Cryptographic key generation using handwritten signature

M. Freire-Santos[a], J. Fierrez-Aguilar[a], J. Ortega-Garcia[a]

[a]ATVS-Biometrics Research Lab., Escuela Politecnica Superior,
Universidad Autonoma de Madrid, E-28049 Madrid, Spain

ABSTRACT

Based on recent works showing the feasibility of key generation using biometrics, we study the application of handwritten signature to cryptography. Our signature-based key generation scheme implements the cryptographic construction named fuzzy vault. The use of distinctive signature features suited for the fuzzy vault is discussed and evaluated. Experimental results are reported, including error rates to unlock the secret data by using both random and skilled forgeries from the MCYT database.

Keywords: Biometrics, on-line signature, cryptography, key generation

1. INTRODUCTION

Cryptography is one of the fundamental building blocks of computer security. Unfortunately, cryptographic security is conditioned by an authentication step typically based on long pseudo-random keys (of at least 128 bits in symmetric encryption), which are almost impossible to remember. As a result, cryptosystems commonly rely on user-generated passwords, which are easy to memorize, in order to release the pseudo-random keys. This fact eases the work of an eventual attacker, remarkably decreasing the overall security of the data being protected.[1]

On the other hand, biometric systems use physiological or behavioral traits such as fingerprint, face, iris, speech or handwritten signature to authenticate users.[2] The use of biometrics enhances both the usability (e.g., avoiding the use of multiple passwords) and the security (e.g., non-repudiation) of password-based authentication systems. Biometric signals are also harder to copy or steal, and cannot be forgotten or lost.

Biometric cryptosystems, or crypto-biometric systems, combine cryptographic security with biometric authentication.[1] The integration of biometrics with cryptography can be done broadly at two different levels. In biometrics-based key release, a biometric matching between an input biometric signal and an enrolled template is used to release the secret key. In biometrics-based key generation, the biometric signals are monolithically bounded to the keys.

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.[2] This work is focused on dynamic signature verification, i.e., the time functions of the dynamic signing process are available (e.g., position trajectories, or pressure versus time). Different approaches are considered in the literature in order to extract relevant information from on-line signature data;[3,4] they can coarsely be divided into: *i*) feature-based approaches, in which a holistic vector representation consisting of global features is derived from the acquired signature trajectories,[5] and *ii*) function-based approaches, in which time sequences describing local properties of the signature are used for recognition (e.g., position trajectory, velocity, acceleration, force, or pressure).[6]

The focus of this work is on biometrics-based key generation using handwritten signature. In particular, we study the feasibility of a crypto-biometric system using local features (i.e., maxima and minima in the signature dynamics). We propose an implementation of the cryptographic construction named *fuzzy vault*,[7] following the previous work in fingerprint-based key generation by Uludag *et al.*[8]

Further author information: (Send correspondence to M.F.-S.)
M.F.-S.: E-mail: m.freire@uam.es
J.F.-A.: E-mail: julian.fierrez@uam.es
J.O.-G.: E-mail: javier.ortega@uam.es

This paper is structured as follows. Previous works in signature-based crypto-biometrics, as well as other fuzzy vault applications, are studied in Sect. 2. In Sect. 3 the fuzzy vault construction proposed by Juels and Sudan[7] is sketched. Our implementation of the fuzzy vault for written signatures is described in Sect. 4. Preliminary experiments are reported in Sect. 5. Finally, some conclusions and further work are drawn in Sect. 6.

2. RELATED WORKS

A number of recent works have proposed signature-based crypto-biometric systems. Vielhauer *et al.* studied a biometric hash extractor based on global parameters of the on-line signatures.[9] They achieved an average of 7.05% of FRR and 0% FAR (skilled forgeries) over a reduced set of 11 users.

A similar approach was proposed by Yip *et al*,[10] presenting a method for replacing compromised keys using biometrics hashes obtained from on-line signatures. In their worst scenario (stolen token) the EER is $< 6.7\%$ for 40 users. Also worth noting, Fairhurst *et al.* studied the feasibility of biometric key generation based on partitions of feature subspaces.[12]

The fuzzy vault scheme used in this contribution has already been applied to fingerprint-based key generation.[8,11] These pioneer works identified the important problem of automatic template alignment for biometrics-based key generation, and then manual alignment of the biometric templates was used. In this contribution, where a more variable trait is evaluated (i.e., written signature), automatic alignment of the biometric templates is used in all the experiments reported.

3. FUZZY VAULT OVERVIEW

The *fuzzy vault* cryptographic construction provides encryption and decryption of secret data using a fuzzy key, which is an unordered set of points. In this system, Alice places a secret S in a vault using an unordered set A. Bob will only be able to reveal the secret if his set B is in a high degree similar to A.

This scheme is a good candidate to implement a crypto-biometric system to replace the use of passwords, where the vault key (i.e., the biometric sample) is not stable, and the secret S is the random key that will be used afterwards in the traditional cryptosystem (e.g., AES).[13]

We follow the approach by Uludag *et al.* to adapt the fuzzy vault for biometrics.[8] To solve the problem of the error-correction, they concatenate the CRC-16 (IBM version)[14] of the secret S to the input while encoding. After reconstructing the secret (decoding), they determine whether the secret is the original one or not by checking the CRC. If the check succeeds, the extracted secret is considered genuine with an error probability of 2^{-16}. Otherwise, the secret is not released. Note that all operations take place in Galois field $GF(2^{16})$.[15]

3.1. Encoding

Encoding (Fig. 1) operates two input values: a random K-bit value, which is the data that is going to be protected by the vault (secret S), and a template feature vector T, extracted from the biometric trait, which plays the role of the vault key. In order to hide a secret that would become the key in a traditional cryptosystem such as AES,[13] a size of 128 bits in the secret S is appropriate for current security needs, therefore $K = 128$. The template feature vector T is formed by N 16-bit units t_i, $i = 1, \ldots, N$.

So that to allow the secret reconstruction while decoding, the CRC-16 of the secret is concatenated to S, thus becoming a 144-bit value. A polynomial is then constructed with degree $D = K/16$ (e.g., 8 for 128-bit secret), being its coefficients the 16-bit packets in S: $p(x) = S_1 x^n + ... + S_D x^1 + S_{D+1}$. We then calculate the polynomial projections for all the points in the template feature vector, forming the set of genuine points $\mathrm{G} = \{(t_1, p(t_1)), ..., (t_N, p(t_N))\}$. We also generate a random set C of M chaff points, $\mathrm{C} = \{(c_1, d_1), ..., (c_M, d_M)\}$, where $d_i \neq p(c_i)$. The number of chaff points used in the reported experiments is $M = 200$.

The genuine G and the false C sets are scrambled, thus obtaining the final set $\mathrm{V} = \{(v_1, w_1), ..., (v_{N+M}, w_{N+M})\}$, which is the fuzzy vault. The impossibility to distinguish the genuine and chaff points makes infeasible the reconstruction of the secret polynomial, while knowing enough genuine points ($D + 1$ or more) enables the reconstruction of the polynomial and therefore the release of the secret data.

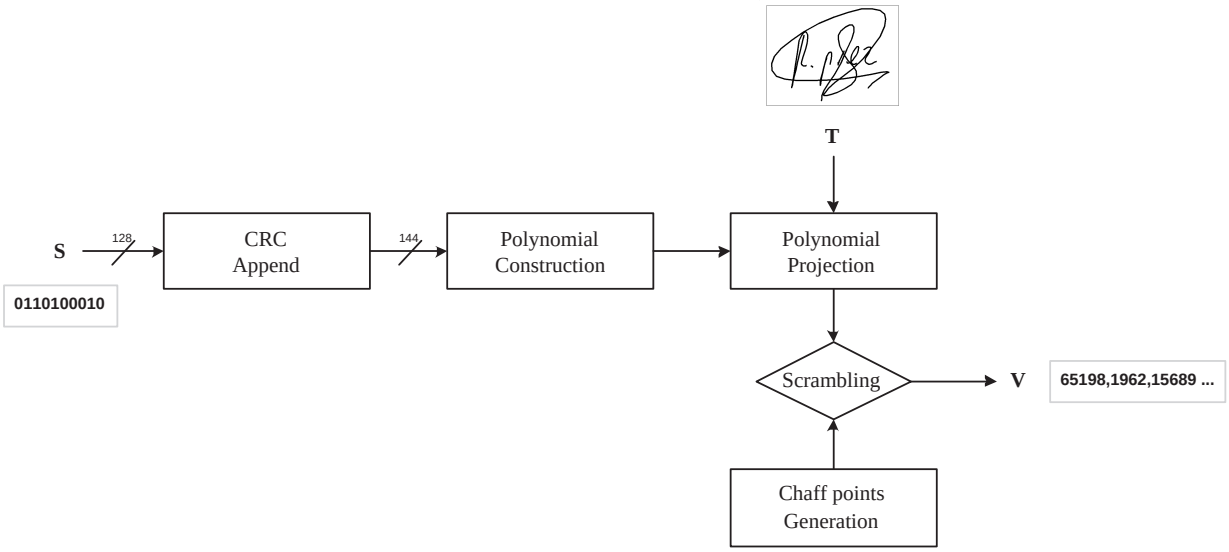


Figure 1. Fuzzy Vault Encoding for a 128-bit secret S.

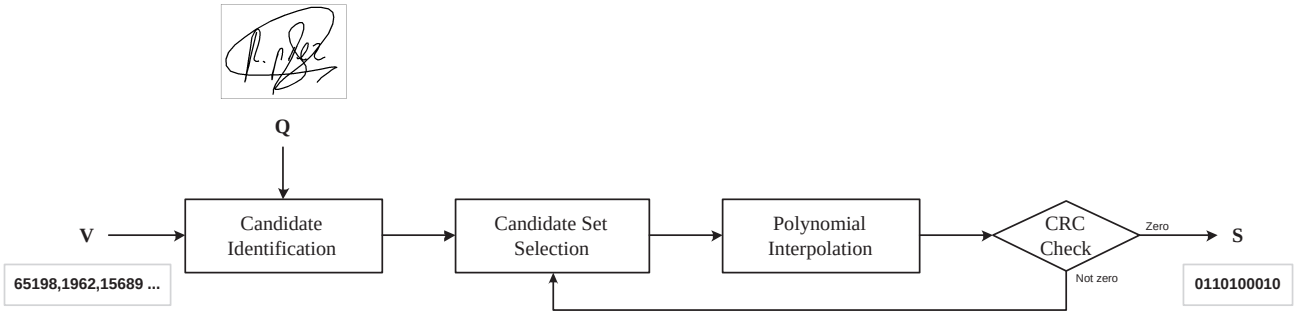


Figure 2. Fuzzy Vault Decoding.

3.2. Decoding

Decoding (Fig. 2) starts from an encoded fuzzy vault, V, and a query feature vector Q. The number of query points is the same as the template points used while encoding.

For each query point q_i in Q that is present in the abscissa of the vault V in the position j (i.e., $q_i = v_j$), the pair (v_j, w_j) is selected as a candidate to reconstruct the secret polynomial. There must be at least $D + 1$ candidates in order to reveal the secret S.

We then interpolate the secret polynomial with all combinations of $D + 1$ candidates using the Lagrange method.[8] This way, we obtain a number of new polynomials p_k. If the remainder of the division of p_k with the CRC-16 primitive polynomial is zero, the secret is valid with an error probability of 2^{-16}. Otherwise, the secret is not valid.

4. FUZZY VAULT FOR ON-LINE SIGNATURE

We study the feasibility of a crypto-biometric system based on written signature. In particular, we implement the fuzzy vault for signatures using on-line local parameters (maxima and minima).

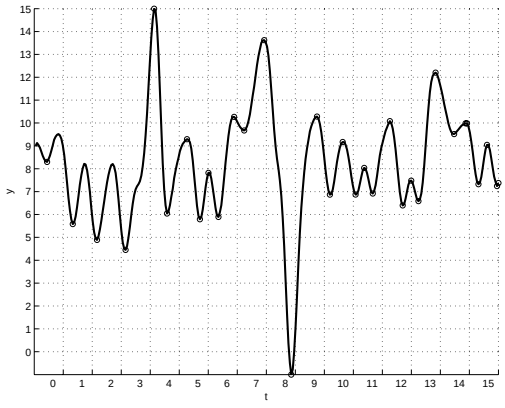


Figure 3. Example of 4×4-bit quantization for 32 selected maxima and minima in $y(t)$ for one signature from MCYT.

The feature vector in the fuzzy vault is formed by 16-bit units. We study a combination of the values in time, horizontal position, vertical position and pressure signals of the signature. We calculate these maxima and minima in the y-axis of the signature with respect to time, $y(t)$.

The signatures are preprocessed automatically. In particular, no manual pre-alignment is used. In this respect, this work is one of the first fully automated key-generation schemes using biometrics. This preprocess consists of: *i*) the elimination of first and last 10% of the signature, as it was observed that these regions are highly unstable, *ii*) geometric normalization, based on the center of mass and the standard deviations of the function values, and *iii*) function smoothing in order to reduce noise effects.

We then extract N maxima and minima in $y(t)$, where N is determined a priori as the number of points used to encode and decode the vault. In order to select only N points, we progressively smooth $y(t)$ until N maxima and minima are found: $\mathrm{M} = \{m_1, m_2..., m_N\}$, where m_i is the position in time of the maximum or minimum.

For each point in M, we construct the vault input point $\{m_i|x(m_i)|y(m_i)|z(m_i)\}$, where $x(m_i)$, $y(m_i)$, and $z(m_i)$ stand for displacement in x- and y-axis and pressure at the maximum or minimum, respectively. Each vault input point is quantized into a 16-bit unit, therefore reducing signature variability. The effects of this quantization are studied in the reported experiments. An example of 4×4-bit quantization is shown in Fig. 3.

In order to overcome the high variability present in the written signature, in the reported experiments we study the simultaneous use of multiple signatures for encoding. When more than one written signature is used for encoding, we extract the feature vectors from all of them, and then we select the N most repeated points. In this case, no smoothing for obtaining an exact number of point is applied. If more than N maxima/minima are repeated in all the training signatures to be used for encoding, then N of them are chosen randomly.

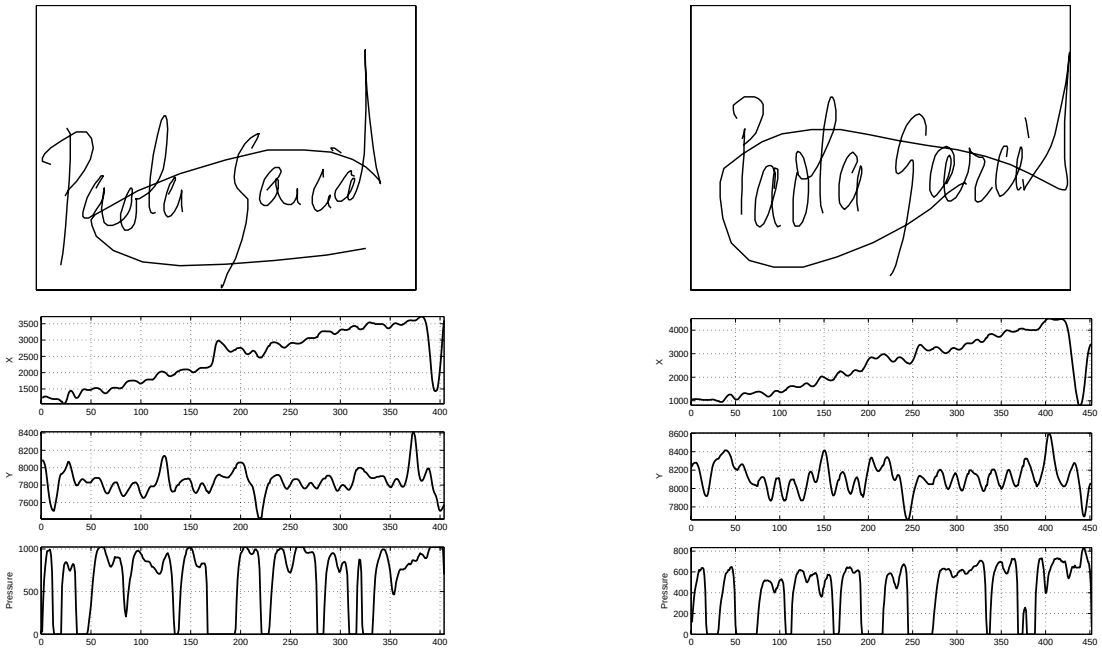
5. EXPERIMENTS

5.1. Database and experimental protocol

The MCYT database signature corpus is used for the experiments.[16] This database contains 330 users with 25 genuine signatures and 25 skilled forgeries per user. Forgers were asked to imitate after observing the static image of the signature to imitate, trying to copy them at least 10 times, and then signing naturally (without breaks or slowdowns). An example genuine signature and a forgery are presented in Fig. 4.

A subset of users (MCYTX) has been defined, consisting of the users with medium to high signature complexity, i.e., those able to encode/decode the fuzzy vault. The threshold for rejecting low complexity signers was established a priori in $N \geq 32$ maxima and minima in y. This way, 126 users out of 330 were selected (38% of the total).

Genuine and impostor matches and error rates are calculated as follows. For each user, a secret is encoded using 5 genuine signatures. Then, decoding is attempted using each one of the remaining 20 different genuine



(Genuine signature) (Forgery)

Figure 4. Example of a genuine signature and a forgery from the MCYT database.

signatures and 25 skilled forgeries (20×126 genuine and 25×126 skilled impostor decoding attempts). Random forgeries are also used for measuring system performance. In this case impostors attempt to decode the secret by presenting their own signature (1 signature per user, avoiding symmetric matches, 126×125 random impostor attempts in total).

Note that, in the proposed scheme, $D + 1 = 9$ or more matches allow the polynomial reconstruction, and therefore the successful decoding of the vault.

5.2. Results

Four different experiments are reported:

Function analysis. Statistics are extracted for the number of maxima and minima in the different signals (x, y and z). The full MCYT database is used. Maxima and minima histograms for x, y and z from all genuine signatures in the database are given in Fig. 5 (330×25 signatures). We observe that x and y signals have a similar number of maxima and minima (around 35 in average), while z presents a higher value (49.2).

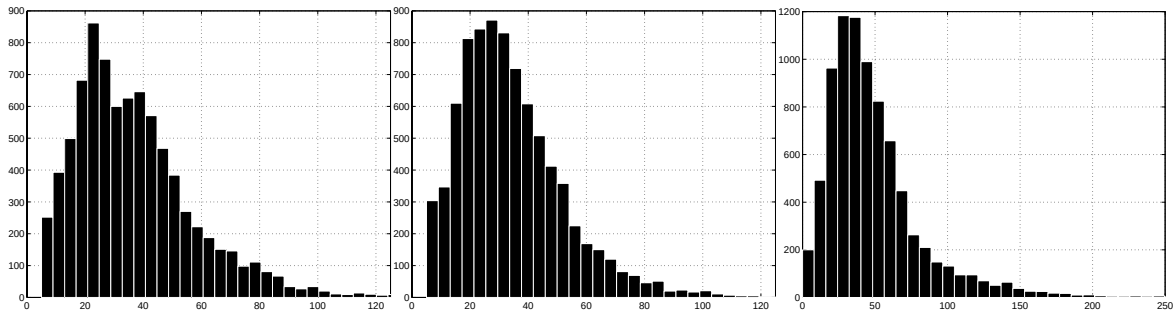


Figure 5. Histograms of the number of maxima and minima for the different signals in the MCYT on-line signature database, from left to right: displacement in x-axis, displacement in y-axis, and pressure z.

Function comparison. Individual behavior of the different signature signals is observed on the MCYTX subcorpus, with a fixed quantization size of 10 bits. Results are presented in Table 1. When considering skilled impostor matches, note that x and t obtain the lowest number of impostor matches (7.60 and 8.64 in average) while y and z result in an average number of matches (9.03 and 11.80, respectively) above the acceptance threshold (9).

Table 1. Average number of matches for genuine signatures and impostors using one signal with a quantization of 10 bits.

Signal	Genuine signatures	Skilled impostors
t	10.08	8.64
x	9.99	7.60
y	11.45	9.03
z	13.03	11.80

Quantization. The effect of the quantization size to the number of matches between template and query vectors on the MCYTX subcorpus is studied. Average number of matches is given in Table 2 for signal x with 8, 10 and 12 bits quantization. We observe that a higher number of quantization steps produces less forger matches (10.41, 7.60 and 5.71 for 8, 10 and 12 bits, respectively), while distinctiveness decreases (difference between genuine and impostor matches progressively decreases from 2.98 for 8 bits to 1.84 for 12 bits).

Table 2. Average matches for genuine and impostors using signal x and quantization sizes of 8, 10 and 12 bits.

Quantization	Genuine	Skilled impostors	Difference (G-S)
8-bit	13.39	10.41	2.98
10-bit	9.99	7.60	2.39
12-bit	7.55	5.71	1.84

Evaluation of the fuzzy vault. Finally, we study different 10-bit configurations for the construction of the fuzzy vault for on-line signature. Results with system error rates and average number of matches are given in Table 3. In the best configuration (3 bits quantization in t and x, 2 bits in y and z), a FRR of **57.30%** is achieved with a FAR of **1.18%** and **0.32%** for skilled and random forgeries, respectively. Note that in order to keep the false acceptance low, the system rejects more than a half of genuine attempts. This can be overcome by using decoding strategies based on multiple input signatures.

Table 3. System performance for different configurations of 10-bit vault input points (MCYTX set - 126 users).

Quantization				Genuine signatures		Skilled forgeries		Random forgeries	
t	x	y	z	Avg. matches	FRR	Avg. matches	FAR (*Skilled*)	Avg. matches	FAR (*Random*)
3	3	2	2	8.06	**57.302%**	2.32	**1.175%**	1.53	**0.318%**
3	4	2	1	9.53	39.841%	3.71	7.778%	2.20	1.473%
4	4	2	0	7.58	61.786%	2.51	2.318%	1.79	0.813%
3	4	3	0	6.43	74.167%	2.12	0.921%	1.43	0.152%

6. CONCLUSIONS

An automatic signature-based key generation system has been proposed, using local features applied to the cryptographic construction named *fuzzy vault*. An analysis of the suitability of different signature parameters has also been conducted.

Preliminary experimets reveal the feasibility of crypto-biometric systems based on signature dynamics. Results reveal high false rejection but considerably low false acceptance to decode the secret data (57.3% FRR, 1.2% and 0.3% FAR for skilled and random forgeries, respectively). Therefore work is currently being done in enhancing automatic pre-alignment and more invariant feature extraction.

In the future we plan to study the fusion of global and local features in signature-based crypto-biometric systems,[5] and the application of signature-based cryptosystems to new scenarios like Tablet PC.[17]

ACKNOWLEDGMENTS

This work has been supported by Spanish MCYT TIC2003-08382-C05-01 and by European Commission IST-2002-507634 Biosecure NoE projects. J. F.-A. is supported by a FPI scholarship from Comunidad de Madrid.

REFERENCES

1. U. Uludag, S. Pankanti, P. S., and A. Jain, "Biometric cryptosystems: Issues and challenges," *Proceedings of the IEEE* **92**, pp. 948–960, June 2004.
2. A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. Circuits Syst. Video Techn.* **14**(1), pp. 4–20, 2004.
3. R. Plamondon and G. Lorette, "Automatic signature verification and writer identification - The state of the art," *Pattern Recognition* **22**(2), pp. 107–131, 1989.
4. D.-Y. Yeung, H. Chang, Y. Xiong, S. E. George, R. S. Kashi, T. Matsumoto, and G. Rigoll, "SVC2004: First International Signature Verification Competition," in *Proc. ICBA*, *Lecture Notes in Computer Science* **3072**, pp. 16–22, Springer, 2004.
5. J. Fierrez-Aguilar, L. Nanni, J. Lopez-Peñalba, J. Ortega-Garcia, and D. Maltoni, "An on-line signature verification system based on fusion of local and global information," in *Proc. AVBPA*, *Lecture Notes in Computer Science* **3617**, pp. 523–532, Springer, 2005.
6. J. Fierrez-Aguilar, S. Krawczyk, J. Ortega-Garcia, and A. K. Jain, "Fusion of local and regional approaches for on-line signature verification," in *Proc. IWBRS*, *Lecture Notes in Computer Science* **3617**, pp. 188–196, Springer, 2005.
7. A. Juels and M. Sudan, "A fuzzy vault scheme," in *Proc. IEEE Int. Symp. Information Theory*, p. 408, 2002.
8. U. Uludag, S. Pankanti, and A. K. Jain, "Fuzzy vault for fingerprints," in *Proc. AVBPA*, *Lecture Notes in Computer Science* **3546**, pp. 310–319, Springer, 2005.
9. C. Vielhauer, R. Steinmetz, and A. Mayerhöfer, "Evaluating biometric encryption key generation," in *Proc. ICPR*, pp. 123–126, 2002.
10. K. W. Yip, A. Goh, D. N. C. Ling, and A. T. B. Jin, "Generation of replaceable cryptographic keys from dynamic handwritten signatures," in *Proc. ICB*, *Lecture Notes in Computer Science* **3832**, pp. 509–515, Springer, 2006.
11. T. C. Clancy, N. Kiyavash, and D. J. Lin, "Secure smartcard-based fingerprint authentication," in *WBMA '03: Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications*, pp. 45–52, ACM Press, (New York, NY, USA), 2003.
12. 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*, pp. 93–96, 2005.
13. NIST, "Advanced encryption standard (AES)," 2001.
14. T. V. Ramabadran and S. S. Gaitonde, "A tutorial on CRC computations," *IEEE Micro* **8**(4), pp. 62–75, 1988.
15. S. Lin and D. J. Costello, *Error Control Coding, Second Edition*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2004.
16. 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), pp. 395–401, 2003.
17. F. Alonso-Fernandez, J. Fierrez-Aguilar, and J. Ortega-Garcia, "Sensor interoperability and fusion in signature verification: A case study using Tablet PC," in *IWBRS*, *Lecture Notes in Computer Science* **3781**, pp. 180–187, Springer, 2005.