Improved Privacy Protection in Authentication by Fingerprints

Joep A. de Groot Jean-Paul M.G. Linnartz
Technische Universiteit Eindhoven - TU/e
Signal Processing Systems, Electrical Engineering
Den Dolech 2, 5612 AZ, Eindhoven, the Netherlands
j.a.d.groot@tue.nl j.p.m.g.linnartz@tue.nl

Abstract

We show the feasibility of implementing a new zero leakage quantization scheme on biometric templates. In particular, we investigate the challenge that a system designer has to model the feature distributions prior to actually enrolling participants. For a Gaussian a priori model, we calculate the capacity and privacy leakage for the commonly used biometric dataset FVC2000 and compare it to other verification schemes. We show that the zero leakage scheme achieves an equal error rate of approximately 3.5% while model mismatch leads to an average leakage of less than 0.1 bit per dimension. It outperforms quantization index modulation, both in terms of prevention of leakage and equal error rate, and binary quantization in terms of verification errors that can be tolerated.

1 Introduction

Biometric templates can be misused for identity theft, cross matching or function creep. Preventing this has become a topic of interest. This is mainly because of the threats associated with storing biometric templates and the associated risks of leakage. Template protection schemes with helper data are known to mitigate this issue [1].

Recently [2] we have discovered an extension of these schemes to ensure zero leakage by a nonlinear pre-distortion prior to the quantization scheme, to give the biometric a uniform probability density. However, a disadvantage is that it requires the Cumulative Distribution Function (CDF) of the entire population to be known –or at least to be estimated– before the first user enrolls. The pre-distortion yields an equally distributed secret and uniformly distributed helper data after quantization and thus zero leakage.

We will explore whether a system designer can approximate the pre-distortion as a Gaussian CDF to ensure small leakage, without affecting authentication performance substantially, compared to the known leaky Quantization Index Modulation (QIM) with helper data [1]. As biometric data we use the features extracted from the FVC2000 fingerprint database as described in [3, 4]. Since the a priori model imperfectly approximates the distribution, an attacker who knows the precise distribution of all enrolled participants, gains from some information leakages, but less than from QIM systems.

2 Methods

2.1 Feature extraction

The fingerprint dataset used is the same dataset as in [3, 4]. The data consists of the concatenation of horizontal and vertical squared directional fields [5] and the output of 4 Gabor responses with different angles [3]. The output of both methods is smoothed

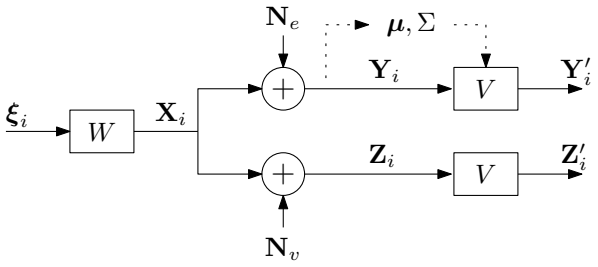


Figure 1: Biometric model, in which the hypothetical variable $\boldsymbol{\xi}_i$ generates biometric $\mathbf{X}_i$, that leads to enrollment sample $\mathbf{Y}_i$ and verification sample $\mathbf{Z}_i$.

by a low-pass Gaussian window before sampling, which was set for an 8 pixel inter-sample distance. A 16×16 sample grid was used, which results in 1536 features for the 6 feature extraction methods. The sample grid is centered at the core point of the fingerprints, i.e. the uppermost point of the innermost curving ridge.

2.2 Fingerprint capacity

In communication theory, the concept of capacity is often used as a theoretical maximum for the amount information that can be reliably send over a channel. A biometric verification scheme can also be considered as a communication channel, in which we try to send a message from enrollment to verification. Moreover it has been shown that this capacity directly relates to the logarithm of the number of users that can be identified uniquely. Therefore it would be interesting to estimate the capacity of the biometric and compare it achieved capacity of the system.

Given the model in Fig. 1 we use Willems' result [6] for identification capacity

$$C = I(\mathbf{Y}_i; \mathbf{Z}_i), \tag{1}$$

with $\mathbf{Y}_i$ the enrollment sample and $\mathbf{Z}_i$ the verification sample of user i. In this model the 'hypothetical' biometric $\boldsymbol{\xi}_i$, a vector of i.i.d. Gaussian variables, is transformed by an matrix W, such that the features $\mathbf{X}_i \in \mathbb{R}^k$ become correlated. Moreover, each measurement $\mathbf{Y}_i$ and $\mathbf{Z}_i$ is subject to uncorrelated additive Gaussian noise $\mathbf{N}_e$ and $\mathbf{N}_v$ respectively.

The mutual information between enrollment and verification sample can be expressed as

$$I(\mathbf{Y}_i; \mathbf{Z}_i) = h(\mathbf{Y}_i) + h(\mathbf{Z}_i) - h(\mathbf{Y}_i, \mathbf{Z}_i) \tag{2}$$

in which $h(\mathbf{X})$ is the differential entropy function. The first term of equation (2) can be calculated as follows

$$h(\mathbf{Y}_i) = \frac{1}{2} \log_2((2\pi e)^k |\Sigma_X + \Sigma_{N_e}|) \tag{3}$$

for $\mathbf{X} \sim \mathcal{N}_k(\boldsymbol{\mu}, \Sigma)$ distributed variables [7]. Since we have assumed the biometric $\mathbf{X}_i$ to be uncorrelated with the additive noise $\mathbf{N}_e$ and $\mathbf{N}_v$ we can add the two corresponding covariance matrices $\Sigma_Y = \Sigma_X + \Sigma_{N_e}$.

The third term can be calculated in a similar way

$$h(\mathbf{Y}_i, \mathbf{Z}_i) = \frac{1}{2} \log_2((2\pi e)^{2k} |\Sigma_{YZ}|) \tag{4}$$

in which

$$\Sigma_{YZ} = \begin{bmatrix} \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 & \mathbb{E}[YZ] - \mathbb{E}[Y]\mathbb{E}[Z] \\ \mathbb{E}[ZY] - \mathbb{E}[Z]\mathbb{E}[Y] & \mathbb{E}[Z^2] - \mathbb{E}[Z]^2 \end{bmatrix} = \begin{bmatrix} \Sigma_X + \Sigma_{N_e} & \Sigma_X \\ \Sigma_X & \Sigma_X + \Sigma_{N_v} \end{bmatrix} \tag{5}$$

Combining equations (3) and (4) according to equation (2) yields

$$I(\mathbf{Y}_i;\mathbf{Z}_i) = \frac{1}{2}\log_2((2\pi e)^k|\Sigma_X+\Sigma_{N_e}|) + \frac{1}{2}\log_2((2\pi e)^k|\Sigma_X+\Sigma_{N_v}|) \tag{6}$$

$$- \frac{1}{2}\log_2((2\pi e)^{2k}|\Sigma_{YZ}|) \tag{7}$$

$$= \frac{1}{2}\log_2\left(\frac{|\Sigma_X+\Sigma_{N_e}||\Sigma_X+\Sigma_{N_v}|}{|\Sigma_{YZ}|}\right) \tag{8}$$

To simplify calculation of this value we diagonalize the covariance matrices by choosing a proper orthogonal transformation V for the $\mathbf{Y}_i$ and $\mathbf{Z}_i$ samples as depicted in Figure 1. This can be achieved by choosing the columns of V equal to the Eigenvectors of Σ_X and multiplying as follows

$$\mathbf{Y}'_{i,j} = V^T \cdot \mathbf{Y}_{i,j} \qquad \text{and} \qquad \mathbf{Z}'_{i,j} = V^T \cdot \mathbf{Z}_{i,j} \tag{9}$$

This operation is similar to a principal component analysis based on the covariance matrix Σ_X. Since we have assumed to noise to be uncorrelated with the biometrics and with itself now both $\Sigma_{X'}$ and $\Sigma_{T'}$ will only have non-zero elements on their diagonal. Therefore we can write

$$I(\mathbf{Y}'_i;\mathbf{Z}'_i) = \frac{1}{2}\sum_{i=0}^{k}\log_2\left(\frac{(\Sigma_{X'}+\Sigma_{N'_e})\cdot(\Sigma_{X'}+\Sigma_{N'_v})}{(\Sigma_{X'}+\Sigma_{N'_e})\cdot(\Sigma_{X'}+\Sigma_{N'_v})-\Sigma_{X'}\Sigma_{X'}}\right)_{i,i} \tag{10}$$

which simplifies to a intuitive form in case one could observe $\mathbf{X}_i$ exactly, for instance by enrolling the average of multiple fingerprints samples, such that $\Sigma_{N'_e} \to 0$, hence

$$I(\mathbf{Y}'_i;\mathbf{Z}'_i) = \frac{1}{2}\sum_{i=0}^{k}\log_2\left(1+\frac{\Sigma_{X'}}{\Sigma_{N'_v}}\right)_{i,i} \tag{11}$$

These equations for correlated variables extend the results of [8] for i.i.d. Gaussian variables.

Based on our model it is not possible to observe covariance matrix Σ_X directly, since our observed samples $\mathbf{Y}$ and $\mathbf{Z}$ are always subject to additive noise. An estimate can be obtained as follows. First calculate mean values over M observations and N enrolled users. Subsequently calculate the average within-class covariance matrix Σ_w and between-class covariance matrix Σ_b

$$\Sigma_w = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{M-1}\sum_{j=1}^{M}(\mathbf{Y}'_{i,j}-\boldsymbol{\mu}_i)(\mathbf{Y}'_{i,j}-\boldsymbol{\mu}_i)^T\right], \quad \text{with } \boldsymbol{\mu}_i = \frac{1}{M}\sum_{j=1}^{M}\mathbf{Y}'_{i,j} \tag{12}$$

$$\Sigma_b = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{\mu}_i-\boldsymbol{\mu})(\boldsymbol{\mu}_i-\boldsymbol{\mu})^T, \qquad \text{with } \boldsymbol{\mu} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mu}_i \tag{13}$$

respectively. We expect the noise contribution to be reduced by averaging over all samples of a user, yielding our best estimate of the underlying biometric value, i.e. $\mathbf{X}_i \approx \boldsymbol{\mu}_i$. Using this average in turn to calculate the covariance gives an approximation of the biometric covariance, i.e. $\Sigma_{X'} \approx \Sigma_b$. Similarly, the noise covariances can be represented by $\Sigma_{N_e} \approx \Sigma_{N_e} \approx \Sigma_w$.

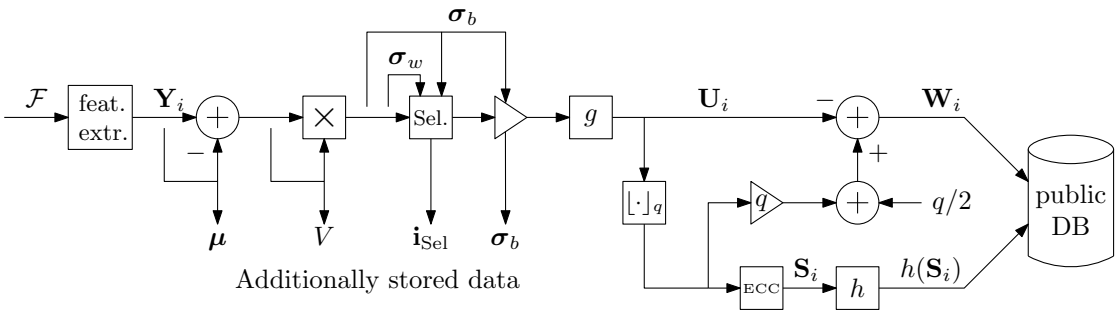


Figure 2: Enrollment of the fingerprints

2.3 Enrollment

During enrollment we made the features zero mean and apply the same orthogonal transformation as described in section 2.2. However, there is one difference, we would not be able to observe the verification samples under practical circumstances, therefore we exclude them from calculation of the covariance matrix. To be able to compare the same transformed features during verification we also store the subtracted mean value and transformation matrix in our database and re-apply it prior to verification.

From the transformed features, a selection of K features is made based on their signal to noise ratio, which is defined as between-class variance divided by within-class variance, hence $\text{SNR}_i = (\Sigma_b)_{i,i}/(\Sigma_w)_{i,i}$. Again this selection is stored to be used during verification.

Finally, the between-class standard deviation σ_b is used to scale the features to unit variance, which is required for the pre-distortion function g. For this work we have set the non-linear pre-distortion function equal to the standard $(0, 1)$ Gaussian CDF. The undistorted features are used in a binary and QIM quantization scheme with helper data for comparison. The entire enrollment is depicted in Figure 2.

2.4 Verification

Prior to verification we apply the same transformation, feature selection, scaling and non-linear pre-distortion. Subsequently the user specific helper data is added and the secret extracted by quantizing the value. An additional saturation is required to prevent the observation of undefined secret values. This can happen due to the addition of helper data, which might push the transformed verification sample outside the unit interval. The entire verification is depicted in Figure 3.

2.5 Leakage analysis

We define leakage as mutual information between helper data and secret, hence

$$I(W;S) = H(S) - H(S|W) \tag{14}$$

$$= -\sum_{n=0}^{1} \mathrm{P}(S=n)\log_2 \mathrm{P}(S=n) + \int_{-q/2}^{q/2} \sum_{n=0}^{1} f_W(w|S=n)\mathrm{P}(S=n)\log_2 \frac{f_W(w|S=n)}{f_W(w)}\mathrm{P}(S=n)\, dw. \tag{15}$$

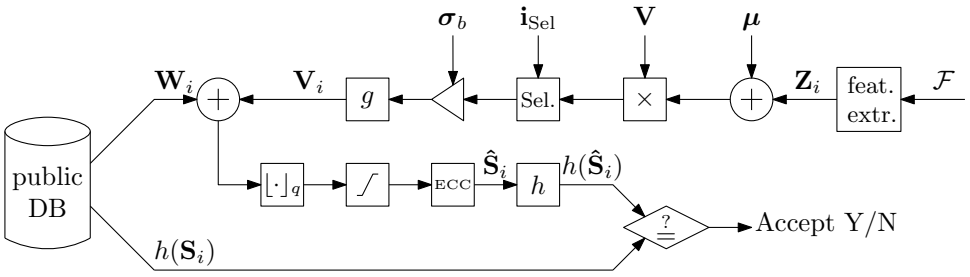


Figure 3: Verification of the fingerprints

In our experiment with a biometric database, of course, we do not encounter a probability density for the helper data, but wee see a set of discrete, measured values. However, we might assume that the within-class distribution follows a Gaussian distribution. Therefore we estimate f_W by modeling it based on a mixture of N Gaussian distributions, each having the mean value $(\boldsymbol{\mu}_i)_n$ and variance $(\boldsymbol{\sigma}_i)_n$ that corresponds to the measured average and spread in the within-class distribution for that particular user. The density function for the samples is estimated as

$$f_{y_n}(y_n) = \sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}(\boldsymbol{\sigma}_i)_n^2} \exp\left(-\frac{(y-(\boldsymbol{\mu}_i)_n)^2}{2(\boldsymbol{\sigma}_i)_n^2}\right) \tag{16}$$

The helper data density function can be derived from this function for both the QIM and zero leakage quantization scheme, namely for the QIM scheme

$$f_W(w,s)\mathrm{P}(s) = \sum_{j=-\infty}^{\infty} f_{y_n}\left(w + q\left(s + 2j + \frac{1}{2}\right)\right) \qquad \text{for } -q/2 \le w < q/2 \tag{17}$$

While for the zero leakage quantization scheme we have to pay attention to the change of variables by the distortion function, hence

$$f_W(w,s)\mathrm{P}(s) = f_{y_n}\left(g^{-1}\left(w + \frac{s}{2} + \frac{1}{4}\right)\right) \cdot \frac{d}{dw}\, g^{-1}\left(w + \frac{s}{2} + \frac{1}{4}\right) \tag{18}$$

$$= f_{y_n}\left(\Phi^{-1}\left(w + \frac{s}{2} + \frac{1}{4}\right)\right) \cdot \sqrt{\frac{\pi}{2}} \exp\left(\frac{1}{2}\Phi^{-1}\left(w + \frac{s}{2} + \frac{1}{4}\right)^2\right), \tag{19}$$

in which $\Phi^{-1}(x) = \sqrt{2}\,\mathrm{erf}^{-1}(2x-1)$, the inverse Gaussian cumulative distribution function.

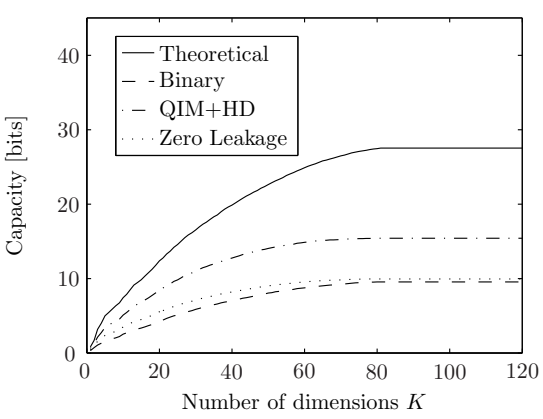
Finally, the unconditional density function can easily be derived by using

$$f_W(w) = f_W(w|S=0)\mathrm{P}(S=0) + f_W(w|S=1)\mathrm{P}(S=1) \tag{20}$$

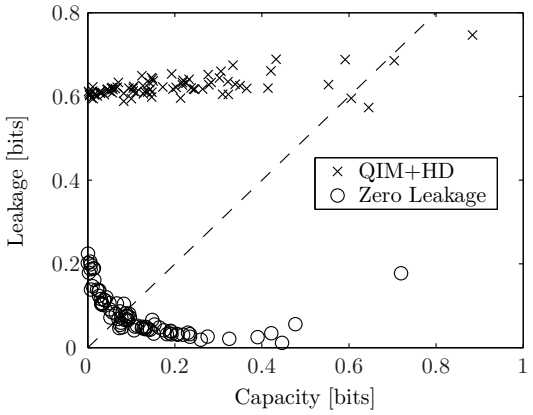
3 Results

3.1 Preliminaries

We use the FVC2000 fingerprint database [9]. The database consists of 8 fingerprints images from 110 subjects. From these fingerprint images we use the directional field and Gabor response features as described in [5] and [3] respectively.



(a) Theoretical and achieved capacity



(b) Capacity versus leakage for the first 81 dimensions

Figure 4: Capacity and leakage results

During experiments we found many of the users to be badly enrolled, i.e. the sampling grid was partly outside the filter responses output. This resulted in row and columns entirely filled with zeros, which cannot be considered as a biometric feature. Therefore we decided to reject all samples of a users that had at least one sample with a row or column at either side filled entirely with zeros. This resulted in rejection of almost half of the available users, so 59 users with 8 correct sampler remained.

To maintain a larger sample space, we first cropped the sample grids by 2 pixels, so a sample grid of 12 was left. This did not significantly influence classification performance. However, the crop followed by rejection of users with at least one bad enrollment, resulted in a better populated sample space with 82 users, each having 8 correct fingerprint samples.

Experiments were carried out with 6 enrollment and 2 verification samples, which resulted in 28 possible splits, which were all used to assess classification performance. This is a similar approach as in [4].

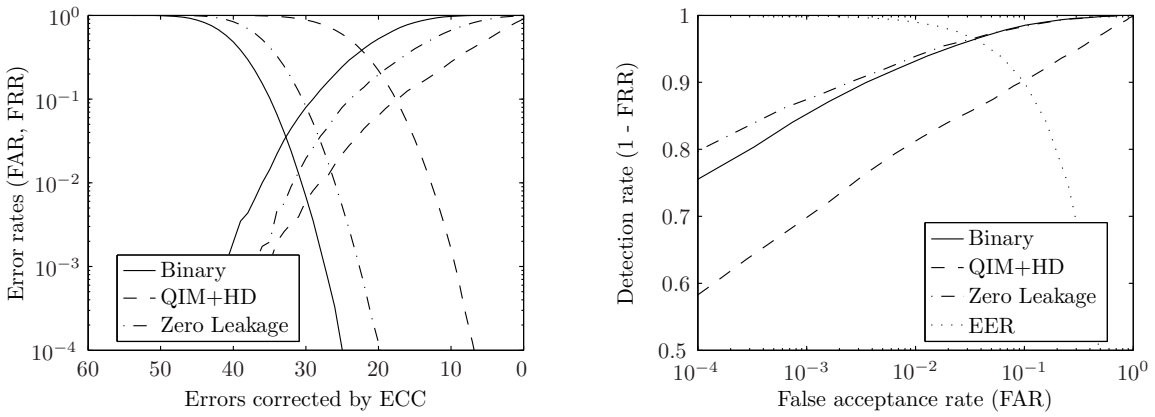
A small experiment showed that basing the PCA on all samples, instead of the enrollment samples only, did not make a significant difference. Still we have used the enrollment samples only to create the transformation matrix V.

3.2 Capacity

Figure 4(a) shows the cumulative verification capacity for an increasing number of dimensions sorted by Eigenvalue. The method based on the diagonals (10) gives gives the most optimistic estimate of capacity. Based on this result one could state that under these circumstances one could achieve a capacity of approximately 27.5 bits from 81 dimensions.

The results based on estimated error probability, $C = 1 - h(P_e)$, are however much lower. Both binary and zero leakage quantification achieve a capacity just below 10 bits, approximately 9.5 and 9.9 bits respectively.

The QIM with helper data (QIM+HD) scheme cannot really be compared. Any arbitrary value of capacity can be achieved by increasing the quantization width q, but this comes at the cost of a high false acceptance and leakage, as can be seen in figure 5(a) and 4(b). The latter depicts the leakage results based on the estimation described in section 2.5 versus the capacity based on the estimated error rate P_e. For our choice of $q = 2.5$ we achieve a capacity of approximately 15 bits.



(a) Error rates

(b) Receiver operating curves

Figure 5: Classification results for the various methods

3.3 Classification

Classification results are depicted in figure 5. Since we consider binary classification schemes the extracted bit will either be equal or different in each dimension. Therefore detection performance and false acceptance will depend on the number of errors the that the system can handle by near exact matching (not privacy protecting) or by helper data and an error correcting code (ECC). The false acceptance rate (FAR, dashed) and false rejection rate (FRR, solid) for a decreasing number of tolerated errors is depicted in 5(a). Based on the capacity results, the total number of features is set to 81.

False rejection for the methods with helper data is clearly lower, but this gain is canceled out by the higher false acceptance. Especially the QIM method has a very high false acceptance. However, both methods reach their equal error rate (EER) at a lower number of tolerated errors, which makes the enrolled secret larger.

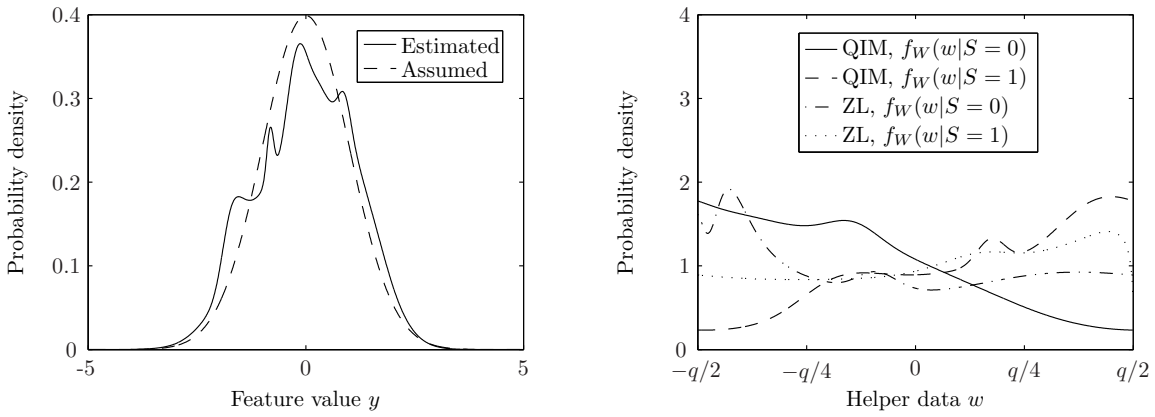
Combining both FAR and FRR as receiver operating curves, depicted in Figure 5(b), put the classification performance of the compared methods into perspective. The binary classification and zero leakage scheme perform almost equally at 33 and 31 tolerated errors and an EER (dotted line) of 3.6% and 3.5% respectively. The QIM scheme achieves an EER of 9.8% at 19 tolerated errors.

3.4 Leakage

Finally we will assess the leakage in both helper data schemes. The zero leakage scheme should ideally not leak any information about the enrolled secret. However, in a practical situation, where a system designer has to pick a pre-distortion prior to having seen the actual population, there will be some leakage. This leakage is caused by the mismatch between the pre-distortion function g and the actual distribution of the features. An example for the first principal component is given in Figure 6(a), in which the feature's density estimation is based on the method presented in section 2.5.

The distortion mismatch causes the conditional density functions of the helper data to be different, i.e. $f_W(w|S=0) \neq f_W(w|S=1)$, which on its turn causes the leakage. Figure 6(b) gives an example of the estimated conditional density functions of the helper data. Although the conditional density functions for the zero leakage scheme (ZL) are not equal, they is a smaller difference between them than between those of the QIM scheme.

Figure 4(b) depicts the capacity and leakage per dimension. The sensible system region is $I(w;s) \ll C$, so dimensions in the lower right corner are the most useful. We see that QIM has seemingly higher capacity, but this is largely caused by the leaky



(a) Example of an estimation of the feature density function.

(b) Example of the helper data conditional density functions.

Figure 6: Examples of the density estimation

helper data. As can be seen from this figure, the leakage of the zero leakage scheme is in practice also not zero, but the features have a leakage of 0.08 bit on average. This is much lower than the leakage of the QIM scheme, of which the average leakage is more than 0.6 bit, which is a factor 7.5 higher.

4 Remarks

This first attempt to implement the zero leakage quantifier scheme proved to reduce leakage, but it also showed that leakage did not become zero. The estimation of the pre-distortion function strongly determines the performance and achieved reduction of leakage. Therefore it is clear that simply assuming the features to be Gaussian did not yield to most optimal results. We will continue searching for a better implementation of the pre-distortion.

Also the unexpectedly large false acceptance rate for the zero leakage scheme could most likely be explained by the mismatch between the applied pre-distortion function and the actual distribution of the features. If one would compared this to theory, in which we have assumed an i.i.d correlated biometrics model [2] with $\rho = 0$ for an impostor and $\rho \approx .9$ for a genuine user, not only FRR, but also FAR should be lower, because an impostor is more likely to cause errors in the zero leakage scheme than in a binary quantization scheme.

5 Conclusions

We have shown that the zero leakage scheme maintains leakage to be below 0.1 bit on average, even if the distribution of participants to be enrolled has to be modelled in advance. With this quantization scheme we achieved an EER $\approx 3.5\%$. Although the binary quantization scheme can reach a similar EER, it is outperformed by both the zero leakage and QIM scheme in a practical regime when only a few errors in a genuine user verification can be tolerated. Moreover we have introduced a method to estimate verification capacity of a biometric and a method to estimate leakage in a biometric quantification scheme when only a limited number of biometric samples is available.

References

[1] J.-P. Linnartz and P. Tuyls, "New shielding functions to enhance privacy and prevent misuse of biometric templates," in *Audio- and Video-Based Biometric Person Authentication*, Springer, 2003.

[2] J. de Groot and J.-P. Linnartz, "Zero leakage quantization scheme for biometric verification," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process*, May 2011. In Press.

[3] A. Bazen and R. Veldhuis, "Likelihood-ratio-based biometric verification," *IEEE T Circ Syst Vid*, vol. 14, pp. 86 – 94, 2004.

[4] P. Tuyls, A. Akkermans, T. Kevenaar, G.-J. Schrijen, A. Bazen, and R. Veldhuis, "Practical biometric authentication with template protection," in *Audio- and Video-Based Biometric Person Authentication*, pp. 436–446, Springer, 2005.

[5] A. Bazen and S. Gerez, "Systematic methods for the computation of the directional fields and singular points of fingerprints," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, pp. 905 –919, July 2002.

[6] F. Willems, T. Kalker, J. Goseling, and J.-P. Linnartz, "On the capacity of a biometrical identification system," in *Information Theory, 2003. Proceedings. IEEE International Symposium on*, p. 82, 2003.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc., second ed., 2005.

[8] J.-P. Linnartz, P. Tuyls, and B. Skoric, *A Communication-Theoretical View on Secret Extraction*, ch. 4, pp. 57–77. Security with Noisy Data, Springer, 2007.

[9] D. Maio, D. Maltoni, R. Cappelli, J. Wayman, and A. Jain, "Fvc2000: fingerprint verification competition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, pp. 402 –412, Mar. 2002.