ELECTRONIC FRONTIER FOUNDATION

# Biometrics: Who's Watching You?

September 2003

## Introduction

Among the many reactions to the September 11 tragedy has been a renewed attention to biometrics. The federal government has led the way with its new concern about border control. Other proposals include the use of biometrics with ID cards and in airports, e.g. video surveillance enhanced by facial-recognition technology.

The purpose of this document is to sketch out EFF's concerns about biometrics. In today's public arena, biometric technologies are being marketed as a "silver bullet" for terrorism; however, very little independent, objective scientific testing of biometrics has been done. Deploying biometric systems without sufficient attention to their dangers makes them likely to be used in a way dangerous to civil liberties. This document is very much a work in progress and we welcome comments.

## What Are Biometrics?

Biometrics refers to the automatic identification or identity verification of living persons using their enduring physical or behavioral characteristics. Many body parts, personal characteristics and imaging methods have been suggested and used for biometric systems: fingers, hands, feet, faces, eyes, ears, teeth, veins, voices, signatures, typing styles, gaits and odors.

## Our Major Concerns

- **Biometric technology is inherently individuating and interfaces easily to database technology, making privacy violations easier and more damaging.** If we are to deploy such systems, privacy must be designed into them from the beginning, as it is hard to retrofit complex systems for privacy.
- **Biometric systems are useless without a well-considered threat model.** Before deploying any such system on the national stage, we must have a realistic threat model, specifying the categories of people such systems are supposed to target, and the threat they pose in light of their abilities, resources, motivations and goals. Any such system will also need to map out clearly in advance how the system is to work, in both in its successes and in its failures.
- **Biometrics are no substitute for quality data about potential risks.** No matter how accurately a person is identified, identification alone reveals nothing about whether a person is a terrorist. Such information is completely external to any biometric ID system.
- **Biometric identification is only as good as the initial ID.** The quality of the initial "enrollment" or "registration" is crucial. Biometric systems are only as good as the initial identification, which in any foreseeable system will be based on exactly the document-based methods of identification upon which biometrics are supposed to be an improvement. A terrorist with a fake passport would be issued a US visa with his own biometric attached to the name on the phony passport. Unless the terrorist A) has already entered his biometrics into the database, and B) has garnered enough suspicion at the border to merit a full database search, biometrics won't stop him at the border.
- **Biometric identification is often overkill for the task at hand.** It is not necessary to identify a person (and to create a record of their presence at a certain place and time) if all you really want to know is whether they're entitled to do something or be somewhere. When in a bar, customers use IDs to prove they're old enough to drink, not to prove who they are, or to create a record of their presence.
- **Some biometric technologies are discriminatory.**A nontrivial percentage of the population cannot present suitable features to participate in certain biometric systems. Many people have fingers that simply do not "print well." Even if people with "bad prints" represent 1% of the population, this would mean massive inconvenience and suspicion for that minority. And scale matters. The INS, for example, handles about 1 billion distinct entries and exits every year. Even a seemingly low error rate of 0.1% means 1 million errors, each of which translates to INS resources lost following a false lead.
- **Biometric systems' accuracy is impossible to assess before deployment** Accuracy and error rates published by biometric technology vendors are not trustworthy, as biometric error rates are intrinsically manipulable. Biometric systems fail in two ways: false match (incorrectly matching a subject with someone else's reference sample) and false non-match (failing to match a subject with her own reference sample). There's a trade-off between these two types of error, and biometric systems may be "tuned" to favor one error type over another. When subjected to real-world testing in the proposed operating environment, biometric systems frequently fall short of the performance promised by vendors.
- **The cost of failure is high.** If you lose a credit card, you can cancel it and get a new one. If you lose a biometric, you've lost it for life. Any biometric system must be built to the highest levels of data security, including transmission that prevents interception, storage that prevents theft, and system-wide architecture to prevent both intrusion and compromise by corrupt or deceitful agents within the organization.

Despite these concerns, political pressure for increasing use of biometrics appears to be informed and driven more by marketing from the biometrics industry than by scientists. Much federal attention is devoted to deploying biometrics for border security. This is an easy sell, because immigrants and foreigners are, politically speaking, easy targets. But once a system is created, new uses are usually found for it, and those uses will not likely stop at the border.

With biometric ID systems, as with national ID systems, we must be wary of getting the worst of both worlds: a system that enables greater social surveillance of the population in general, but does not provide increased protection against terrorists.

### Some Current Biometric Initiatives

Sec. 403(c) of the **USA-PATRIOT Act** specifically requires the federal government to "develop and certify a technology standard that can be used to verify the identity of persons" applying for or seeking entry into the United States on a U.S. visa "for the purposes of conducting background checks, confirming identity, and ensuring that a person has not received a visa under a different name."

The recently enacted **Enhanced Border Security and Visa Entry Reform Act of 2002**, Sec. 303(b)(1), requires that only "machine-readable, tamper-resistant visas and other travel and entry documents that use biometric identifiers" shall be issued to aliens by October 26, 2004. The Immigration and Naturalization Service (INS) and the State Department currently are evaluating biometrics for use in U.S. border control pursuant to EBSVERA.

Even prior to September 11, however, large-scale civilian biometric identification systems were being pushed. Both the **Personal Responsibility and Work Opportunity Act of 1995** (PRWOA), a welfare reform law, and the **Immigration Control and Financial Responsibility Act of 1996** (ICFRA), an immigration reform law, called for the use of "technology" for identification purposes.

The PRWOA requires the states to implement an electronic benefits transfer program "using the most recent technology available . . . which may include personal identification numbers, photographic identification . . . and other measures to protect against fraud and abuse." This law covers, for example, the Food Stamps program.

The ICFRA requires the President to "develop and recommend . . . a plan for the establishment of a data system or alternative system . . . to verify eligibility for employment in the United States, and immigration status in the United States for purposes of eligibility for benefits under public assistance programs . . . or government benefits." This system "must be capable of reliably determining with respect to an individual whether . . . the individual is claiming the identity of another person."

The **Illegal Immigration Reform and Immigrant Responsibility Act of 1996** (IIRAIRA) requires the INS to include on alien border crossing cards "a biometric identifier (such as the fingerprint or handprint of the alien) that is machine readable." The State Department collects fingerprints and photographs of aliens for these cards.

The **Truck and Bus Safety and Regulatory Reform Act of 1988** (TBSRRA) requires "minimum uniform standards for the biometric identification of commercial drivers."

### EFF's concerns about biometrics

Why be concerned about biometrics? Proponents argue that: A) biometrics themselves aren't dangerous because all the real dangers are associated with the database behind the biometric information, which is little different from problems of person-identifying information (PII) databases generally; B) biometrics actually promote privacy, e.g., by enabling more reliable identification and thus frustrating identity fraud.

But biometric systems have many components. Only by analyzing a system as a whole can one understand its costs and benefits. Moreover, we must understand the unspoken commitments any such system imposes.

### Surveillance

The chronic, longitudinal capture of biometric data is useful for surveillance purposes. Our Surveillance Monitor page highlights some of these issues. Biometric systems entail repeat surveillance, requiring an initial capture and then later captures.

Another major issue relates to the "voluntariness" of capture. Some biometrics, like faces, voices, and fingerprints, are easily "grabbed." Other biometrics, at least under present technology, must be consciously "given." It is difficult, for instance, to capture a scan of a person's retina or to gather a hand geometry image without the subject's cooperation. Easily grabbed biometrics are a problem because people can't control when they're being put into the system or when they're being tracked. But even hard-to-grab biometrics involve a trust issue in the biometric capture device and the overall system architecture.

### Databases

To be effective, a biometric system must compare captured biometric data to a biometric database. Our National ID System page highlights issues surrounding database abuse, which has both static and dynamic dimensions.

The static issues surrounding databases are mainly about safeguarding large and valuable collections of personally identifying information. If these databases are part of an important security system, then they (and the channels used to share PII) are natural targets for attack, theft, compromise, and malicious or fraudulent use.

The dynamic issues surrounding databases mainly concern the need to maintain reliable, up-to-date information. Databases that seek to maintain accurate residence information must be updated whenever one moves. Databases that are used to establish eligibility for benefits must be updated so as to exclude persons no longer eligible. The broader the function of the system, the more and broader the updating that is required, increasing the role of general social surveillance in the system.

It may seem that one of the issues that plagues token-based ID systems (like ID cards) -- the security or integrity of the token itself -- does not apply for biometric systems, because "you are your ID." But the question of the reliability of the token is really a question about trust. In an ID card system, the question is whether the system can trust the card. In biometric systems, the question is whether the individual can trust the system. If someone else captures your signature, fingerprint, or voice, for instance, what prevents it from being used by others? Any use of biometrics with a scanner run by someone else involves trusting someone's claim about what the scanner does and how the captured information will be used.

Vendors and scanner operators may say that they protect privacy in some way, perhaps by hashing the biometric data or designing the database to enforce a privacy policy. But the end user typically has no way to verify whether such technical protections are effective or implemented properly. End-users should be able to verify any such claims, and to leave the system completely if they are not satisfied. Exiting the system, of course, should at least include the expungement of the end-user's biometric data and records.

### Linking

An oft-noted risk of biometric systems is the use of biometrics as a linking identifier. This risk, of course, depends to some extent on standardization. Consider, for instance, the use of the Social Security number as a linker across disparate databases. While the private sector would not have been able to develop anything like the SSN on its own, once the government created this identifier, it became a standard way of identifying individuals. Standardization therefore creates new privacy risks because information gathered for one purpose can be used for completely unrelated, unconsented-to purposes.

Currently, Automated Fingerprint ID Systems (AFIS) are heavily used by the government in connection with law enforcement, but there is at present little standardization within the AFIS industry. If law enforcement and private industry were to unify their fingerprint databases under one common standard, such as under a national ID system, this would potentially put one's entire life history in interoperating databases that are only a fingerprint away.

### Tracking

By far the most significant negative aspect of biometric ID systems is their potential to locate and track people physically. While many surveillance systems seek to locate and track, biometric systems present the greatest danger precisely because they promise extremely high accuracy. Whether a specific biometric system actually poses a risk of such tracking depends on how it is designed.

Why should we care about perfect tracking? EFF believes that perfect tracking is inimical to a free society. A society in which everyone's actions are tracked is not, in principle, free. It may be a livable society, but would not be our society.

EFF believes that perfect surveillance, even without any deliberate abuse, would have an extraordinary chilling effect on artistic and scientific inventiveness and on political expression. This concern underlies constitutional protection for anonymity, both as an aspect of First Amendment freedoms of speech and association, and as an aspect of Fourth Amendment privacy.

Implemented improperly, biometric systems could:

- increase the visibility of individual behavior. This makes it easier for measures to be taken against individuals by agents of the government, by corporations, and by our peers.
- result in politically damaging and personally embarrassing disclosures, blackmail and extortion. This hurts democracy, because it reduces the willingness of competent people to participate in public life.
- increase the 'circumstantial evidence' available for criminal prosecution. This might dramatically affect the existing balance of plausible -sounding evidence available to prosecutors, and hence increase the incidence of wrongful conviction. Many criminal cases are decided by plea bargaining, a process that is sensitive to the perceived quality of evidence. Even ambiguous or spurious evidence generated by complex technical systems may be difficult for overburdened public defenders to challenge.
- enable the matching of people's behavior against pre-determined patterns. This could be used by the government to generate suspicion, or by the private sector to classify individuals into micro-markets, the better to manipulate consumer behavior.
- aid in repressing readily locatable and trackable individuals. While the public's concern is usually focused on the exercise of state power, these technologies may also greatly empower corporations. If proper privacy safeguards are not constructed into such systems, they would prove useful in dealing with such troublesome opponents as competitors, regulators, union organizers, whistleblowers, and lobbyists, as well as employees, consumer activists, customers and suppliers.

### The Attributes of Biometric Systems

### How Do Biometrics Compare to Other Types of ID?

Currently other than personally recognizing someone, or having a trusted third party personally swear to their identity, the only other technique for identifying a person is through the use of a "token." These tokens, which are in essence representations of the oath of a trusted third party, come in two basic forms:

- **Knowledge tokens**, such as passwords, secret PINs (Personal Identification Numbers), or knowledge of personal data (knowing one's mother's maiden name, e.g.), or
- **Physical tokens** such as ID cards, passports, chip cards, or plain old keys.

Token IDs offer certain advantages over biometric identification. Security against "false acceptance" of impostors can be raised by increasing the complexity of the token used for identification. Also, in the event of loss or compromise, the token, be it a password, PIN, key, or ID card, can be revoked, changed or reissued, a biometric measurement cannot.

The advantage of biometrics is that unlike tokens, biometrics cannot be lost, loaned, or forgotten. Token-based systems must verify that the presenter is the authorized user, not an unauthorized person who has come to possess the token.

Used carefully, biometrics may be combined with token-based systems to mitigate the vulnerability of ID tokens to unauthorized use.

### Functions of Biometric Systems

One useful way of thinking about biometrics is that they are used for one of two purposes: A) To prove that you are who you say you are (positive ID), or B) To prove that you are not who you say you are not (negative ID).

In a positive ID situation, the subject asserts that she is Jane Doe and submits a "live" sample (a fingerprint, for example) to the system. The system then checks its database of previously enrolled or registered samples to see if the live sample matches the reference sample. A positive ID system is designed to prevent more than one person from using a single identity.

In a negative ID situation, John Roe claims *not* to be someone already known to the system. Here, the system checks its database to see that Roe is not on the watchlist of suspected criminals and terrorists, whose biometrics are already in the system. A negative ID system is designed to prevent one person from using more than one identity.

When biometrics are employed to effect negative identification, one need not be enrolled. The only persons who must be "in" the database are those whom the operator is trying to keep out or catch.

Biometrics alone cannot establish "true identity." A biometric system cannot prevent someone from furnishing fake credentials when they first enter the system. They can only prevent them from using another identity once enrolled.

### Common Aspects of All Biometric Systems:

All biometric technology systems have certain aspects in common. All are dependent upon an accurate reference or "registration" sample. If a biometric system is to identify a person, it first must have this sample, positively linked to the subject, to compare against. Modern biometric identification systems, based on digital technology, analyze personal physical attributes at the time of registration and distill them into a series of numbers. Once this reference sample is in the system, future attempts to identify a person are based on a comparison of a "live" sample and the reference sample or samples.

A perfect system would recognize a person 100% of the time, and reject an impostor 100% of the time. However, biometric samples are gathered from people in environmental conditions that are uncontrollable, over equipment that may slowly be wearing out, and using technologies and methods that vary in their level of precision. Consequently, the accuracy of biometric systems is assessed in light of these confounding variables via its tendency to experience either a "false match" (also called a "false accept") or a "false non-match" ("false reject"). The point at which these two rates intersect is called the equal error rate or crossover point.

Biometric systems may be "tuned" to diverge from the equal error rate to provide a match threshhold that satisfies the designer's requirements. If a system compares a large number of persons against a small number of samples, and the consequence of a false match is low, (for example, at a border crossing or airport that is looking for a short list of criminals) a system biased towards a higher "false accept" or "false match" rate may be desirable. The advantage to biasing a system in this manner is that it is likely to err on the side of safety, and less likely to let a criminal slip through undetected. The disadvantage is that the system will falsely associate innocent people with criminals. If other safeguards are in place and the system operators understand the system's bias towards false match, the result can be a relatively trivial loss of convenience due to increased scrutiny (extra inspection of luggage, questioning, etc.) Biasing such a system towards a high "false non-match" or "false reject" rate will result in fewer passengers slowed down at the gate, but at the cost of possibly losing the sought-after criminals.

When assessing the utility or the cost of a biometric system, it's important to bear the common features of all such systems in mind. How is the reference sample to be gathered and catalogued? How it the live sample going to be gathered? Can a live sample be captured without the subject's knowledge and cooperation? What are the implications of all four possible outcomes (true match, true non-match, false match, false non-match)? What is the value of a successful system, and what is the cost, to all parties, should it fail?

Further, we should not assess failure simply from the perspective of the core biometric technology itself. Even an ideal system can be defeated easily if it is incorporated into an insecure or poorly-designed overall system architecture. Any biometric system, especially one that involves a component of telecommunication, **must** be very carefully designed to prevent the loss or interception of user biometrics. Any deployed system must incorporate safeguards to prevent the interception of biometric data while it is being communicated. If a user's biometric is intercepted, criminals may be able to replicate either the sample itself or the string of binary data produced by a successfully-matched sample. Armed with such intercepted biometric data, a criminal would be able to effect a potentially very damaging identity theft.

### Types of Biometrics

A proper assessment is built not only on a general understanding of biometrics, but also on an understanding of specific technologies. An understanding of both biometrics in general and specific biometric technologies is a necessary condition for a solid understanding of the larger social implications of biometrics.

### Signature

The biometric most familiar to us is the signature. Our ability to judge by sight if one signature matches another has made this a time-proven and legally-binding biometric. However, by sight alone, most of us cannot recognize the pressure of the pen on the paper or the speed and rhythms of its traverse of the page. Computers can do all these things, and quantify, analyze and compare each of these properties to make signature recognition a viable biometric technology. Being based on things that are not visible (pen pressure and velocity, for example), signature-based biometric technology, offers a distinct advantage over regular signature verification -- in addition to mimicking the letter forms, any potential forger has to fabricate a signature at the same speed, and with the same pen weight, as his victim.

Signature biometrics pose a couple of unique problems. The first is the comfort with which people are already willing to use their signature as a form of identification. While this high level of consumer acceptance is viewed as a strength by vendors of such systems, this bears with it a strong downside. Without proper notification, a person may sign an electronic signature pad and unwittingly also be surrendering a reference or live biometric sample. Since the custom of leaving a signature as one's "official mark" is based on the presumption of irreproducibility (i.e., that a forger would be hard-pressed to imitate a signature just by looking at it), people are willing to provide a signature without giving its potential for reproduction a second thought. However, electronic data is easy to copy and transmit. And so, a forger posing as a delivery man might fraudulently secure a signature biometric by presenting a victim with a "gift" box, requesting a signature to confirm delivery, and making off with the victim's biometric data.

The second unique property of signature biometrics is that unlike all other biometrics, which either establish an identity (identification) or confirm an identity (authentication), a signature can convey *intent* (authorization). In other words, a traditional signature on paper is taken both to authenticate the signator, and to convey the signator's legal authority. An electronic system that solicits a user's non-signature biometric must provide a separate step to convey the user's legal authorization for any binding transaction. A signature-based biometric system could mimic our current legally customary acceptance of a signature to simultaneously convey both identity and authority.

### Keystroke Dynamics

The rhythms with which one types at a keyboard are sufficiently distinctive to form the basis of the biometric technology known as keystroke dynamics. While distinct, keystroke dynamics are not sufficiently unique to provide identification, but can be used to confirm a user's identity.

Keystroke dynamics, unlike other biometric technologies, is 100% software-based, requiring no sensor more sophisticated than a home computer. Because of this, deployment is occurring in fairly low-stakes, computer-centric applications, such as content filtering (Net Nanny owns BioPassword, the leading keystroke dynamics vendor) and digital rights management, in which passwords to download music are bolstered with by keystroke dynamic verification, to prevent password-sharing. As a general rule, any method involving home or office computers is inherently insecure, as these devices leave a lot more room for experimentation than devices like ATMs or entry systems, and the information they use tends to travel over unsecured communication lines.

### Hand Geometry

Perhaps the most ubiquitous electronic biometric systems are hand geometry based. Hand-geometry-based systems require the subject to place his or her hand (usually the right hand) on a plate where it is photographically captured and measured. Made of 27 bones and a complex web of interconnected joints, muscles, and tendons, the human hand presents a sufficiently peculiar conformation of anatomical features to enable authentication, but is not considered sufficiently unique to provide full identification. Further, the geometry of the hand is variable over time, as hand shape may be altered due to injury, disease, aging, or dramatic weight swings. A simple hand-geometry system will measure length and thickness of digits, width of the palm at various points, and the radius of the palm. This results in a relatively simple identification that can be expressed in a very simple, compact string of data. Efforts have been made to improve the accuracy of hand geometry, including three-dimensional sampling (i.e., a second camera measuring the thickness of the hand from the side), and a patented system (owned by the British concern, Neusciences) that measures the pattern of the veins of the hand. Neusciences claims that their system provides a high degree of accuracy and that the hand vein feature is unique and relatively invariable, changing little over a person's lifespan.

In deployment, traditional hand geometry systems have found acceptance in applications requiring verification of an identity, rather than a full proof or establishment of an identity. Airports, prisons, and factories have successfully employed hand-geometry-based systems to restrict access to runways, to prevent walk-out escapes during visits, and to ensure that time cards are being punched only by the worker, and not by that worker's pal on his or her behalf. In all these instances, the subject is attempting to prove or disprove his or her membership in a relatively small group of people (authorized runway personnel, prisoners/visiting family, factory workers). When stakes are high, these systems are not relied on exclusively to confirm identity; rather, they are used to provide an additional layer of security above and beyond that provided by existing security systems.

Since they must accommodate the largest of hands, any hand geometry or hand vein system must be somewhat bulky, and requires the user to perform an obtrusive task (placing his or her hand on the platen for sampling). Because of this obtrusiveness, hand-based biometrics represent less of a privacy threat than some other systems: subjects cannot have their biometric features sampled without their knowledge, and the sampling method is unambiguous in its intent.

**Fingerprint**

Fingerprinting is a highly familiar and well-established biometric science. The traditional use of fingerprinting, of course, has been as a forensic criminological technique, used to identify perpetrators by the fingerprints they leave behind them at crime scenes. Scientists compare a latent sample left at a crime scene against a known sample taken from a suspect. This comparison uses the unique features of any given fingerprint, including its overall shape, and the pattern of ridges, valleys, and their bifurcations and terminations, to establish the identity of the perpetrator.

In the context of modern biometrics, these features, called fingerprint minutiae, can be captured, analyzed, and compared electronically, with correlations drawn between a live sample and a reference sample, as with other biometric technologies. Fingerprints offer tremendous invariability, changing only in size with age, are highly resistant to modification or injury, and very difficult to "forge" in any useful way. Although the development of some sort of surreptitious sensor is not inconceivable, the reality is that sensors remain obtrusive, requiring a willful finger pressure to gather a useful sample. Unlike other systems, based on cameras and high-tech sensors, fingerprint sampling units are compact, rugged, and inexpensive, with commercially available systems from multiple vendors offering very good accuracy. Next-generation scanners can analyze below the surface of the skin, and can add pore pattern recognition in addition to the more obvious minutia of the fingerprint.

**Facial Recognition**

Facial recognition sprung into the national spotlight during the 2001 Super Bowl, when Tampa police scanned the faces of game fans without their knowledge for the purpose of spotting terrorists in the crowd. While this proved a public relations nightmare in January 2001, the use of this technology in New Orleans at the post-9/11 Super Bowl of 2002 generated little controversy. Facial recognition remains one of the more controversial biometric technologies because of its very *un*obtrusiveness. With good cameras and good lighting, a facial recognition system can sample faces from tremendous distances without the subject's knowledge or consent.

Most facial recognition technology works by one of two methods: facial geometry or eigenface comparison. Facial geometry analysis works by taking a known reference point (for example, the distance from eye to eye), and measuring the various features of the face in their distance and angles from this reference point. Eigenface comparison uses a palette of about 150 facial abstractions, and compares the captured face with these archetypal abstract faces. In laboratory settings, facial recognition results are excellent, but critics have questioned the effectiveness of the technology in real-world circumstances. Nevertheless, the accuracy of facial recognition has been good enough for casinos to have put the the technology to use since the late 1990s as a means to spot banned players. Facial recognition technology proponents claim good performance even against disguises, weight changes, aging, or changes in hairstyle or facial hair.

**Eye biometrics: Iris/Retina**

The human eye offers two features with excellent properties for identification. Both the iris (the colored part visible at the front of the eye) and the veins of the retina (the thin film of nerve endings inside the eyeball that capture light and send it back to your brain) provide patterns that can uniquely identify an individual. Retinal scanning is the older technology, and requires the subject to look into a reticle and focus on a visible target while the scan is completed. It's definitely one of the more intrusive biometric technologies, with some subjects reporting discomfort at the scanning method. Iris recognition has an advantage in ease of use, in that it merely requires the subject to look at a camera from a distance of three to ten inches. The iris scanner illuminates the iris with invisible infra-red light, which shows details on darker -colored eyes that are not visible to the naked eye. The pattern of lines and colors on the eye are, as with other biometrics, analyzed, digitized, and compared against a reference sample for verification.

Iridian Technologies, who hold the patents on iris recognition, claim that the iris is the most accurate and invariable of biometrics, and that their system is the most accurate form of biometric technology. Iridian's system also has the benefit of extremely swift comparisons. The compay claims that it can match an iris against a database of 100,000 reference samples in 2-3 seconds, whereas a fingerprint search against a comparable database might take 15 minutes.

### Voice Verification

None of us finds it remarkable when a friend recognizes our voice on the telephone. However, what we find easy to do is still a very hard problem for computers, especially when their job is to identify someone positively. The prospect of accurate voice verification offers one great advantage, which is that it would allow a remote identification using the phone system, an infrastructure that's already been built and thus has zero client-side cost: no special reader needs to be installed in your home. Even without the phone system, the sampling apparatus, a microphone, remains far cheaper than competing, largely optically-based biometric technologies.

But voice recognition technology is still not good enough to be used as a front-line biometric technology. Simply put, voice verification systems have to account for a lot more variables than do other systems, starting with the inevitable compression of a voice captured by cheap microphones (especially those found on phone handsets), discriminating a voice from background noise and other sonic artifacts, and the human voice's tremendous variability, due to colds, aging, and simple tiredness. Also, just as a voice can be surreptitiously recorded over the telephone or face-to-face, a person's voice can be captured surreptitiously by a third party (either by tapping or bugging) and replayed, or a person's voice might be biometrically sampled remotely without consent during a fake door-to-door or telephone sales call. Because of these difficulties, commercial deployments of voice verification have been limited to "backup" status, systems in which there are other token-based methods of identification, with voice verification providing an added layer of protection.

### Characterizing Different Biometrics

Different biometric features have characteristics that make them more or less useful for particular applications. Dr. James Wayman, director of the National Biometric Test Center at San Jose State University, categorizes biometric features in terms of five qualities:

- **Robustness**: repeatable, not subject to large changes.
- **Distinctiveness**: there are wide differences in the pattern among the population.
- **Accessibility**: easily presented to an imaging sensor.
- **Acceptability**: perceived as non-intrusive by the user.
- **Availability**: a user may present a number of independent measurable features.

Dr. Wayman explains these qualities by comparing fingerprinting to hand geometry.

> *"Fingerprints are extremely distinctive, but not very robust, sitting at the very end of the major appendages you use to explore the world. Damaging your fingerprints requires less than a minute of exposure to household cleaning chemicals. Many people have chronically dry skin and cannot present clear prints. Hands are very robust, but not very distinctive. To change your hand geometry, you'd have to hit your hand very hard with a hammer. However, many people (somewhat less than 1 in 100) have hands much like yours, so hand geometry is not very distinctive. Hands are easily presented without much training required, but most people initially misjudge the location of their fingerprints, assuming them to be on the tips of the fingers. Both methods require some "real-time" feedback to the user regarding proper presentation. Both fingerprints and the hand are accessible, being easily presented. In the 1990 Orkand study*, only 8% of customers at Department of Motor Vehicle offices who had just used a biometric device agreed that electronic fingerprinting "invades your privacy." Summarizing the results of a lengthy survey, the study rated the public acceptance of electronic fingerprinting at 96%. To our knowledge, there is no comparable polling of users regarding hand geometry, but we hypothesize that the figures would not be too different. With regard to availability, our studies have shown that a person can present at least 6 nearly-independent fingerprints, but only one hand geometry (your left hand may be a near mirror image of your right)."*
>
> **Orkand Corporation, "Personal Identifier Project: Final Report", April 1990, State of California Department of Motor Vehicles report DMV88-89, reprinted by the U.S. National Biometric Test Center.*

### Characterizing Biometric Applications

Dr. Wayman suggests characterizing biometric applications in terms of seven variables:

- **Cooperative vs. Non-Cooperative** This refers to the behavior of the "threat" or would-be deceptive user. Is the "threat" trying to cooperate with the system? If the threat is trying to enter a restricted area, she either cooperates with the positive ID system to try to fool it into thinking she's allowed in, or deceptively tries not to cooperate with a negative ID system so as not to trigger the alarm. One implication of this variable is the scope of database search. In cooperative applications, users may first identify themselves with a card or PIN, so that the system need only match against the claimed identity's template. In non-cooperative applications, users can't be trusted to identify themselves correctly, so the entire database may need to be searched.
- **Overt vs. Covert** Is the user aware that the biometric sampling and identification is occurring?
- **Habituated vs. Non-Habituated** Is the intended user expected to be experienced in the use of the system?
- **Attended vs. Non-Attended** Will the intended user be supervised when using the system?
- **Standard vs. Non-Standard Environment** How controlled are the environmental conditions for operation?
- **Public vs. Private** Will users be customers (public) or employees (private)?
- **Open vs. Closed** Will the application be required to exchange biometric data with other systems or not?

Dr. Wayman explains that the positive biometric identification (hand geometry) of users of the Immigration and Naturalization Service's Passenger Accelerated Service System (INSPASS) for rapidly admitting frequent travelers into the United States:

> *"can be classified as a cooperative, overt, non-attended, non-habituated, standard environment, public, closed application. The system is cooperative because those wishing to defeat the system will attempt to be identified as someone already holding a pass. It will be overt because all will be aware that they are required to give a biometric measure as a condition of enrollment into this system. It will be non-attended and in a standard environment because collection of the biometric will occur near the passport inspection counter inside the airports, but not under the direct observation of an INS employee. It will be non-habituated because most international travelers use the system less than once per month. The system is public because enrollment is open to any frequent traveler into the United States. It is closed because INSPASS does not exchange biometric information with any other system." (emphases added)*

**Characterizing Biometric ID Systems**

Dr. Wayman suggests that biometric ID systems should be viewed in terms of a generic biometric system made up of five basic components or subsystems, depending on the application: data collection, transmission, signal processing (which comprises feature extraction, quality control, pattern matching), storage, and decision.

- **Data collection** Biometric systems involve at least two discrete data collection steps. First, any biometric system must contain a biometric characteristic deemed "true" or canonical from the system's viewpoint. The term "enrollment" or "registration" refers to the first entry of biometric data into the database. Second, the system must compare a later-submitted "sample" (often called a "live sample") to the sample in the database. (Scale is crucial to the enrollment step, sometimes for quite mundane reasons. In the context of biometric visa issuance by the State Department, for instance, the government has been looking at whether or not U.S. consulates around the world have room to handle the additional equipment and physical traffic needed for data collection.)
- **Transmission** Many biometric systems collect data at one location but store and/or process it at another. Such systems require data transmission.
- **Signal processing** Once a biometric is acquired, it must be prepared for comparison. There are three basic tasks here: feature extraction, quality control, and pattern matching. A fourth task in large-scale systems is pattern classification.
  - *Feature extraction* involves finding the true biometric pattern amid noise and signal degradation, preserving the critical information, and discarding redundant or unnecessary data. Dr. Wayman gives the example of a text-independent speaker-recognition system. A properly implemented system isolates "features that depend only on the speaker and not on the words being spoken." At the same time, the system focuses on features that do not change "even if the speaker has a cold or is not speaking directly into the microphone."
  - *Quality control* involves checking to see if the signal is of good quality. Ideally, it should be possible to make a quick determination so that another measure can be taken if the signal is inadequate.
  - *Pattern matching* involves comparing the live sample to the reference sample in the database. If the user claims to be Jane Doe, the pattern-matching process may only need to compare the sample to Jane Doe's stored template. In other situations, the sample must be compared to multiple templates. The pattern-matching process generates a quantitative "distance" measure of the comparison -- how close are they? Even for the same person, the distance is rarely if ever zero.
  - *Pattern classification* is a technique aimed at reducing the computational overhead of pattern matching. In large-scale systems, it can be computationally taxing to match each sample against all stored templates in the database. If biometric patterns can be categorized, then it may be possible to perform the match against only the stored templates in that category. This is sometimes referred to as "binning." A different technique with the same goal is "filtering," which involves partitioning the database based on information not contained in the biometric itself. If you know the person is a man, you don't need to check against women's biometrics. Both of these techniques introduce additional error possibilities; if binning or filtering is erroneous, then the true template is not used and a false non-match results.
- **Decision** This subsystem implements the biometric ID system's actual policy with regard to matching. In general, lowering the number of false non-matches raises the number of false matches, and vice versa. The signal processing subsystem yields a quantitative "distance" measure, but "how close or far is enough?" is a matter of policy. In a high-security application where the cost of a false acceptance could be high, system policy might prefer very few false acceptances and many more false rejections. In a commercial setting where the cost of a false acceptance could be small and treated as a cost of doing business, system policy might favor false acceptances in order not to falsely reject and thereby inconvenience large numbers of legitimate customers. The inevitable existence of these errors means that any biometric ID system must also have well-designed policies for exception handling.
- **Storage** Biometric reference samples must be stored somewhere for matching purposes. For systems only performing "one-to-one" matching, the database may be distributed on cards carried by each enrolled user. The user simply presents his or her biometric and the system checks to see if it matches the template stored on the card. Depending upon system policy, no central database need exist, although in this application a centralized database can be used to detect counterfeit cards or to reissue lost cards without re-collecting the biometric pattern.

  In other cases, centralized storage is necessary because the system must match the live sample to multiple templates. As the number of templates grows, speed becomes an increasingly significant issue. One technique is to partition the database (i.e., binning or filtering) so that any sample need only be matched to the templates in one partition. This increases system speed and decreases false matches at the expense of increasing the false non-match rate owing to partitioning errors. System error rates thus change with increasing database size and ID systems do not linearly scale.

Full biometric patterns cannot be reconstructed from the stored reference samples if these are stored as templates, which reduce data richness dramatically. Templates themselves are often created using the system vendor's proprietary feature extraction algorithms. Whether stored templates themselves can be used to "spoof" the system internally is entirely dependent on the security of the system architecture.

Biometric ID systems may store not only the templates but also raw data. One reason to do so would be to allow changes to the system or to change system vendors without having to re-collect data from all enrolled users. Full raw data storage is a riskier practice in that new templates may be extracted from the data or the raw data itself may be used against the system.

### The State of Scientific Testing of Biometric ID Systems

According to Dr. Wayman,

> *"Testing of biometric devices requires repeat visits with multiple human subjects. Further, the generally low error rates mean that many human subjects are required for statistical confidence. Consequently, biometric testing is extremely expensive, generally affordable only by government agencies. Few biometric technologies have undergone rigorous, developer/vendor-independent testing to establish robustness, distinctiveness, accessibility, acceptability and availability in 'real-world' (non-laboratory) applications."*

An in-depth discussion of the statistical methodology in testing biometric ID systems is beyond the scope of this discussion. We recommend Dr. Wayman's website to interested persons. Note, however, that it is very difficult to generalize from test results. At this time, scientists have no way of accurately estimating how large a test is needed to adequately characterize any biometric device in any application, even with advance knowledge of theoretical error rates.

### Media Coverage/Resources

Check out this page for helpful resources including lots of media coverage and links related to the topic.

### Acknowledgements:

This document was written and compiled by William Abernathy and Lee Tien with editorial assistance from Sarah Granger and technical assistance from Johnson Hor.

Want to learn how you can defend free speech, stand up for privacy, fight for government transparency, support consumer rights, and protect your right to innovation in the digital world? Visit **http://eff.org/fight** to find ways to help.