspects [254]). And even if the  
 villain wasn’t framed, he can always claim that he was (and the jury might  
 believe him).

In the USA, the Supreme Court in its Daubert judgment held that trial

judges should screen the principles and methodology behind forensic evidence to  
 ensure it is relevant and reliable [516]. The judge ought to consider the refereed  
 scientiﬁc literature – and in the case of ﬁngerprints this has been lacking, as law  
 enforcement agencies have been generally unwilling to submit their examination  
 procedures to rigorous double-blind testing. A number of Daubert hearings

relating to forensic ﬁngerprint evidence have been held in US trials, and the  
 FBI has generally prevailed [761]. However, the bureau’s traditional line that  
 ﬁngerprint examination has a zero error rate is now widely ridiculed [1809].

**17.5** **Iris codes**

We turn now from the traditional ways of identifying people to the modern and  
 innovative. Recognizing people by the patterns in the irises of their eyes has far  
 and away the best error rates of any automated biometric system when mea-  
 sured under lab conditions. The initial research was funded by the Department  
 of Energy, which wanted the best possible way of securing entry to premises  
 such as plutonium stores, and the technology is now used in applications from  
 immigration to welfare. The international standards for machine-readable travel  
 documents mandate the use of photographs, and permit both ﬁngerprints and  
 irises.

So far as is known, every human iris is measurably unique. It is fairly easy to

detect in a video picture, it does not wear out, and it is isolated from the external  
 environment by the cornea (which in turn has its own cleaning mechanism). The  
 iris pattern contains a large amount of randomness, and appears to have many  
 times the number of degrees of freedom of a ﬁngerprint. It is formed between the  
 third and eighth month of gestation, and (like the ﬁngerprint pattern) appears  
 to be under limited genetic inﬂuence; the mechanisms that form it appear to  
 be chaotic. The patterns are different even for identical twins (and for the two  
 eyes of a single individual), and they appear to be stable throughout life.

Leonard Flom and Aran Saﬁr patented the idea of an iris identiﬁcation sys-

tem in 1987, observing that every iris is different. In 1993, John Daugman

ﬁgured out how to make the idea work, developing signal-processing techniques  
 that extract the information from an image of the iris into a 256 byte *iris*  
 *code*. This involves a circular wavelet transform taken at a number of con-

centric rings between the pupil and the outside of the iris (Figure 17.2). The  
 resulting iris codes have the neat property that two codes computed from the  
 same iris will typically match in 90% of their bits [517]. This is much simpler  
 than in ﬁngerprint scanners where orienting and classifying the minutiae is a ﬁd-  
 dly computational task. The speed and accuracy of iris coding, and the expiry  
 of the Daugman patents, have led to a number of commercial iris recognition  
 products [1996]. Iris codes provide the lowest false accept rates of any known  
 veriﬁcation system – zero, in tests conducted by both the US Department of

|  |  |  |
| --- | --- | --- |
|  |  |  |

Energy and the NPL [1217]. The equal error rate has been shown to be better  
 than one in a million, and if one is prepared to tolerate a false reject rate of one  
 in ten thousand then the theoretical false accept rate would be less than one in  
 a trillion. In practice, the false reject rate is signiﬁcantly higher than this; many  
 things, from eyelashes to hangovers, can cause the camera to not see enough of  
 the iris. The US Department of Defense found a 6% false reject rate in its 2002  
 ﬁeld trials [1258]; a UK Passport Office trial found 4% for normal users and  
 9% for disabled users [1920]. A further problem is failure to enrol; the Passport  
 Office trial failed to enrol 10% of participants, and the rate was higher among  
 black users, the over-60s and the disabled.

Figure 17.2: – an iris with iris code (courtesy John Daugman)

One practical problem with iris scanning used to be getting the picture

cheaply without being too intrusive. The iris is small (less than half an inch)  
 and an image with several hundred pixels of iris is needed. A cooperative subject  
 can place his eye within a few inches of a video camera, and the best standard  
 equipment will work up to a distance of two or three feet. All current iris

scanning systems use infrared light, and some people feel uncomfortable when  
 this is shone in their eyes. Given more sophisticated cameras, with automatic  
 facial feature recognition, pan and zoom, it is now possible to capture iris codes  
 from airline passengers covertly as they walk along a corridor [1240], and the  
 cost came down after the key patent ran out in 2011.

The ﬁrst large-scale deployment was in the United Arab Emirates, which

wanted to track people expelled from the country, particularly for prostitution  
 offences. Expellees would turn up again some weeks later with new and com-  
 pletely valid passports from certain Asian countries, obtained by corruption.  
 Since its deployment in 2003, this has led to the detention of over 330,000 peo-

|  |  |  |
| --- | --- | --- |
|  |  |  |

ple attempting to enter the country despite a ban or with false papers.

The largest deployment is the Aadhaar system in India, under which all

residents had their ﬁngerprints and irises scanned. They get an Aadhaar card  
 with a 10-digit number that enables a veriﬁer to look up their proﬁle in a  
 database. The initial motivation for the project was to enable the 300 million  
 Indians who live below the poverty line and get welfare, to move into the cities  
 to seek work. Previously welfare was only available in their towns or villages.  
 The system enrolled a billion people between 2011 and 2016, and all iris codes  
 were checked against each other for uniqueness. Aadhaar is now mandatory for  
 many purposes, and the collected ﬁngerprints are also made available to the  
 police for crime scene forensics.

Possible attacks on iris recognition systems include – in unattended operation

at least – a simple photograph of the target’s iris. There are terminals that will  
 detect such simple fakes, for example by measuring *hippus* – a natural ﬂuctuation  
 in the diameter of the pupil that happens at about 0.5 Hz. But the widely-sold  
 cheap terminals don’t do this, and if liveness detection became widespread then  
 no doubt attackers would try more sophisticated tricks, such as printing the  
 target’s iris patterns on a contact lens.

The system in active use the longest is the UAE’s system for detecting de-

portees who return with false papers. A typical attack was for the returning  
 deportee to take atropine eyedrops on the plane, dilating her pupils; nowadays  
 such travelers are held in custody until their eyes return to normal. As for

Aadhaar, the main abuses and disputes happen around the system rather than  
 through it. In 2019, a hot issue is the authorities’ reluctance to register Mus-  
 lims in Assam and other border regions, part of a larger policy of trying to  
 portray them as illegal immigrants. The Supreme Court of India has ruled that  
 services should not be withheld from people who are not registered, but this  
 has not stopped registration being a requirement in practice for opening a bank  
 account, buying a phone or SIM card, and school enrolment.

Despite the difficulties, iris codes are in some sense the most powerful bio-

metric as they can, in the correct circumstances, assure you that the individual  
 in front of you is the same human as the one whose iris was initially regis-  
 tered. They alone can meet the goal of automatic recognition with zero false  
 acceptances.

**17.6** **Voice recognition and morphing**

*Voice recognition* – also known as *speaker recognition* – is the problem of iden-  
 tifying a speaker from a short utterance. While *speech recognition* systems are  
 concerned with transcribing speech and need to ignore speech idiosyncrasies,  
 voice recognition systems need to amplify and classify them. There are many  
 subproblems, such as whether the recognition is text-dependent or not, whether  
 the environment is noisy, whether operation must be real time and whether one  
 needs only to verify speakers or to recognize them from a large set.

As with ﬁngerprints, the technology is used for both identiﬁcation and foren-

sics. In *forensic phonology*, the task is usually to match a recorded telephone

|  |  |  |
| --- | --- | --- |
|  |  |  |

conversation, such as a bomb threat, to speech samples from a number of sus-  
 pects. Typical techniques involve ﬁltering and extracting features from the

spectrum; for more details see [1058]. A more straightforward biometric au-  
 thentication objective is to verify a claim to identity in some telephone systems.  
 These range from telephone banking to the identiﬁcation of military personnel,  
 with the NSA maintaining a standard corpus of test data for evaluating speaker  
 recognition systems. In the UK, asylum seekers are required to ring in several  
 times every week [1902]. Such systems tend to use caller-ID to establish where  
 people are, and are also used for people like football hooligans who’re under  
 court orders not to go to certain places at certain times. The only system I’ve  
 used personally is run by one of the banks I use, and authenticates you to their  
 phone app when you change your phone. But a major UK bank was embar-  
 rassed when it ﬁelded a voice biometric system in a phone app in 2016, only to  
 have it broken the following year by a BBC reporter who got his non-identical  
 twin to mimic his voice [1744].

Quite apart from the possibility that a relative or a villain might some-

how manage to imitate you, there are some powerful attacks. In [730] there  
 is a description of a 1990s system ﬁelded in US EP-3 aircraft that breaks up  
 intercepted messages from enemy aircraft and ground controllers into quarter  
 second segments that are then cut and pasted to provide new, deceptive mes-  
 sages. That was primitive compared with what can now be done two decades  
 later. There are now many videos online of public ﬁgures appearing to say in-  
 appropriate things, and ‘Deepfake’ editing software now enables such voice and  
 image morphing to be done in near real time. Most recently, criminals used AI  
 to impersonate a chief executive’s voice and order a payment of e220,000: the  
 victim of that deception wasn’t even a machine, but another executive [1841].  
 This may be the ﬁrst case of voice morphing software being used in a real fraud;  
 we can be sure it won’t be the last.

**17.7** **Other systems**

Many other biometric technologies have been proposed [1315]. Typing patterns,  
 were used in products in the 1980s but don’t appear to have been successful  
 (typing patterns, also known as keystroke dynamics, had a famous precursor  
 in the wartime technique of identifying wireless telegraphy operators by their  
 *ﬁst*, the way in which they used a Morse key). Vein patterns have been used in  
 one or two systems but don’t seem to have been widely sold (in the NPL trials,  
 the vein was the worst of the lot [1217]). Hand geometry was used for a while  
 in some airports, and has a historic predecessor in the system of Bertillonage,  
 whereby the French police in the 19th century identiﬁed criminals by a system  
 of physical measurements.

There has been growing interest recently in *stylometry*, the science of iden-

tifying authors, whether of text or of code, from their writing styles. This goes  
 back at least a century; as a young man, the famous cryptologist William Fried-  
 man was hired along with his wife Elizebeth by an eccentric millionaire to study  
 whether Bacon wrote Shakespeare. (They eventually debunked the idea but got  
 interested in cryptography in the process.) Computers make it possible to run

|  |  |  |
| --- | --- | --- |
|  |  |  |

ever more subtle statistical tests, and modern applications range from trying  
 to identify people who post to cybercrime markets and extremist web forums  
 to the detection of plagiarism by college students [3]. Researchers have shown  
 that people can change their writing styles enough to defeat simple stylometry  
 if they try [318]. But most people don’t, and with a bit more work, the fact of  
 an attempted obfuscation can usually be detected [28]. Stylometry also extends  
 to code; programmers can be recognised from their coding style [370].

Other proposals include *facial thermograms* (maps of the surface tempera-

ture of the face, derived from infrared images), the shape of the ear, gait, lip  
 prints and electrocardiograms. Bertillon used the shape of the ear in nineteenth  
 century Paris. And perhaps the huge investment in developing digital noses for  
 quality control in the food and drink industries may lead to personal devices  
 that recognize their master by scent.

One ﬁnal biometric deserves mention – DNA. This has become a valuable

tool for crime scene forensics and for determining parenthood in child support  
 cases, but it is way too slow and expensive for real-time applications. Being  
 genotypic rather than phenotypic, its accuracy is limited by the incidence of  
 monozygotic twins: about one white person in 120 has an identical twin. There’s  
 also a privacy problem in that it is possible to reconstruct a growing amount of  
 information about an individual from their DNA sample. There have been major  
 procedural problems, with false matches resulting from sloppy lab procedure.  
 And there are also major data quality problems; the UK police have the biggest  
 DNA database in the world, with records on almost six million people, but  
 got the names misspelled or even wrong for about half a million of them [878].  
 They also had court judgments against them for retaining the DNA of innocent  
 people, from acquitted suspects to bystanders [102]. The processes that work  
 for local policing don’t always scale nationally – small errors from mistyped  
 records, to suspects giving false names that were never discovered because they  
 weren’t prosecuted, accumulate along with lab errors until the false-positive rate  
 becomes a serious operational and political issue. In this context, many were  
 concerned when in 2019, a Florida detective managed to get a warrant to search  
 all million records held by a private DNA testing company GEDmatch [899].  
 It will be interesting to see whether this undermines the business of the larger  
 consumer DNA ﬁrms, such as 23andMe and ancestry.com, enough for them to  
 lobby for stronger privacy laws.

**17.8** **What Goes Wrong**

As with other aspects of security, we ﬁnd the usual crop of failures due to bugs,  
 blunders and complacency. In section 3.4.9 I noted a report that the ﬁrm which  
 supplies biometric building entry control systems to 5,700 organisations in 83  
 countries left its database unprotected online. And the second time Uber lost  
 its London operating licence, it was because they failed to stop banned drivers  
 re-registering, thanks to a photo checking bug [310]. And the main problem  
 faced by DNA typing was an initially high rate of false positives, due to careless  
 laboratory procedure. This led to disputed court cases and miscarriages of

justice. As with ﬁngerprints, any system that’s believed to be infallible will

|  |  |  |
| --- | --- | --- |
|  |  |  |

make its operators careless enough to break it.

Biometrics are also like many other physical protection mechanisms (alarms,

seals, tamper sensing enclosures, ...) in that environmental conditions can cause  
 havoc. Noise, dirt, vibration and unreliable lighting conditions all take their  
 toll. Some systems, like speaker recognition, are vulnerable to alcohol intake  
 and stress. Changes in environmental assumptions, such as from closed to open  
 systems, from small systems to large ones, from attended to stand-alone, from  
 cooperative to recalcitrant subjects, and from veriﬁcation to identiﬁcation, can  
 all break things.

Many interesting attacks are more speciﬁc to biometric systems and apply

to more than one type of biometric.

*•* Forensic biometrics often don’t tell as much as one might assume. Apart  
 planted by the police, it may just be old. The age of a ﬁngerprint can’t  
 be determined directly, and prints on areas with public access say little.  
 A print on a bank door says much less than a print in a robbed vault.  
 So in premises vulnerable to robbery, cleaning procedures may be critical  
 for evidence. If a suspect’s prints are found on a bank counter, and he  
 claims that he had gone there three days previously, he may be convicted  
 by evidence that the branch counter is polished every evening. Putting  
 this in system terms, freshness is often a critical issue, and some quite  
 unexpected things can ﬁnd themselves inside the ‘trusted computing base’.

*•* Another aspect of freshness is that most biometric systems can, at least  
 tacks on voice recognition, attacks on iris scanners by photos on a contact  
 lens, and moulds of ﬁngerprints. Even simpler still, in countries like South  
 Africa where ﬁngerprints are used to pay pensions, there are persistent  
 tales of ‘Granny’s ﬁnger in the pickle jar’ being the most valuable prop-  
 erty she bequeathed to her family. The lesson to be learned here is that  
 unattended operation of biometric authentication devices is tricky. At-  
 tacks aren’t always straightforward; although it’s easy to make a mould  
 from a good ﬁngerprint [406], the casual prints that people leave lying  
 around on doorknobs, beer glasses and so on are often too smudged and  
 fragmentary to pass an identiﬁcation system. But attacks are deﬁnitely  
 possible, and deﬁnitely happen. Defences are also possible; voice recog-  
 nition systems can demand that you read out an unpredictable challenge  
 to thwart recordings, while one version of the app that EU citizens use  
 to apply for residence in the UK post-Brexit took a video of your face as  
 colours change on the phone screen in front of you.

*•* Most biometrics are not as accurate for all people, and some of the pop-  
 elderly, and manual workers, often have damaged or abraded ﬁngerprints;  
 there’s a tradition of hardcore criminals doing this deliberately. People  
 with dark eyes, and large pupils, give poorer iris codes. Disabled people  
 with no ﬁngers, or no eyes, risk exclusion. (That’s one reason Aadhaar  
 uses both irises and ﬁngerprints.) Illiterates who make an ‘*X*’ are more at  
 risk from signature forgery.

|  |  |  |
| --- | --- | --- |
|  |  |  |

Biometric engineers sometimes refer to such subjects dismissively as ‘goats’,  
 but this is foolish and discriminatory. A biometric system that is (or is  
 seen to be) socially regressive – that puts the disabled, the poor, the old  
 and ethnic minorities at greater risk of impersonation – should meet with  
 principled resistance. It might be defeated by legal challenges [1552]. It  
 may also be defeated by villains who pretend to be disabled. And some-  
 times the lack of heed for minority population groups is so offensive as to be  
 unlawful. For example, in 2019 the UK Home Office deployed a passport  
 app despite knowing that it didn’t work properly for black people [1950].

*•* A point that follows from this is that systems may be vulnerable to col-  
 money from it; Alice then complains of theft and produces a watertight  
 alibi. Quite apart from simply letting Betty take a rubber impression of  
 her ﬁngertip, Alice might voluntarily decrease handwriting ability; by giv-  
 ing several slightly different childish sample signatures, she can force the  
 machine to accept a lower threshold than usual. She can spend a couple  
 of weeks building a wall in her garden, and wear her ﬁngerprints ﬂat, so  
 as to degrade registration in a ﬁngerprint system. She might register for  
 a voice recognition system when drunk.

*•* The next issue is compulsion. If you get arrested in China, and since  
 August 2020 in Hong Kong, the police will hold your ﬁnger to your phone  
 to unlock it. If it uses face recognition they’ll pin your head and point  
 your phone at you; if you want to resist you have to close your eyes and  
 scrunch up your face [1348].

*•* The statistics are often not understood by system designers, and the birth-  
 example, there are about 50,000,000 pairs. So even with a false-accept  
 rate of only one in a million, the likelihood of there being at least one false  
 match will rise above one-half as soon as there are somewhat over a thou-  
 sand people enrolled4. So identiﬁcation is a lot tougher than veriﬁcation.  
 The practical consequence is that a system designed for authentication  
 may fail when you try to rely on it for evidence.

*•* Another aspect of statistics comes into play when designers assume that by  
 will often improve either the false accept rate or the false reject rate, while  
 making the other worse. If you install two different burglar alarms at your  
 home, then the probability that they will be simultaneously defeated goes  
 down while the number of false alarms goes up.

*•* The statistics are often somewhat uneven, so that as well as so-called  
 there may be ‘lambs’ who are particularly easy to impersonate, and ‘wolves’  
 who are particularly good at impersonating others. So it is vital to test

|  |  |  |
| --- | --- | --- |
| 4More precisely, 1177: a false match pairing in a database of N people becomes likelier  p than not as soon as *N >* 1*.*386*/f* where *f* is the single false-match rate, here 10*�*6 [519]. Check: 1177 people make 1177 x 1176 / 2 = 692,076 pairings, and the probability that none of these makes a false match is: 0*.*999999692*,*076 = 0*.*500 | | |
| **Security Engineering** |  |  |

systems thoroughly on substantial and diverse populations before deploy-  
 ment.

*•* Many vendors have claimed that their products protect privacy, as what’s  
 template that’s derived from it, somewhat like a one-way hash, and from  
 which you can’t be identiﬁed. It’s been argued from this that biometric  
 data are not personal data, in terms of privacy law, and can thus be passed  
 around without restriction. These claims were exploded by Andy Adler  
 who came up with an interesting *hill-climbing attack* on face recognition  
 systems. Given a recogniser that outputs how close an input image is

to a target template, the input face is successively altered to increase the  
 match. With the tested systems, this led rapidly to a recognizable image of  
 the target – a printout of which would be accepted as the target’s face [24].  
 He then showed how this hill-climbing technique could be used to attack  
 other biometrics, including some based on ﬁngerprints [25].

*•* It’s worth thinking what happens when humans and computers disagree.  
 is derived from phase information to which the human eye is not sensitive.  
 But what happens when a guard and a program disagree on whether a  
 subject’s face matches a ﬁle photo? Psychologists advise that biometric  
 systems should be used in ways that support and empower human cogni-  
 tion and that work within our social norms [586]. Yet we engineers often  
 ﬁnd it easier to treat the users as a nuisance that must adapt to our tech-  
 nology. This may degrade the performance of the humans. For example  
 when an automated ﬁngerprint database pulls out what it thinks is the  
 most likely print and presents it to the examiner: is he not likely to be  
 biased in its favour? Would it not perhaps be better for the computer  
 to test the examiner’s alertness constantly by giving him the three best  
 matches plus two poor matches, or would that be too annoying?

*•* Finally, Christian fundamentalists are uneasy about biometrics. They ﬁnd  
 small and great, rich and poor, free and slave, to receive a mark on their  
 right hand or on their foreheads, and that no one may buy or sell except  
 one who has the mark or the name of the beast, or the number of his  
 name.’

So there are some non-trivial problems. But biometrics have now gone main-

stream, and a good security engineer needs to know how to use them appropri-  
 ately.

**17.9** **Summary**

Biometric measures of one kind or another have been used to identify people  
 since ancient times, with handwritten signatures, facial features and ﬁngerprints  
 being the traditional methods. Three systems are now deployed at scale: ﬁn-  
 gerprint recognition on our phones, iris recognition in India and the Middle

|  |  |  |
| --- | --- | --- |
|  |  |  |

East, and facial recognition – which has become rapidly more accurate thanks  
 to the neural network revolution. These systems have different strengths and  
 weaknesses, and the statistics of error rates can be deceptively difficult.

When a biometric becomes very widely used, there may be an increased risk

of forgery in unattended operation: photographs of irises, ﬁngerprint moulds  
 and even good old-fashioned forged signatures must all be thought of in system  
 design. Context matters; even a weak biometric like handwritten signature

veriﬁcation can be effective if it is well embedded in the social and legal matrix.

Biometrics are usually more powerful in attended operation, where with

good system design the relative strengths and weaknesses of the human and  
 the machine may complement one another. Forensic uses are problematic, and  
 courts are much less blindly trusting of even ﬁngerprint evidence than they were  
 ten years ago. Historically, many biometric systems achieved most of their effect  
 by deterring criminals rather than actually identifying them. And although

there’s now the prospect of identifying people at scale from face recognition,  
 and authoritarian countries like Russia and China are doing it, there’s now  
 serious debate about whether we should allow the large-scale routine use of this  
 technology in democracies.

**Research Problems**

Many practical research problems relate to the design, or improvement, of bio-  
 metric systems. The hot topic in 2019 is the scalability of mass surveillance  
 CCTV systems, and the policy questions this raises about privacy, autonomy  
 and sovereignty. Given that facial recognition technology is still improving

rapidly and ﬁnding new applications, the debate is likely to run for some time  
 and to drive technical research on related topics.

One idea I thought up while writing this chapter for the ﬁrst edition in 2000

was instrumenting a car so as to identify a driver by the way in which he operated  
 the gears and the clutch. If your car thinks it’s been stolen, it phones a control  
 center which calls you to check. There is now research showing that users of  
 haptic systems can be recognised by the way in which they use tools [1478]. So  
 here’s another idea. Can we identify humans, and AI/ML systems, by other  
 learned skills? For example, the quote at the head of this chapter – where the  
 Ephraimites were spotted and killed for their inability to say the Hebrew letter  
 ‘shin’ – is actually about a skill that people learn when young or, with more  
 difficulty, as an adult. The ability to speak a language ﬂuently in the local

dialect is one of the most universal and visceral ways of identifying the in-group  
 from the out-group. The cool crowd speak the latest slang and dance the latest  
 dance. Now that robots, as well as humans, have skills that are acquired only  
 with effort, does this lead anywhere interesting?

**Further Reading**

The standard British history of ﬁngerprints is by Commander G.T.C. Lam-  
 bourne [1120], while the history in India is told by Chandak Sengoopta [1701].

|  |  |  |
| --- | --- | --- |
|  |  |  |

The McKie case is described in a book by Ian McKie and Michael Russella [1273].  
 A good technical reference on automated ﬁngerprint identiﬁcation systems is the  
 book by Davide Maltoni, Dario Maio, Anil Jain and Salil Prabhakar [1213]. As  
 for facial recognition, see Guodong Guo and Na Zhang [834]. The standard  
 work on iris codes is by John Daugman [517]. For speaker recognition forensics,  
 see Richard Klevans and Robert Rodman [1058].

As for the future, the US Department of Homeland Security is building

a new Homeland Advanced Recognition Technology (HART) database which  
 will include multiple forms of biometrics, from face recognition to DNA, and  
 consolidate records on both US residents and foreigners; there’s a description  
 and a discussion of the policy implications by the EFF [1196]. And the errors  
 in biometric forensics are mirrored in other forensic techniques; a 2009 report  
 from the US National Research Council showed that apart from DNA analysis,  
 most forensic methods were unreliable in various ways, relating not only to the  
 underlying science and technology but also to the fragmented nature of forensic  
 practice, the lack of standards and poor governance [1413]. As a recent example,  
 Sophie Nightingale and Hany Farid found that a common method of identifying  
 denim clothes by seam patterns was nowhere near as reliable or reproducible as  
 forensic examiners had claimed for many years [1447].

|  |  |  |
| --- | --- | --- |
|  |  |  |