WHITE PAPER

What Are Biometrics?

AWARE

What Are Biometrics?

Identity and Trust

There are countless things about us that, in concert, make each of us unique, such as our physical attributes, home address, birthdate, relationships, and our knowledge. The uniqueness of our physical embodiment and personal story is represented by what we commonly think of as our "identity." In today's interconnected, computer-powered world, there is ever-increasing utility in 1) correctly attributing digital information to an individual, and in 2) asserting our identity in a way that can be communicated and trusted. Our identity might be used simply to properly attribute information about us that is useful for some purpose in the future (e.g. a medical or financial record). But these types of records also enable us to demonstrate a historical pattern of behavior towards establishing trust, and in doing so compel personal accountability. We leverage this trust and accountability to earn privileges such as access to an asset, facility, or country. For the purpose of access, the utility of identity is twofold: first, to communicate our trustworthiness and accountability, and later—upon attempting to transact upon our earned "trust capital"—to assert that we are in fact the same person with whom trust was earlier established.

"Biometrics are our most unique physical (and behavioral) features that can be practically sensed by devices and interpreted by computers so that they may be used as proxies of our physical selves in the digital realm. In this way we can bond digital data to our identity with permanency, consistency, and unambiguity, and retrieve that data using computers in a rapid and automated fashion."

Conversely, our identity might be challenged in order to counter fraudulent representation, or used by someone else to assert mistrust upon us.

Our names and personal numbers offer a time-tested and relatively efficient means to represent our identity. Importantly, they can be interpreted not only by people but also by computers to bind digital information and trust- or mistrust-earning attributes to us, and this is clearly useful for many applications. A school transcript, a speeding ticket, and a credit history all serve this purpose. But our names and numbers are effective only to the degree that they are 1) unique, 2) permanent, 3) consistent, and 4) unambiguously bonded to our physical selves. We know they are not necessarily unique (e.g. John Smith), or permanent (e.g. Judy Smith née Johnson), and they are clearly not unambiguously bonded with us physically (e.g. a forehead tattoo). This is where modern biometrics are useful. Biometrics are our most unique physical (and behavioral) features that can be practically sensed by devices and interpreted by computers so that they may be used as proxies of our physical selves in the digital realm. In this way we can bond digital data to our identity with permanency, consistency, and unambiguity, and retrieve that data using computers in a rapid and automated fashion.

Biometric Modalities

Much is made about the breadth of biometric modalities, and indeed research into new, exotic biometrics (ear, gait, odor, etc.) is compelling. But the modalities that are field-proven in large-scale deployments are fingerprint, face, iris, and voice. These happen to be the biometric modalities that, today, best meet our tests for uniqueness, permanence, and consistency while also being conducive to capture using sensing devices in an ergonomically and economically practical way. Proprietary[1] techniques that have also been deployed include vascular (palm, finger vein), and hand geometry.

Biometrics are largely statistical in nature, so it follows that:

a) the more data we have in a biometric sample (or set of samples), the more likely that it is unique,

b) there is always some likelihood that two different individuals will generate very similar or equivalent biometric samples, and

c) there is always some likelihood of false match or false non-match (Type I or Type II error) results from a biometric comparison.

Some biometric modalities are less permanent over time than others, and some are more difficult to present and capture consistently. Some are more prone to sample quality problems.

There is no perfect biometric modality; each has advantages and disadvantages for a given use case. For example, perhaps the most differentiating feature of fingerprints as a modality is that they leave behind evidence at a crime scene as "latents" (e.g. fingerprints on a glass). Irises are perhaps the most consistent, information-dense, "barcode-like" of the modalities. Facial images stand out because they are the biometric modality that humans excel at comparing, and so we can integrate complementary human-and machine-based recognition. Additionally, facial images are abundant in the digital realm, and also can be collected covertly from a distance. Voice is notable for being behavioral as well as physical, and thus the samples available from a given individual are abundant.

Even when our biometric samples are unique, permanent, consistent, and physically bonded to us, the sensors and algorithms we have devised to acquire and analyze them are imperfect. Sensors introduce optical and electrical distortion. Information is lost as sample data is converted from analog to digital form, and then again when the digital signal is compressed. Sampling rates (spatial resolution in the digital domain) significantly impact the quality of biometric samples. The algorithms designed to extract computer-matchable "templates" from a sample vary dramatically in precision and performance, as do algorithms and systems used by computers to rapidly assess their similarity. Machines are good at very fast, reasonably accurate, automated signal processing and template comparison, but they lack a human's ability to visually perceive, analyze and characterize the similarity of two samples. Nevertheless, our physical selves provide many features that are well-suited for biometric comparison and search, and advances in modern sensing and computing technologies continue to improve the ability of a machine to perform biometric identification extremely quickly and accurately.

Biometric Processes

Biometric systems rely on several discrete processes: enrollment, live capture, template extraction, and template comparison. The purpose of enrollment is to collect and archive biometric samples and to generate numerical templates for future comparisons. By archiving the raw samples, new replacement templates can be generated in the event that a new or updated comparison algorithm is introduced to the system. Practices that facilitate enrollment of high-quality samples are critical to sample consistency, and improve overall matching performance, which is particularly important for biometric identification by "one-to-many" search.

We can differentiate "live capture" from enrollment as the process of collecting live biometric "probe" samples upon an access or identification attempt and comparing them to a "gallery" of previously enrolled templates.

1) "Proprietary" here means that the capture and matching software and capture hardware peripheral are inextricably interdependent.

Template extraction requires signal processing of the raw biometric samples (e.g. images or audio samples) to yield a numerical template. Templates are typically generated and stored upon enrollment to save processing time upon future comparisons. Comparison of two biometric templates applies algorithmic computations to assess their similarity. Upon comparison, a match score is assigned. If it is above a specified threshold, the templates are deemed a match.

Typically, biometric template extraction and comparison algorithms are proprietary (different and secret) and so can't be used with those from different vendors in the same system (e.g. to compare templates generated by different products, or use a matching algorithm from one company to compare templates generated by algorithms of another). Exceptions are MINEX-certified, minutiae-based fingerprint template generators and matching algorithms. This category of templates and matchers have been specifically designed, tested, and independently certified by NIST to be interoperable for one-to-one verification and so are ideal for compact storage on smart cards or travel documents.

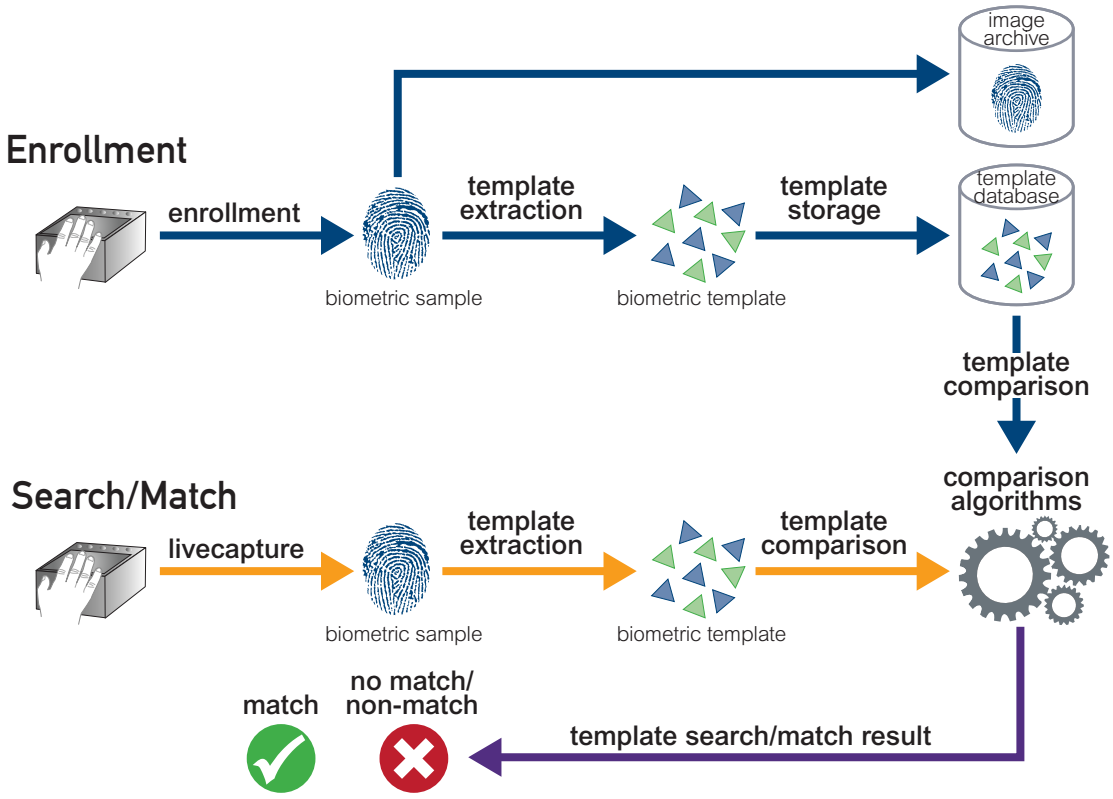

Figure 1 - A biometric system

Biometric System Accuracy Testing

The accuracy of a biometric system is quantified most typically by a "receiver operating characteristic", or "ROC curve" plot indicating its "false match rate (FMR)" and "false non-match rate (FNMR)" against some biometric sample gallery. The false match rate is the frequency with which biometric samples from different sources are erroneously assessed to be from the same source. The false non-match rate is the frequency with which samples from the same source are erroneously assessed to be from different sources. A well-performing biometric system is characterized by prompt results and low rates of false matches and false non-matches. The accuracy of a system falls on a point on the ROC curve whose location is a function of the matching "threshold" applied. A higher match threshold reduces false match rate but increases false non-match rate (higher security, lower convenience). A lower match threshold reduces the false non-match rate but increases false match rate (higher convenience, lower security; See Figure 3). Higher quantities of data (e.g. more fingerprints) and higher-quality (highly consistent) samples are required for one-to-many search processes as compared to one-to-one matching for verification.

It is important to recognize that biometric system accuracy is highly dependent on the nature of the biometric data in the system. Every different biometric data gallery against which a set of probe samples is searched will yield a different accuracy ROC curve. There are biometric galleries in the public domain, and they serve to provide common benchmarks to compare different matching algorithms. But algorithms can be "trained" to work better on known databases, which is analogous to seeing the questions on a test before taking it. Doing so will improve their comparative accuracy on known databases, but does not necessarily indicate the performance of the system on unknown data, as is the case in a real-world scenario. So the best way to predict how a biometric system will behave in a real-world deployment is to test its performance on data to which it has not been explicitly trained.

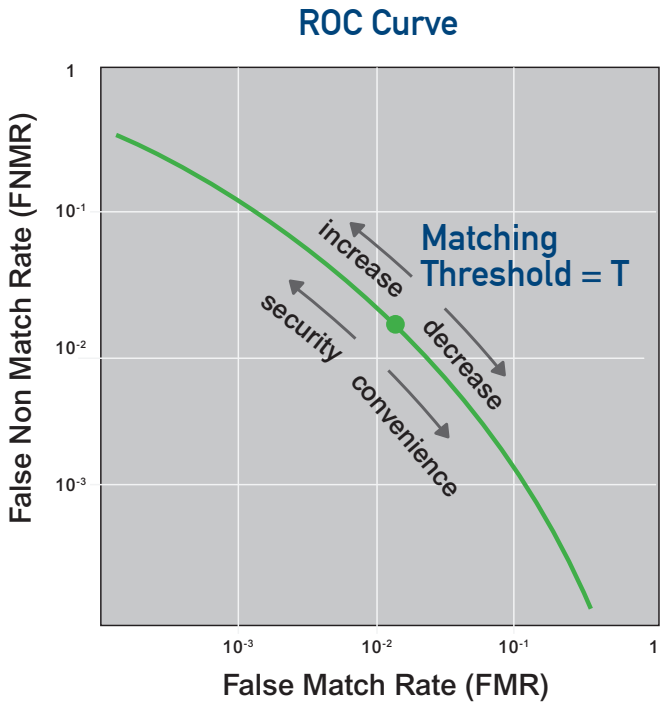

Figure 2 - An ROC curve for a given biometric matching system and dataset

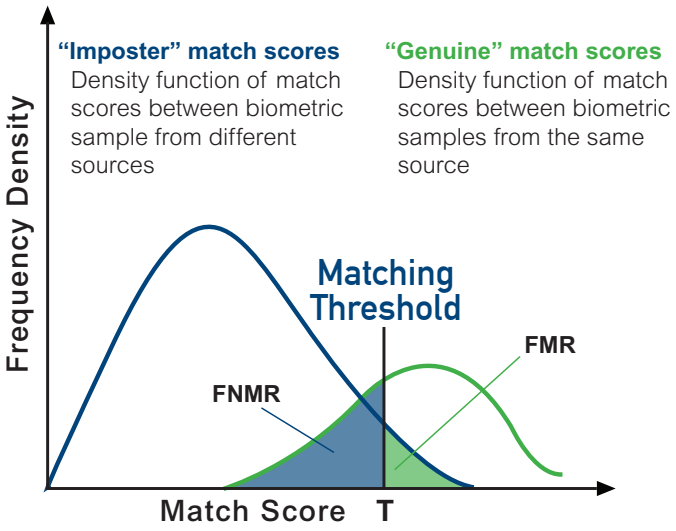

Figure 3 - Density functions of comparison scores between a) samples from different sources and b) samples from the same sources, illustrating FMR and FNMR.

Biometric Applications

The first application of biometrics was the use of fingerprints towards identifying a suspect in a criminal investigation. With the help of modern image capture technologies and powerful computing, this process that was once paper-based and labor-intensive is today largely digital and highly (but not completely) automated. New technology has lent biometric search to other applications; namely "authentication" for various physical and logical access control applications, as well as near real-time biometric identification and watch list search for border control and other applications where results are needed extremely quickly.

Biometric applications can be classified into three purposes: 1) verification, 2) identification, and 3) duplicate checking:

Verification involves performing a "one-to-one" biometric comparison towards securing access to either a physical asset, such as a room or building, or to a digital asset, such as a computer application or database. For this application, biometrics are used much like passwords and PIN codes to enhance access control by performing a comparison of an individual's live biometric sample to a single trusted stored sample. This stored sample might reside either in a central database, smart phone, or as a token on a credential such as a smart card ID. In this way, we can "authenticate" the assertion of a person's identity, answering the question "are you to the person to whom this token was issued?" and using the comparison result to either grant or deny their access. Use of biometrics for access control is of particular interest for commercial or personal security applications. Biometric verification can be offered as a more convenient alternative or enhancement to a PIN or password, in which case the user is offered to use it but can opt to bypass it for PIN or password entry at their discretion. This is the usage model employed by the Apple iPhone 5S, for example.

Identification is a very different and more demanding process (in terms of biometric algorithm and computing performance) that serves to assess whether an individual's biometrics are present in a database, or "gallery". A gallery may contain tens of millions of templates or many more. In this process, an individual's live biometrics are captured and submitted to a biometric search system for "one-to-many" comparison. The system mathematically compares the template from the live probe sample to all the templates in the gallery. In doing so, biometrics help identify an individual even if they are not truthfully identifying themselves. Identification is performed most often for public-sector applications where trusted identity is critical to public safety, including criminal investigation and law enforcement, visa issuance and border management, background checks for employment screening, defense and intelligence.

Duplicate checking is yet another biometric process performed to determine whether there are individuals represented more than once in a database. This might be performed to detect fraud, such as in the case where an individual has enrolled multiple times in a social benefits program. This process involves matching the biometric template of each record in the database to every other, in a process called "biometric deduplication."

Devices and Sensors

Devices and sensors are any mechanical or electronic system used to enroll and capture raw biometric samples in a form that can be digitized and converted to a biometric template. For fingerprints, face, iris, and voice, these are fingerprint sensors, digital cameras, iris cameras, and microphones, respectively. Most fingerprint sensors are based upon either optical or capacitive techniques, but light emitting sensors and multispectral approaches are gaining adoption. Capacitive sensors can be either full-finger or swipe. It is critical to the performance of matchers for fingerprint images to be captured at sufficient resolution (500 ppi) and contrast, be compressed properly with WSQ, and be free of distortion. An optical sensor uses a prism, light source, and light sensor to capture images of fingerprints. Capacitive sensors are based on a silicon chip that detects electrical currents when the finger ridges make contact. Swipe sensors do not generate image quality sufficient for one-to-many identification. Generally speaking, the quantity and consistency of the biometric samples required is a function of the size of the database that must be searched.

Capture of facial images is performed using consumer-grade digital SLRs, pocket cameras, and webcams. Low-cost sensor technology has improved dramatically recently, making biometric facial capture with smart phones also viable. Digital facial images traditionally require an interocular resolution of about 60 pixels for

one-to-one matching and 90 pixels for more accurate one-to-many matching. The more critical and challenging factor affecting facial matcher performance is consistency; achieving consistent pose, head angle, and facial expression of the subject, and brightness, contrast, sharpness, and background clutter of the full image.

Iris biometrics have also benefitted from dramatic improvements in sensors. Iris matching differs from face in that it requires an infrared image of the iris to optimize the image contrast so as to facilitate machine based analysis. The degree to which a pure infrared image can be captured (with minimal "pollution" from visible light), the better matching performance is achievable. This is why off-the-shelf cameras aren't yet used for iris image capture, and a special camera is required; a system must illuminate the iris with infrared light and then filter out other wavelengths.

Their audio capabilities and ubiquity make smart phones a particularly viable means to deploy large-scale voice biometrics for one-to-one verification. Voice biometrics are impeded by the same challenges as facial biometrics in that the capture environment can be unpredictable and inconsistent; as with facial images background noise can interfere with the capture and matching process.

Modes of Use & System Architecture

An "owner-based" biometric application is one by which a single individual uses one-to-one biometric verification to secure access to their own assets, such as a smart phone. A "permission-based" system involves the controller of an asset granting self-access to that asset, (e.g. a company using biometrics to grant employees access to their data). "Operator-based" applications require an authorized, trained operator of the device to collect biometrics from the individual providing the biometric sample, such as in a law enforcement application. "Kiosk-based" applications require capture to be performed by the subject without any training or experience and minimal instruction, such as in automated border control.

The location of the previously enrolled biometric template or templates to which template from a live-captured sample is compared can reside in any of several locations, including within a smart phone, on a credential such as an ID card chip or printed bar code, on a mobile biometric capture device, or on a central server. The location of the enrolled templates and the location where the match is performed are a function of the use case, performance, and security of the application. One-to-one biometric comparison can even be performed entirely on the chip of a smart card.

Privacy

Governments collect personal information about its citizens, typically in the interest of improved social, medical, or physical security of some kind. Not everyone agrees on how much of this personal information is too much, and biometrics tend to epitomize the type of personal information considered by some to be too much. The historical use of biometrics by government law enforcement agencies as a tool for criminal booking and investigation perpetuates their association with the forfeit of personal rights. In some parts of the world there is a history of abuse of personal information that has forged a strong aversion to its possession by governments. Although private corporations today possess, utilize, and transact upon vastly greater quantities of personal data, we tend to perceive it to be more innocuous, and that we are getting something in return, such as use of their products.

More recently, the proliferation of the Internet, digital cameras, smart phones, and social media has introduced the era of "Big Data", and with it has come an exponential increase in the availability of personal data and potential for its abuse. We are learning that in this new era, privacy is a very personal choice; some individuals choose to minimize the amount of personal information they share, while others "overshare" enthusiastically. In either case, biometrics have the potential to provide a more convenient and secure means to improve privacy through better control of access to an increasingly vast abundance of personal information, particularly when used in conjunction with other traditional security mechanisms such as PINs and passwords.

The abundance of facial images on the Internet presents an opportunity to abuse them as biometrics. It is conceivable that through a process of "identity resolution," facial images and their associated data (e.g. name, school, associates, etc.) can be bonded via

biometric facial matching with information from different websites and databases where the facial images are stored. Identity resolution is a process by which otherwise disparate, "siloed" data is aggregated into a "digital identity" that comprises a more comprehensive view of a person than exists from any individual data source. Where small and scattered amounts of personal information had been made available—each with a particular use and audience in mind—the aggregation of this personal data from multiple sources made possible with a biometric facial search can pose a privacy threat. It should be noted that it is not clear whether this has actually ever been done in such a way that has impacted someone's privacy. Furthermore, this process is more traditionally (and perhaps more effectively) performed using text-based data, and so the potential threat exists with or without the presence of facial images. It's also worth noting that other biometric modalities do not pose the same risk as facial images for this type of process because they don't exist abundantly in the public domain. In assessing the impact of biometrics on privacy, it is critical to consider them in a larger context of all text- and signal-based identity data; this includes data that is held by government agencies, by private entities, available on the Internet, and from other open sources.

Security

There have been very few documented accounts of successful fraudulent defeat of biometric security measures in a real-world system either to avoid identification or gain unauthorized access. Attempts are occasionally simulated by journalists and widely publicized, and so there tends to be an outsize perception of the threat of security holes posed by biometrics.

The first threat scenario is where an individual in some way obfuscates their biometric samples in order to avoid identification, such as by fingerprint mutilation or iris dilation. These are not terribly effective because they are highly detectable, and in the case of mutilation, irreversible.

The second scenario is one where a biometric sample is covertly obtained or fabricated by an impostor and somehow faked or "spoofed" to fraudulently gain entry or access to the rightful owner's assets, just as they might through use of a stolen PIN, password, or credential. But while passwords can be changed and reissued to the genuine user, the inherent permanency of biometrics precludes them from being changed, and so the secure use of that biometric modality in the future is conceivably compromised, at least until the impostor is so identified.

Spoofing or obfuscating a biometric requires skill and effort and is extremely difficult to achieve without detection. While it is conceivably possible, it is particularly difficult, unreliable, and ineffective in situations where biometric capture is multi-sample, multimodal, attended by an operator, or used with other security mechanisms. Improvements in "liveness detection" (e.g. blood flow, blink, and pupillary pulsation detection) and other anti-spoofing techniques make most failure modes virtually impossible. Another technique is to issue "revocable" biometrics, which are encoded and matched only in an encrypted domain. They are secure and can be regenerated if compromised.

Virtually every security mechanism can be defeated with some degree of skill and effort, and biometrics are no exception. The security of biometrics should be considered in the context of their application in relative terms to other alternative security mechanisms. It is also important to use biometrics in concert with other security measures; no security mechanism should break down under a single mode of failure.

About Aware, Inc.

Aware is a leading provider of biometrics software products and development services to government departments, system integrators, and solution suppliers globally. Our products include SDKs, software components, workstation applications, and a modular, centralized, service-oriented platform. They fulfill a broad range of functions critical to biometric authentication and search using fingerprints, face, and iris, including sample autocapture, image quality assurance, abstraction of capture hardware peripherals, centralized data processing and workflow, subsystem connectivity, and biometric matching algorithms. The products are used to enable identity-centric security solutions with biometrics for applications including law enforcement, border management, credentialing and access control, and defense and intelligence. Aware is a publicly held company (NASDAQ: AWRE) based in Bedford, Massachusetts.

AWARE

Please contact Aware or visit our website for additional information:

sales@aware.com
www.aware.com