Biometrics and Privacy

Stelvio Cimato*, Roberto Sassi and Fabio Scotti

Dipartimento di Tecnologie dell’Informazione, Università degli Studi di Milano, Via Bramante 65, 26013 Crema, Italy

Received: May 7, 2008; Accepted: May 26, 2008; Revised: June 3, 2008

Abstract: By now, biometric systems find already home in a broad range of commercial and institutional applications. They are typically employed to authenticate users before allowing access to restricted areas or services. The interest in biometrics is mainly due to their potential advantages over traditional authentication procedures. Such advantages extend across: security, accuracy, reliability, usability, and friendliness. Unfortunately, with the benefits also possible risks arise. A main concern in the design of biometric authentication systems is the protection of the biometric templates. In fact, potential threats to users’ privacy might derive from the abuse of biometric information. These privacy concerns are the subject of current discussions and often prevent the adoption of biometric systems on a large scale. The aim of this paper is to present the latest advances in the area of biometric authentication systems comprising the construction of biometric templates from actual samples. The discussion will cover research results published in the open literature and patents recently filed with inventions related to the field.

Keywords: Biometric identification, privacy, secure sketch, fuzzy extractor, multimodal biometrics, user authentication.

1. INTRODUCTION

Today, the diffusion of biometric technologies is blooming. This might accounted to two main needs. First, the large diffusion of e-government, e-banking and e-commerce applications and the possibility of malicious behaviours leading to economic loss and frauds set forward the expectation of more stringent methodologies to identify customers. Second, the recent global escalation of terrorist actions brought the national governments to search for secure, fast and non-intrusive identification techniques as core components of homeland security policies. Biometrics seems to be a natural candidate to efficiently address both these needs since it is commonly accepted (even if actually object of discussion) that biometric data are unique for each person and strictly associated to its owner. They are irrevocable, in the sense that the association cannot be changed during the human life and in many cases they are hard to forge. As a consequence biometric features are increasingly used for authentication and identification purposes in a broad variety of institutional and commercial systems.

The usage of fingerprints for identification is several centuries old, but other biometric traits have been taken into account over the years. Traits might be physiological or behavioural and some are more suited than others given a certain application. Many research efforts have tackled the problem of devising practical systems for personal identification and verification relying on biometric data [1]. But given the potential market connected with such technologies, researchers are generally reluctant to reveal many relevant details. Along the line, companies have been proposing commercial systems but the proposed implementations usually do not provide security guarantees in explicit form and rely on proprietary procedures which are generally not disclosed. Moreover to reach a higher degree of confidence in authentication systems, more than one biometric trait might be required for each user to complete the authentication: such systems are termed “multi-modal”.

A sign of the growing interest for biometric technologies coming also from the institutions is the central source of information on biometrics-related activities set up by the USA federal government, available at “http://www.biometrics.gov”. Such website provides a repository of biometrics-related public information, such as documents, reports and guidelines, with the aim to improve collaboration and sharing of information for the development of biometric activities among government agencies and commercial entities involved in both national and international organizations. A catalogue of international commercial companies involved in biometric technologies is also provided in a twin website (http://www.biometricscatalog.org) where more than 180 entries are listed, presenting or selling products and systems for the deployment of biometrics.

Unfortunately, the adoption of biometric technologies brings along new threats, this time to the privacy of the users. Indeed, biometrics such as fingerprints, voice and face are permanently associated with the user and can therefore obviate the need to carry tokens or remember passwords and keys. But, the strict association between each user and his biometric templates raises concerns on possible uses and abuses of such kind of sensible information. For all of these traits privacy is a critical issue since, in the case their digital representations are lost or stolen, they cannot be replaced or modified in any way; the protection of the biometric data is absolutely essential [2]. Given such irrevocability, many people are not generally keen on giving out biometric traits with little assurance that they cannot be stolen or used without an expressed consent. The abuse of biometric information would in fact again expose to the same kind of malicious behaviours they were intended to prevent. Therefore, the adoption of biometric systems in practical situations is object of discussion and often prevented on a...
large scale by the national privacy authorities. Multi-modal systems augment the worries about the protection of biometric information and their diffusion is even more limited. The growing concerns for privacy recently led the National Science & Technology Council Subcommittee on Biometrics, which is the principal means within the executive branch to coordinate science and technology policy across the diverse entities, to release a document dealing with biometrics privacy [3].

Trying to move one step further, a broad and rapidly growing literature is now focused on multidisciplinary approaches, encompassing biometric and cryptographic techniques, in order to satisfy the requirements of practical applications [4, 5]. The task is not straightforward though. The higher variability within different biometrics readings makes them unsuitable to be directly used to secure data. Cryptographic keys need to have zero uncertainty and single-bit difference (in the key or in the message) spoils the possibility of accessing the encrypted data. Secondarily, biometric features cannot be considered as "secrets" since it is possible to capture them to create real or digital artefacts suitable to attack a biometric system [6].

In this work, we present a set of methods and patents dealing with the creation of biometric identifiers for humans capable to effectively protect the user privacy. The outline of the work is as follows. First, in the next Section, we discuss some preliminary issues concerning biometric authentication and privacy problems induced by the adoption of biometric systems. A few design guidelines will be sketched. Then, in Section 2 we present the approaches traditionally used to protect the templates of a biometric system. Next, in Section 3 we briefly present a set of patents which can offer a broad view of the state of the art in privacy protection of biometric templates. Finally, some conclusions are sketched in Section 4.

2. GENERAL BACKGROUND

In this section, we introduce the preliminary concepts needed to deal with biometric applications, discussing then the problem of privacy protection of biometric templates. The section ends with a brief description of the current guidelines, enabling an effective protection of the user privacy when a biometric system is deployed.

2.1. Biometric Authentication and Verification Basics

The USA National Institute of Standards and Technology (NIST) defined a biometric system as a system exploiting “automated methods of recognizing a person based on physiological or behavioural characteristics” (biometric identifiers, also called features). Physiological biometrics is based on data derived from direct measurement of a body part (i.e. fingerprints, face, retina, iris), while behavioural biometrics is based on measurements and data derived from a human action (i.e. gait and signature).

The biometric approach for authentication then differs from traditional methods, which are usually token-based or knowledge-based and rely on secret passwords and PINs, ID cards, keys, passes etc. Biometric methods are based on the elaboration of personal and unique characteristics of individuals which cannot be easily forged or shared. The accuracy of a biometric system can be evaluated by classical techniques [7], even if peculiarities are present: indeed to effectively test biometric systems a great number of volunteers is required or a large database of biometric records must be accessed. During the experiments, the data maintainer has to cope with important problems related to the security and privacy of the biometric records.

A number of attributes can be used to characterize a biometric system and must be taken into account in its design: uniqueness, universality, permanence, measurability, user friendliness, acceptability and circumvention [8,9]. For a given biometric feature, uniqueness refers to the fact that such feature must be unique: an identical feature should not belong to two different persons. Universality means that the feature type is present/occurs in as many people as possible. The Permanence property is related to the need that the feature does not change over time, or at least, its variation occurs slowly. Measurability concerns the capability of measuring the feature with relatively simple technical instruments. User friendliness requires that the measure should be easy and comfortable to be done, and acceptability refers to the people’s acceptance of the measure in their daily lives. Finally, circumvention concerns the toughness to deceive the system by fraudulent methods.

A biometric system is inherently composed of few main parts. The first one, the biometric sensor, fulfils the task of acquiring the sample of the biometric trait itself, generally by mean of a digital picture. The picture is then transmitted to the signal processing section. During the process, sampling, noise and frequency response characteristics of the sensor and transmission channel might alter the signal collected. The signal processing module represents the core of the biometric system and is generally composed by sub-modules implementing: pre-processing functions (i.e. image filtering and enhancement) and features extraction.

Usually, the picture of the trait is not retained after processing, but more typically, a biometric system uses and stores only a mathematical representation of the information extracted by the signal processing module. Such representation will be used to construct or compare against the templates registered during the enrolment phase: the biometric feature. Examples of features are minutiae coordinates and iris-codes. In many cases, the feature extraction is followed by a modelling phase, in which user’s features are analyzed or combined in order to build a template, which can be stored into a biometric system and used for future identification or verification (matching). The reverse process, i.e. reconstructing the features from a given template, can be very hard, and in this case this process gives some guarantees for user’s privacy. On the other hand, some systems store the complete sample data in the storage unit. Storing samples is often deprecated in the literature due to privacy and security issues that will be discussed in the next sections.

The overall structure of a biometric system is depicted in Fig. (1). The scheme is composed of two basic modules: the enroll module creates the template starting from the biometric samples. The verification module performs the verification process starting from the novel biometric
readings and the previously stored template. Each biometric system defines a measure of similarity between features derived from a sample and a stored template. The measure produces a typical index called matching score. Hence, a match/non-match decision may be made according to whether this score exceeds a decision threshold or not. The term transaction refers to an attempt by a user to validate a claim of identity or non-identity by consecutively submitting one or more samples, as allowed by the system decision policy.

In the scheme just described the system checks, the claim of identity of a specific individual, thus it compares a single template previously enrolled and a number of fresh readings. In such configuration, the system is said to perform “verification” of the user. When instead a fresh sample is compared to a whole database of templates, the system is said to “identify” the user through its biometric traits.

2.2. Privacy Issues in Biometrics

In this section, we discuss the privacy issues concerning the practical usage of biometric systems, considering both the perception of the users and the real risks which they could be exposed to. Different perspectives about privacy can also be given with respect to (i) the application context in which biometrics is exploited; (ii) the particular methodology used for the collection of biometric data and (iii) the specific biometric traits employed. Fig. (2) plots a taxonomy of the main privacy issues found when biometric solution are encompassed in real applications.

2.2.1. User Perception and Real Risks

Users commonly perceive biometric authentication and identification techniques as potential threats for their privacy. Few aspects enforce this perception [6]. The acquisition of the biometric traits is considered as an exact and permanent filing of the user’s behaviour, which might lead to the possibility of tracking his activities, even in the far future. Commonly this issue is associated to a sort of “big brother” phobia, in which a superior entity is capable of observing and acquiring knowledge on each activity of the user. A few misunderstandings support such idea. For example, it is very common the thought that most biometric system has 100% identification accuracy (which is rarely the case!). Or that the biometric samples, once sent over a network, can be easily collected by an eavesdropper and used in some malicious applications, stealing the identity of the unaware user.1.

Very interestingly, users often overlook others privacy related problems arising when biometric systems are deployed. A first case concerns the possible usage of biometric information for operating Proscription Lists. For example, a user can be classified on the basis of his previous behaviours or activities in a specific class, and then - as a consequence of this classification - some services and accesses can be denied to him. Examples of this profiling are the black lists collected in call centres and service providers especially designed to identify and to manage users considered as “offending” or “not-collaborative”. Other examples are the “bad-credit” lists filled by many investors and mutual founds. Indeed, proscription lists can be employed also without the adoption of biometric systems (and actually they are), but the deployment of biometric

1Actually, the latter is a well founded concern. In fact, while it should be granted to the user that the biometric information collected should not be used for any other activities than the ones expressly declared, in some cases it is harder to grant this aspect, especially if the biometric samples themselves need to be sent over a network.
technologies could make such situation more and more dramatic. A second point relates to the fact that many biometric features can be used to obtain personal information about the enrolled users, such as medical information of past illnesses or the current (and future) clinical trends. For example, the retinal pattern acquired by biometric systems can produce valuable information about hypertension, diabetes and others illnesses [10]. Much more personal information can be extracted from DNA samples [11]. Finally, for part of the population, the usage of a biometric system is also perceived as uncomfortable or dangerous. For example, the fingerprint sensor - when previously used by other persons and not properly cleaned - can be considered as unpleasant or disgusting. Or face and iris acquisition systems might induce apprehension for the risk of damaging the eyes caused by lasers and/or IR sources.

2.2.2. Applicative Contexts

The real risks for privacy can be analyzed in more detail with respect to both the final application which the biometric system is dedicated to and the biometric traits which are involved [8]. Biometric covert applications (such as the surveillance systems without explicit authorization from the users) are considered to be more risky for users’ privacy. On the other hand, biometric systems for identification or verification that are optional are considered more privacy compliant. In this case, users can decide to avoid the check by the biometric system, and adopt a different identification/verification system. The distinction between behavioural and physiological traits is also relevant. Physiological data (such as fingerprints or iris templates) can be used in applications without the explicit consent of the user. That is related to the fact that physiological traits are most stable in time and they are characterized by very high verification/identification accuracy. On the other side, behavioural traits tend to be less accurate, and, most of the time, they request the user collaboration.

Privacy is considered to be exposed to a greater risk when the biometric system performs identification instead of a simpler verification task. That is related to the fact that the identification process encompasses a “1-to-many” comparison, which, in most cases, is not carried out in the same place of the acquisition (typically, the biometric data is sent through a network to a database for the comparisons).

Also the duration of the retention of the biometric data and the storage method bring risks. The privacy risks are lower in applications where the individuals retain usage rights over the data (for example when the personal biometric information is stored only on a smart card belonging to the user). The worst case is when they are all stored in a central database, out of the users’ control. Best practice notions require that every project which encompasses biometric data retention should always explicitly state its duration, as the risks increase with time. The storage format is relevant: templates are usually carrying less information than the original sample/images. While they are less powerful when used as direct identifiable data, they are privacy-invasive and template protection is an important issue for protecting users’ privacy.

Different risks are present with respect to the sector of application: the biometric setups in the public sector are considered to be more susceptible to privacy invasiveness than the same installations in the private sector as they encompass a larger number of persons.

Also the role of the individuals that use the biometric system has great impact on the privacy risk which increases with: individual, customer, employee, citizen. The most relevant privacy invasion is related to the association of the fundamental rights of the individual to a biometric identity test.
2.2.3. System Design and Data Collection

Another useful taxonomy concerns the design of a biometric system. The International Biometric Group classifies four different classes concerning the privacy protection: Protective, Invasive, Neutral, and Sympathetic [8].

A privacy-protective system is designed to protect or limit the access to personal information, providing a means for an individual to establish a trusted identity. In this case, the biometric systems use biometric data to protect personal information which might otherwise be copied, stolen or misused.

A privacy-sympathetic system limits access/usage to personal data. A privacy-sympathetic approach encompasses the specific design of elements able to protect biometric data from unauthorized access and usage. Also the storage and the transmission of biometric data must be informed, if not driven, by privacy concerns.

In a privacy-neutral system, privacy aspects are not important or the potential privacy impact is slight. Privacy-neutral systems are designed to be difficulty misused with regards to privacy issues, but they do not have the capability to protect personal privacy.

A privacy-invasive system facilitates or enables the usage of personal data in a fashion which is contrary to privacy principles. In privacy-invasive systems personal data are used for purposes broader than what originally intended. Systems which facilitate the linkage of personal data without an individual’s consent, and those in which personal data are loosely protected belong to this class.

The different biometric technologies associated with each biometric trait can inherently produce various levels of privacy risk. The privacy related aspects are summarized by taking into account the following four most significant aspects of the technology itself: (i) capability of processing searches in databases of biometric records (the higher this capability, the higher the privacy risk); (ii) suitability to work in a covert mode (a face recognition system can be more likely used in a covert manner than a classical fingerprint system. The higher this suitability, the higher the privacy risk); (iii) stability in time of the traits employed (the higher the variability in time, the lower the privacy risk) and (iv) databases interoperability and presence of numerous and/or large available databases to process comparisons (the higher the interoperability and the presence of available databases, the higher the privacy risk).

2.3. Guidelines for the Biometric Template Protection

From the discussion in the previous sections, we might summarize that biometric features, samples and templates can hardly be considered as “secrets” since it is possible to capture them with relatively low effort. But, in any case, their protection is absolutely essential both for security and privacy reasons [2]. For what concern privacy, the design and the usage of a biometric system should always respect strict guidelines. These requirements encompass four main points [8]: (i) scope and capabilities of the system; (ii) data protection; (iii) user control of personal data; (iv) disclosure, auditing and accountability of the biometric system. In the following discussion, we refer to two main classes of actors: the users and the operators who manage the biometric system.

2.3.1. Scope and Capabilities of the System

The scope and the functionalities of the system should not be expanded without the explicit and informed consensus of all the users. The termination date of all system functionalities should be provided, or, at least, the temporal retention period must be communicated to the user. In addition, also the amount of biometric information retained into the template should be limited to the minimal needed level. While enrolment data need to be retained, verification data should always be deleted. Only templates should be recorded: any raw data, images and recordings should be deleted as soon as possible during the functioning. Also the collection of other information should not be pursued: no other personal data should be integrated into the biometric template.

2.3.2. Data Protection

The use of proper techniques to protect the biometric data should always be considered. Suitable examples are the adoption of encryption primitives and private networks which must be designed and managed using the state-of-the-art best practices. Systems should also be hosted in secure and controlled areas. These conditions must be ensured during all the life cycle of the biometric system. It is important to note that also the result of the matching phase (the “match”, “non-match”, and errors cases) must be protected and considered as private information. The access to the biometric data should be limited to a well-defined and numerically small group of operators.

Biometric templates should be subject to randomization transformations in order not to suffer from information leakage. Such feature could lead to a wider acceptance from who is worried about the protection of personal information. Furthermore, such system could overcome the legal issues connected to the respect of privacy protection laws, currently ruling in several countries.

2.3.3. User Control of Personal Data

The user must keep the control on her/his biometric data. The biometric system should be used voluntarily, and, in any case, the system must ensure to the user the possibility to be unrolled. In addition, the user should be always able to correct and modify her/his personal data.

2.3.4. Disclosure, Auditing and Accountability

The exact purpose of the biometric system should be explained to the operators and the enrollees. In particular, it must be explained if the biometric acquisition is optional or compulsory. It is important to let users know when the biometric system is used, especially when enrolment and verification or identification phases are carried on. The guidelines suggest also that each operator should be made accountable for the possible misuses/errors perpetrated

---

Footnote:

3For example, a face acquisition can be used for multiple search in different databases with relatively low efforts. Similarly, many - and large - databases of fingerprints templates exist and they can be queried using fingerprints taken with different sensor and techniques.
during the working activities. Also suitable procedures must be considered in order to solve disputes concerning the usage of the biometric system. The owner of the biometric system and the operators must also be able to provide a clear and effective process of auditing when an institution or a third party entity need to perform a critical review of all the modules which compose the biometric system.

The possibility to effectively adopt methods of the protection of the biometric template is related to many features and considerations depending on the kind of the biometric traits which are involved and on the final application of the system. In the following different approach developed for the protection of biometric templates.

3. BIOMETRIC TEMPLATE PROTECTION

Recently, much work has been devoted to the construction of techniques for the protection of biometric templates in biometric based authentication schemes. Several academic research projects, government guided initiatives, standardization activities and commercial products development have been started with the aim of meeting the increased security requirements when biometrics are used in several critical applications.

The storage of biometric templates during the enrolment phase in a more or less secured database (for the successive identification of verification process) as well as their transmission over an unsecured channel has a number of risks for users’ privacy. In case of attack, an adversary could steal the biometric template and from then impersonate the legitimate user: this way the adversary could access private information, or run applications accessing sensible resources. Such situation is much more worrying if the irreversibility of the biometrics is taken into account. The stealing of the biometric templates opens the door to possible abuses of such kind of sensible information. The loss of biometric data is then an important security issue which directly affects the valuation of a biometric authentication system and should be carefully considered to prevent thefts of identity [2].

Several organizations and research groups, such as the European Biometric Identification Technology Ethic (BITE Project) [12], are investigating the legal aspects of biometric technologies in order to define and consider bioethical issues arising from such emerging technologies. In several countries strict rules have been adopted in order to limit the impact of biometric technologies on the privacy of citizens. Biometric authentication schemes often have to face the legal constraints imposed by such directives considering the risk of function creep and data misuse.

3.1. Academic Research

A first approach to protect biometric templates and users’ privacy is the application of a given transformation to the template extracted during the enrolment phase. The transformed template can then be stored in a database and the authentication process can be correctly performed, as depicted in Fig. (3). The transformation usually relies on a known secret or a password. In this way unauthorized access to the stored templates leaves the adversary with a small and unusable amount of data on the biometrics of the attacked user. Such an approach in some way replicates the typical password based authentication schemes where users’ passwords are typically stored in their hashed form. Due to the mono-directionality of the used hash functions, the knowledge of the hashes does not give any information; so if the database has been corrupted the passwords are not compromised.

The authentication process is performed by applying the same transformation to the fresh template and then comparing the result with the stored transformed template. The properties of the hash functions guarantee that the two resulting hashes match if and only if the input are the same \( h(x) = h(x') \) if and only if the input are the same and it is very hard to find a collision, i.e., two different input values holding the same hash. When dealing with biometric templates, things are more complicated since the higher variability within different readings of biometric data makes them unsuitable to be directly used as input for hash functions or as cryptographic keys. Indeed the features extracted for the same individual, will be similar to the data enrolled and give a positive match even if the biometric readings will be not exactly the same.

3.1.1. Hash Based Transformations

Hash based biometric authentication schemes rely on modified hash functions which ensure a “robustness” property so that small changes in the input biometric samples do not change the resulting hash value and authentication can be correctly performed. Such an approach has been uses by Davida et al. in [13, 14], where “robust” hash functions have been devised to protect the sensitive biometric template. The one-way transformation is designed as a combination of various Gaussian functions which at the end behave as a robust hash scheme. Then, the hash function is used to cryptographically secure the biometric templates stored in the database. Different kinds of comparison between the hashed templates are used in the one-way transformation combined with a secure cryptographic hash function.

Hash based techniques have been applied taking into account different biometric traits. In [15], a hash based technique has been defined for signatures. In the described application, a pen-based PDA is used to collect a signature which is transformed into a hash value. Then, the hash value is used to create a key which can be used for securing a data communication channel. Differently from usual, authentication is performed by comparing a vector of the resulting hash values, which are obtained considering 24 features computed from the signature. Such method uses a statistical approach: during the enrolment phase, four signatures per user are required to build a personal interval matrix which will be stored in the database. The final decision is made by comparing the fresh hash values in the vector with the stored interval matrix of each individual present in the database. Hash based technique for palmprint biometrics has been proposed in [16]. The features of palmprints are extracted from the palm images, and then the Fisher Discriminant Analysis is applied to select the most significant ones producing a reduction of the space dimensionality. This set of features is then combined with a random number by the “PalmHashing” algorithm achieving a discretization process. This algorithm projects the biometric input into an
The Biohashing technique has been introduced in [17] and relies on the usage of a two-factor authenticator combination of pseudo-random numbers and a biometric binarized feature. The random key used to perform the transformation must be kept secret and presented during the authentication. Using different keys, it is possible to obtain several templates starting from the same user biometrics. The main disadvantage of the BioHashing method is that if the key is compromised, the impostor holding both the key and the transformed template can recover the original template and be authenticated [5]. The usage of a multi-modal biometric authentication system where one or two biometric features have been “biohashed” is shown to reduce the effect of this drawback, but the proposed technique increases the overall Equal Error rate. In [18], a biohashing approach is used to produce the Facehashing algorithm. In this case, the face images are pre-processed using the Fourier-Mellin wavelet transformation in order to obtain a low-frequencies face representation. The resulting representation is more robust with respect to facial expressions and small occlusions. Then, a discretization process is defined, achieved by a repeated inner-product of the used data and an orthonormal base obtained with a secret number (the token) using the Gram-Schmidt process. The final hashed data are considered to be a zero-knowledge representation of the user input. In [19] the face is used to produce a non-reversible binary template by using recognition of fiducial points (eyes, nose, eyebrows) and the application of a set of Gabor filters to the face images. The quantization of the extracted features is then processed using a comparison between the obtained features vector from the face and the mean features vectors present in the database. Every bit in the binary template is associated with a reliability estimate based on the standard deviation of its corresponding feature. The most reliable components of the vector after quantization are used to compose the final binary template. The matching function has been designed using a correlation quantifier.

### 3.1.2. Non Invertible Transformations

An approach similar to the one above described, relies on a set of “non invertible transformations”, which operate a transformation on the original biometric template, still preserving similarity in the transformed domain. Such transformations have the characteristics that they are computationally hard to be inverted, and in this aspect are similar to the hash based transformation above described. In literature, the main difference reported with respect to the hash based approach, is that even knowing the key or the parameters used in the transformation, an adversary is not able to recover the original biometric template. Such an approach has been used mainly for transforming fingerprint minutiae. A general scheme has been proposed in [20] to produce a non-invertible function capable to transform a point pattern using high-order polynomials. Such scheme has been applied both to the minutiae set present in a fingerprint and to the frequency-amplitude parameters of a speech pattern. Another non invertible transformation and a correlated matching algorithm for fingerprint have been proposed in [21]. The transformation is based on geometric translations of the minutiae coordinates and their angles. Such transformation depends on a key and is considered not-reversible. Changing the key, it is possible to produce a new transformed template from the same fingerprint. Unfortunately, the study does not provide a complete analysis of the security of the scheme, focusing only on the error rates. A deeper insight on geometrical and functional transfor-
mations in fingerprint biometrics is given in [22]. The study compares the capability of the cartesian, radial and functional transformations in producing cancellable biometrics. This approach provides flexibility to change the transformation from one application to another to ensure the security and privacy of biometric data. The paper demonstrates the non-reversibility by proving that it is computationally hard to recover the original biometric identifier from a transformed version. A similar approach has been proposed in [4] to achieve a biometric system for offline verification of certified, cryptographically secure documents. The presented technique can produce printable IDs obtained from an extracted and compressed iris feature and an arbitrary text.

In most of the presented approaches, rigorous security analysis is missing. In particular, it is not clear the real robustness of these schemes once the hash values/function are also compromised (or the transformed-templates/transformation-algorithm for the second approach), as well as the related keys and parameters.

3.1.3. Cryptography Based Techniques

Variability of biometric templates and protection of personal data have been faced in literature by a wide range of techniques based on the combination of biometrics and cryptography. A comprehensive survey of different approaches and of the related problems can be found in [1].

The main challenge for such techniques is to devise methods for the generation of cryptographic keys from biometric templates. To cope with variability of templates, an error tolerant representation of the biometric features must be constructed, such that authentication can be performed correctly as usual, when the claimant template is within a given threshold from the original template. In other cases a distance preserving robust transformation operating on the biometric template has been selected.

The transformation of biometric templates in a suitable representation which can be efficiently treated, for example in a metric space, is itself an active research area [23]. IrisCode [24] and Fingercode [25] are techniques for the extraction of a binary string from iris and fingerprint templates, respectively. A string representation is extracted from the considered biometric feature and successively a non invertible transformation is applied in order to securely store the biometric template. The same transformation is applied to the fresh biometric templates acquired during the authentication phase, and the biometric match succeeds if the two obtained transformations are equal or sufficiently close. The non invertibility of the transformation ensures that an adversary does not get any valuable information even if he gets or steals the stored (transformed) template.

Another recently developed approach relies on the extraction of helper data during the enrollment phase which is stored together with the hashed form of the biometrics. Such data can be made publicly available and is used in the authentication phase in combination with fresh biometric features in order to reconstruct the derived secret as depicted in Fig. (4). The recently introduced fuzzy cryptographic primitives [26], secure or fuzzy sketch and fuzzy extractor build on this principle and allow the secure extraction of a uniformly random string from the (biometric) input in a noise-tolerant way. Such constructions usually rely on a metric space and most of the constructions have been given considering Hamming distance. However, set difference and edit distance metrics have also been considered, referring to the size of the symmetric difference of two input sets in the first case and to the number of insertions and deletions needed to convert one string into the other, in the second case. Based on these primitives, recently several constructions for devising practically usable biometric authentication systems have been proposed [27, 28].

Such techniques, in order to cope with the variability of biometric templates use error correction codes aiming to extract a unique associated feature from each different biometric reading: the different readings are treated as corrupted codewords and are accordingly decoded. During the verification phase, the feature retrieved by a biometric reading is given as input to a hash function, and compared with the hash value stored during the enrolment phase. One of the first applications of such technique has been developed by Juels et al [29].

Juels and Sudan proposed a “fuzzy vault scheme” in [30] relying on the polynomial interpolation technique in order to cope with variability of the stored biometric template. With such technique the problem of having an order invariant representation of the biometric template is overcome. The basic idea is to lock a secret in a vault using an unordered set. The secret could be successfully retrieved using another unordered set which substantially overlaps with the first used set. More in detail, the secret is encoded using the evaluation of a polynomial over a given set of points using the Reed Solomon encoding scheme, i.e., such points represent a codeword. To increase the security, a set of chaff points are added to the first set in order to form the vault. To reconstruct the codeword, the user has to provide a set of points which overlaps with the original set.

The fuzzy vault construction has been successfully applied by Uludag and Jain using fingerprint templates [31]. Clancy et al. proposed a construction of a biometric identification schema using a secure smartcard to store the vault [32]. Their construction however has been slightly modified in order to cope with real life parameters. Finally the problem of the selection of chaff points, avoiding that the attacker get enabled to distinguish between chaff and real points has been considered by Chang and Li [33]. Some bounds on the entropy loss have also been introduced.

An important step towards the realization of personal identification system based on cryptographic key derived from biometrics features has been recently done by Dodis et al. [26]. In their work, novel primitives such as secure or fuzzy sketch and fuzzy extractor have been introduced. Fuzzy sketches resolve the problem of error tolerance, enabling the computation of a public string $P$ from a biometric reading $r$, such that from another reading $r'$ sufficiently close to $r$ it is possible to reconstruct the original one. Furthermore, the knowledge of $P$, should not reveal too much information on the original reading $r$, i.e. the entropy on $r$ is high-enough to be useful even if $P$ is public. Fuzzy extractors address the problem of non-uniformity by associating a random uniform string $R$ to the public string $P$ still keeping all the properties of fuzzy sketches. Indeed, fuzzy extractors can be built out
of fuzzy sketches and enable the recovering of the secret uniform random string $R$, from the knowledge of the public string $P$ and a reading $r'$ sufficiently close to $r$.

Since the introduction of the fuzzy primitives, many researchers have proposed several authentication schemes based on the applications of such techniques. A general framework to design and analyze a secure sketch for biometric templates is presented in [27], where the face biometrics has been used as example. Interestingly, the paper shows that theoretical bounds have their limitations in practical schemes. In particular, it has been shown that the entropy loss of the template can not be considered a complete description of the robustness level of the scheme in practical application, while the analysis of the FAR and FRR should be always envisioned. In [34], a near-optimal error-correcting code is discussed (based on a two-dimensional iterative min-sum decoding algorithm) for application with iris biometrics in a fuzzy sketches scheme. The paper produces also an explicit estimation of the upper bounds on the correction capacity of Fuzzy Sketches on iris-based biometrics. A fuzzy based construction for fingerprint biometrics has been discussed in [28], where the string representation of the biometric templates relies on FingerCodes.

### 3.2. Recent Patents

Often, the different academic research endeavours described in the previous section led to the registration of corresponding patent. But also many unrelated initiatives were undertaken. In this subsection, we will give a short insight to patents and technologies useful to actually implement the protection of the templates. First, core biometric technologies will be taken into account: templates creation and matching algorithms. Then, a collection of patents of effective schemes for templates protection are described.

#### 3.2.1. Biometric Templates

Along fingerprint and iris codes, a major biometric technique which is gaining importance is facial recognition\(^1\). Algorithms meant to perform biometric authentication through face recognition were developed since the late ’80. Among those, a few exploited a three dimensional characterization of the face in a quest for higher accuracy. Several techniques are available today for 3-D face reconstruction; among them the projection of a structured light is one of the more reliable, but others are available, such as stereovision and shape-from-shadow (SFS). In the latter, given a picture of the subject, the aim is to reconstruct the surface shape at each pixel in the image from the gradual variation of shading in the image itself. One of the key points is that SFS algorithms assume the direction of the light source to be known. While the estimation of such direction is generally feasible, it requires some assumptions about the surface shape and its orientation [35]. It is reasonable to expect that a greater accuracy in the light source direction estimation might lead to a higher precision in 3-D face reconstruction and thus on subsequent identification procedures.

The patent [36, 37] describes a technique which is meant to increase such accuracy. A target device with particular coating and predefined shape is inserted within the image so

---

\(^1\)The decision of ICAO, the International Civil Aviation Organization to specify face biometrics as required for international travel documents, busted the efforts in this direction.
that once detected the target in the picture, the known characteristics of its surface reasonably facilitate the estimation of light diffusion and direction. Secondly, if two or more target devices are employed concurrently and it their positions are calibrated, the shadows on the targets might be used to triangulate the position of the light source relative to each point on the face of the subject. The generated contour map is a 3-D description of the shape of the face and might be used as input in a face recognition system (which is not described in the patent itself). The inventor also suggests to “tag” each acquired image with the illumination data to ensure that such information might be available in further analysis.

Biometric traits are typically acquired by means of digital pictures (exceptions obviously apply as for the case of the voice recognition). Biometric features are then obtained by processing such pictures and personal identity is verified by comparing the sets of features. A somewhat simpler approach is that of comparing the pictures directly by means of image correlation. The inventors of [38] claim that such comparison might be made both faster and more robust with respect to noise by first integrating the image along one direction. Integration might be performed simply by summing up the intensity levels along each line or column of the picture. If two pictures are identical also their integrals will be so and in turn the correlation among them will be symmetrical with respect to the peak C(θ). The inventors suggest the usage of a metric called "symmetry function" to discriminate if two similar pictures depict the same biometric sample. Once considered C(θ−Δ) and C(θ+Δ), the symmetry function is computed as the ration of the absolute value of C(θ−Δ)-C(θ+Δ) with respect to C(θ). Unfortunately, the inventors failed to demonstrate how such technique might be rendered independent from rotation, a typical problem encountered with fingerprints, e.g. rotation on the finger on the sensor. Other patents, such as [39], and [40], are similarly focused on the correlation of biometric images.

3.2.2. Biometrically Protected Storage Devices

The idea of protecting personal data stored on chips and cards by using biometric approaches is common to many patents in which the biometrics sensor (chip) is miniaturized and incorporated with the biometrics authentication device [40-43].

A different approach is that of having the biometric system as a stand alone module which is used only to rule the access to protected data on a storage unit. A drawback of such an approach is the high cost of the units (for example 32-bits RISC processors or DSP are required). Patents [44,45] follow this approach; in both a device for protecting data stored in a memory block by way of fingerprint identification is presented. The disclosed identification architecture appears an independent fingerprint identification device in which the sample is captured and identified in the same device.

A way to cut on the cost of the unit is suggested by patent [46] where the biometric software is installed on a terminal computer (host). Such a solution produces the relevant drawback to be not general (platform dependency) and it requires a proper and safe installation procedure. In the cited patent, the host sends the biometric template collected at verification back to the biometric module through an interface. The biometric module compares the user minutiae with a template stored in the memory module at enrolment, and enables a private block of the memory module to be accessed by the host according to a matching result.

A similar solution is envisioned from patent [47] where it is preferred to use the microprocessor of the host to execute the biometric image processing and authentication in order to reduce the overall cost effectively. The biometrics application programs (encompassing the image processing, the authentication and encrypting/decrypting operations, and the biometrics matching program) are run or executed in the host system in order to achieve a sort of plug-and-play functionality with respect to the employed biometric sensor. In addition, the cited patent allows for providing a storage device and a method of protecting data stored therein capable of enabling the storage device from being intercepted. The method is patented in [48,49].

Finally, a slightly different approach is followed in [50] where more submodules are specified. The modular system comprises: a host interface connected to a host terminal, a controller, a fingerprint sensor, a storage device interface, and a storing unit. In this setup, drives and applications are loaded from the storage unit and installed in the host. The matching phase is carried out by a traditional biometric matching algorithm, then, in case of positive match, the specific block hidden in the storage unit will be given in output to the host by means of encryption/decryption algorithms.

It is worth noting that in all these patents, the biometric features are used to protect private data in storage devices, not the biometric features themselves.

3.2.3. Biometric Authentication Systems

Juels and Wattenberg proposed in [51] the “fuzzy commitment” scheme, where a secret message is protected using a biometric template. Such scheme is also patented [29]. Using the described fuzzy cryptographic primitive, they show also how to construct biometric authentication systems. In their proposal, an error correcting code is used in order to associate a codeword c with a person and to compute an offset ($\delta = \oplus x$) for the biometric template x. The encrypted message (the fuzzy commitment) is then represented by the pair ($\delta, h(c)$), where $h(c)$ is a one way hash function. It is worth to notice that neither the biometric feature, nor the associated codeword are publicly stored. The authentication process is correctly performed if a fresh biometric reading y allows the computation of a binary string c'=$\delta \oplus y$ sufficiently close to c so that the code decodes it to c and the comparison between their hash values succeeds. A similar construction has been proposed by Hao et al., with the application of an iris code feature extraction algorithm and the combined use of Hadamard and Reed-Solomon codes [52].

The authentication system presented in [53] enables the use of mobile terminals in order to perform individual authentication based on the biometrics. The system relies on the combination of: i) an IC card, issued by the financial institution and holding the biometric information; ii) a terminal acting as an automated cash transaction, a biometric
input device and a Felica-compatible terminal capable of communication with the mobile terminal; iii) a mobile terminal, capable of long-range communication with the mobile network, of short-range communication with the Felica-chip, of reading/writing to and from a contact-free IC-card, and of executing i-mode applications previously downloaded.

When the authentication is started, a biometrics data registration process is performed. During this phase, the biometric i-mode application program is downloaded from the application server to the portable communication terminal and an authenticated biometric information area is created. The mobile terminal then creates an authentication service area in the common area and registers it in the Felica chip. The biometric authentication program performs mutual authentication with the Felica reader/writer and biometric functions are provided in order to allow the comparison of the biometric information stored in the common area against the biometric information obtained via the Felica reader/writer.

To perform authentication, the user carrying the portable terminal where the authentication program is registered, uses an ATM and inserts there the IC card. At the same time the user employs the facilities of the ATM to provide a fresh biometric reading. The read out data is verified against the data registered in the IC card. If the authentication is successful, user information is transferred using the short range communication from the portable terminal to the ATM terminal. Such information is compared with the information stored in the host and in case of successful verification, the ATM starts a biometrics-validated transaction informing the server of the financial institutions.

The authors of the patent, discuss the basic setting for the biometric authentication system based on palm vein patterns, but easily adaptable to other biometrics, such as finger vein patterns, blood vessel of the back of the hand, palmprints, or facial features. The application of the biometric system can also be extended to automated vending equipment, as well as identification systems to access restricted areas.

Patent [54] presents a method for user authentication, concealing a key inside a structure, such that the key can be recovered only if the structure or its representation is retrieved. The idea is to embed the key, for example, in an original image, by generating false image points interpolated among a chosen subset of the true image points. The true image points form a master template, while the false image points form a transient template. In the case, for example, of fingerprint based authentication, the key is concealed in additional false minutiae, which are mixed with the true minutiae constituting the representation of user biometric data. To perform the authentication, a claimant image is used to produce corresponding true image points of the original image held in the transient template. Then the residual image points are treated as false image point and used to retrieve the key iteratively. Such a key is used to decrypt the cipher-text and compare the resulting plain-text with the known original plain-text. Such step is repeated, until a successful authentication is achieved or all the attempts to generate the key fail, or a policy (on the number of attempts) limitation is reached. Different applications of the system are proposed for key decoding and authentication. The transient template can be stored in a smart-card together with the pair of plain-text and cipher-text and protected also through traditional mechanisms, such as passwords or PINs. Transient templates can also be stored in distributed databases and used as credentials when required to perform use authentication.

A multimodal biometric authentication system is described in patent [55]. Combining multiple (at least two) biometric readings of different traits, the systems issues a personal identifier. The ID while permitting a successive biometric authentication procedure, guarantees that the templates on which it is based are not disclosed, so that even in case of loss or steal of the document, privacy is guaranteed. Such system would permit to exploit the benefit of multimodality (increased reliability and robustness) while minimizing the threats to privacy. The construction is based on an extension of the “fuzzy commitment” scheme just described. By increasing the number of biometric traits needed for the identification, the authentication system can be tuned in order to provide different levels of security according to the protection degree requested by the application.

4. CURRENT & FUTURE DEVELOPMENTS

Biometric authentication systems are being widely deployed in order to build reliable identity management systems. The advantage in perspective is to build automatic authentication mechanisms with low error rates attesting in an incontrovertible way the physical presence of the authenticating person. However, the security of such systems must be carefully analyzed. The irrevocable association between a user and its biometrics imposes a strict evaluation of the risks for users’ privacy coming from the stealing of biometric templates.

Both protection of biometric templates and biometric authentication system development are promising research fields, where academic and industrial researchers are focussing their efforts. In this work, we reviewed the most important techniques developed in the open literature and presented also the most recent patents related to such matter, found after a search in public databases using opportune keywords, yet not intending to give a full coverage.

Some of the examined patents deal with the study and the creation of innovative user templates. Other patents present biometric authentication systems effectively deployable and matching the requirements we sketch in the first part of the paper.

At the moment lack of standardization and of security analysis seem to be the most important problems which stop a wider diffusion of biometric systems. As regards biometric template creation, in literature interesting solutions are now available by using fingerprint and iris biometric technologies (sensors, algorithms and templates). Applicative results are expected also from innovative face templates by using high-resolution multiple stills and 3D images. Facial traits are considered one of the most widely accepted by the users and current international independent vendor competitions show that the performances of the systems are very promising.

Promising results are also expected from the application of the recently introduced cryptographic primitives, which
enable the construction of robust biometric crypto-systems. Such systems should finally provide enhanced guarantees for the protection and the privacy of biometric templates.

REFERENCES


