Face Recognition with Renewable and Privacy Preserving Binary Templates

T.A.M. Kevenaar, G.J. Schrijen, M. van der Veen, A.H.M. Akkermans
Philips Research
Prof. Holstlaan 4, 5656 AA, Eindhoven , the Netherlands
{tom.kevenaar,geert.jan.schrijen,michiel.van.der.veen,ton.h.akkermans}@philips.com

F. Zuo
Technical University Eindhoven
Den Dolech 2, 5600 MB Eindhoven
f.zuo@tue.nl

Abstract

This paper considers generating binary feature vectors from biometric face data such that their privacy can be protected using recently introduced helper data systems. We explain how the binary feature vectors can be derived and investigate their statistical properties. Experimental results for a subset of the FERET and Caltech databases show that their is only a slight degradation in classification results when using the binary rather than the real-valued feature vectors. Finally, the scheme to extract the binary vectors is combined with a helper data scheme leading to renewable and privacy preserving facial templates with acceptable classification results provided that the within-class variation is not too large.

1 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 [6] 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. Besides the MRTD's we see applications in the area of device personalization, access control, video surveillance and the like.

Although this widespread introduction and use of facial recognition would enable a more convenient way of living, it also introduces potential privacy threats when templates originating from one person are used in multiple locations:

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, as described in [11], 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.

2. *Cross matching*: Especially in networked environment 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.

We refer to the previous issues as the *privacy problem* of biometrics.

Recently the privacy problem was recognized by several authors and techniques were proposed that can be used to solve the problem. The fuzzy commitment scheme [7] allows a commitment to a noisy binary vector by using a noise-robust parametrized one-way hash function. In [10] the authors propose 'cancelable biometrics' by applying a parametrized one-way distortion function to a biometric before storing it. The fuzzy vault [8] allows fuzzy (un)locking of a secret using unordered data sets (applied in [15] to fingerprint minutiae) while the fuzzy extractor [3, 4] extracts a uniformly random string from biometric input in an error-tolerant way.

In this paper we address this problem by using so-called Helper Data Systems (HDSs) to generate multiple 'anonymous' derivatives from a single biometric template.

The theoretical concepts of HDSs in general were covered in [2, 5, 9] whereas a practical HDS applied to binary feature vectors derived from fingerprint data was demonstrated in [14]. In this paper we evaluate the possibility of using this HDS on facial biometrics. Therefore we will first derive binary feature vectors from subsets of the FERET and Caltech databases, study their statistical properties and give classification results. Secondly we will use these binary vectors in the HDS of [14] that allows for privacy protection and renewability and discuss the final classification results.

This paper is organized as follows. In Section 2 we describe the face recognition algorithm and explain how we generate binary feature vectors. Next in Section 3 we discuss the statistical properties of the derived feature vectors while in Section 4 we use the binary feature vectors in a HDS and discuss the influence of the HDS on the classification results. We draw conclusions in Section 5.

2 Feature Vector Generation

The process of deriving a feature vector from a biometric measurement is called feature vector generation. In this section it is explained how binary feature vectors are derived from facial images such that the privacy of these feature vectors can be protected using the HDS introduced in [14].

2.1 Face Recognition Algorithm

In order to demonstrate the feasibility of template protection for facial biometrics, we concentrate on one particular algorithm for face recognition described in [19, 20].

A schematic of this face recognition algorithm is presented in Figure 1. First a camera is used to register the facial image F. In following steps, the face is localized, the image is normalized and from the normalized facial image, six key objects are identified: left and right eye, left and right eyebrow, mouth and 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 $\vec{X} \in \mathbb{R}^k$, texture information is derived for every fiducial point using the Gabor kernels taken from [18]:

$$\phi_{\nu,\mu}(\vec{p}) = \frac{k_m^2}{4\pi^2} e^{\left(-\frac{k_m^2 p^2}{2\sigma^2}\right)} \left[e^{j(\vec{k}_m \cdot \vec{p})} - e^{-2\pi^2}\right], \text{ where}$$

$$\vec{k}_m = \begin{pmatrix} k_\nu \cos\phi_\mu \\ k_\nu \sin\phi_\mu \end{pmatrix}, k_\nu = 2^{-\frac{\nu+2}{2}}\pi, \phi_\mu = \mu\frac{\pi}{8}, m = 8\nu+\mu. \quad (1)$$

We use 5 frequencies and 8 orientations for the kernels by specifying $\nu = 0, 1..4$ and $\mu = 0, 1..7$, respectively, resulting in 40 kernels. A small image patch centered around every fiducial point is then convolved with every Gabor kernel $\phi_{\nu,\mu}(\vec{p})$ and the modulus of the result is taken as one component $(\vec{X})_i$ in the feature vector $\vec{X}$.

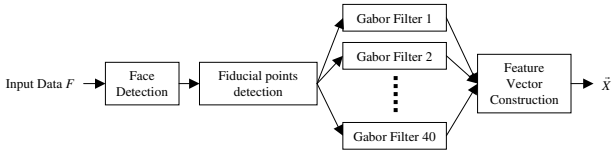

Figure 1. Schematic for extracting feature vectors $\vec{X}$ from facial images.

2.2 Binary Feature Vector Extraction

To protect the privacy of the facial template using the HDS presented in [14] we require binary feature vectors (binary strings). In this section we will explain how the feature vector $\vec{X} \in \mathbb{R}^k$ is transformed into a binary string $Z \in \{0,1\}^K$ with $K \leq k$, by discussing the individual blocks in Figure 2. We assume that, during an enrollment phase, we

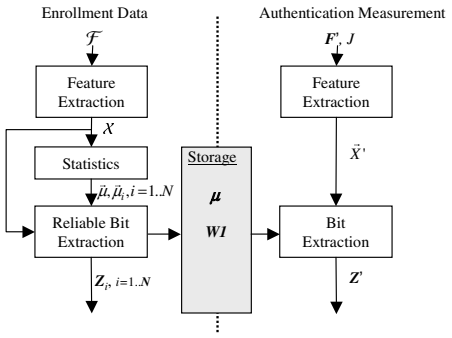

Figure 2. Binary feature vector extraction with N enrolled users.

collected a set of facial images $\mathcal{F} = \{F_{i,j}\}_{i=1..N, j=1..M}$ where $F_{i,j}$ denotes the j-th image of the i-th user such that we have N users and M facial images per user and $\mathcal{F}$ consists of NM images.

Feature Extraction In the Feature Extraction block, feature vectors $\vec{X}$ are extracted from the images in $\mathcal{F}$, according to the method described in Section 2.1. The set of the resulting NM feature vectors is denoted as $\mathcal{X} = \left\{\vec{X}_{i,j}\right\}_{i=1..N, j=1..M}$, where $\vec{X}_{i,j} \in \mathbb{R}^k$ has components $(\vec{X}_{i,j})_t, t = 1..k$.

Statistics This block estimates the mean feature vector $\vec{\mu}_i$ of person i and the mean $\vec{\mu}$ over all enrollment feature vectors as follows,

$$\vec{\mu}_i = \frac{1}{M}\sum_{j=1}^{M}\vec{X}_{i,j}\ , \quad \vec{\mu} = \frac{1}{N}\sum_{i=1}^{N}\vec{\mu}_i \quad . \tag{2}$$

Reliable Bit Extraction For every user i we derive a binary string $Q_i \in \{0,1\}^k$ by quantising $\vec{\mu}_i$ around $\vec{\mu}$ such that for $t = 1..k$

$$(Q_i)_t = \begin{cases} 0 & \text{if } (\vec{\mu_i})_t \leq (\vec{\mu})_t \\ 1 & \text{if } (\vec{\mu_i})_t > (\vec{\mu})_t \ . \end{cases} \tag{3}$$

Next, the *reliability* or *robustness* of a bit $(Q_i)_t$ is determined as follows.

1. Using $(\vec{X}_{i,j})_t, j = 1..M$ we estimate the variance $s_{i,t}^2$ of the t-th component of user i as

$$s_{i,t}^2 = \frac{1}{M-1}\sum_{j=1}^{M}((\vec{X}_{i,j})_t - (\vec{\mu}_i)_t)^2. \tag{4}$$

2. The reliability $R_{i,t}$ of bit t of user i is

$$R_{i,t} = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{|(\vec{\mu}_i)_t - (\vec{\mu})_t|}{\sqrt{2s_{i,t}^2}}\right)\right) \tag{5}$$

where erf is the error function. $R_{i,t}$ is the probability that, for a new measurement $\vec{X}_{i,M+1}$ of the feature vector of user i, we have $(\vec{X}_{i,M+1})_t \leq (\vec{\mu})_t$ if $(Q_i)_t = 0$ or $(\vec{X}_{i,M+1})_t > (\vec{\mu})_t$ if $(Q_i)_t = 1$, assuming a normal distribution for a feature vector component with mean $(\vec{\mu}_i)_t$ and variance $s_{i,t}^2$ (see Fig. 3). Using this definition, bits with a higher reliability have a higher discriminating power due to the larger difference between $(\vec{\mu}_i)_t$ and $(\vec{\mu})_t$ relative to the standard deviation $s_{i,t}$.

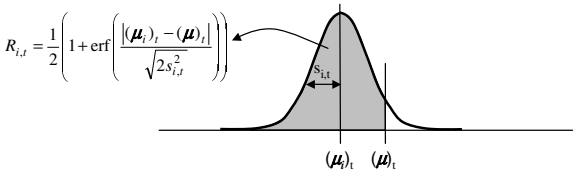

Figure 3. An illustration of the reliabilitiy $R_{i,t}$ of the t-th component of user i.

We finally generate the binary string $Z_i \in \{0,1\}^K$ containing the K most reliable components of Q_i and the vector $\vec{W}1_i$ containing the indices of the reliable bits in Q_i.

During the authentication phase of a claimed identity J, a noisy biometric F' is measured that must be transformed in a binary string Z'. On F' the following computations are performed. i) Features are extracted from F' (see Section 2.1) and a feature vector $\vec{X}'$ is obtained. ii) A bit string Q' is derived by comparing the value of each component $(\vec{X}')_t$ with the mean value $(\vec{\mu})_t$ according to Eq. 3 (where $\vec{\mu_i}$ is replaced by $\vec{X}'$ and Q_i is replaced by Q'). iii) Using the indices in $\vec{W}1_J$, K components from Q' are selected yielding a bit string Z'. This selection procedure will henceforth be denoted using the operator $\circ$ such that we have $Z' = \vec{W}1_J \circ Q'$.

3 Statistical Properties of the Binary Feature Vectors

By quantizing the feature vectors $\vec{X}$ around the population mean $\vec{\mu}$ and selecting a limited number of components, some information in $\vec{X}$ is discarded. It is therefore important to study the statistical properties of the binary feature vectors Z derived from real face databases as well as the classification properties of the feature vectors $\vec{X}$.

First we use a subset of the FERET database [12] containing 237 persons each with at least four images. We randomly selected four images for persons that have more than four images. During feature extraction, 51 fiducial points (see Section 2.1) were used resulting in feature vectors $\vec{X}$ with 2040 components ($k = 2040$).

The classification results of the feature vectors $\vec{X}$ averaged over all possible $3-1$ splits of 3 training measurements and 1 test measurement of the four available images are depicted in Figure 4 (left) using a correlation classifier where the correlation C between two feature vectors $\vec{X}$ and $\vec{Y}$ is defined as

$$C = \frac{(\vec{X}-\vec{\mu})^T(\vec{Y}-\vec{\mu})}{\left\|\vec{X}-\vec{\mu}\right\| \cdot \left\|\vec{Y}-\vec{\mu}\right\|}. \tag{6}$$

It can be seen that the EER = 1.5% for $C = 0.17$.

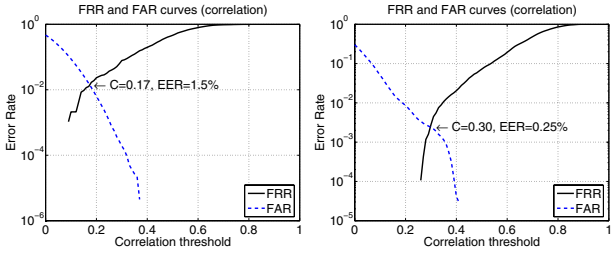

Figure 4. Classification results based on correlation for the FERET database (left) and the Caltech database (right).

In order to get a measure for the statistical independence of the binary feature vectors Z_i, we first follow the approach

by Daugman (e.g. [1]) before looking at the classification results of the vectors Z_i.

For every person i we use all four images to derive Q_i, Z_i and $\vec{W}1_i$ (see Section 2.2 where we chose $K = 511$). Next, for $i = 1..N, l = 1..N, i \neq l$ we determine the Hamming distance HD between Z_i and $\vec{W}1_i \circ Q_l$ resulting in 55932 comparisons. The results in terms of the fractional Hamming distance FHD are given in Figure 5. From the mean of the distribution it follows that $p_e = 0.48$ and from the width we find $\sigma = 0.06$. Also depicted in Figure 5 is the quantile-quantile (Q-Q) plot [17] showing a good resemblance between the observations and an exact binomial distribution. From $\sigma = \sqrt{p_e(1-p_e)/N_{DOF}}$ and p_e we estimate $N_{DOF} = 66$ degrees-of-freedom. It is further clear that if we allow less than 30% errors, the probability on an impersonation is very small.

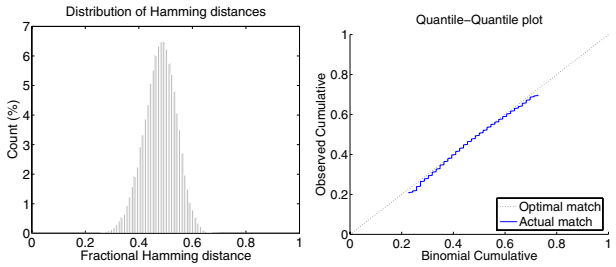

Figure 5. Fractional Hamming distance between-class distribution for the FERET database (left) and the Q-Q plot (right).

The quality of a feature vector with respect to classification does not only depend on the distribution of feature vectors derived from different faces (between-class) but also on the distribution of different measurements of a single face (within-class). The within-class statistical properties for the FERET database are determined as follows. For every person i we take three of the four images to determine Z_i and $\vec{W}1_i$. Next the Hamming distance between Z_i and $\vec{W}1_i \circ Q_i$ is determined where Q_i is derived from the fourth measurement only. This procedure is repeated for all possible $3-1$ splits resulting in 948 comparisons. The results are given in Figure 6 (left) and shows a rather wide distribution, partially overlapping the between-class distribution. Apparently, the variations in the derived feature vectors Z_i due to variations in the measurements in the FERET database is quite large and consequently this will not lead to satisfactory classification results. A second explanation for the wide distribution is that the number of images per user is too small to derive an accurate estimate for the reliability of feature vector components.

If we look at the classification results of the vectors Z_i obtained in a similar manner as the classification results of

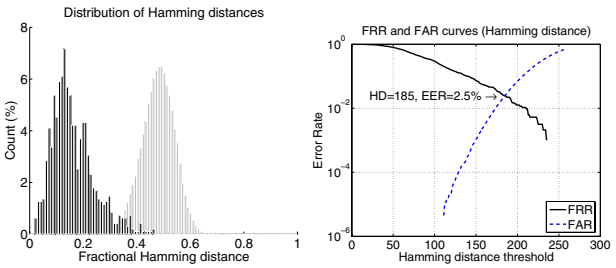

Figure 6. Decision environment for the FERET database (left) and FAR-FRR plots (right).

the vectors $\vec{X}$ but using Hamming distance rather than correlation as a similarity measure, we get the results in Figure 6 (right) with EER = 2.5% for a classification boundary at a Hamming distance of 185 bits (or $FHD = \frac{185}{511} = 0.36$). Compared to the EER=$1.5\%$ obtained for the correlation classifier (see Figure 4 (left)) this means a slight degradation in the classification results. It can further be seen that the EER occurs at a rather high value for HD due to the wide within-class distribution which corresponds to the relatively low value of the correlation classifier for which the EER is obtained.

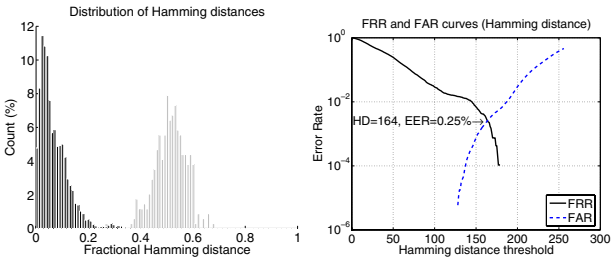

Figure 7. Decision environment for the Caltech database (left) and FAR-FRR plots (right).

We also investigated a face database from Caltech [16] containing 19 persons with at least 11 images per person. The classification results based on correlation using all possible $8-3$ splits are in Figure 4 (right) where we achieve EER = 0.25% for $C = 0.30$. Figure 7(left) gives the results for the between-class and within-class distribution of the binary feature vectors Z_i where it can be seen that the within class distribution of Z_i is less wide than for the FERET database. The classification results are given in Figure 7(right) with EER=0.25% occurring at a Hamming distance of 164 bits (or $FHD = \frac{164}{511} = 0.32$). From these figures it follows that for the Caltech database using the binary vectors Z_i rather than $\vec{X}$ results in no degradation in terms of EER.

4 Helper Data Schemes (HDSs) for Privacy Protection

As explained in Section 1 the privacy and renewability of biometric templates is important if one wants to prevent identity theft and at the same time design flexible biometric systems. In this section we therefore apply a Helper Data System (HDS) as discussed in [14] to the binary feature vectors Z_i.

4.1 A Practical System

The overall system comprising the extraction of the binary vectors Z_i and the HDS is given in Figure 8. Note that the top half is equal to Figure 2 while the lower half is the HDS. The HDS for every user stores helper data $W2_i$

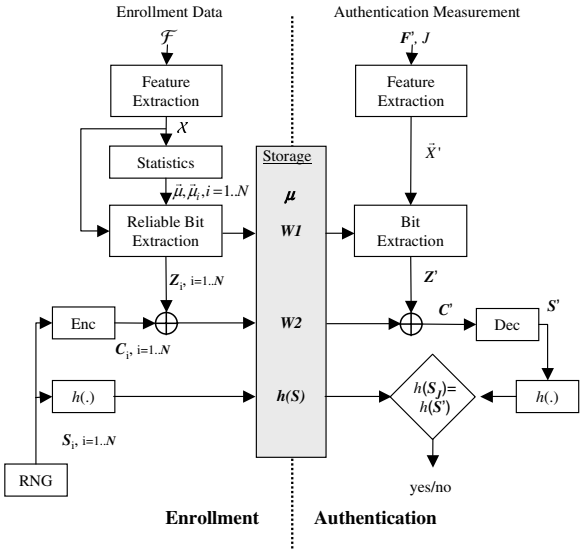

Figure 8. Binary feature vector extraction and template protection with N enrolled users.

and the one-way cryptographic hash[1] $h(S_i)$ of a random string S_i. In order to create helper data $W2_i$ for user i the system uses an error correcting code (ECC) with parameters (K, s, d) where K denotes the length of the code words, s the number of information symbols and d the number of errors that can be corrected. The random string $S_i \in \{0,1\}^s$ is generated by the Random Number Generator (RNG) block in Fig. 8, and encoded into the codeword C_i. The helper data $W2_i$ is then given by $W2_i = C_i \oplus Z_i$ where $\oplus$ stands for bitwise XOR. Knowing $W2_i$ and $h(S_i)$ it is not possible to retrieve $Z_i = C_i \oplus W2_i$ because S_i (and hence C_i) can not be retrieved from $h(S_i)$ due to the one-way property of h. Furthermore, a different S_i leads to a different pair $(W2_i, h(S_i))$ which makes it possible to derive different pairs for a single face thus allowing renewability.

During authentication of a claimed identity J we obtain Z' and set $C' = Z' \oplus W2_J = C_J \oplus (Z_J \oplus Z')$. Finally S' is obtained by decoding C' and $h(S')$ is compared with $h(S_J)$ stored in the database. If both values match authentication is successful.

Because $C' = C_J \oplus (Z_J \oplus Z')$ the success of an authentication completely depends on the Hamming distance between Z_J and Z' through the error correcting capability d of the ECC: for Hamming distances larger than d authentication will fail.

4.2 Discussion

The main motivation behind the use of a HDS is that it is difficult to retrieve Z_i, or worse, $\vec{X}_i$ from the information $W2_i$ and $h(S_i)$ that is stored in, for example, a database and therefore it must be difficult to retrieve S_i from $h(S_i)$ because $Z_i = C_i \oplus W2_i$ and C_i follows directly from S_i by ECC encoding. To prevent an exhaustive search on $h(S_i)$ the size of S_i must be at least 50 bits and this limits the amount of errors d that the ECC can correct and the threshold on the Hamming distance between Z_J and Z'. If we choose a BCH code with $K = 511$ and $s = 58$, we have $d = 91$. From Figures 6 and 7 it is then clear that for both databases the EER cannot be achieved because this would require too many errors to be corrected. The Caltech database allows for an acceptable FRR = 3.5% and FAR $\approx 0\%$ for a maximum of 91 corrected errors but the FERET database leads to an unacceptable FRR = 35% and FAR $\approx 0\%$ due to the large within-class variation and the limited number of measurements per person making it hard to correctly select reliable components. Although a BCH code might not be the best code available, there are fundamental limits on the fraction of errors that can be corrected by ECCs (e.g. [13]).

5 Conclusions

In this paper, binary feature vectors were derived from face data that can be used in Helper Data Systems (HDSs) that allow privacy protection and renewability of biometric templates. We studied the statistical properties of the binary feature vectors derived from the FERET database and the Caltech face database and showed that there is only a slight degradation of classification results in terms of EER when using Hamming distance as a similarity measure for binary strings rather than a correlations measure on the real-valued feature vectors (see Figure 9).

[1] A one-way cryptographic hash function h is a function that is easy to evaluate but hard to invert. Thus, given $h(s)$ it is hard to obtain s.

	EER (for $\vec{X}$)	EER (for Z)	FAR/FRR (for HDS)
FERET	1.5%	2.5%	0% / 35%
Caltech	0.25%	0.25%	0% / 3.5%

Figure 9. Summary of the classification results for the FERET and Caltech databases.

When using the binary feature vectors in a HDS we are constrained by the number of errors that an Error Correcting Code can correct. Consequently the within-class variation should not be too large to obtain satisfactory classification results and sufficient enrollment measurements should be available when deriving the binary feature vectors. For this case the presented a system allows for privacy protected and renewable face templates.

References

[1] J. Daugman. The importance of being random: statistical principles of iris recognition. *Pattern Recognition* 36, pp279-291, 2003.

[2] M. van Dijk, P. Tuyls. Robustness, reliability and security of biometric key distillation data in the information theoretical setting. *26th Benelux Symposium on Information Theory*, Brussels 2005.

[3] 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.

[4] 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/.

[5] J. Goseling, P. Tuyls. Information-Theoretic Approach to Privacy Protection of Biometric Templates Manuscripts. *Proc. IEEE International Symposium on Information Theory (ISIT2004)*, p172, 2004.

[6] International Civil Aviation Organization (ICAO). *http://www.icao.int/*

[7] A.Juels, M. Wattenberg. A Fuzzy Commitment Scheme. *In G. Tsudik, Ed., 6th ACM Conf. Computer and Communication Security*, pp28-36, 1999.

[8] A.Juels, M. Sudan. A Fuzzy Vault Scheme. *Proc. Int'l Symp. Inf. Theory*, A Lapidoth, E.Teletar, Eds., pp408, 2002.

[9] J.-P. Linnartz, P. Tuyls. New shielding functions to enhance privacy and prevent misuse of biometric templates. *Proc. 4th Int. Conf. on Audio- and Video-Based Biometric Person Authentication (AVBPA 2003)*, Springer LNCS 2688, pp393-402, 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] B. Schneier. Inside risks: The use and abuse of biometrics. *Communications of the ACM*, Vol. 42, pp.136.

[12] P.J. Phillips, H. Moon et al. The FERET evaluation methodology for face recognition algorithms. *IEEE Trans. PAMI*, Vol.22, pp1090-1104, 2000.

[13] M. Purser. *Introduction to Error-Correcting Codes*. Artech House, Boston, 1995.

[14] P. Tuyls, A. Akkermans, T. Kevenaar, G.J. Schrijen, A. Bazen, R. Veldhuis. Practical biometric template protection system based on reliable components. *Proc. 5th Int. Conf. on Audio- and Video-Based Biometric Person Authentication (AVBPA 2005)*, Springer LNCS 3546, pp436-446, 2005.

[15] 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 3546, pp310-319, 2005.

[16] M. Weber. Frontal face dataset. *http://www.vision.caltech.edu/html-files/archive*, California Institute of Technology, 1999.

[17] S.B. Vardeman. *Statistics for engineering problem solving*. IEEE Press, 1994.

[18] L. Wiskott, J.M. Fellous et.al. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19 , Issue 7, pp775-779, July 1997.

[19] 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, Vol. 5150, pp1857-1865 (2003).

[20] F. Zuo, P.H.N. de With. Fast facial feature extraction using a deformable shape model with Haar-wavelet based local texture attributes. *Proc. 11th Int'l Conf. on Image Processing (ICIP2004)*, pp1425-1428, 2004.