A protection scheme for enhancing biometric template security and discriminability

Marco GRASSI [a,1], Marcos FAUNDEZ-ZANUY [b,2]
[a] *D.I.B.E.T. Università Politecnica delle Marche, Ancona, Italy*
[b] *Escola Universitària Politècnica de Mataró (Adscrita a la UPC) Matarò, Spain*

Abstract. As biometric techniques are being used in a growing number of applications, biometric template security is becoming of great concern in biometric systems. Unlike passwords and tokens, compromised biometric templates cannot be revoked and reissued. In this paper we propose a biometric authentication system based on face recognition using 2D Discrete Cosine Transform and neural networks. A discriminability criterion is used to select the DCT coefficients that make up the biometric template and a user dependent pseudo-random ordering of the DCT template coefficients is applied to provide template security. Experimental results show how the application of such techniques leads to both, enhanced recognition performance and system security, because user authentication process merges biometric recognition with the knowledge of a secret key.

Keywords. Biometrics, Security, Template protection, Frequency selection, Face recognition

Introduction

In recent years, biometric techniques have been used in a wide spread of applications, ranging from large crime and terrorism prevention to ambient intelligence, from secure payment systems to physic access control. Such growing use of biometrics has lead to rising concerns about the security and privacy of the biometric data itself, i.e. of the set of extracted biometric features stored in a central database or in a smartcard that constitutes the biometric template. In fact, unlike passwords and tokens, compromised biometric templates cannot be revoked and reissued. Therefore in the case it is illegally acquired by an attacker, not only the security of the system is at risk but the privacy and the security of the user may be compromised forever too. In fact, the raw biometric data can be recovered from the biometric template and exploited to attacks also other systems that rely on the same biometric feature of the user. Granting template security represents therefore one of the most important issue of a practical biometric system.

[1]Department of Biomedical, Electronic and Telecommunication Engineering Università Politecnica delle Marche, Ancona, Italy. m.grassi@univpm.it . htt://www.3medialabs.org.
[2]Escola Universitària Politècnica de Mataró (Adscrita a la UPC) 08303 MATARO (BARCELONA), Spain faundez@tecnocampus.com. http://www.eupmt.es/veu.

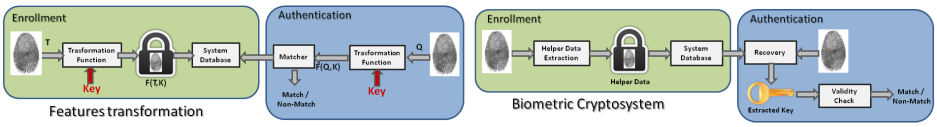

Figure 1. Template protecion schemas: features transformation and biometic cryptosystems

In this paper, we propose a novel template protection schema based on the combination of a discrimanability criterion for the selection of the template coefficients and of user dependent pseudo-random permutation in their ordering. As proof of concept, we show how the application of the proposed techniques in securing the DCT template of a face authentication system not only leads to a security enhancement but also to an improvement of the biometric verification performances.

1. Biometric template protection techniques

An ideal biometric template protection scheme should possess the following properties:

1. *Diversity*: the secured template must not allow cross-matching across databases, thereby ensuring the user's privacy.
2. *Revocability*: it should be straightforward to revoke a compromised template and reissue a new one based on the same biometric data.
3. *Security*: it must be computationally hard to obtain the original biometric template from the secure template. This property prevents a hacker from creating a physical spoof of the biometric trait from a stolen template.
4. *Performance*: the biometric template protection scheme should not degrade the recognition performance (identification or verification rates).

To overcome this problem, the commonly proposed approach is not to store the original biometric template but rather a transformed version of the original template. Standard encryption techniques are not useful for securing biometric templates. Apart from the gradual erosion of the computational difficulty of the mathematical problems on which cryptology is based, due to developments in computation which would increase the vulnerability of biometric systems [2], they require to decrypt the stored template before to perform matching with the query template, leaving the template exposed during every authentication attempt. In addition encryption is not a smooth function and a small difference in the values of the feature sets extracted from the raw biometric data would lead to a very large difference in the resulting encrypted features. For these reasons, many algorithms have been developed on purpose to provide template security, most of which make can be classified in two categories: features transformation and biometric cryptosystems. In the following we present an introductory overview of such techniques. A more exhaustive presentation of state of the art in template protection schemas goes beyond the purpose of this paper and can be found in [1].

1.1. Features transformation

In the enrollment phase, a transformation function *f(x)* is applied to the biometric information and only the transformed template is stored in the database. The parameters of

the transformation function are typically derived from a random key or password. In the authentication process, the same transformation function is applied to query features (Q) and the transformed query (F(Q;K)) is directly matched against the transformed template (F(T;K)).

Feature transformation techniques can be further divided into two categories according to the property of the transformation function:

- *Salting trasformations*. The biometric features are transformed using an invertible function defined by a user-specific key or password, which must be kept secret. The introduction of a secret key ensures revocability. In fact, in case a template is compromised, it's easy to revoke it and replace it with a new one generated by using a different user-specific key. By the way, if the user-specific key is compromised, the template is no longer secure, because the transformation is usually invertible. Different type of salting function have been used for template transformation, as symmetric hash functions (polynomials) [3] or Gaussian functions [4].
- *Non-invertible transformation* The biometric template is secured by applying a non-invertible transformation function that is "easy to compute" but "hard to invert". Even if the key and/or the transformed template are known, it is computationally hard for an adversary to recover the original biometric template. This provides better security than the salting approach. Also diversity and revocability can be achieved by using application-specific and user specific transformation functions, respectively. The main drawback of this approach is the trade-off between discriminability and noninvertibility of the transformation function, which in general leads to a decreese of the recognition performances. Biohashing [5] and random multispace quantization (RMQ) [6] algorithms are examples of this approach.

1.2. Biometic cryptosystems

In the biometric cryptosystems some public information about the biometric template, called *helper data* , is stored. The helper data does not reveal any significant information about the original biometric template but it is needed during matching to extract a cryptographic key from the query biometric features. Matching is performed indirectly by verifying the correctness of the extracted key. Error correction coding techniques are typically used to handle intra-user variations. Biometric cryptosystems offer high security but are not designed to provide diversity and revocability.

Even, biometric cryptosystems can be split into two groups, depending on how the helper data is obtained:

- *Key-binding biometric cryptosystem*. The helper data is obtained by binding a key, that is independent of the biometric features with the biometric template. It's computationally hard to decode the key or the template without any knowledge of the user's biometric data. Error correction coding techniques are used to provide tolerance to intra-user variations in biometric data due. However, their application precludes the use of sophisticated matcher leading to a reduction in the matching accuracy.

- *Key generation biometric cryptosystems*. The helper data is derived only from the biometric template and the cryptographic key is directly generated from the helper data and the query biometric features. Direct key generation from biometrics is an appealing template protection approach which can also be very useful in cryptographic applications. Anyway, it is difficult to generate a key that could grant at the same time high stability and entropy due to intra-user variation in the template. It's hard to develop a scheme that generates the same key for different templates of the same person and at the same time very different keys for different persons.

Two of the most popular techniques for biometric cryptosystem are the fuzzy commitment scheme [7], which treats the biometric template itself as a corrupted codeword, and fuzzy vault scheme [8], where the fuzzy vault is a cryptographic construction that can be thought of as a form of error-tolerant encryption operation where keys consist of sets of biometric features.

Interesting techniques also exist that try to combine features transformation and biometric cryptosystem. For example, [9] propose an hybrid algorithm based on random projection, discriminability-preserving (DP) transform, and fuzzy commitment scheme.

2. Pseudo-random permutation

In our proposal, we use a feature transformation approach that belongs to salting functions. A different random permutation for each user is applied to template coefficients [10], which are constituted by the DCT components of the two dimensional transform of the face image. The permutation order it is given by a Key, which must be kept secret. During the authentication process the user declare his identity to the system and the corresponding ordering is applied to his collected coefficients, which are compared with the stored template, without need to decrypt it before comparison. The proposed schema leads to an increase of privacy because it is computationally very hard to obtain the face image without knowing the permutation order, considering that for a feature vector template of N coefficients the number of possible permutations is equal to N!. There is also an enhancement of the recognition rates because an impostor cannot access the system without knowing the correct permutation order of the template. In addition, even if an impostor enter in possess of a secret key and he is able to sort his features vector according to permutation order of an genuine user, the system security is not compromised but it's still provided by biometric authentication.

3. Frequency selection

DCT (Discrete Cosine Transform) is very used in image processing due to its energy compaction properties. Most of the image information is in fact located in the upper left corner of the transformed image allowing dimensionality reduction. Typically, biometric template based on DCT transform are obtained selecting a top left square region of the transformed image. In order to improve this selection mechanism which is only based on energy, we propose a novel discriminability criteria based on a measure of the inter and intra class variation of the frequencies. The goal is to pick up those frequencies that yield

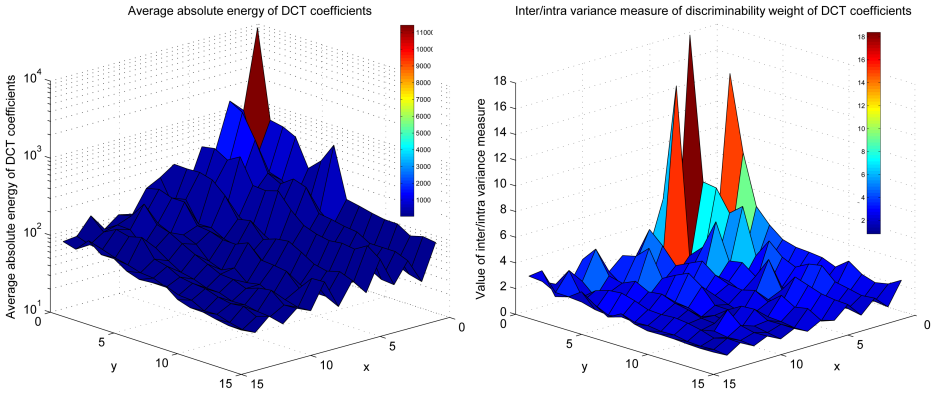

Figure 2. Template protecion schemas: features transformation and biometic cryptosystem

a low intra-class variation and high inter-class variation. Defined P and F respectively the number of people inside the database and the number of face image for each of them, $I_{p,f}(f_1, f_2)$ the f-th transformed image of the p-th person and $m(x, y)$ is the average value of each frequency obtained from the whole training subset images, we propose the following measure:

$$M_1(x, y) = \frac{\sigma^2_{inter}(x, y)}{\sigma^2_{intra}(x, y)} \tag{1}$$

where,

$$\sigma^2_{inter}(x, y) = \frac{1}{P \times F} \sum_{p=1}^{P} \sum_{f=1}^{F} \left(I_{p,f}(x, y) - m(x, y)\right)^2 \tag{2}$$

$\forall p = 1, ..., P$ is the variance of every frequency is the variance of each frequency evaluated over the whole training subset

$$\sigma^2_{intra}(x, y) = \frac{1}{P} \sum_{p=1}^{P} (\sigma^2_{p,f}(x, y)) = \frac{1}{P} \sum_{p=1}^{P} (\frac{1}{F} \sum_{f=1}^{F} \left(I_{p,f}(x, y) - m(x, y)\right)^2) \tag{3}$$

$\forall p = 1, ..., P$ is the average of the variance of each frequency for each person.

Experimental results conducted over the AR database show how the application of the proposed frequency selection criterion leads to a significant improvement in biometric recognition performances both in identification and in verification, both using Radial Basis Function (RBF) and Multilayer Perceptron (MLP) neural network (Figure 3).

4. The Face Authentication System

As proof of concept, we applied the proposed protection schema to an identity verification system based on biometric face recognition. In the system, the frequency selection

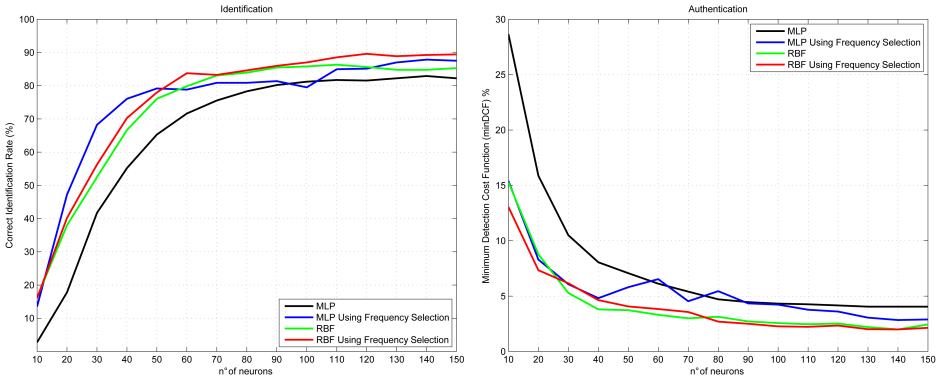

Figure 3. Identification and Authentication Rates (%) using RBF and MLP neural networks with and without frequency selection

algorithm presented in above has been employed to select the most discriminant features used to compose the biometric template, then psedo-random permutation have been applied in securing the obtained template, as shown in Figure 4. An RBF (Radial Basis Function) neural network is employed for classification.

The AR database [11] has been used to test our system performances. AR is a publicly available database of 126 individuals, with 26 images of each, taken in two different sessions at a time distance of two weeks, varying the lighting and the facial expression. We have used 10 of the 26 excluding the ones overexposed and the ones in which the face was partially occluded by sunglasses or scarves, of 117 of the 126 individuals, those of the rest being either not complete or not available. All the images have been cropped and normalized semi-automatically to 64x74 grey ones.

We considered both the presence of genuine users (59 users, 32 man and 27 women) and of impostors (58 impostors, 31 man and 27 women) trying to violate the system, also taking into account the possibility that they have stolen the template.

The RBF neural network has an input layer with 100 neurons, equal to the number of coefficients of the DCT template and an output layer of 59 neurons, equal to the number of users. In our simulation we consider an hidden layer with a number of neurons varying from 10 to 200 and a spread varyng between 1 and 4. We use 5 of the 10 images of each user (the other 5 for test) for training the net according to the following procedure: if the input is a genuine face, belonging to the i-th person of the DB, the net will respond with a 1 value for the i-th output neuron and -1 for the others. In the tests, if the highest value of the output correspond with the declared identity corresponds and such value is greater than an opportune threshold the user is authenticated otherwise is refused.

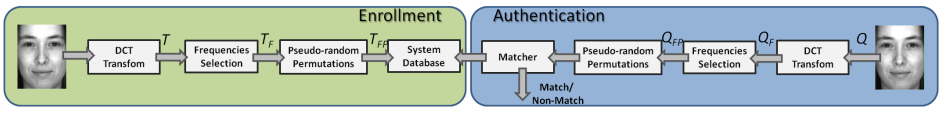

Figure 4. Template protection schema for the face authentication System

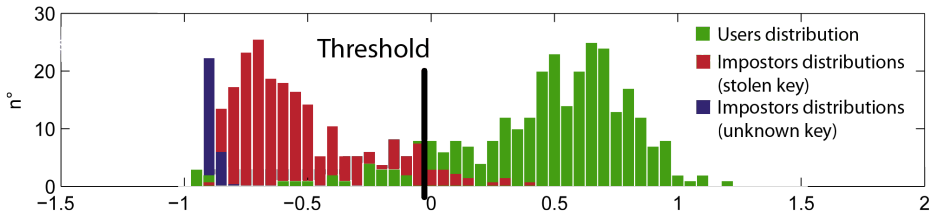

Figure 5. Setting the threshold

We use the first 29 "train impostors", considering 10 images each, to set the threshold for the net output in order to discriminate users and impostors and the others 29 Test Impostors, with 10 images each, for testing the system. There is an obvious thread-off between false accept and false reject and the threshold has been set considering the value that minimizes the total error (false accept + false reject) for the train impostors set as show in Figure 4.

We compared system performances using different DCT templates of 100 coefficients: the original, with frequencies selection, with pseudo-random permutation and with both frequencies selection and pseudo-random permutation. Experimental results highlight how the combination of frequencies selection and pseudo-random permutations leads to the best results. In particular, no impostor has access to the system without knowing the secret key and even knowing the key the probability that an impostor accesses the system remain minor than 8%.

5. Conclusions

In this paper we have presented a novel technique for the template protection, based on a discriminability criterion for the selection of the DCT coefficients that make up the biometric template and on a user dependent pseudo-random of their ordering. Besides to

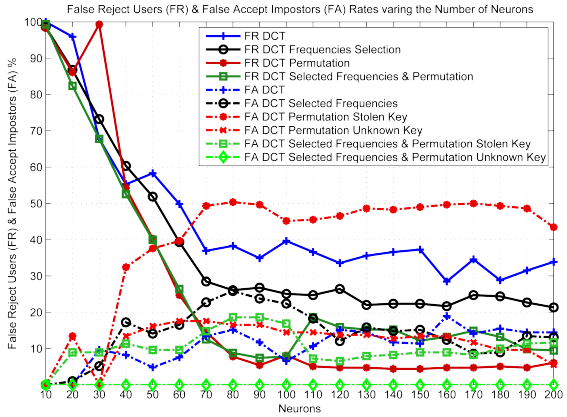

Figure 6. Total Error Rate (%) using different templates and a RBF neural network as classifier

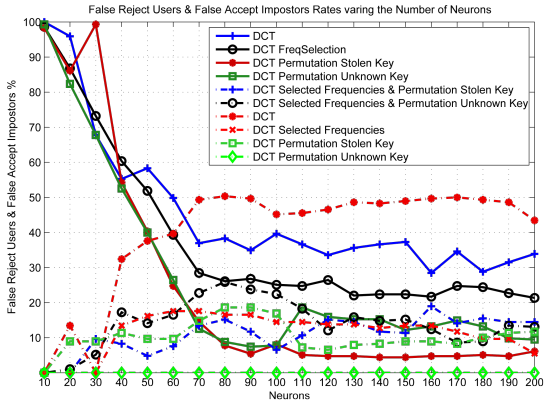

Figure 7. False Accepts and False Rejects Rates (%) using different templates and a RBF neural network as classifier

security enhancement, the application of the proposed technique leads to an improvement in the biometric verification performances.

The security of the system requires the secrecy of a key, that however results computationally very hard to crack taking into account that if the key is kept secret the number of different permutations equals 100!. In addition, even if an impostor enters in possess of the secret key, the system security is not completely compromised but it's still granted by biometric authentication and the probability of the impostor to access the system remains under 8 %.

References

[1] Jain, A. K., Nandakumar, K., Nagar, A., "Biometric template security". Eurasip journal on Advances in Signal Processing. Special issue on Biometrics. Pp.1-20, January 2008.

[2] Faundez Zanuy, M., "Biometric security technology" IEEE Aerospace and Electronic Systems Magazine, Vol.21 nž 6, pp.15-26, June 2006.

[3] S. Tulyakov, V. Chavan, and V. Govindaraju, "Symmetric hash functions for fingerprint minutiae," in Proc. Int.Workshop Pattern Recognition for Crime Prevention, Security, and Surveillance, 2005, pp. 30-38.

[4] [43] Y. Sutcu, H. Sencar, and N. Nemon, "A secure biometric authentication scheme based on robust hashing," in Proc. SeventhWorkshop Multimedia and Security, 2005, pp. 111-116.

[5] A. Goh and D. C. L. Ngo, "Computation of cryptographic keys from face biometrics," in Proc. 7th IFIP TC6/TC11 Conf. Commun. Multimedia Security, 2003, vol. 22, pp. 1-13.

[6] A. Teoh, A. Goh, and D. Ngo, "Random multispace quantization as an analytic mechanism for biohashing of biometric and random identity inputs," IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 12, pp. 1892-1901, Dec. 2006.

[7] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in Proc. Sixth ACM Conf. Comp. and Commun. Security, 1999, pp. 28-36.

[8] A. Juels and M. Sudan, "A fuzzy vault scheme," in IEEE Int. Symp.Information Theory, 2002, p. 408.

[9] Yi C. Feng, Pong C. Yuen, Anil K. Jain "A Hybrid Approach for Generating Secure and Discriminating Face Template" IEEE.

[10] Marco Grassi, Marcos Faúndez-Zanuy: "Protecting DCT Templates for a Face Verification System by Means of Pseudo-random Permutations". IWANN (1) 2009: 1216-1223

[11] Aleix M. Martinez "Recognizing Imprecisely Localized, Partially Occluded, and Expression Variant Faces from a Single Sample per Class IEEE Transaction On Pattern Analysis and Machine Intelligence, Vol.24, N.6, pp. 748-763, June 2002