Face Biometrics with Renewable Templates

Michiel van der Veen, Tom Kevenaar, Geert-Jan Schrijen, Ton H. Akkermans, Fei Zuo

Philips Research, Prof. Holstlaan 4, 5656 AA,Eindhoven
Corresponding author: michiel.van.der.veen@philips.com

1. Abstract

In recent literature, privacy protection technologies for biometric templates were proposed. Among these is the so-called helper-data system (HDS) based on reliable component selection. In this paper we integrate this approach with face biometrics such that we achieve a system in which the templates are privacy protected, and multiple templates can be derived from the same facial image for the purpose of template renewability. Extracting binary feature vectors forms an essential step in this process. Using the FERET and Caltech databases, we show that this quantization step does not significantly degrade the classification performance compared to, for example, traditional correlation-based classifiers. The binary feature vectors are integrated in the HDS leading to a privacy protected facial recognition algorithm with acceptable FAR and FRR, provided that the intra-class variation is sufficiently small. This suggests that a controlled enrollment procedure with a sufficient number of enrollment measurements is required.

2. Introduction

Nowadays, there is an emerging interest in the application of biometric authentication and identification. In this paper we concentrate on face recognition technology because this is a biometric modality being used in an increasing number of applications. For example, the ICAO [7] recently standardized the biometric modalities and the corresponding template formats for Machine Readable Travel Documents (MRTD). As the main modality, facial recognition is recommended for integration in electronic traveling documents possibly supplemented by iris and fingerprint recognition. A recent example is the German biometric enabled ePassport, which was introduced on November 1st 2005. A second trend is towards the use of biometrics as a convenience factor. Everything becomes personalized, from preferences in CE equipment to driver seat positioning in the automotive field. The number of PIN codes we need to remember is increasing to a point where especially the elderly can no longer cope. Replacing these access mechanisms by biometrics provides not just better security, but essentially more ease-of-use.

In all of the above cases, privacy concerns become apparent, and some of the technology projects have seen already a heavy backlash from individual consumers, organized privacy organizations or governmental institutions. One example is the biometric ePassport, which was (and still is) under heavy discussion by several privacy associations. This has led to a situation where crosschecking of identity against centralized databases, the so-called "three-way-check" is prohibited in Europe, even though it would provide an additional layer of security. It is evident that a widespread deployment of biometrics will only increase these privacy and security concerns.

When discussing privacy and security in biometric systems, we distinguish at least two threat scenarios: (i) the spoofing problem in which fake biometrics are used to circumvent the authentication mechanism, and (ii) the threats associated with the digital storage of biometric information (i.e. the biometric templates). The first problem is addressed with liveness detection, which we do not cover in this work. Related to the storage of biometric templates, we identify the following risks:

1. *Identity theft*: A human only has a limited number of biometrics (i.e. 10 fingers, 2 iris, and 1 face). Storage of the templates in multiple locations (e.g. databases) increases the probability of theft and abuse and, this would mean a "theft of identity". Due to the limited set of biometrics available per person, this also means that once the template is compromised it is compromised forever: it cannot be revoked, reissued or even destroyed. This is in contrast to authentication systems that make use of e.g. passwords or PIN codes. When these are lost or

forgotten, they can simply be renewed. In analogy with these systems, it is essential to also deploy renewable biometric templates. Then, in case of theft, reissuing biometric templates becomes possible

2. *Cross matching:* Especially in networked environments attacks on biometric databases form a serious threat. As soon as identical templates are deployed in multiple databases it would be possible to perform cross matching between them. In this way the privacy of the user is not guaranteed.
3. *Sensitive information:* In some literature it is reported that biometrics may reveal sensitive medical information. It speaks for itself that this data must be protected in order to prevent misuse. This becomes particular relevant when using DNA like information as a biometric tool.
4. *Legislation:* In the European Union, the EU Data Protection Directive states that individuals have the right to control the collection and use of their personal data. Although biometrics are not specifically mentioned in the Directive, it is important to realize that the Directive protects the privacy rights of individuals who can be "identified ... by one or more factors specific to his [or her] ... physical or physiological ... identity". Consequently, the information biometric technology provides will likely be considered as protected personal data, thus impacting biometric deployment. More concrete this could mean a discussion on the use of external databases with 'large' collection of biometric data.

A known (apparent) solution to prevent misuse of biometrics is to encrypt all the reference templates that are stored on the smart card or in a database. Using proper encryption algorithms and key-management systems, this indeed yields a secure storage of the reference templates. However, the actual deployment of the encrypted templates in a real application introduces a potential security weakness. This becomes evident when we replay the authentication protocol:

1. A person claims his identity after which his corresponding reference template is extracted from the database or smart card.
2. The person provides his biometric.
3. The actual authentication is done by comparing the encrypted reference template with the newly measured one

The weakness of this protocol is in the third action. In the authentication mechanism, the reference template needs to be matched against the measured template. However it is not possible to compare directly the encrypted reference template with the non-encrypted measurement. One apparent solution would be to encrypt also the measurement data and compare the templates in the encrypted domain. Since the biometric measurements are inherently noisy, and encryption functions are very sensitive to noise, this approach is not possible in a straightforward manner. An alternative is to decrypt the reference measurement and perform a matching in the unencrypted domain. However, the decryption functionality makes it possible that the verifier has access to all the templates in the database or on the smart card and that the privacy issues remain. In recent literature, several approaches were proposed for protecting the biometric templates in a slightly different manner. Juels and Wattenberg [5] are one of the first to provide a technological solution based on so-called fuzzy commitment schemes. They present the analogy with the protection of an ordinary password P. In the computer system, a commitment of P is stored in the password file in the form of a cryptographic hash $h(P)$. In the definition of cryptographic commitment we have the following protocol: the sender sends an encrypted version $h(P)$ of P such that the receiver cannot learn P from $h(P)$, whereas the sender should be able to prove to the receiver that $h(P)$ is indeed an hashed version of P. In [5], the authors propose a theoretic framework for adapting this scheme for noisy input data (i.e. the biometric template) by making use of error correction codes. Ratha et al [10] introduce the so-called cancelable biometrics, which make use of non-invertible signal transforms. The biometric image (e.g. fingerprint, face or iris etc) is first transformed with e.g. a scrambling code operating on a block-by-block basis. In a second step, the feature vectors are extracted. Security and performance analysis are not presented in this work. Yet another alternative is the fuzzy vault approach presented by Juels and Sudan [6]. It allows fuzzy (un)locking of a secret using unordered data sets (applied in [16] to fingerprint minutiae) while the fuzzy extractor [1, 2] extracts a uniformly random string from biometric input in an error tolerant way.

In line with the approach of Juels and Wattenberg [5], we recently proposed a new class of template protection technology defined as Helper Data Systems (HDS). In [3], [4], [9], and [13], we outlined the fundamentals of these systems, including: basic principles, performance analysis, security analysis etc. A practical translation of the HDS framework was done for fingerprinting [14], acoustic ear recognition [15] and faces [8]. In this paper we concentrate on

facial biometrics and explore the operational characteristics when integrating a HDS in a facial recognition system. Besides the system performance we also demonstrate some usage scenarios. This paper is organized as follows. In section 3, we briefly review the key components of the HDS. The face recognition algorithm and its integration with a HDS is shown in section 4 and in section 5 we present examples of how these privacy protection systems can be adopted in various business scenarios.

3. Template Protection with HDS

3.1. Algorithm Overview

In this section we discuss the basic principles of the helper data scheme (HDS) introduced in the previous section. Conceptually, it works in the following manner. We distinguish an enrollment (Figure 1a) and a verification phase (Figure 1b). In the enrollment, one or more biometric measurements are done resulting in an enrollment feature vector $\vec{X}$. We then randomly select a codeword vector $\vec{C}_X$, which is derived from an Error Correction Code (ECC) in a high-dimensional space. In Figure 1, codewords are indicated as solid dots. The difference between the codeword vector and the biometric measurement is defined as the so-called helper data signal:

$$\vec{W}_X = \vec{C}_X - \vec{X} . \qquad [1]$$

Using the error correction decoding function, the codeword vector $\vec{C}_X$ is mapped to a random vector *S*. This vector, or a secure derivative thereof, is finally used for matching purposes. Since we are working with an ECC, we can define a sphere with radius δ around every codeword such that the decoding function maps every point within the sphere onto the corresponding codeword. In the context of helper data schemes this is referred to as a δ-contracting function. This is a requirement when we consider the verification measurement of the same person. In general this can be modeled as:

$$\vec{Y} = \vec{X} + \vec{N} , \qquad [2]$$

in which *N* represents the typical measurement noise of a given biometric. If this noise is small enough (i.e. $\left|\vec{N}\right| \le \delta$), then *Y* is mapped inside the δ-region around $\vec{C}_Y$, such that it can be contracted to $\vec{C}_Y$ in a bit-exact manner.

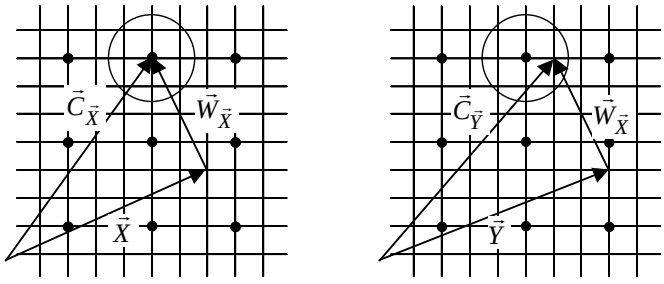


Figure 1: Schematic of the underlying principles of helper-data systems: (a) enrollment procedure, and (b) corresponding verification phase. Dotted grid points correspond to code-words *C* of an error correction code.

When translating these concepts to a practical biometric system with template protection, we have to consider at least the following issues:

- The matching performance (e.g. in terms of FAR, FRR or EER) should be comparable to biometric systems with traditional classifiers.
- At a given EER, FAR or FRR, the HDS system should be able to operate with a given Error Correction Code
- The information leakage should be minimized: i.e. the helper data W should leak a minimum number of bits on C, and also W should leak a minimum amount of information on X. A theoretical foundation of this is provided by Linnartz and Tuyls [9].

In this work we mainly concentrate on the first two issues. In the following we present, inspired by the work in [15], a general framework for implementing the concepts of Figure 1. This helper data scheme is based on so-called reliable components and is given in Figure 2. Using the face recognition algorithm of Zuo and De With ([20],[21]), we briefly review the essential steps in the enrollment and the authentication phase.

3.2. Enrollment with a HDS system

A first step in the enrollment is the measurement of a number of facial images F and the construction of the corresponding feature vectors X. From such an input facial image, the face is localized and normalized. Following that, six key facial feature objects are identified: (i) left eye, (ii) right eye, (iii) left eyebrow, (iv) right eyebrow, (v) mouth and (vi) nose. The shape of each object is modeled by a piecewise linear contour defined by a set of locations or 'fiducial' points. These fiducial points thus form a shape description of the six key objects in the face. In order to generate a real-valued feature vector $\vec{X}$, local texture information around each fiducial point is derived by image convolution with a set of Gabor kernels taken from [19]. The modulus of the convolution result is taken as one component X_i in the feature vector $\vec{X}$. Since we use 5 frequencies and 8 orientations resulting in a total of 40 Gabor kernels, we have 40 components for describing each fiducial point

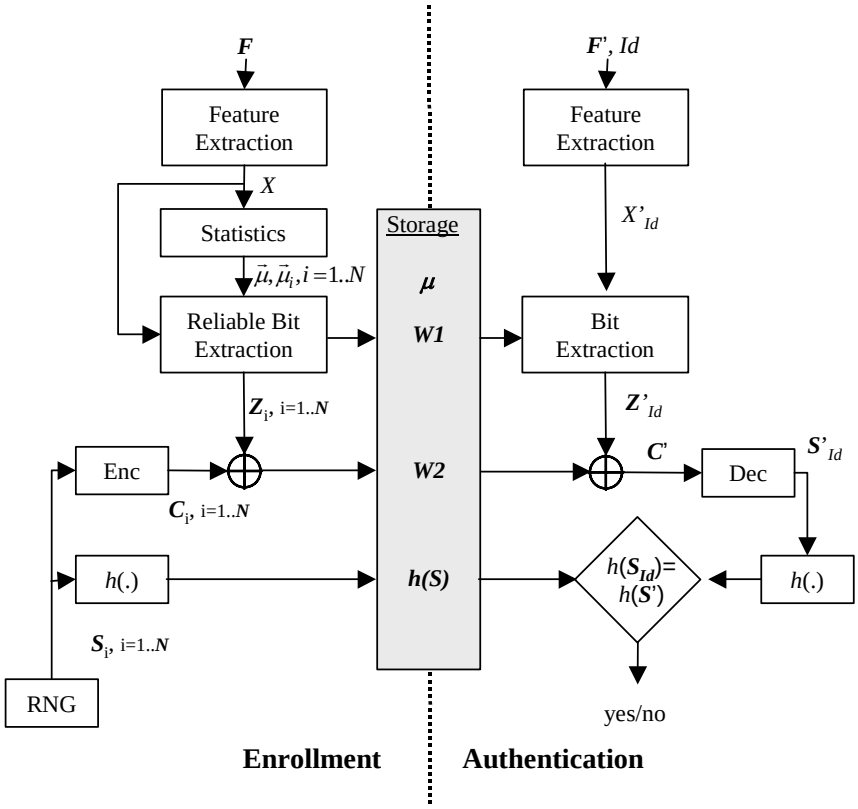


Figure 2: Helper Data System concept of Figure 1 based on reliable component selection. *F* and *F'* represent input measurements for enrollment and authentication; whereas μ, W_1, W_2 and $h(S)$ indicate mean of the interclass distribution, helper data signals 1 and 2, and the cryptographically hashed random number *S*, respectively.

3.2.1. Extracting binary feature vectors

The reliable component HDS of Figure 2 requires a translation of real-valued feature vectors X to binary ones (i.e. binary strings). Therefore, quantizing the feature vectors forms an important step. During the enrollment phase we assume to have recorded a set of M images per users over a total of N users. This results in a total of $N \cdot M$ images: $\{F_{i,j}\}_{i=1..N,\, j=1..M}$. The corresponding set of feature vectors is denoted as $\{\vec{X}_{i,j}\}_{i=1..N,\, j=1..M}$, where $\vec{X}_{i,j} \in \Re^k$ have components $(\vec{X}_{i,j})_t$, $t = 1..k$. For this application we use a binary quantizer to translate the feature vectors to a binary string. This quantization is done using both the mean ($\vec{\mu}_i$) over the intra-class variability and the mean ($\vec{\mu}$) over all enrollment feature vectors:

$$\vec{\mu}_i = \tfrac{1}{M}\sum_{j=1}^{M} \vec{X}_{i,j} \text{ , and } \vec{\mu} = \tfrac{1}{N}\sum_{i=1}^{N} \vec{\mu}_i \text{ .} \qquad [3]$$

For each user i, a binary feature vector Q_i is extracted using:

$$(Q_i)_t = \begin{cases} 0 & \text{if } (\vec{\mu}_i)_t \le (\vec{\mu})_t \\ 1 & \text{if } (\vec{\mu}_i)_t > (\vec{\mu})_t \end{cases}, \qquad [4]$$

with t again indicating the component number.

3.2.2. Reliable component selection

In principle, the binary feature vector Q_i may be deployed for further processing. In our previous work we identified that there is a significant variability in stability of the individual components of Q_i. This becomes evident in cases where the intra-class mean ($\vec{\mu}_i$) is close to the inter-class equivalent ($\vec{\mu}$). In these cases, small variations on the input data (F and thus X) can lead to bit errors in the quantized feature vector. In addition, from a classification point of view, these components are less discriminative for class i as compared to the rest of the users. To minimize these effects, we use a mechanism to extract the most 'reliable' or 'robust' components within one feature vector. In this context, reliable relates to the components for which the probability of assigning a wrong bit in Q_i is lowest. For this work we select these components based on their statistical properties. For one user i, the variance ($s^2_{i,t}$) of the t-th component is given by:

$$s^2_{i,t} = \tfrac{1}{M-1}\sum_{j=1}^{M} \left((\vec{X}_{i,j})_t - (\vec{\mu}_i)_t \right)^2 . \qquad [5]$$

We found that the variability over the components can be modeled as a Gaussian (not shown in this paper). Under this assumption, we can make use of standard error functions to estimate the reliability $R_{i,t}$ of bit t of user i:

$$R_{i,t} = \tfrac{1}{2}\left(1 + erf\left(\frac{(\vec{\mu}_i)_t - (\vec{\mu})_t}{\sqrt{2 s^2_{i,t}}}\right)\right), \qquad [6]$$

where *erf* is the error function. We can interpret this as follows: $R_{i,t}$ is the probability that, for a new measurement $\vec{X}_{i,M+1}$ for user i, we have (i) $(\vec{X}_{i,M+1})_t \le (\vec{\mu})_t$ if $(Q_i)_t = 0$ or (ii) $(\vec{X}_{i,M+1})_t > (\vec{\mu})_t$ if $(Q_i)_t = 1$. In other words it is the probability that the new measurements results in the same bit which was already assigned in previous enrollment measurements. Using this definition, bits or components with a higher reliability $R_{i,t}$ have a higher discriminating power due to the larger difference between $(\vec{\mu})_t$ and $(\vec{\mu}_i)_t$ relative to the standard deviation $s_{i,t}$. The final binary feature vector $Z_i \in \{0,1\}^K$ contains the K most reliable components of Q_i. In the system overview of Figure 2, it is shown that we use a first helper data signal W1 for carrying the indices of the reliable bits in Q_i.

3.2.3. Template Protection

Up to this point, the binary templates Q_i are not protected and do not facilitate the functionalities like renewability and versatility, such as described in Section 2. The lower part of Figure 2 illustrates the schematic that enables these features. It makes use of an error correcting code with parameters (*K*, *s*, *d*) where *K* denotes the length of the codewords, *s* the number of information symbols and *d* the number of errors that can be corrected. For each user *i* a user-specific binary random sequence $S_i \in \{0,1\}^s$ is generated by a Random Number Generator (RNG). This sequence is encoded into the codeword C_i, corresponding to the codeword used in Figure 1. The second helper data signal W2 is now defined as:

$$W2_i = C_i \oplus Z_i , \qquad [7]$$

where $\oplus$, represents the bitwise XOR operation.

The net result of the entire enrollment procedure is that the following information is extracted and used for in the authentication phase: (i) the mean $\vec{\mu}$ of the inter-class distribution, (ii) the first and second helper data signals *W1* and *W2*, and (iii) a cryptographically hashed version of the random secret *S* (i.e. *h*(*S*), with *h* a given one-way-hash function). Knowing the helper data signal $W2_i$ and $h(S_i)$, it is not possible to retrieve $Z_i = C_i \oplus W2_i$ because S_i and hence C_i cannot be extracted from $h(S_i)$. The versatility is introduced by the possibility of choosing different values for S_i, which will lead to a different pair ($W2_i$,$h(S_i)$) This feature makes it possible to derive from one single facial biometric templates, multiple secure biometric derivatives.

3.3. Authentication with a HDS scheme

The authentication procedure follows the general lines of the enrollment system of Figure 2. A person claims his identity *Id* and at the same time provides his facial biometric measurement F'_{iD}. In a following step, the feature extraction algorithm of [20] is used to derive the corresponding real-valued feature vector X'_{iD}. Depending on the exact application (see also section 5), the template information (μ, $W1_{iD}$, $W2_{iD}$, $h(S_{iD})$) is extracted from a smartcard or a database. First the mean μ the quantization rule of Equation 4 and the first helper data $W1_{iD}$ are used to derive the robust binary feature vector Z'_{iD}. We can then determine:

$$C'_{iD} = Z'_{iD} \oplus W2_{iD} = Z'_{iD} \oplus (Z_{iD} \oplus C_{iD}) . \qquad [8]$$

Finally, $S'_{iD,}$ and hence $h(S'_{iD})$, can be derived by decoding C'_{iD}. The authentication is achieved by comparing both hashed values:

$$h(S_{iD}) \stackrel{?}{=} h(S'_{iD}) . \qquad [9]$$

If both values are identical then a person is authenticated.

Compared to the original facial recognition scheme, which is based on a normal correlation, we used a Hamming distance classifier. The success of the authentication completely depends on the Hamming distance between Z'_{iD} and the enrolled template Z_{iD}. Given an ECC with parameters (*K*, *s*, *d*), a positive authentication will occur when this Hamming distance is smaller or equal than *d*. This is in contrast to the traditional classifier in which we could set a threshold for authentication at arbitrary values. Given these constraints we identify the following challenges for integrating HDS in facial recognition:

i. Maximize the secret size *s*. Since *h*(*S)* is stored in a database, it must be practically infeasible to retrieve *S* from *h*(*S)*. To prevent an exhaustive search on *h*(*S)*, the size of *S* must be maximized and preferably be at least 50 bits.
ii. A contradicting requirement is to allow for a large number of errors *d* to be corrected. If we choose a BCH code [12] with *K*=511, *s*=58, we have an error correction capability of *d*=91. Increasing the secret size to e.g. *s*=94 or 130 bits, lowers this capability to *d*=62 and *d*=55, respectively. Therefore a challenge is to setup a configuration that allows for the maximum number of bits to be corrected while achieving a sufficiently large secret *S*.

In Section 4, we address this challenge and evaluate the performance of integrating a HDS in a practical facial recognition system.

4. Results of integrating HDS with Facial Biometrics

4.1. The data sets

For our evaluation we have used two different databases: (i) a subset of the FERET database [11] and (ii) a face database from Caltech [18]. Key information of both datasets is presented in Table 1, whereas some typical images are given in Figure 3. The first database is characterized by a higher intra-class variability in the recordings. The subset contains a total of 237 persons each with at least four images. In our experiments we randomly selected 4 pictures (M=4) for persons that have more than four images. During feature extractions, 51 fiducial points were used resulting in feature vectors $\vec{X}$ with 2040 components (i.e. k=2040).

The Caltech dataset is characterized by a smaller intra-class variation (see e.g. Figure 3) and a larger number of measurements per person. From the dataset we selected 19 persons for which at least 11 measurements (M=11) per person are available.

Table 1: Key characteristics and experimental results of facial databases FERET and Caltech

	FERET	**Caltech**
Picture size (pixels)	512x768	150x180
Resolution (dpi)	96	96
No Persons	237	24
Avg. No. of Measurements per person	4.9	14.4
No. of fiducial points	51 (8 eyes; 8 eyebrows; 8 mouth; 11 nose)	45 (8 eyes; 5 eyebrows; 8 mouth; 11 nose)
Feature vector size (k)	2040	1800
Variability on images	high	medium-low
EER for $\vec{X}$	1.5 %	0.25 %
EER for binary feature vector Z	2.5 %	0.25 %
FAR and FRR for HDS implementation	FAR = 0 % FRR = 35 %	FAR = 0% FRR = 3.5 %

4.2. Correlation-based classification results

For reference purposes we first examined the classification results using the traditional classification techniques. For the FERET database, classification results are averaged over all possible 3-1 splits of 3 training and 1 test measurement of the four available images. For this test we used a 'traditional' correlation classifier, where the correlation C between the two feature vectors $\vec{X}$ and $\vec{Y} = \vec{X}'$ is defined as:

$$C = \frac{\left(\vec{X} - \vec{\mu}\right)^T \left(\vec{Y} - \vec{\mu}\right)}{\left\|\vec{X} - \vec{\mu}\right\| \cdot \left\|\vec{Y} - \vec{\mu}\right\|}. \qquad [10]$$

The results are given in Table 1 and Figure 4a. It can be seen that the Equal Error Rate (EER) is equal to 1.5 % at a threshold correlation value of C=0.17. The experiment was repeated for the Caltech database (Figure 4b) for a configuration in which the correlation is derived for all possible 8-3 splits. For this data, the effective EER was measured at 0.25 % at a correlation threshold of C=0.30.

The results show that the variability in the measurement data, as shown in Figure 3, clearly affects the recognition performance. The more conditioned Caltech database results in significant better EER compared to the FERET one. Another important observation is that for both systems, the FAR and FRR can be varied by changing the correlation

threshold. For example, for high-security applications we could increase C, whereas convenience-type of applications usually require a better FRR corresponding to a lower correlation threshold C. In the remainder of this paper we will use these EER results as a benchmark for the equivalent system with integrated privacy protection.

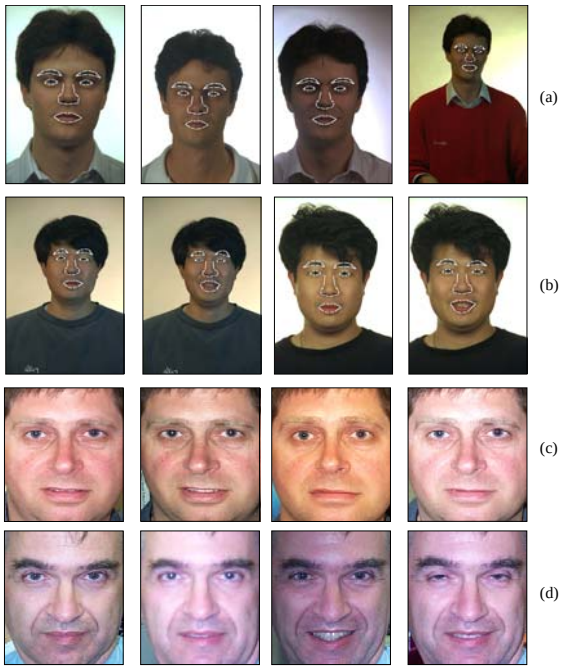
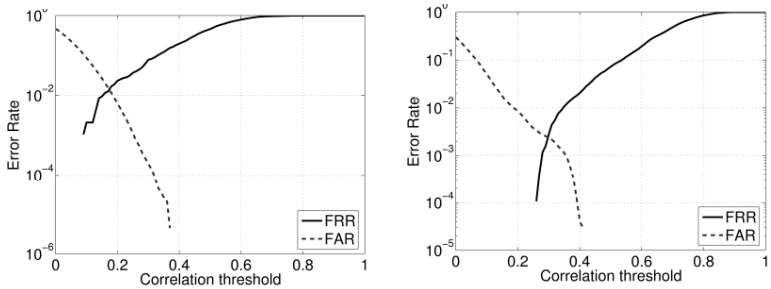


Figure 3: Typical facial images from the FERET (a and b) and Caltech (c and d) databases.



(a) (b)

Figure 4: Correlation-based classification results of (a) FERET and (b) Caltech database. EER for both datasets are 1.5% and 0.25%, respectively.

4.3. Inter and Intra-class variability of binary feature vectors

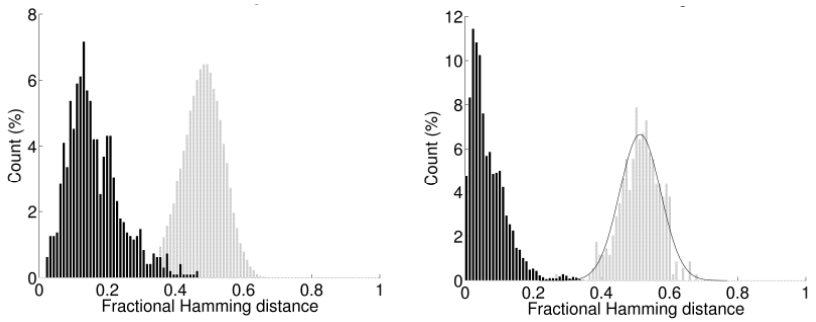
When integrating a HDS approach in the face recognition system, the classification results should be in the same order as the original results shown in Figure 4. We evaluate this by examining both the inter- and intra-class variability of the binary feature vectors indicated in Figure 2.

For the FERET database this is done for every person i, by using all four images, the entire binary feature vector Q_i, the helper data signal $W1_i$ with indices to the most reliable components and the final binary feature vector Z_i. In the

experiments we choose K=511, corresponding to the number of reliable components. For all inter-class comparisons we determined the Hamming distance HD between Z_i and $W1_i \circ Q_p$, where p refers to all other indices than i (i.e. $i \neq p$), and $W1_i \circ Q_p$ means selecting components from Q_p according to indices in $W1_i$. This resulted in a total of 55932 comparisons. The results are presented in Figure 5a (grey area) and are plotted in terms of fractional Hamming distance (FHD). The mean and standard deviation were found to be μ_{FHD}=0.48 and σ=0.06, respectively.

The intra-class distribution is determined in the following manner. For every person i, we take three of the four images to determine Zi and W1i. Next the Hamming distance between Z_i and $W1_i \circ Q_i$ is derived from the fourth measurement only. This procedure is repeated for all possible 3-1 splits resulting in 948 comparisons. Also these results are plotted in Figure 5a (black area) and show a rather 'wide' distribution, partially overlapping the inter-class distribution. Apparently, the variations in the measurements in the FERET database are significant and possibly affect the overall classification performance. A second explanation for this shape is the relative limited number of images per user. This may lead to an inaccurate estimate for the reliability of the feature vector components.

In analogy with the FERET database, both the inter- and intra-class variability were also determined for the Caltech data (Figure 5b). The inter-class distribution is comparable in mean and variance, whereas the intra-class variability is much smaller than the FERET case. This is fully in-line with our observations in Figure 4.



(a) (b)

Figure 5: Inter- (grey area) and Intra-class (black area) distributions for the binary feature vectors Z_i of (a) FERET database and (b) Caltech database.

4.4. Classification results

The inter- and intra-class distributions of Figure 5 give a qualitative indication of the discriminating performance for both databases. We used these data to calculate the FAR and FRR (Figure 6). For the FERET case we observe an EER=2.5 % for a classification boundary at a Hamming distance of 185 bits (corresponding to a $FHD = \frac{185}{511} = 0.36$). Compared to the EER=1.5 % obtained for the correlation classifier (Figure 4a) this means a slight degradation in the classification results. We further observe that the EER occurs at a relatively high HD=185 bits. This is caused by the relatively wide intra-class variation observed in Figure 5a.

The classification results for the Caltech database are given in Figure 6. Here we achieve EER=0.25% for a Hamming distance of 164 bits (corresponding to $FHD = \frac{164}{511} = 0.32$). Comparing this with the correlation-based classification results (Figure 4b) we observe that the same EER can be reached. Similar to the FERET case this EER occurs at a relatively large Hamming distance.

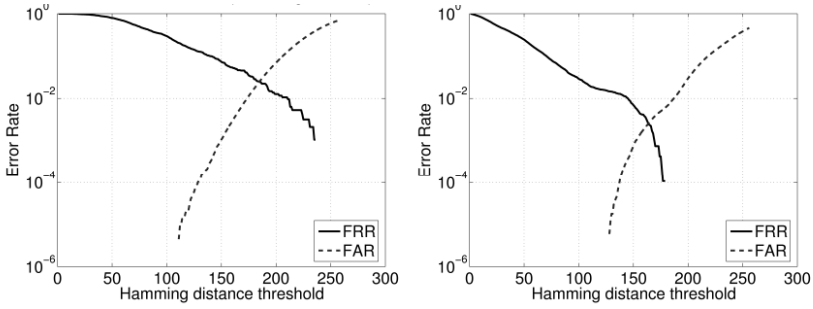


Figure 6: Classification results for the binary feature vectors Z_i of (a) FERET database and (b) Caltech database

4.5. Integrating Error Correction Coding

Although the initial results of Figure 6 look promising in terms of achievable FAR, FRR and EER we have to consider also the feasibility of implementing an Error Correction Code that is capable of correcting the given bit-errors. As discussed in section 3.3, it is the goal that, given a feature vector of length K=511, the secret size s is maximized while also the error correction capability d should be as large as possible. If we choose a BCH code [12], then we can only make use of discrete set of (K, s, d) values. For (511, 58, 91) the Caltech database allows for an acceptable FRR=3.5 %, and a FAR≈0, whereas the FERET database leads also to a FAR≈0 but also to an unacceptable FRR=35%.

The fact that EER is reached at relatively large Hamming distance is mainly caused by the large intra-class variation and the limited number of measurements per person. This makes it more difficult to reliably select the most robust feature components. Although we did not verify it, we anticipate that a larger number of enrollment measurements (e.g. extracted from a video sequence), measured under 'controlled' conditions will enable a better reliable component selection. Hence we expect that the results presented here to be even better when applied in examples like the ones sketched in section 5.

5. Application Example: Three-Way-Check in Biometric ePassport

In this section we go beyond the algorithm descriptions and propose a few 'business' models in which template protection such as HDSs can be deployed. In all of the scenarios, including the ones presented in this section, we use the following information signals discussed in the previous section:

- $h(S)$ - the hashed version of a random number
- $W=(W1, W2, \mu)$ – user and/or application specific 'helper' data which is required to extract the random sequence S' in the authentication phase
- X - the biometric measurement of a person

Management of this information can be tuned according to the specific application. In the following we will explain this in more detail using the example of the so-called three-way-check.

The rollout of the first generation ePassport, which is foreseen in 2006, is fixed. The ICAO (International Civil Aviation Organization) dictates the standards for Machine Readable Travel Documents (MRTD) including the new ePassport. Among the recommendations is the so-called three-way-check for secure verification at the border. It involves comparing data originating from (i) the biometric sensor, (ii) the biometric facial image stored on the MRTD, and (iii) biometric data stored in external (centralized) databases. At this moment the necessary functionality is not deployed because of EU privacy regulations on the storage of biometric data in these centralized databases.

The template protection technology gives the opportunity to enhance the currently specified system with backwards compatible, privacy preserving three-way-checking: in addition to the JPEG templates stored on the smartcard, their

secure versions (i.e. hashed templates) are also stored in a third-party database. The three-way check is then performed by matching the hashed templates from the database, the smartcard and the biometric measurements. The raw JPEG images are not required in this process. This is also better explained in Figure 7. Take as an example a border-control scenario in which there is a biometric kiosk, a border-control authority, and a party managing a secure biometric database. Border passage now involves the following stages:

1. At a kiosk, a user claims his identity *ID*, and measures his biometric *Y* (e.g. facial image, fingerprint or iris).
2. The identity information is sent to the border control authority and can be used to extract from a third-party database the corresponding hashed derivative of the biometric template $h(S)$.
3. The helper data *W* is transmitted to the kiosk
4. The helper data *W* and the biometric measurement *Y* are combined to derive the template protected equivalent $h(S')$ of the measurement.
5. The JPEG image of the iris, face or fingerprint is extracted from the smartcard and used together with the helper data W to also derive the template protected equivalent $h(S'')$. This derivation can be avoided by already storing $h(S'')$ on the smartcard, alongside the JPEG images.
6. Both the measured and the stored templates ($h(S')$ and $h(S'')$) are transmitted to the border-control authority and verified against the database version $h(S)$. A positive authentication is achieved when all versions are bit-exact.

What we achieve with this scenario is a backward compatible extension of the ePassport functionality using a three-way-check in a privacy-preserving manner: i.e. JPEG images are not stored in the external database. Since we are dealing here with an inherently secure database, there is no absolute need for setting up complicated encryption and key-management protocols. Moreover, this storage mechanism possibly does not suffer from stringent privacy legislations. Effectively this means an easy-to-implement and easy-to-use management system for handling privacy sensitive biometric information.

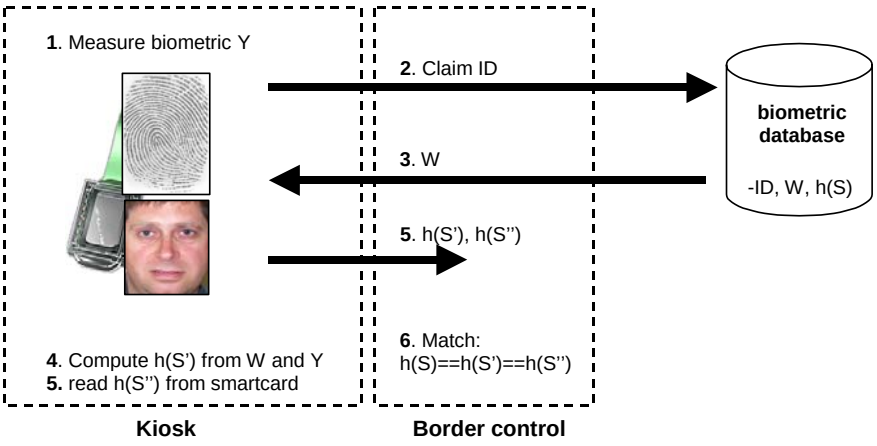


Figure 7: Example of a HDS privacy protection system in use for three-way-check in the ePassport scenario. The individual steps in the authentication process correspond to the numbers described in the paper.

6. Conclusions

In this paper we integrated the so-called Helper Data System in a facial recognition system such that (i) the biometric templates are privacy protected and (ii) the templates can be renewed. Extracting binary feature vectors from the real-valued originals is a key processing step. We showed that a proper quantization of the feature vectors does not significantly lower the classification performance. For the FERET database, the EERs increased from 1.5% to 2.5% for a correlation-based and a Hamming distance based classification, respectively. The Caltech databases showed EER of 0.25% for both classifications. In a practical system these EER cannot be reached due to the limited error correction capability of commonly used ECCs. We found that HDS for face recognition system typically can operate at FAR≈0

and FRR=3.5% providing the intra-class variation is sufficiently small. This suggests a well-conditioned enrollment procedure in which the variability between the measurements is minimized and a sufficiently large number of measurements are taken.

7. References

1. Y. Dodis, L. Reyzin, A. Smith. Fuzzy Extractors: How to generate strong secret keys from biometrics and other noisy data. Advances in Cryptology, Eurocrypt2004, LNCS 3027, pp523-540, 2004.
2. Y. Dodis, L. Reyzin, A. Smith. Fuzzy extractors and cryptography, or how to use your fingerprints. Cryptology ePrint Archive, Report 2003/235, 2003. http://eprint.iacr.org/.
3. M. van Dijk and P. Tuyls, Robustness, reliability and security of biometric key distillation data in the information theoretical setting, 26th Benelux Symposium on Information Theory, Brussels 2005.
4. J. Goseling, P. Tuyls, Information-Theoretic Approach to Privacy Protection of Biometric Templates Manuscripts, Proc. IEEE International Symposium on Information Theory 2004 (ISIT2004), p172.
5. A. Juels, M. Wattenberg. A Fuzzy Commitment Scheme. In G. Tsudik, Ed., 6th ACM Conf. Computer and Communication Security, pp28-36, 1999.
6. A.Juels, M. Sudan. A Fuzzy Vault Scheme. Proc. Int'l Symp. Inf. Theory, A Lapidoth, E.Teletar, Eds., pp408, 2002.
7. ICAO, International Civil Aviation Organization (ICAO), http://www.icao.int
8. Tom Kevenaar, Geert-Jan Schrijen, Ton H. Akkermans, Michiel van der Veen and Fei Zuo, Face Recognition with Renewable and Privacy Preserving Binary Templates, Automatic Identification Advanced Technologies (AutoID2005), Buffalo, New York, USA.
9. J.-P. Linnartz and P. Tuyls, New shielding functions to enhance privacy and prevent misuse of biometric templates, 4th International conference on audio and video based biometric person authentication (AVBPA), 2003.
10. N.K. Ratha, J.H. Connell, R. Bolle. Enhancing Security and Privacy of Biometric-based Authentication Systems. IBM Systems Journal, Vol. 40, No. 3, 2002.
11. P.J. Phillips, H. Moon et al. The FERET evaluation methodology for face recognition algorithms, IEEE Trans. PAMI, Vol.22, pp1090-1104, 2000.
12. M. Purser, Introduction to Error-Correcting Codes, Artech House, Boston, 1995.
13. P. Tuyls, J. Goseling, Capacity and Examples of Template-Protecting Biometric Authentication Systems, Biometric Authentication Workshop 2004, Prague Czech Republic, LNCS 3087, pp.158-170.
14. P. Tuyls, E. Verbitsky, T. Ignatenko, D. Schobben and T.H. Akkermans, Privacy Protected Biometric Templates: Acoustic Ear Identification Proc. SPIE Vol. 5404, p. 176-182, Biometric Technology for Human Identification; Anil K. Jain, Nalini K. Ratha; Eds., Aug. 2004.
15. P. Tuyls, A. Akkermans, T. Kevenaar, G.J. Schrijen, A. Bazen, R. Veldhuis, Practical biometric template protection system based on reliable components, AVBPA 2005.
16. U. Uludag, S. Pankanti, A.K. Jain. Fuzzy Vault for Fingerprints. Proc. 5th Int. Conf. on Audio- and Video-Based Biometric Person Authentication (AVBPA 2005), Springer LNCS
17. M. Weber, Frontal face dataset, http://www.vision.caltech.edu/html-files/archive, California Institute of Technology, 1999.
18. M. Weber, Frontal face dataset. http://www.vision.caltech.edu/html-files/archive, California Institute of Technology, 1994.
19. L. Wiskott, J.M. Fellous et.al, Face recognition by elastic bunch graph matching, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 19, Issue 7, (July 1997), pp775 - 779.
20. F. Zuo, P.H.N. de With, Towards fast feature adaptation and localization for real-time face recognition systems, Visual Communications and Image Processing 2003. Edited by Ebrahimi, Touradj; Sikora, Thomas. Proceedings of the SPIE, Volume 5150, pp. 1857-1865 (2003).
21. F. Zuo, P.H.N. de With, Fast facial feature extraction using a deformable shape model with Haar-wavelet based local texture attributes, Proc. ICIP, pp1425-1428, 2004.