On the applicability of off-line signatures to the fuzzy vault construction

Manuel R. Freire, Julian Fierrez, Marcos Martinez-Diaz, Javier Ortega-Garcia

ATVS - Biometric Recognition Group
Escuela Politecnica Superior, Universidad Autonoma de Madrid
C/ Francisco Tomas y Valiente 11, E-28049 Madrid, Spain
{m.freire,julian.fierrez,marcos.martinez,javier.ortega}@uam.es

Abstract

In the present contribution, the applicability of off-line handwritten signatures to the fuzzy vault construction is studied. Feature extraction is based on quantized maxima and minima from upper and lower envelopes of the signature. Baseline results are reported for skilled and random forgeries of the MCYT off-line signature database, showing that the proposed scheme is suitable for signers with good separability between genuine signatures and skilled forgeries.

1. Introduction

The goal of biometrics is to infer the identity of people based on anatomical or behavioral data (e.g., fingerprint, face, signature or voice) [7]. In contrast with classical knowledge or token-based identification, biometric recognition deals with information inherent to the user, which therefore cannot be forgotten or stolen. The advances in biometrics research in the last years have led to the growth of a series of new applications. Among them, one that is attracting an increasing research effort is *crypto-biometrics*, i.e., the application of biometrics to cryptography.

In a world of ubiquitous networked devices, cryptography plays a leading role in security and privacy. The classical password-based systems limit the security of cryptosystems to the complexity of the password (i.e., the weakest link). This lack of security, as well as the appearance of new scenarios such as keyboard-less interfaces, suggest alternative approaches to cryptography. *Biometric cryptosystems*, where typically a biometric trait is used instead of a password, have been proposed to overcome limitations of classical cryptosystems in this context [15].

Within biometrics, automatic signature verification has been an intense research area because of the social and legal acceptance and widespread use of the written signature as a personal authentication method [12]. Signature recognition systems are typically divided into two major groups: off-line signature recognition, where the still image of the signature is used, and on-line signature recognition, where dynamic information of the realization of the signature is available.

The present work studies the feasibility of crypto-biometrics using off-line signatures, which finds application in a number of important scenarios like document and identity management based on handwritten signatures. The proposed system uses upper and lower signature envelopes, which have provided encouraging results in off-line signature verification [9, 1].

This paper is structured as follows. In Sect. 2 we summarize related works dealing with crypto-biometrics. Sect. 3 briefly describes of the fuzzy vault construction, which is used in the proposed scheme presented in Sect. 4. Experimental results are reported in Sect. 5. Finally, conclusions are given in in Sect. 6.

2. Related work

A review of the state of the art of biometric cryptosystems is reported in [15]. It establishes a commonly accepted classification of biometric cryptosystems, namely: (*i*) *key release*, where a secret key and a biometric template are stored in the system, the key being released after a valid biometric match, and (*ii*) *key generation*, where a template and a key are combined into a unique token, such that it allows reconstructing the key only if a valid biometric trait is presented. This last scheme has the particularity that it is also a form of cancelable biometrics (i.e., the key can be changed) and is secure against system intruders since the stored token does not reveal information from neither the key nor the biometric.

Several schemes of key generation biometric cryptosystems have been proposed. The *fuzzy vault* scheme [8] establishes a framework for biometric cryptosystems. In this

construction, a secret (typically, a random session key) is encoded using an unordered set of points A, resulting in an indivisible vault V. The original secret can only be reconstructed if another set B is presented and overlaps substantially with A. The fuzzyness of this construction fits well with the intra-variability of biometrics. Uludag et al. [14] proposed a biometric cryptosystem for fingerprints based on the fuzzy vault, where the encoding and decoding sets were vectors of minutiae data. Other works have followed this approach for on-line signature data [5].

In [16], Vielhauer et al. propose a biometric hashing scheme for statistical features of on-line signatures. Their work is based on user-dependent helper data, namely an Interval Matrix. Another biometric hashing scheme has been presented by Fairhurst et al. [3]. This work identifies the problematic of intra-variability and proposes a key generation algorithm based on vector quantization of feature subspaces. Other interesting approach to crypto-biometrics using handwritten signature is biohashing, where pseudo-random tokens and biometrics are combined to achieve higher security and performance [6, 10].

Information-theoretical approaches to crypto-biometrics have also been presented. One of the most interesting proposals is that of Dodis et al. [2], where a theoretical framework is presented for cryptography with fuzzy data (here, biometrics). They propose two primitives: a *secure sketch*, which produces public information about a biometric signal that does not reveal details of the input, and a *fuzzy extractor*, wich extracts nearly uniform randomness from a biometric input in an error-tolerant way helped by some public string. Also, they propose an extension of the fuzzy vault that is both more evaluable theoretically and secure than the original scheme. An implementation of a fuzzy extractor is proposed in [13], where the fuzzy vault for fingerprints is enhanced with helper data extracted from orientation field flow curves.

3. Overview of the fuzzy vault construction

The cryptographic construction *fuzzy vault* defines the operations of encoding and decoding using a fuzzy key [8]. An encoded secret can only be decoded if the key overlaps substantially with the one used to encode. This characteristic makes fuzzy vault a suitable construction for a biometric cryptosystem.

Encoding. Given a field F (e.g., GF(2^q), in which the elements can be coded with q bits), a secret $\kappa \in F^k$ is encoded using a set $A = \{a_i\}_{i=1}^t$, where $a_i \in F$ and a_i's are distinct for $t \geq k$. The secret κ is mapped to the coefficients of a polynomial p of degree $k-1$. A genuine set G is calculated as $G = \{(a_i, p(a_i))\}_{i=1}^t$. A chaff set C is computed as $C = \{(c_i, d_i)\}_{i=1}^m$, where $d_i \neq p(c_i)$. Finally, $V = G \cup C$ conforms the encoded vault.

Decoding. Given the field F used in encoding and the encoded vault V, the original secret is claimed using the set $B = \{b_i\}_{i=1}^t$. The candidate set Q is computed as $Q = \{(b_i, x_i)\}_{i=1}^t$, where $(b_i, x_i) \in V$. Then a polynomial is reconstructed. A secret κ' is reconstructed using an error-correction algorithm, such as Reed-Solomon. If error correction is successful, $\kappa' = \kappa$. Otherwise $\kappa' =$ 'null', and therefore the secret is not recovered. Note that k is the minimum number of points from the genuine set G necessary to reconstruct the polynomial p, and therefore necessary to reconstruct the secret.

Uludag et al. adapted this scheme to a fingerprint-based biometric cryptosystem [14]. They propose to append the CRC-16 of the secret to the input while encoding. In decoding, candidate secret CRC is checked. If the check succeeds, the extracted secret is considered genuine with high probability. Otherwise, the secret is not released.

A generic fuzzy vault scheme for biometrics is presented in Fig. 1.

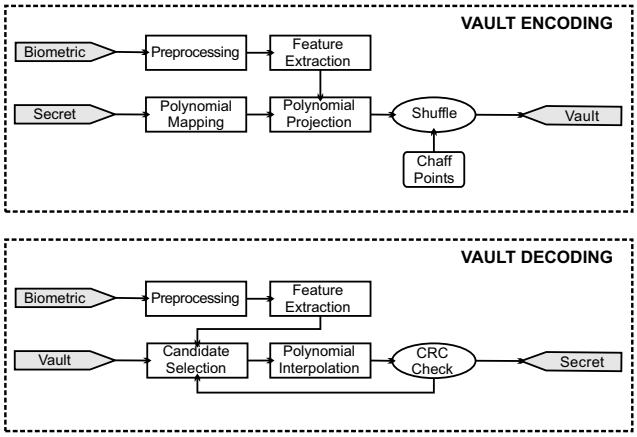

Figure 1. Generic fuzzy vault scheme for biometrics.

4. Fuzzy vault for off-line signatures

Given the fuzzy vault construction, the challenge of an implementation based on any biometric trait is to extract stable information from the biometric. Therefore, the objective of a fuzzy vault system for off-line signatures is to extract a set of fixed-size chunks of information from the signature image that are both stable for multiple realizations of the same user and different from random and skilled forgeries.

4.1. Preprocessing

Input signature images are preprocessed in three consecutive stages as follows [4]:

Binarization: Input images are first binarized by using Otsu's histogram-based global thresholding. A morphological closing operation with a 3×3 squared structuring element is then applied to the binarized image.

Segmentation: Signature is then segmented by using a bounding box based on vertical and horizontal pixel counts. Left and right height-wide blocks having all columns with signature pixel count lower than threshold T_p, respectively top and bottom width-wide blocks having all rows with signature pixel count lower than T_p are discarded (T_p = 15 in the reported experiments).

Normalization: Database used for experiments have been acquired on a restricted size grid, so intra-user rotation variability is expected to be low and no rotation normalization is applied. Segmented signatures are resized in order to have a width of 512 pixels while maintaining the aspect ratio.

4.2. Feature extraction

In order to apply the fuzzy vault scheme to off-line signatures, we use the upper and lower envelopes of the signature [9, 1]. Envelopes are extracted as follows: for each column in the normalized binary image, first non-zero pixel is assigned to the upper envelope, and last non-zero pixel to the lower envelope. As a result, two 1D signals are obtained of the smoothing of the envelopes, using the moving average method with span 35.

Following the approach in [5], maxima and minima of the two signals are extracted and quantized into $N \times M$ steps, where N and M are the number of steps for the quantization of x and y of each maximum or minimum. The number of bits per point is therefore $\log_2(N \times M)$. Finally, the fuzzy vault input key is formed by the set of quantized maxima and minima.

5. Experimental results

5.1. Database description and experimental protocol

A subcorpus of the larger MCYT bimodal database [11] is used for the experiments. MCYT database encompasses fingerprint (optical and capacitive acquisition) and on-line signature data (x, y, pressure, azimuth and altitude trajectories of the pen) of 330 contributors from 4 different Spanish sites. In case of signature, high skilled forgeries are also available (forgers are provided the signature images of the clients to be forged and, after training with them several times, they are asked to imitate the shape with natural dynamics, i.e., without breaks or slowdowns).

Signature information was acquired in MCYT project by using an inking pen and paper templates over a pen tablet (each signature is written within a 1.75×3.75 cm^2 frame), so signature images were available on paper. Paper templates of 75 signers (and their associated skilled forgeries) have been randomly selected and digitized with an scanner at 600 dpi (dots per inch) [4]. Resulting subcorpus comprises $2,250$ signature images, with 15 genuine signatures and 15 forgeries per user (contributed by 3 different user-specific forgers). The 15 genuine signatures were acquired at different times (between 3 and 5) of the same acquisition session. At each time, between 1 and 5 signatures were acquired consecutively.

Two different experimental protocols have been followed for a training scheme based on a single signature. The first protocol, referred to as *intra-set protocol*, has been defined as follows. Genuine matchings are computed within each set, avoiding symmetric matches. Impostor matchings are computed between each genuine signature and both the available skilled forgeries (15 per signer) and the first genuine signature of all the other users avoiding again the symmetric matches (random forgeries).

The second protocol will be called *inter-set protocol*, and has been defined as follows. For genuine matches, each signature is matched against all the signatures of the remaining sets of the signer at hand, avoiding symmetric matches. Skilled and random forgeries are computed as in the intra-set protocol.

5.2. Results

An example of a raw signature of the MCYT database and the result of preprocessing is presented in Fig. 2.

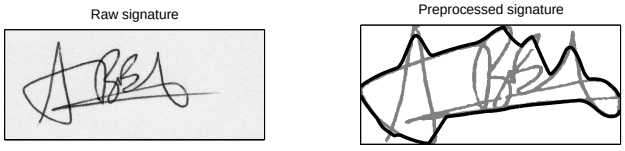

Figure 2. Example of a raw signature (left) and the preprocessed image (right, in grey) with the smoothed envelopes (right, in black).

Input point size for the fuzzy vault was set to 6 bits, 3 bits for the x coordinate and 3 for the y coordinate of each

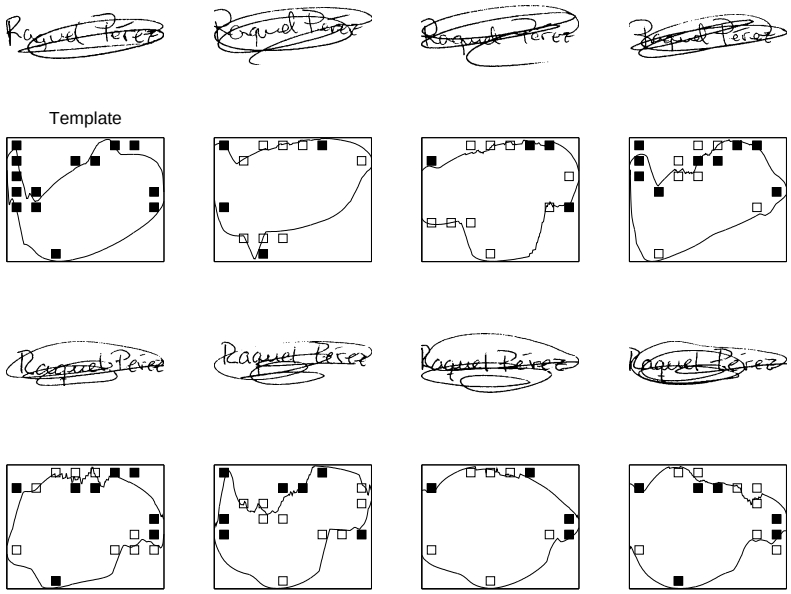

Figure 3. Example user. Each plot corresponds to the signature above. First signature is used as the template biometric, with the squares being the quantized maxima and minima of the envelopes. Signatures 2-4 are three different genuine realizations, where filled squares represent matches with the template. Signatures 5-8 are four skilled forgeries.

maximum or minimum, based on related experiments with on-line signature quantization for fuzzy vault, as those reported in [5].

An example of how the vault input is extracted from the signatures is presented in Fig. 3 for several genuine signatures and skilled forgeries of the same user. We observe that the represented user has a good separability between genuine and impostor realizations.

Average number of genuine matching points was measured for the intra-set and inter-set evaluation protocols (see Fig. 4). We observe that average matches range from 3 to 18, being the majority of the users among 6 and 10. This value has an important effect in the implementation of the fuzzy vault construction, since the number of genuine matching points determine the system threshold k, i.e. the number of matching points necessary to release the secret. From Fig. 4 we conclude that not all the signers are suitable for using this scheme to protect a medium size secret. We also observe that inter-set evaluation is only slightly worse than intra-set evaluation. This fact shows that envelopes are robust to inter-session variability for the given system.

Also, it is interesting to see the average distance between genuine and impostor vault input vectors for the same user, presented in Fig. 5. This gives a measure of how separable are the signatures for the given system. We observe that distance for skilled forgeries ranges from 6 (good separability) to 0 (impossible to distinguish). A better separability is observed for random forgeries, where best signature achieve an excellent separability distance of 12. In this figure we again observe that not all the signatures are well-suited for this scheme.

6. Discussion

A study of the applicability of off-line signature to the fuzzy vault construction has been presented. A baseline scheme based on quantized maxima and minima of upper and lower envelopes of the signature image has been proposed. Preliminary results with the MCYT off-line signature database show the feasibility of the proposed scheme for a number of users with high separability between genuine signatures and skilled forgeries.

A number of findings are extracted from the reported experiments. First, the variability between different users make it difficult to settle a unique operating point k of the fuzzy vault, which is related to the size of the secret that can be locked in the vault. This could be overcome using user-dependent thresholds. Also, we observed that the proposed scheme is not suitable for a portion of users. Finally, in the future it will be interesting to study the effect of helper data, which is not explored in this work and can enhance system performance, as observed with other biometric traits [13].

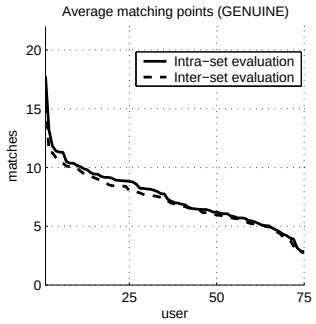

Figure 4. Average number of matching points for genuine signatures for the users in the database, sorted from highest to lowest.

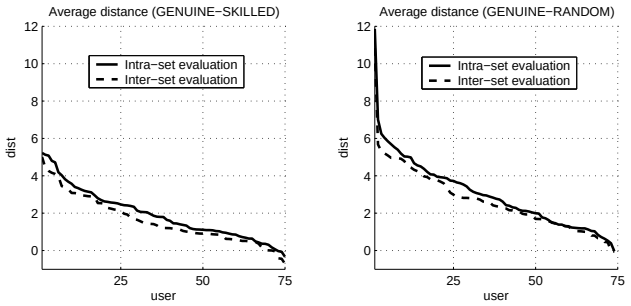

Figure 5. Average distance of vault input vectors between genuine signatures and forgeries for the users in the database, sorted from highest to lowest.

Acknowledgments

This work has been supported by Spanish Ministry of Education and Science under project TEC2006-13141-C03-03 and BioSecure NoE (IST-2002-507634). M. R. F. is supported by a FPI Fellowship from Comunidad de Madrid. J. F. is supported by a Marie Curie Fellowship from European Commission.

References

[1] S. Chen and S. N. Srihari. Use of exterior contours and shape features in off-line signature verification. In *ICDAR*, pages 1280–1284, 2005.

[2] Y. Dodis, L. Reyzin, and A. Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In *Advances in Cryptology - EUROCRYPT 2004*, volume 3027 of *Lecture Notes in Computer Science*, pages 523–540. Springer, 2004.

[3] M. Fairhurst, S. Hoque, G. Howells, and F. Deravi. Biometric hash based on statistical features of online signatures. In *Proc. of the COST-275 Workshop: Biometrics on the Internet*, pages 93–96, 2005.

[4] J. Fierrez-Aguilar, N. Alonso-Hermira, G. Moreno-Marquez, and J. Ortega-Garcia. An off-line signature verification system based on fusion of local and global information. In *Proc. European Conf. on Computer Vision, Workshop on Biometric Authentication, BIOAW*, Springer LNCS-3087, pages 295–306, Prague, Czech Republic, May 2004.

[5] M. Freire-Santos, J. Fierrez-Aguilar, and J. Ortega-Garcia. Cryptographic key generation using handwritten signature. In *Proc. SPIE*, volume 6202, pages 225–231, 2006.

[6] A. Goh. Random multispace quantization as an analytic mechanism for biohashing of biometric and random identity inputs. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(12):1892–1901, 2006.

[7] A. K. Jain, A. Ross, and S. Pankanti. Biometrics: A tool for information security. *IEEE Transactions on Information Forensics and Security*, 1(2):125–143, June 2006.

[8] A. Juels and M. Sudan. A fuzzy vault scheme. *Des. Codes Cryptography*, 38(2):237–257, 2006.

[9] A. Kholmatov. Biometric identity verification using on-line and off-line signature verification. Master's thesis, Sabanci University, 2003.

[10] A. Lumini and L. Nanni. An improved biohashing for human authentication. *Pattern Recognition*, 40(3):1057–1065, 2007.

[11] J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzalez, M. Faundez-Zanuy, V. Espinosa, A. Satue, I. Hernaez, J.-J. Igarza, C. Vivaracho, D. Escudero, and Q.-I. Moro. MCYT baseline corpus: A bimodal biometric database. *IEE Proceedings Vision, Image and Signal Processing*, 150(6):395–401, 2003.

[12] R. Plamondon and S. N. Srihari. On-line and off-line handwriting recognition: A comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):63–84, 2000.

[13] U. Uludag and A. Jain. Securing fingerprint template: Fuzzy vault with helper data. In *CVPRW '06: Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop*, page 163, Washington, DC, USA, 2006. IEEE Computer Society.

[14] U. Uludag, S. Pankanti, and A. K. Jain. Fuzzy vault for fingerprints. In *AVBPA*, pages 310–319, 2005.

[15] U. Uludag, S. Pankanti, S. Prabhakar, and A. K. Jain. Biometric cryptosystems: Issues and challenges. *Proceedings of the IEEE*, 92(6):948–960, June 2004.

[16] C. Vielhauer, R. Steinmetz, and A. Mayerhoefer. Biometric hash based on statistical features of online signatures. In *ICPR '02: Proceedings of the 16 th International Conference on Pattern Recognition*, volume 1, pages 123–126. IEEE Computer Society, 2002.