Biometric System Security

Andy Adler

Systems and Computer Engineering
Carleton University, Ottawa, Canada adler@sce.carleton.ca

Security is "freedom from risk or danger", while computer and data security is "the ability of a system to protect information and system resources with respect to confidentiality and integrity". Defining biometrics system security is difficult, because of the ways biometric systems differ from traditional computer and cryptographic security [40]. Implicit in all definitions is the concept of an attacker; however, biometrics should always be assumed to operate in an (at least somewhat) hostile environment – after all, why should one test identity if all can be trusted? The ability of a biometric system to stand up to "zero-effort" attackers is measured by the false accept rate (FAR). Attackers may then change makeup, facial hair and glasses, or abrade and cut fingerprints in order to avoid being recognized; attackers prepared to try harder may use spoofing . This chapter deals with attacks which are not spoofing, but those that target processing within the biometric system.

We define biometric system security by its absence. Since biometrics is "automated recognition of individuals based on their behavioral and biological characteristics", a vulnerability in biometric security results in incorrect recognition or failure to correctly recognize individuals. This definition includes methods to falsely accept an individual (template regeneration), impact overall system performance (denial of service), or to attack another system via leaked data (identity theft). Vulnerabilities are measured against explicit or implicit design claims.

1 Biometrics Security Overview

The key design challenge for biometric algorithms is that people's biometric features vary, both with changes in features themselves (cuts to fingers, facial wrinkles with age) and with the presentation and sensor environment (moisture on fingerprints, illumination and rotation of a presented iris). A biometric algorithm must reject "natural" and environmental changes to samples, while focusing on those which differ between individuals. This chapter concentrates on system vulnerabilities which are a consequence of this core biometric challenge. Since biometric systems are implemented on server computers, they are vulnerable to all cryptographic, virus and other attacks which plague modern computer systems [15]; we point out these issues, but do not cover them in detail.

Maltoni *et al.* [27], classify biometric system vulnerabilities as follows:

- *Circumvention* is an attack which gains access to the protected resources by a technical measure to subvert the biometric system. Such an attack may subvert the underlying computer systems (overriding matcher decisions, or replacing database templates) or may involve replay of valid data.
- *Covert acquisition (contamination)* is use of biometric information captured from legitimate users to access a system. Examples are spoofing via capture and playback of voice passwords, and lifting latent fingerprints to construct a mold. This category can also be considered to cover regenerated biometric images (Sec. 3). For example, a fingerprint image can be regenerated from the template stored in a database (and these data can be captured covertly [16]). Covert acquisition is worrisome for cross-application usage (eg. biometric records from a ticket for an amusement park used to access bank accounts).
- *Collusion and Coercion* are biometric system vulnerabilities from legitimate system users. The distinction is that, in collusion, the legitimate user is a willing (perhaps by bribe), while the coerced user is not (through a physical threat or blackmail). Such vulnerabilities bypass the computer security system, since the biometric features are legitimate. It may be possible to mitigate such threats by automatically detecting the unusual pattern of activity. Such attacks can be mounted from both administrator and user accounts on such a system; attacks from user accounts would first need to perform a privilege escalation attack [15].
- *Denial of Service (DoS)* is an attack which prevents legitimate use of the biometric system. This can take the form of slowing or stopping the system (via an overload of network requests) or by degrading performance. An example of the latter would be enrolling many noisy samples which can make a system automatically decrease its decision threshold and thus increase the FAR. The goal of DoS is often to force a fall back to another system (such as operator override) which can be more easily circumvented, but DoS may be used for extortion or political reasons.
- *Repudiation* is the case where the attacker denies accessing the system. A corrupt user may deny her actions by claiming that their biometric data were "stolen" (by covert acquisition or circumvention) or that an illegitimate user was able to perform the actions due to the biometric false accept. Interestingly, biometric systems are often presented as a solution to the repudiation problem in the computer security literature [15]. One approach to help prevent repudiation would be to store presented images for later forensic analysis, however, this need must be balanced against user privacy concerns [7].

Another class of biometric vulnerabilities are those faced by the system user, which impact the user's privacy and can lead to identity theft or system compromise [33].

- *Biometrics are not secret:* Technology is readily available to image faces, fingerprints, irises and make recordings of voice or signature – without subject consent or knowledge [40][23]. From this perspective, biometrics are not secret. On the other hand, from a cryptography [6] or privacy [7] perspective, biometric data are often considered to be private and secret. This distinction is important, as our understanding of computer and network security is centered around the use of secret codes and tokens [15]. For this reason, cryptographic protocols which are not robust against disclosure of biometric samples are flawed. One proposed solution is revocable biometrics (Sec 4.1), although the vulnerability of such systems is not well understood.
- *Biometrics cannot be revoked:* A biometric feature is permanently associated with an individual, and a compromised biometric sample will compromise all applications that use that biometric. Such compromise may prevent a user from re-enrolling [40]. Note, however, that this concern implies that biometrics are secret, contradicting the previous consideration.
- *Biometrics have secondary uses:* If an individual uses the same biometric feature in multiple applications, then the user can be tracked if the organizations share biometric data. Another aspect to this problem is *secondary use* of ID cards. For example, a driver's license is designed with the requirements to prove identity and driver certification to a police officer, but it is used to prove age, name and even citizenship. Similarly, biometric applications will be designed with a narrow range of security concerns, but may be used in very different threat environments.

Biometric systems form part of larger security systems and their risks and vulnerabilities must be understood in the context of the larger system requirements. An excellent review of the security of biometric authentication systems is [23]. Each assurance level from "passwords and PINs" to "Hard crypto token" is analyzed to determine which biometric devices are suitable.

2 Vulnerabilities in Biometric Systems

In order to classify biometric security vulnerabilities, it is typical to study each subsystem and interconnection in a system diagram (Fig. 1). Early work is presented in [34], with later contributions comming from [9][44][46]. We consider each system module in turn:

2.1 Identity Claim (A):

Identity claims are not biometric properties, but form an essential part of most biometric security systems An example of an exception is in verifying a season ticket holder; the person's identity doesn't matter, and long as he is

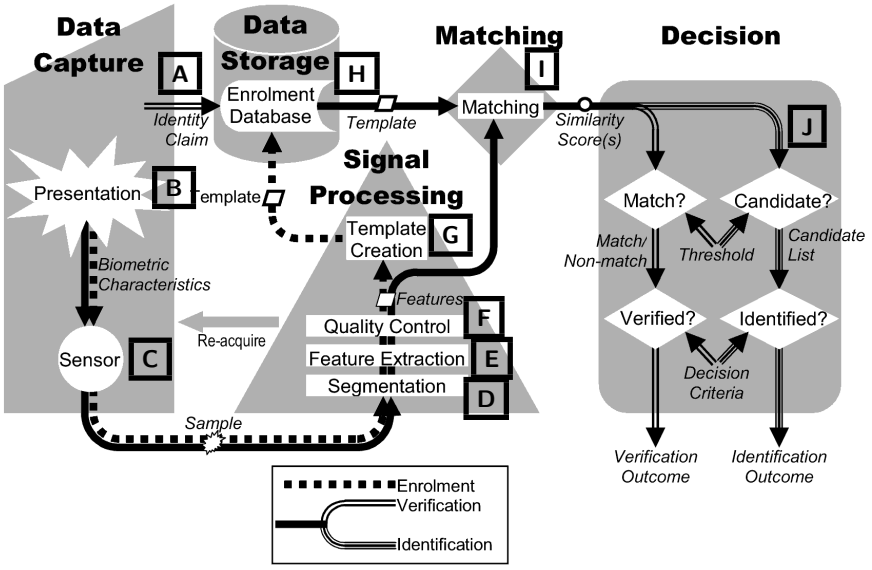

Fig. 1. Biometric System Block Diagram (from [24]). Steps A – H are analyzed in sec. 2. Each presented sample (B) is acquired by a sensor (C) processed via segmentation (D) and feature extraction (D) algorithms. If available, a sample quality (E) assessment algorithm is used to indicate a need to reacquire the sample. Biometric features are encoded into a template, which is stored (H) in a database, on an identity card or in secure hardware. For biometric encryption (Sec. 4.2) systems, a code or token is combined with the biometric features in the template. During enrollment, biometric samples are linked to a claimed identity (A), and during subsequent verification or identification, samples are tested against enrolled samples, using a matching algorithm (I) and an identity decision (J) is made, either automatically, or by a human agent reviewing biometric system outputs.

the one who paid. Identity claims are primarily based on links to government issued identity documents, and are thus vulnerable to all forms of fraud of such documents. This is a problem even for highly secure documents, such as passports, which are often issued on the basis of less secure "breeder documents" [37] such as birth certificates issued by local government, hospital or even religious authorities.

2.2 Presentation (B):

An attacks on the biometric sensor provides false biometric sample into the system. Such attacks are designed to either avoid detection (false negative) or masquerade as another (false positive). The latter attack is typically called spoofing. Clearly, avoiding detection is easier than masquerading, since fea-

tures simply need to be changed enough to confuse the segmentation or feature extraction module. Changing makeup, facial hair and glasses or abrading or wetting fingers is often successful; although recent progress in biometric algorithms has dramatically reduced the effectiveness of such techniques. Knowledge of the details of algorithms can make such attacks easier; for example, rotating the head will confuse many iris algorithms that do not expect image rotation of more than a few degrees.

2.3 Sensor (C):

Attacks on the biometric sensor include any technique which subverts or replaces the sensor hardware. In some cases subverting the sensor allows complete bypassing of the biometric system. For example, in some biometric door locks, the sensor module includes the entire biometric system including a Wiegand output or relay output to activate the solenoid in a door lock. Subverting such a system may be as simple as physically bypassing the biometric system.

In many cases, an attack on the sensor would take the form of a replay. The connection between the biometric sensor and the biometric system is subverted to allow input of arbitrary signals, and images from legitimate users are input into the system. In order to obtain the signals, several strategies may be employed. Eavesdropping requires hiding the recording instruments and wiring of the sensor. For biometrics using contactless smart cards such eavesdropping becomes more feasible (see [16]). Another approach is to record signals from a sensor under the control of the attacker.

2.4 Segmentation (D):

Biometric segmentation extracts the image or signal of interest from the background, and a failure to segment means the system does not detect the presence of the appropriate biometric feature. Segmentation attacks may be used to escape surveillance or to generate a denial of service (DoS) attack. For example, consider a surveillance system in which the face detection algorithm assumes faces have two eyes. By covering an eye, a person is not detected in the biometric system. Another example would be where parts of a fingerprint core are damaged to cause a particular algorithm to mis-locate the core. Since the damaged area is small, it would not arouse the suspicion of an agent reviewing the images.

2.5 Feature Extraction (E):

Attacks of the feature extraction module can be userd either to escape detection or to create impostors. The first category is similar to those of Sec. 2.4. Knowledge of the feature extraction algorithms can be used to design special features in presented biometric samples to cause incorrect features to be calculated.

Characterizing feature extraction algorithms: In order to implement such an attack, it is necessary to discover the characteristics of the feature extraction algorithm. Are facial hair or glasses excluded (face recognition)? How are the eyelid/eyelash regions detected and cropped (iris recognition)? Most current high performing biometric recognition algorithms are proprietary, but are often based on published scientific literature, which may provide such information. Another approach is to obtain copies of the biometric software and conduct offline experiments. Biometric algorithms are likely susceptible to reverse engineering techniques. It would appear possible to automatically conduct such reverse engineering, but we are not aware of any published results.

Biometric "zoo": There is great variability between individuals in terms of the accuracy and reliability of their calculated biometric features. Doddington *et al.* developed a taxonomy for different user classes [14]. *Sheep* are the dominant type, and biometric systems perform well for them. *Goats* are difficult to recognize. They adversely affect system performance, accounting for a significant fraction of the FRR. *Lambs* are easy to imitate – a randomly chosen individual is likely to be identified as a lamb. They account for a significant fraction of the FAR. *Wolves* are more likely to be identified as other individuals, and account for a large fraction of the FAR. The existence of lambs and wolves represents a vulnerability to biometric systems. If wolves can be identified, they may be recruited to defeat systems; similarly, if lambs can be identified in the legitimate user population, either through correlation or via directly observable characteristics, they may be targets of attacks.

2.6 Quality Control (F):

Evaluation of biometric sample quality is important to ensure low biometric error rates. Most systems, especially during enrollment, verify the quality (Chap. **??**) of input images. Biometric quality assessment is an active area of research, and current approaches are almost exclusively algorithm specific. If the details of the quality assessment module can be measured (either though trial and error or through off-line analysis) it may be possible to create specific image features which force classification in either category. Quality assesment algorithms often look for high frequency noise content in images as evidence of poor quality, while line structures in images indicate higher quality. Attacks on the quality control algorithm are of two types: classifying a good image as poor, and classifying a low quality image as good. In the former case, the goal of the attack would be to evade detection, since poor images will not be used for matching. In the latter case, low quality images will be enrolled. Such images may force internal match thresholds to be lowered (either for that image, or in some cases, globally). Such a scenario will create "lambs" in the database and increase system FAR.

2.7 Template Creation (G):

Biometric features are encoded into a template, a (proprietary or standards-conforming) compact digital representation of the essential features of the sample image. One common claim is that, since template creation is a one-way function, it is impossible or infeasible to regenerate the image from the templates [20]. Recent research has shown regeneration of biometric samples from images to be feasible (see Sec. 3).

Interoperability: Government applications of biometrics need to be concerned with interoperability. Biometric samples enrolled on one system must be usable on other vendor systems if a government is to allow cross-jurisdictional use, and to avoid vendor lock-in. However, recent work on interoperability has revealed it to be difficult, even when all vendors are conform to standards. Tests of the International Labour Organization seafarer's ID card [22] showed incompatibilities with the use of the minutiae type "other" and incompatible ways to quantize minutiae angles. Such interoperability difficulties present biometric system vulnerabilities, which could be used to increase FRR or for a DoS attack.

2.8 Data Storage (H):

Enrolled biometric templates are stored for future verification or identification. Vulnerabilities of template storage concern modifying the storage (adding, modifying or removing templates), copying template data for secondary uses (identity theft), or modifying the identity to which the biometric is assigned.

Storage may take many forms, including databases (local or distributed), on ID documents (into a smart card [16] or 2D barcode [22]) or on electronic devices (a hardened token [23], laptop, mobile telephone, or door access module). Template data may be in plaintext, encrypted or digitally signed. In many government applications, it may be necessary to provide public information on the template format and encryption used, in order to reassure citizens about the nature of the data stored on their ID cards, but this may also increase the possibility of identity theft. Vulnerabilities of template storage are primarily those of the underlying computer infrastructure, and are not dealt with in detail here.

Template transmission: The transmission medium between the template storage and matcher is similarly vulnerable to the template storage. In many cases, attacks against template data transmission may be easier than against the template storage. This is especially the case for passive eavesdropping and recording of data in transit for wireless transmission (such as contactless ID cards). Encrypted transmission is essential, but may still be vulnerable to key discovery [16].

2.9 Matching (I):

A biometric matcher calculates a similarity score related to the likelihood that two biometrics samples are from the same individual. Attacks against the matcher are somewhat obscure, but may be possible in certain cases. For biometric fusion systems (Chap. **??**) extreme scores in one biometric modality may override the inputs from other modalities. Biometric matchers which are based on Fisher discriminant strategies calculate global thresholds based on the between class covariance, which may be modified by enrolling specifically crafted biometric samples.

2.10 Decision (J):

Biometric decisions are often reviewed by a human operator (such as for most government applications). Such operators are well known to be susceptible to fatigue and boredom. One of the goals of DoS attacks can be to force operators to abandon a biometric system, or to mistrust its output (by causing it to produce a sufficiently large number of errors) [15].

2.11 Attack Trees

Complex systems are exposed to multiple possible vulnerabilities, and the ability to exploit a given vulnerability is dependent on a chain of requirements. Vulnerabilities vary in severity, and may be protected against by various countermeasures, such as: supervision of enrollment or verification, liveness detection, template anonymization, cryptographic storage and transport, and traditional network security measures. Countermeasures vary in maturity, cost, and cost-effectiveness. In order to analyze such a complex scenario, the factors may be organized into *attack trees*. This analysis methodology was developed by Schneier [39] and formalized by Moore *et al.* [29]. In [39], the example attack "Open Safe", is analyzed to occur due to "Pick Lock", "Learn Combo", "Cut Open Safe" or "Install Improperly". "Learn Combo" may, in turn, occur due to "Eavesdrop", "Bribe" or other reasons, which in turn depend on further factors. The requirements for each factor can be assessed (Eavesdropping requires a technical skill, while Bribing requires an amount of money). Attack trees may be analyzed by assigning each node with a feasibility, the requirement for special equipment, or cost.

Attack tree techniques for biometric system security have been developed by Cukic and Barlow [9]. Figure 2 shows a fraction of the attack tree of [9] for image regeneration from templates [46].

3 Biometric Template Security

Biometric templates carry the most important biometric information, and thus present an important concern for privacy and security of systems. The

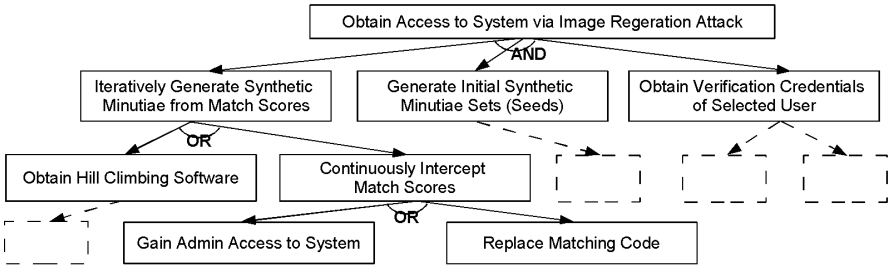

Fig. 2. Attack tree fraction adapted from [9] (dotted blocks represent removed tree portions) to implement the template regeneration attack of [46]. AND/OR nodes indicate that *all*/*one* of the sub-blocks are/is required, Further analysis of the attack tree may be performed by assigning each block a parameter (feasibility, required technical skill, expense) and calculating the cost for the overall attack.

basic concern is that templates may be used to spoof the owner of the document, or for identity theft to another system. Biometric algorithm vendors have largely claimed that it is impossible or infeasible to regenerate the image from the templates [20]; thus biometric templates are sometimes considered to be effectively non-identifiable data, much like a password hash. These claims are supported by: 1) the template records features (such as fingerprint minutiae) and not image primitives, 2) templates are typically calculated using only a small portion of the image, 3) templates are small – a few hundred bytes – much smaller than the sample image, and 4) the proprietary nature of the storage format makes templates infeasible to "hack". In this section, we consider two pathways to regenerate images from templates: 1) from the template directly, based on a knowledge of the features, and 2) from match score values from a biometric algorithm.

3.1 Image Regeneration from Templates

The goal of image regeneration from a biometric template is to compute an image which best matches the feature values in the template. In order to regenerate images in this way, it is necessary for templates to be available in unencrypted form. Thus, encryption of template data storage does impede this vulnerability; however, templates must be available in unencrypted form to perform matching, and are vulnerable at that point.

Published work on image regeneration from templates is for fingerprints, for the reason that regeneration is trivial for most iris and face recognition templates, in which the template features are based on subspace image transforms. If feature vector, $\mathbf{y}$, is computed from an image, $\mathbf{x}$ using a transform that can be approximated by $\mathbf{y} = \mathbf{H}\mathbf{x}$ for a convolution matrix, $\mathbf{H}$, then a reconstructed image, $\hat{\mathbf{x}}$, can be computed from $\hat{\mathbf{x}} = \mathbf{H}^{\dagger}\mathbf{y}$ using a pseudo-inverse $\mathbf{H}^{\dagger}$.

Hill [19] developed an ad-hoc approach to calculate an image from the template of an unspecified fingerprint system vendor. Software was designed to create line pattern images which had a sufficient resemblance to the underlying ridge pattern to be verified by the match software. This work also devised a simple scheme to predict the shape (class) of the fingerprint using the minutiae template. The algorithm iterated over each orientation, core and delta position keeping the image with the best match score. It is worth noting the line patterns do not visually resemble a fingerprint, although these images could be easily improved manually or automatically.

More recently, Ross *et al.* [36] have demonstrated a technique to reconstruct fingerprint images from a minutiae description, without using match score values. First, the orientation map and the class are inferred based on analysis of local minutiae triplets and a nearest neighbor classifier, trained with feature exemplars. Then, Gabor-like filters were used to reconstruct fingerprints using the orientation information. Correct classification of fingerprint class was obtained in 82% of cases, and regenerated images resembled the overall structure of the original, although the images were visually clearly synthetic and had gaps in regions which lacked minutiae. Another valuable contribution of this work is calculation of the probability density fields of minutiae; such information could be used to attack fingerprint based biometric encryption schemes (Sec. 4.2).

3.2 Image Regeneration from Match Scores

Image regeneration from match score values does not require access to the template, and, therefore, template encryption is not a countermeasure. Instead, the requirements are: the ability to present arbitrary images for matching against a target, and access to calculated match scores. The goal is to: 1) determine an image which matches against the target for the specific biometric algorithm, and 2) determine a good estimate of the original image. Clearly, if one can test arbitrary images, one could mount a brute force attack. Given a biometric database of sufficient quality and variety, it should be possible to attain the first goal in approximately 1/FAR attempts. A brute-force attack would be guaranteed to succeed in the second goal, but the size of image space is extremely large.

Brute force searches would only be necessary if biometric image space were random, and nothing could be learned from the output of previous tests. Soutar *et al.* [41] first proposed the possibility of "hill-climbing" in order to practically regenerate images from match score data. A hill-climbing algorithm functions as follows:

1. *Initial image selection*: Choose an initial image estimate (IM). Typically, a sample of initial biometric patterns are tested and the one with the largest match score, MS, is selected.
2. *Iterative estimate improvement*:

(a) Modify IM (to get IM_{test}) in a random, but biometrically reasonable way (details below).
(b) Calculate MS_{test} for IM_{test}.
(c) If $MS_{test} > MS$, set $IM = IM_{test}$ and $MS = MS_{test}$.
(d) End iterations if MS is no longer increasing.

The only difficulty to a practical implementation of this algorithm is to implement "biometrically reasonable" modifications. For face images, Adler [4] added a small factor times a PCA (eigenface) component to the face image. For fingerprint minutiae, Uludag and Jain [46] made modifications to perturb, add, replace, or delete an existing minutiae point at each step. The key constraint is that such modifications attempt to maintain "biometric feasibility" in the search space. Other image modifications, such as changing random pixels in the image, do not converge under hill-climbing.

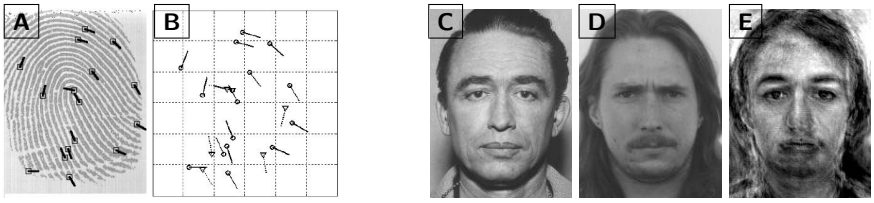

Fig. 3. Regenerated images using hill climbing techniques. A – B: Regenerated fingerprint minutiae (from [46]). The target fingerprint with labeled minutiae (A), and regenerated minutiae positions (B). C – E: Regenerated face images (from [4]). The target face image (C); the initial selected image for hill climbing (D), and regenerated face image (E).

In fact, "hill-climbing" algorithms are simply one type of multi-dimensional optimization algorithm. Other methods for unconstrained minimization (or maximization) such as the Nelder-Mead simplex perform equally or better than hill-climbing (unpublished observations).

In order to protect against regeneration of biometric images, Soutar *et al.* [41] suggested that match score output be quantized to a limited set of levels. The idea is that small image modifications are unlikely to push the MS up by one quantum, so that the hill-climbing algorithm will not see the effect of its changes. This recommendation is maintained in the BioAPI specification [5]. However, by an appropriate modification of the algorithm, Adler showed that hill-climbing could still function [3]. Each hill-climbing iteration is applied to a quadrant of IM. Before each calculation, noise is added to the image in the opposite quadrant, in order to force the match score to a value just below the quantization threshold. This means that the quantized match score is brought into a range where it provides useful information. Images were successfully regenerated for quantization levels equal to a 10% change in FAR.

These results suggest that biometric images can generally be regenerated if: 1) arbitrary images can be input into the biometric system, and 2) raw or quantized match score values are output. The images calculated are of sufficient quality to masquerade to the algorithm as the target, and give a good visual impression of the biometric characteristics. In order to prevent this attack, it is necessary to either limit image input, or to provide only Match/Non-match decisions.

4 Encoded Biometric Schemes

Classical biometric systems require access to enrolled templates in uncoded form. This differs from traditional computer security systems, where a raw password need never be stored. Instead, a cryptographic hash (one-way function) of the password is stored, and each new test password is hashed and compared to the stored version. Since such cryptographic techniques provide important protections, there is great incentive to develop analogous methods for biometric systems. Encoded encryption techniques are designed to avoid these problems by embedding the secret code into the template, in a way that can be decrypted only with an image of the enrolled individual [11][42]. Since the code is bound to the biometric template, an attacker should not be able to determine either the enrolled biometric image or secret code, even if they have access to the biometric software and hardware. Such technology would enable enhanced privacy protection, primarily against secondary use of biometric images [7][45]. It would also reduce the vulnerability of network protocols based on biometrics [23]. Biometrically enabled computers and mobile phones currently must hide passwords and keys in software; biometric encryption would protect against this vulnerability. Another interesting application is for control of access to digital content with the aim of preventing copyright infringement. Biometrically encoded digital documents are subject to attacks, especially since both the documents and the software to access them will be widely distributed [21]. Currently, to the best of our knowledge, biometric encryption systems are not widely deployed; research systems still suffer from high error rates and slow processing speed. However, such systems offer some compelling benefits for many applications, and research is active (eg. [47] [18] [26] [33] [38]).

4.1 Revocable Biometrics

Revocable biometrics are encoded with a distortion scheme that varies for each application. The concept was developed by Ratha *et al.* [35] (and clarified in [34][33]), to address the privacy and security concerns that biometrics are not secret and cannot be canceled. During enrollment, the input biometric image is subjected to a known distortion (Fig. 4) controlled by a set of distortion parameters. The distorted biometric sample can then be processed

with standard biometrics algorithms, which are unaware that the features presented to them are distorted. During matching, the live biometric sample must be distorted in exactly the same way, otherwise it cannot match the enrolled sample. This distortion must also satisfy the constraint that multiple different distortion profiles cannot match. Thus, the revocable nature of this scheme is provided by the distortion, in that it is not the user's "actual" biometric which is stored, but simply one of an arbitrarily large number of possible permutations. One key advantage of this scheme is that it is independent of the biometrics matching algorithm.

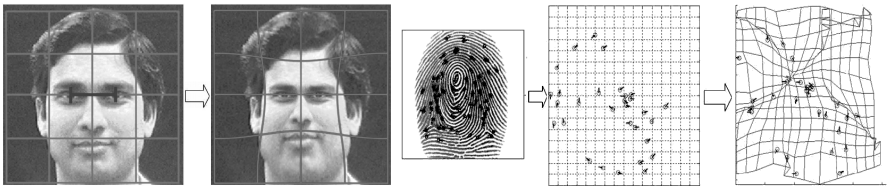

Fig. 4. Distortions of images to implement revocable biometrics. *Left:* a distorted face image centered at the eyes (from [34]), and *Right:* a fingerprint minutiae set distorted spatially and in minutiae angle (from [33]).

For faces, the distortion takes place in the raw image space [34], since face recognition feature sets are not standardized. This places tight constraints on the nature of the distortion, since severely distorted faces will not be recognized and properly encoded by the algorithms (note that the face image in Fig. 4 was not part of an implemented system). A different approach is taken by Savvides *et al.* [38] in which the revocable distortion is tied to a face recognition algorithm based on correlation filters. Enrolled and test face images are distorted with a random kernel calculated from a key to generate an encrypted correlation filter. Since the same convolution kernel is present for both images, its effect is mathematically canceled in the correlation filter. This scheme is somewhat similar to the biometric encryption approach of Soutar *et al.* [42].

A theoretical approach to revocable biometrics uses shielding functions [26], to allow a verifier to check the authenticity of a prover (user wanting to be verified) without learning any biometric information, using proposed δ-contracting and ϵ-revealing functions. The proposed system was based on simple Gaussian noise models and not tested with an actual biometric system. Unfortunately, it is unclear how practical functions can be found which account for the inherent biometric feature variability.

The cancellable fingerprint templates of [33] use the minutiae rather than the raw image, since this allows both minutiae position and angle to be permuted (increasing the degrees of freedom of the transformation), and since distortion will interfere with the feature extraction process. The distortion is modeled on the electric field distribution for random charges. Results show

a small impact on biometric errors (5% increase in FRR) over undistorted features.

While revocable biometrics represent a promising approach to address biometric security and privacy vulnerabilities, we are unaware of security analyses of such schemes, so the security strength of such a transformation is unclear. More significantly, it appears trivial to "undistort" the template given knowledge of the distortion key. Since such keys will presumably not be much better protected than current passwords and PINs, in many application scenarios there is no security advantage of such revocable schemes over an encrypted traditional template.

4.2 Biometrics Encryption

Biometric encryption seeks to use the biometric sample as a key to conduct cryptographic protocols. Normally the biometric template is bound to a secret key which is designed to only be recoverable with a biometric image from the enrolled individual. The primary difficulty in designing biometric encryption systems is the variability in the biometric image between measurements [13]. This means that the presented biometric image cannot itself be treated as a code, since it varies with each presentation. For biometric encryption systems, this variability becomes especially difficult. An algorithm must be designed which allows an image from the enrolled person, with significant differences from the original, to decode the complete secret code. At the same time, an image from another person – which may only be slightly more different from the enrolled image – must not only not decode the secret, it must not be allowed to decode (or "leak") any information at all.

The earliest biometric encryption system was proposed by Soutar [42][43]. Enrollment requires several sample images and a secret code, and creates a template binding the code to the images. During enrollment, an average image f_0 is obtained (with 2D Fourier transform F_0) from multiple samples of the input fingerprint, after suitable alignment. In order to encode the secret, a random code is chosen and encoded as a phase-only function R_0. A Wiener inverse filter is calculated, $H_0 = (F_0^* R_0^*) \,/\, \left(F_0^* F_0 + N^2\right)$, where N^2 is the image noise power. As N increases, an image more dissimilar from the one enrolled can decrypt the code, at the expense of a smaller secret (in bits). In order for biometric encryption to allow for variability in the input image, the secret code must be robustly encoded, using an error correcting code (ECC); [42] uses Hamming distances and majority decision. During key release, a new image, f_1, is acquired. This image is deconvolved with the filter H_0 to calculate $R_1 = sign(imag(H_0 F_1))$, an estimate of R_0. If F_1 is from the same individual as F_0, then R_1 should be a good estimate of R_0; but since $R_1 \neq R_0$, some phase elements will be incorrect. However, if R_1 is sufficiently close, the ECC should allow the correct value of the secret to be obtained.

A somewhat similar scheme was proposed for voice passwords by Monrose *et al.* [28], in which a vector of features is calculated. From this vector each

value is used to select a fraction of the key bits from a table. A correct feature value during key release will select correct key bits while an incorrect value will select a table entry with random data. For features determined to be less reliable, correct key bits are put in all table positions. Reported error rates were FRR = 20%; however, it would seem that such a scheme could make better use of an ECC, since a single feature error will prevent code release.

Hao *et al.* recently proposed a biometric encryption scheme for Iris images based on similar techniques [18]. During enrollment, an encoded key is XORed with the 2048 bit iris code to produce an encrypted code. Variability in the iris is due to background random errors, and to burst errors from undetected eyelashes and specular reflections. The key is encoded with a Hadamard code to protect against background errors, and with a Reed-Solomon code to protect against burst errors. During key release, the encrypted code is XORed with a new iris code sample, and Hadamard and Reed-Solomon decoding are used to correct for errors in the key. Rotation of the iris is handled by iteratively shifting the observed iris codes and attempting decoding. Results show FRR = 0.47% for a key length of 140 bits. In terms of security, the authors note that iris images have significant spatial correlations, which will be preserved in such a linear cryptographic scheme.

Fig. 5. Schematic diagram of the biometric encryption scheme of [8]. *Left:* a raw fingerprint image is enrolled. *Middle:* minutiae points (circles) are used to encode the value of a polynomial representing the secret. *Right:* chaff points (squares), sufficiently far from minutiae, are used to encode random values of the polynomial.

More recent work in biometric encryption has been done in the cryptography community, with much based on the *fuzzy vault* construction of Juels and Sudan [25]. This scheme allows a cryptographic encoding with a variable number of un-ordered data points, which makes it suitable for fingerprint minutiae. This approach has been pursued by Dodis *et al.* [13], who develop the concept of a *fuzzy extractor* which extracts uniformly random and error-tolerant bits from its input, and a *secure sketch* which produces public output that does not reveal the input. Boyen *et al.* [6] further develop this scheme for secure remote authentication. Unfortunately, neither work clarifies how to use these frameworks in a practical biometric application. Not all biometric encryption schemes use a key; for example, in [10][11], the biometric

image forms a unique key, although the results of Linnartz *et al.* [26] suggest encryption schemes based on the biometric only are inherently vulnerable.

Based on [25], Clancy *et al.* [8] designed a fingerprint algorithm which encodes the secret as the coefficients of a Galois field polynomial (Fig. 5). After alignment, minutiae points are encoded as pairs (x_i, y_i) where x_i is a minutiae point, and y_i is a point on the polynomial. Additionally, numerous "chaff" points are encoded, in which the value of y_i is random. During key release, the minutiae of the new fingerprint image are calculated, and the points x_i closest to the minutiae are chosen. The y_i corresponding to these points are used to estimate the polynomial, using a Reed-Solomon ECC framework. If enough legitimate points are identified (equal to the number selected at vault design), the correct polynomial will be obtained and the correct secret decrypted. An interesting generalization of this scheme is given by the "secure sketches" of [13].

Little work has been done to attack biometric encryption schemes, and their security is thus mostly unknown. In their analysis, Uludag *et al.* [47] note that most proposed biometric encryption systems only appear to account for a "limited amount of variability in the biometric representation." In order to quantify this notion, experiments were conducted to estimate the variability in fingerprint minutiae. Matched fingerprint pairs were imaged and minutiae locations identified by a human expert, which was assumed to give an upper bound on system performance. Using these data, the algorithm of [8] was analyzed to estimate the ROC curve during key generation and key release, with an equal error rate of approximately 6.3%. This suggests that biometric encryption systems can be attacked simply via the FAR, by presenting biometric samples from a representative population. A cryptographic attack of biometric encryption was developed by Adler [2], based on using any "leaked" information to attempt a "hill-climbing" of the biometric template, using the quantized *MS* hill-climbing algorithm. This approach was used to reconstruct a face image from a biometric encryption scheme based on [42][43].

Based on the success of these early attacks, we feel that biometric encryption schemes have significant remaining vulnerabilities. Although some schemes offer security proofs (ie. [25][13]) these depend on invalid models of the biometric data. Biometric data inherently has strong internal correlations, many of which cover the entire image. Another important area for attack is the requirement for segmentation and alignment of images before comparison can take place. In a practical system, such as that of Uludag *et al.* [48], carefully selected data are made available to permit alignment with a minimum of "leaked information". Thus, we feel that, in general, current biometric encryption schemes have unknown security value.

4.3 Measures of biometric information content

The information content of biometric samples (or biometric feature entropy) is related to many issues in biometric technology. For example, one of the most common biometric questions is that of uniqueness – "are fingerprints unique?" [31] Such a measure is important for biometric system vulnerabilities, especially as a measure of the strength of cryptosystems and for privacy measures. It also is relevant for applications such as biometric fusion, where one would like to quantify the biometric information in each system individually, and the potential gain from fusing the systems.

Several approaches have been taken to answer this question. Wayman [49] introduced a statistical approach to measure the separability of Gaussian feature distributions using a "cotton ball model". Daugman [12] developed the "discrimination entropy" to measure the information content of iris images. This value has the advantage that it is calculated directly from the match score distributions, but it is unclear how it relates to traditional measures of entropy. Golfarelli *et al.* [17] showed that the most commonly used feature representations of hand geometry and face biometrics have a limited number of distinguishable patterns, on the order of 10^5 and 10^3, respectively, as measured by a theoretical estimate of the equal error rate. Penev *et al.* [32] determined that the dimentionalityh of the PCA subspace necessary to characterize the identity information in faces is in the range 400–700. Biometric encryption studies calculate 46 bits in spoken passwords [28], and 69 bits in fingerprints [8]. Adler *et al.* [1] developed a measure of biometric information in terms of the relative entropy $D(p\|q)$ between the population (inter-class) feature distribution q and the individual (intra-class) distribution p, and calculated an information content for various face recognition feature representations to be between 37 and 45 bits. In this work, the term *biometric information* is defined as the "decrease in uncertainty about the identity of a person due to a set of biometric measurements". Biometric information content is still an open field, with no consensus on techniques used. All cited work measures the information content of a given feature representation, and not that of the biometric sample itself.

5 Discussion

Our understanding of biometrics system security is in its early stages – much more so than many aspects of biometric recognition algorithms. This is perhaps to be expected; people needed to be convinced the technology would work at all, before it was worth trying to understand when it failed.

It is also worth noting that many privacy issues associated with biometric systems are closely related to the security vulnerabilities. Thus, according to Cavoukian [7],

> The threat to privacy arises not from the positive identification that biometrics provide best, but the ability of third parties to access

this data in identifiable form and link it to other information, resulting in secondary uses of the information, without the consent of the data subject.

Based on this understanding, a biometric requirement list was developed to include: original biometric image must be destroyed, biometrics must be encrypted, biometrics used only for verification, fingerprint image cannot be reconstructed, and finger cannot be used as a unique ID. The other significant privacy concern is that "we only have 10 fingers" – biometric data loss is catastrophic in the sense that it cannot be replaced [40]. While there are many promising developments that address these issues, such as biometric encryption (Sec. 4.2), revocable biometrics (Sec. 4.1), or work to de-identify images [30], unfortunately, our analysis in this chapter suggests that currently mature biometric technology is unable to properly address these privacy concerns in the way they are stated.

At the same time, biometric systems are being used in many scenarios with high security value. Vulnerabilities and attack scenarios have been carefully considered and well thought out recommendations are available (eg. [23]). Recent work in standards bodies has given much thought to security standards for biometrics (Chap. **??**). In summary, biometrics system security is challenged by many vulnerabilities, from the biometrics system, the computer infrastructure which supports it, and the users it identifies. However, biometrics can also provide (with careful use) the identity assurance that is foundational to systems security.

6 References

References

1. Adler A, Youmaran R, Loyka S (2005) Information content of biometric features. In Proc. Biometrics Consortium Conference, Washington DC, USA
2. Adler A (2005) Vulnerabilities in biometric encryption systems. In Proc. AVBPA, Tarrytown, NY, USA, LNCS 3546:1100-1109
3. Adler A (2004) Images can be regenerated from quantized biometric match score data. In Proc. Can. Conf. Elec. Comp. Eng. Niagara Falls, Canada, pp. 469–472
4. Adler A (2003) Sample images can be independently restored from face recognition templates. In Proc. Can. Conf. Elec. Comp. Eng. Montréal, Canada, pp. 1163–1166
5. BioAPI Consortium (2001) BioAPI Specification version 1.1
6. Boyen X, Dodis Y, Katz J, Ostrovsky R, Smith A (2005) Secure remote authentication using biometric data. In Advances in Cryptology (EUROCRYPT)
7. Cavoukian A (1999) Privacy and Biometrics. In Proc. Int. Conf. Privacy and Personal Data Protection, Hong Kong, China
8. Clancy TC, Kiyavash N, Lin DJ (2003) Secure smartcard-based fingerprint authentication. In Proc. ACMSIGMM 2003 Multimedia, Biometrics Methods and Applications Workshop pp. 45–52

9. Cukic B, Barlow N (2005) Threats and Countermeasures, In Proc. Biometrics Consortium Conference, Washington DC, USA
10. Davida GI, Frankel Y, Matt BJ, Peralta R (1999) On the relation of error correction and cryptography to an offline biometric based identification scheme. In Proc. Conf. Workshop Coding and Cryptography (WCC99) pp. 129–138
11. Davida GI, Frankel Y, Matt BJ (1998) On enabling secure applications through off-line biometric identification. In Proc. IEEE Symp. Privacy and Security, pp. 148–157
12. Daugman J (2003) The importance of being random: Statistical principles of iris recognition. Pattern Recognition 36:279–291
13. Dodis Y, Reyzin L, Smith A (2004) Fuzzy Extractors and Cryptography, or How to Use Your Fingerprints. In Proc. Eurocrypt'04
14. Doddington G, Liggett W, Martin A, Przybocki N, Reynolds D (1998) Sheep, Goats, Lambs and Wolves: An Analysis of Individual Differences in Speaker Recognition Performance. In Proc. Int. Conf. Auditory-Visual Speech Processing, Sidney, Australia
15. Ferguson N, Schneier B (2003) Practical Cryptography. John Wiley & Sons, NJ, USA
16. The Guardian (17 Nov. 2006) Cracked it!
17. Golfarelli M, Maio D, Maltoni D (1997) On the Error-Reject Tradeoff in Biometric Verification Systems. IEEE Trans. Pattern Analysis and Machine Intel. 19:786–796
18. Hao F, Anderson R, Daugman J (2005) Combining cryptography with biometrics effectively. Technical Report UCAM-CL-TR-640, University of Cambridge, Cambridge, UK
19. Hill C (2001) Risk of Masquerade Arising from the Storage of Biometrics, B.S. Thesis, Austrailian National University
20. International Biometric Group (2002) Generating Images from Templates
21. Kundur D, Lin C-Y, Macq B, Yu H (2004) Special Issue on Enabling Security Technologies for Digital Rights Management. Proc. IEEE 92:879–882
22. International Labour Organization (2005) Biometric Testing Campaign Report (Addendum to Part I). Geneva
23. InterNational Committee for Information Technology Standards (INCITS) (2006) Study Report on Biometrics in E-Authentication, Technical Report INCITS M1/06-0693
24. ISO (2006) Standing Document 2, version 5 – Harmonized Biometric Vocabulary. Technical Report ISO/IEC JTC 1/SC 37 N 1480
25. A Juels, M Sudan (2002) A fuzzy vault scheme, In Proc. IEEE Int. Symp. Information Theory, pp. 408
26. Linnartz J-P, Tuyls P (2003) New shielding functions to enhance privacy and prevent misuse of biometric templates. In Proc. AVBPA, Guiford, UK, LNCS 2688:393–402
27. Maltoni D, Maio D, Jain AK, Prabhakar S (2003) Handbook of Fingerprint Recognition. Springer, Berlin
28. Monrose F, Reiter MK, Li Q, Wetzel S (2001) Cryptographic key generation from voice, In Proc. IEEE Symp. Security and Privacy, Oakland, CA, USA, pp. 202-213
29. Moore AP, Ellison RJ, Linger RC (2001) Attack Modeling for Information Security and Survivability. Technical Report CMU/SEI-2001-TN-001, Carnegie Mellon University, Pittsburgh, PA, USA

30. Newton EM, Sweeney L, Malin B (2005) Preserving Privacy by De-Identifying Face Images. IEEE Trans. Knowledge Data Eng. 17:232–243
31. Pankanti S, Prabhakar S, Jain AK (2002) On the Individuality of Fingerprints. IEEE Trans. Pat. Anal. Mach Intel. 24:1010–1025
32. Penev PS, Sirovich L (2000) The Global Dimensionality of Face Space. In Proc. 4th IEEE Int. Conf. Automatic Face and Gesture Recognition, Grenoble, France, pp. 264–270
33. Ratha N, Connell J, Bolle RM, Chikkerur S (2006) Cancelable Biometrics: A Case Study in Fingerprints. In Proc. Int. Conf. Pattern Recognition, 4:370–373
34. Ratha NK, Connell JH, Bolle RM (2001) Enhancing security and privacy in biometrics-based authentication systems. IBM Systems Journal 40:614–634
35. Ratha N, Connell J, Bolle R (2000) Cancelable Biometrics. In Proc. Biometric Consortium Conference, Washington DC, USA
36. Ross A, Shah J, Jain AK (2005) Towards Reconstructing Fingerprints From Minutiae Points. In Conf. SPIE Biometric Technology for Human Identification II, 5779:68–80
37. Salter MB (2004) Passports, Mobility, and Security: How smart can the border be? International Studies Perspectives 5:71–91
38. Savvides M, Vijaya Kumar BVK, Khosla PK (2004) Cancelable biometric filters for face recognition In Proc. Int. Conf. Pattern Recognition pp. 922–925
39. Schneier B (1999) Attack Trees. Dr. Dobb's Journal
40. Schneier B (1999) The Uses and Abuses of Biometrics. Communications of the ACM 42:136, 1999.
41. Soutar C, Gilroy R, Stoianov A (1999) Biometric System Performance and Security. In Proc. Conf. IEEE Auto Identification Advanced Technol
42. Soutar C, Roberge D, Stoianov A, Gilroy R, Vijaya Kumar BVK (1998) Biometric Encryption using image processing. In Proc. SPIE Int. Soc. Opt. Eng. 3314:178–188
43. Soutar C, Roberge D, Stoianov A, Gilroy R, Vijaya Kumar BVK, (1998) Biometric Encryption: enrollment and verification procedures, In Proc. SPIE Int. Soc. Opt. Eng. 3386:24–35
44. C Tilton (2006) Biometrics in E-Authentication: Threat Model. Biometrics Consortium Conference, Baltimore, MD, USA
45. Tomko G (1998) Privacy Implications of Biometrics - A Solution in Biometric Encryption. In 8th Ann. Conf. Computers, Freedom and Privacy, Austin, TX, USA
46. Uludag U, Jain AK (2004) Attacks on Biometric Systems: A Case Study in Fingerprints. In Proc. SPIE-EI 2004, Security, Steganography and Watermarking of Multimedia Contents VI 5306:622–633
47. Uludag U, Pankanti S, Prabhakar S, Jain AK (2004) Biometric Cryptosystems: Issues and Challenges. Proc. IEEE, 92:948–960
48. Uludag U, Pankanti S, Jain AK (2005) Fuzzy Vault for Fingerprints. In Proc. Conf. AVBPA, Tarrytown, NY, USA, LNCS 3546:310-319
49. Wayman JS (2004) The cotton ball problem. In Proc. Biometrics Consortium Conference, Washington DC, USA