Systematic Construction of Iris-Based Fuzzy Commitment Schemes[*]

Christian Rathgeb and Andreas Uhl

University of Salzburg, Department of Computer Sciences, A-5020 Salzburg, Austria
{crathgeb,uhl}@cosy.sbg.ac.at

Abstract. As a result of the growing interest in biometrics a new field of research has emerged entitled *Biometric Cryptosystems*. Only a small amount of work, which additionally tends to be custom-built according to the specific application context, has been published in this area. This work provides a systematic treatment of how to construct biometric cryptosystems based on iris biometrics. A cryptographic primitive called *Fuzzy Commitment Scheme* is adopted to different types of iris recognition algorithms to hide and retrieve a cryptographic key in and out of a biometric template. Experimental results confirm the soundness of the approach.

1 Introduction

Taking into account today's ever-increasing demand on high security standards, in order to secure any kind of crucial information, the science of cryptography has become even more important. While in generic cryptographic systems authentication is possession based [1], key management is performed introducing alternative authentication mechanisms such as password or PIN.

By introducing biometrics to replace password-based authentication the security of cryptographic systems is improved. Several approaches have been made to combine biometric authentication with key management systems to build up so-called "biometric cryptosystems", which are classified by the way biometric authentication is merged with the respective cryptosystem. The trivial way of introducing biometric authentication into a generic key management system, replacing password/PIN-based authentication through biometric authentication, is called "key release scheme". Key release schemes are easy to implement, still these are not frequently used. Within such schemes biometric templates as well as cryptographic keys, which are not secure, are stored in a database separately. This is a very critical issue because biometric templates and cryptographic keys can be stolen or compromised. Thus, a biometric cryptosystem based on a key release scheme is not appropriate for high security applications. The second class of biometric cryptosystems includes "key generation schemes" and "key binding schemes". Key generation schemes directly derive cryptographic keys from biometric data. However, within key generation schemes a cryptographic key cannot

[*] This work has been supported by the Austrian Science Fund, project no. L554-N15.

M. Tistarelli and M.S. Nixon (Eds.): ICB 2009, LNCS 5558, pp. 940–949, 2009.

be changed if it is compromised once. The most promising types of biometric cryptosystems are key binding schemes. By seamlessly binding a cryptographic key with biometric information via a key binding algorithm, secure templates are provided which do not reveal any information about the biometric data, nor about the cryptographic key. With an appropriate key retrieval algorithm keys are released again.

Juels and Wattenberg [2] proposed a theoretical basis for biometric key binding schemes that they refer to as “fuzzy commitment scheme” (FCS). Since the iris is one of the most accurate biometric characteristic [3,4] it is desirable to apply the fuzzy commitment approach to iris biometrics. However, until now only little literature has been published concerning iris-based biometric cryptosystems. This work will provide a systematic approach of how to build up iris-based FCSs. Furthermore, two different iris recognition algorithms [5,6] will be used to demonstrate the construction of different types of iris-based FCSs.

This paper is organized as follows: first a short summary of previous work concerning iris-based biometric cryptosystems will be given, where the fundamentals of a FCS are examined in detail (Sect. 2). Subsequently a generic approach of how to construct iris-based FCS is presented (Sect. 3) which is then demonstrated by applying it to two different types of iris recognition algorithms (Sect. 4). Finally experimental results are presented and discussed (Sect. 5, 6).

2 Iris-Biometric Cryptosystems

In the past several years some key-papers have been published concerning biometric cryptosystems [7,8,9,10,11,12,13,14,15]. Several biometric characteristics, including fingerprints, voice, etc., have been examined for the extraction of cryptographic keys. Still, only a few of these approaches focus on iris biometrics [8,9,10].

Davida *et al.* [8,9] were the first to create a key generation scheme that they refer to as “private template scheme”. Within their approach a hashed value of preprocessed iris codes and user specific attributes serves as a cryptographic key. The result of these preprocessed iris codes is concatenated with check digits which are part of a linear ECC. This ECC is capable of correcting a fixed number of errors defined at system setup. At the time of authentication the error correction information, which is stored as part of the template, is used to correct faulty bits in the acquired biometric data. Finally the same hash function as in the registration step is applied to generate a hash which can be used as cryptographic key. Unfortunately, performance measurements and test results are renounced.

According to the idea of the private template scheme Wu *et al.* [16] proposed a system in which 256-dimensional feature vectors are extracted out of preprocessed iris images using a set of 2-D Gabor filters. A hash function is applied to this vector to generate a cryptographic key. Additionally, an ECC is generated. During authentication another feature vector is extracted from a biometric input.

This feature vector is error correction decoded and the same hash function like in the encryption phase is used to generate a cryptographic key. The extracted cryptographic key is suggested to be used in a symmetrical cryptosystem. For a total number of over 100 persons a FRR of approximately 5.55% and a zero FAR are reported.

In order to construct biometric cryptosystems based on the key binding approach Juels and Wattenberg [2] combined well-known techniques from the areas of ECCs and cryptography to achieve a type of cryptographic primitive called FCS. Fuzzy commitment is the analogon to "fuzzy logic" in artificial intelligence. In their definition a FCS consists of a function F, which is used to commit a codeword $c \in C$ and a witness $x \in \{0, 1\}^n$. The set C is a set of error correcting codewords c of length n and x represents a bitstream of length n termed witness (in a biometric cryptosystem x represents the biometric data). To enhance security only the difference vector of the codeword and the biometric measurement, $\delta \in \{0, 1\}^n$ where $x = c+\delta$, and a hash value $h(c)$ are stored as the commitment. The commitment, which is nothing else then these two values is termed $F(c, x)$. To deal with the fuzziness of x it is proposed that every x', which is sufficiently "close" to x should be able to reconstruct c. If the system is presented with a witness x' that is near x, the difference vector δ is used to translate x' in direction of x. If the correct codeword c is reconstructed with the use of error correction the hash of c', $h(c')$ will match the stored hash value resulting in a successful authentication. The enrollment and authentication process within a FCS operates as follows: during enrollment a user U presents a witness x to the authentication system S. The system selects a codeword $c \in C$ (in a biometric cryptosystem c represents a cryptogaphic key prepared with error correction information), calculates the fuzzy commitment $F(c, x)$ (the difference vector δ and the hash value of the codeword c, $h(c)$) and stores it in a database. At the time of authentication a user purporting to be U presents a witness x' to S. The system looks up the commitment of user U and checks whether x' yields a successful decommitment, which would lead to a successful authentication. In Fig. 1 the basic working flow of a FCS, with respect to the use in a biometric cryptosystem, is illustrated.

Hoa *et al.*[10] applied the FCS to iris biometrics. In their approach a 140-bit cryptographic key is encoded with a concatenation of ECCs and subsequently XORed with a 2048-bit iris code to generate a secure template. During authentication another iris code is extracted from the person and XORed with the template. Finally error correction decoding is performed on the resulting bitstream to regenerate the cryptographic key. By applying a concatenation of ECC a remarkable FRR of 0.47% and a zero FAR were reported for a total number of 70 different persons. Still their approach does not clarify the construction of a generic FCS with respect to required precondition as well as the choice of ECCs.

Up to the present there are no other achievements concerning iris-based biometric cryptosystem which are worth mentioning. Thus the motivation of this work is to provide a systematic approach of how to construct iris-based key

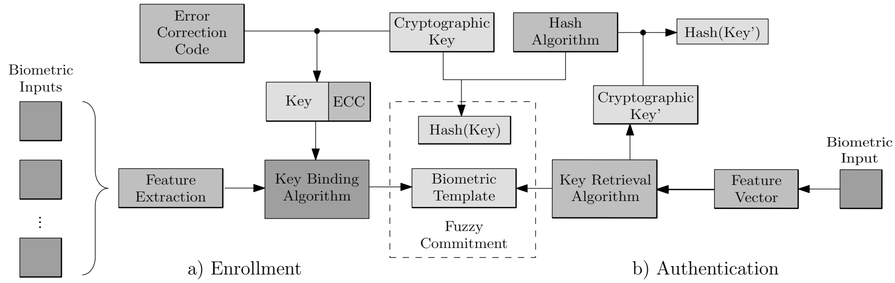

Fig. 1. a) The basic enrollment procedure in a fuzzy commitment scheme. b) The basic authentication procedure in a fuzzy commitment scheme.

binding schemes based on FCSs. In the following section the theoretical basis of constructing iris-based FCSs will be defined.

3 Construction of Iris-Based FCSs

Building up iris-based FCSs first of all preconditions have to be declared: to create a FCS for the use in a biometric cryptosystem the applied iris recognition algorithm should produce an order invariant bitstream (FCSs cannot handle order variant bitstreams). Furthermore, this bitstream should be as long as the cryptographic key concatenated with error correction bits, where the error correction bits provide the information to correct the estimated number of errors between iris data of the same person.

Since ECCs should be able to correct the number of errors occurring between legitimate persons, the next step is to analyze the maximal intra class distance (ICD) of the applied iris recognition algorithm, according to the particular bit block sizes (by analysing ICD of typical and large data sets). Errors between biometric measurements are not distributed uniformly random which implies the usage of block-level ECCs as first layer of error correction is inevitable to handle burst errors. Block level ECCs are codes capable of correcting blocks of bits in which errors occur, while the error correction information is provided by other bit blocks. Experience has shown Reed-Solomon codes are a suitable choice for block-level error correction. Named after I. Reed and G. Solomon, a Reed-Solomon code is defined as $RS(l, n)$ which means that l blocks of of length m are encoded by adding $l - n$ blocks of length m resulting in a total number of $n \cdot m$ bits, where $n \leq 2^{m-1}$ is required. Redundant information is added by oversampling a polynomial constructed from the l information blocks. If less than $(n - l)/2$ blocks are faulty after transmission the polynomial can be reconstructed which means the l information bit blocks can be reconstructed. Further details about Reed-Solomon codes can be found in [18,19].

If the maximal ICD of the applied iris recognition algorithm lies beyond the number of bit blocks, which the applied block-level ECC is capable to correct, another layer of error correction has to be inserted, since simply adding more redundant bit blocks is not possible (the resulting bitstream must have the same size as the extracted iris code). Bit-level ECCs are capable of correcting single bit errors, while the error correction information lies within each codeword. In practical use one type of bit-level ECC has proved worth, namely Hadamard codes. Hadamard codes, which are generated using Hadamard matrices, are ECCs of the type $[2^n, n+1, 2^{n+1}]$, which means bitstreams of length $n+1$ are mapped to codewords of length 2^n while the whole code consists of a total number of 2^{n+1} codewords. A Hadamard matrix H_n of dimension $n \times n$ generates a Hadamard code consisting of $2n$ codewords, each of length n, capable of correcting up to $n/4 - 1$ errors. Further details about Hadamard codes and Hadamard matrices can be found in [17].

In Fig. 2 the encoding/decoding flow of concatenated error correction is illustrated. It is essential that the block-level ECC and the bit-level ECC operate on the same bit blocks, otherwise one faulty block in the bit level could cause several faulty blocks in the block level during decoding. In the encoding step first a block-level ECC is applied and subsequently a bit-level ECC. Thus, in the decoding step, the bit-level ECC corrects single bit errors and the block-level ECC corrects remaining burst errors. If the the maximal ICD is still too large after bit-level error correction, decoding the block-level ECC will not be able to regenerate a hidden key. If this is the case the applied iris recognition algorithm is not adequate to be used in a FCS. Otherwise the parameters of the applied bit-level ECC and the applied block-level ECC have to be adjusted.

In the following section two different iris recognition algorithms will be analyzed and ECCs will be adapted to build up FCSs which are capable of hiding and retrieving cryptographic keys, sufficiently long to be used in generic cryptosystems.

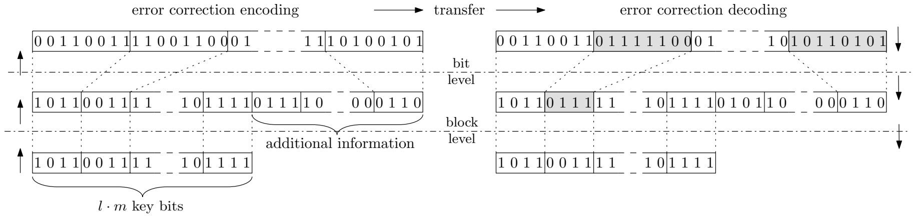

Fig. 2. The concatenation of block-level ECCs and bit-level ECCs (faulty bit blocks are marked gray)

4 Proposed Schemes

For the proposed FCSs two different implementations of iris recognition algorithms are applied: The first implementation is based on a algorithm published by Ma *et al.* which is invariant to translation, scale and rotation. In this approach the iris texture is treated as a kind of transient signal which is processed using wavelet transform. The local sharp variation points, which denote important properties of transient signals, are recorded as features to extract binary iris codes of 1280 bytes. Further details about this algorithm can be found in [5].

The second implementation is based on a algorithm published by Ko *et al.* which uses cumulative sum based change analysis to analyze preprocessed iris textures. Enhanced iris textures are divided into cells out of which mean gray scale values are calculated and furthermore, an iris code $I \in \{0, 1, 2\}^{1000}$ is extracted, using the suggested parameters for the calculation of cumulative sums. Further details about this algorithm can be found in [6].

It will be shown the both types of extracted iris codes are sufficiently long to be used in a FCS. For bit-level error correction Hadamard codes are applied and for block-level error correction Reed-Solomon codes are applied.

To choose an adequate bit block size for block-level error correction the maximal ICDs of both algorithms are estimated, according to the respective block sizes, which are summarized in Table 1. Within the iris code of the algorithm of Ko *et al.* a sequence of 1s indicates an upward slope of cumulative sums of gray scale values and a sequence of 2s indicates a downward slope. This code is simply mapped to a binary code of twice the length. Subsequently the resulting bitstream is rearranged so that the first half of the iris code contains all upward slopes and the second half of the iris code contains all downward slopes. Thus for each part of the resulting code a sequence of 1s suffices to indicate the respective slope. Thereby the number of block errors between different iris codes is minimized. In contrast, Fig. 5 shows the distribution of genuine persons and imposters without the arrangement of the bitstream, according to a blocksize of 8 bit. Analyzing the maximal ICDs of the applied algorithms it is assumable that for the algorithm of Ko *et al.* a single layer of block-level error correction suffices, while for the algorithm of Ma *et al.* a second layer of bit-level error correction is added. This is because within the algorithm of Ko *et al.* mostly burst errors ocurr (the maximal ICD increases only slightly).

Table 1. The maximal intra class distances of the iris codes, using the algorithm of Ma *et al.* and Ko *et al.*, according to the size of faulty bit blocks

Bit Block Size	**Max. ICD Ma *et al.* (%)**	**Max. ICD Ko *et al.* (%)**
1	48.3	14.2
2	53.3	21.2
4	67.2	31.0
6	79.5	38.5
8	85.1	43.6

In view of the maximal ICDs of both algorithms, according to the block sizes, ECCs are adapted. As mentioned above, for the algorithm of Ma *et al.* a concatenation of ECCs is applied. It will be shown that a blocksize of $m = 8$ is a suitable choice. To bind and retrieve a sufficiently long key K, the size of the key is set to $|K| = l \cdot m$, where $l = 16$, so that the key consists of 128 bits (note that the key size has to be a multiple of m). Since the entire encoding procedure should produce a bitstream of length $1280 \cdot 8$ bits output size of the Reed-Solomon code, denoted by RS_{res} which is calculated by $|RS_{res}| = 1280 \cdot 8/2^{m-1}$ because the

Hadamard code maps bit blocks of size m to bit blocks of size 2^{m-1}. Thus, for a cryptographic key K, of length $l \cdot m$, Reed-Solomon block-level error correction is defined by $RS(l, 1280 \cdot 8/2^{m-1})$ which means l m-bit information blocks are encoded using a total number of $1280 \cdot 8/2^{m-1}$ bits, which are mapped to $1280 \cdot 8$ bits applying Hadamard encoding. In summary, for the algorithm of Ma *et al.* first a $RS(16, 80)$ block-level ECC is applied and afterwards a Hadamard code which maps the resulting 80 8-bit blocks to 80 128-bit blocks (=1280·8 bits).

For the algorithm of Ko *et al.* the encoding step is trivial. According to Table 1, the maximal ICD is still less than 44% for a block size of $m = 8$. In other words a single layer of block-level error correction does suffice. By mapping the extracted iris code to a binary code and rearranging it, a bitstream of 2000 bits is extracted. By setting the key size to 128 bit, to generate 250 8-bit blocks (=2000 bits), the block-level ECC is defined by $RS(l, 2000/m)$. A single use of bit-level error correction makes no sense for both schemes. While in the algorithm of Ma *et al.* the maximal ICD lies far beyond 25% in the algorithm of Ko *et al.* occurring errors are not distributed uniformly random. This means, applying Hadamard codes to the algorithm of Ko *et al.* the maximal number of occurring bit errors within 128-bit blocks has to be corrected, since a block size of $m = 8$ would be suitable for mapping a 128-bit key to a sufficiently long bitstream. However, the maximal number of bit errors within 128-bit blocks lies beyond $128/4 - 1 = 31$ bits as illustrated in Fig. 4. For the algorithm of Ma *et al.* two layers of error correction are adequate because, according to Table 1, the maximal ICD lies far beyond 50% for suitable block sizes. In contrast, for the algorithm of Ko *et al.* the maximal ICD is clearly below 50% implying the introduction a layer of bit-level error correction is unnecessary.

In the enrollment procedure of the proposed schemes three iris images are preprocessed as described in [4]. The resulting 512×64 pixel iris textures are slitted from the right side [45^o to 315^o] and from the left side [135^o to 225^o] to get rid of most of the eyelids and eyelashes, according to the idea in [6]. The three extracted iris codes are majority decoded and gaps of 1s and 0s of the resulting bitstream are filled. The entire enrollment procedure is illustrated in Fig. 3. A cryptographic key is error correction encoded and XORed with the enrollment iris code to create the commitment. At the time of authentication a single iris image is preprocessed, slitted and the according iris recognition algorithm is applied. The resulting iris code is XORed with the stored commitment and error correction decoding is performed. If the error correction decoding succeeds a correct key is returned, which is sufficiently long to be used in a generic cryptosystem.

5 Experimental Results

For the performance evaluation of the proposed schemes a subset of the CASIA-IrisV3-Interval database [20] is used, where for each person at least 8 iris images are available, which makes a total number of about 100 different persons. The first three images of each person were used for the enrollment procedure and the remaining five images were tested against the stored templates.

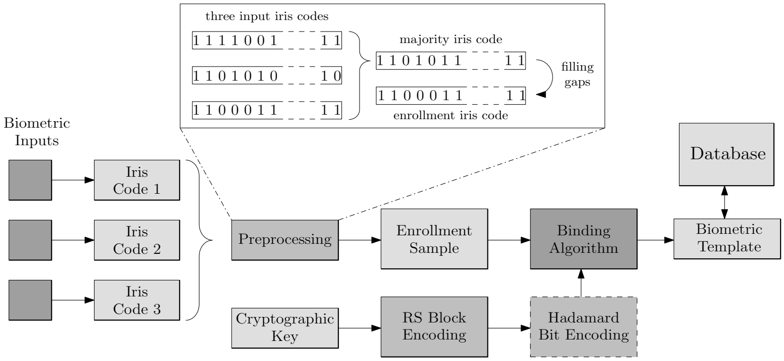

Fig. 3. The entire enrollment procedure: Iris codes are extracted from preprocessed iris images according to the applied algorithm and bound with encoded cryptographic keys

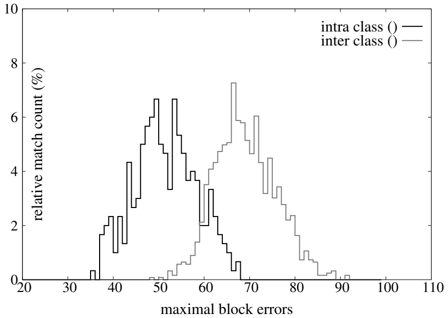

Fig. 4. The distribution of the maximal number of errors within 128-bit block for the algorithm of Ko *et al.*

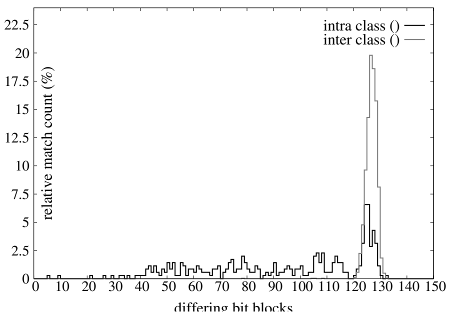

Fig. 5. The distribution of the total number of 8-bit block level errors without rearranging the bits for the algorithm of Ko *et al.*

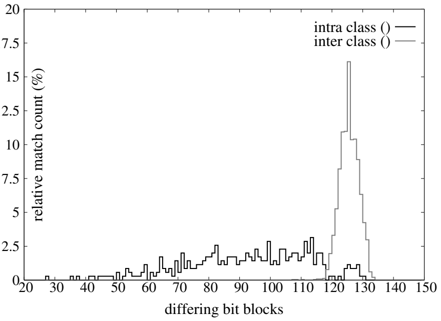

Fig. 6. The distribution of the total number of 8-bit block level errors of the proposed fuzzy commitment scheme, using the algorithm of Ko *et al.*

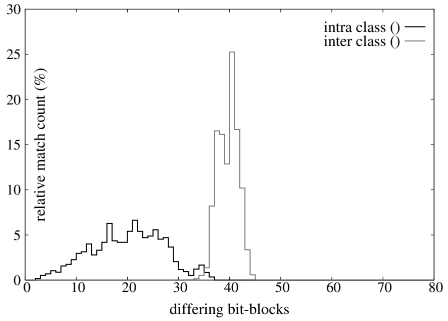

Fig. 7. The distribution of the total number of 8-bit block level errors remaining after Hadamard decoding of the proposed fuzzy commitment scheme, using the algorithm of Ma *et al.*

The performance of each proposed schemes is described by its FRR and FAR. In contrast to common recognition systems, in biometric cryptosystems the FRR is the ratio between truly matching samples for which faulty keys are generated and the total number of tests. By analogy, the FAR describes the percentage of truly non-matching samples for which correct keys are returned. In order to avoid returning faulty keys a hash of the constructed key could be tested against a previously stored hash of the correct key.

In Fig. 7 the intra class and the inter class distance of the FCS, which uses the algorithm of Ma *et al.* is shown, according to the differing bit blocks after Hadamard decoding. While the Hadamard code corrects bit blocks containing less that 25% bit errors, the applied Reed-Solomon code is capable of correcting a total number of (80-16)/2=32 block-level errors resulting in a zero FAR and a FRR of 4.64%. The intra class and the inter class distance according to the differing bit blocks of the FCS, for which the algorithm of Ko *et al.* is applied is shown in Fig. 6 (compare Fig. 5 for the results without bit rearrangement). The Reed-Solomon code is capable of correcting (250-16)/2=117 block-level errors which results in a of FAR 0.08% and a FRR of 6.57%.

6 Summary

Until now only little work has been published concerning biometric cryptosystems which use the iris as biometric characteristic. Additionally, the published literature is mostly custom-built according to the area of application. In contrast, this work shows a generic approach of how to build up iris-based biometric cryptosystems by applying a cryptographic primitive called FCS. Different types of iris recognition algorithms, for which defined preconditions are fulfilled, are used to construct two FCSs in a systematic manner. Experimental results demonstrate the soundness of the proposed approach.

References

1. Uludag, U., Pankanti, S., Prabhakar, S., Jain, A.K.: Biometric cryptosystems: issues and challenges. Proceedings of the IEEE 92(6), 948–960 (2004)
2. Juels, A., Wattenberg, M.: A FCS. In: Sixth ACM Conference on Computer and Communications Security, pp. 28–36 (1999)
3. Bowyer, K., Hollingsworth, K., Flynn, P.: Image understanding for iris biometrics: a survey. Computer Vision and Image Understanding 110, 281–307 (2008)
4. Daugman, J.: How Iris Recognition Works. IEEE Trans. CSVT 14(1), 21–30 (2004)
5. Ma, L., Tan, T., Wang, Y., Zhang, D.: Efficient Iris Recogntion by Characterizing Key Local Variations. IEEE Transactions on Image Processing 13(6), 739–750 (2004)
6. Ko, J.-G., Gil, Y.-H., Yoo, J.-H., Chung, K.-I.: A Novel and Efficient Feature Extraction Method for Iris Recognition. ETRI Journal 29(3), 399–401 (2007)
7. Soutar, C., Roberge, D., Stoianov, A., Gilroy, R., Kumar, B.V.: Biometric Encryption - Enrollment and Verification Procedures. In: Proc. SPIE, Optical Pattern Recognition IX, vol. 3386, pp. 24–35 (1998)

8. Davida, G., Frankel, Y., Matt, B.: On enabling secure applications through off-line biometric identification. In: Proc. of IEEE Symp. on Security and Privacy, pp. 148–157 (1998)
9. Davida, G., Frankel, Y., Matt, B.: On the relation of error correction and cryptography to an off line biometric based identication scheme. In: Proc. of WCC 1999, Workshop on Coding and Cryptography, pp. 129–138 (1999)
10. Hao, F., Anderson, R., Daugman, J.: Combining Cryptography with Biometrics Effectively. IEEE Transactions on Computers 55(9), 1081–1088 (2006)
11. Monrose, F., Reiter, M.K., Wetzel, S.: Password hardening based on keystroke dynamics. In: Proceedings of sixth ACM Conference on Computer and Communications Security, CCCS, pp. 73–82 (1999)
12. Monrose, F., Li, Q., Reiter, M.K., Wetzel, S.: Cryptographic Key Generation from Voice. In: SP 2001: Proceedings of the 2001 IEEE Symposium on Security and Privacy, 12 pages (2001)
13. Dodis, Y., Ostrovsky, R., Reyzin, L., Smith, A.: Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data. In: Cachin, C., Camenisch, J.L. (eds.) EUROCRYPT 2004. LNCS, vol. 3027, pp. 523–540. Springer, Heidelberg (2004)
14. Ratha, N.K., Connell, J.H., Bolle, R.M.: Enhancing security and privacy in biometrics-based authentication systems. IBM Systems Journal 40, 614–634 (2001)
15. Vielhauer, C., Steinmetz, R., Mayerhöfer, A.: Biometric hash based on statistical features of online signatures. In: ICPR 2002: Proceedings of the 16th International Conference on Pattern Recognition (ICPR 2002), vol. 1, p. 10123 (2002)
16. Wu, X., Qi, N., Wang, K., Zhang, D.: A Novel Cryptosystem based on Iris Key Generation. In: Fourth International Conference on Natural Computation (ICNC 2008), pp. 53–56 (2008)
17. Agaian, S.S.: Hadamard Matrix and Their Applications. Lect. notes in math., vol. 1168. Springer, Heidelberg (1985)
18. Reed, I., Solomon, G.: Polynomial codes over certain finite fields. Journal of the Society for Industrial and Applied Mathematics 8, 300–304 (1960)
19. Berlekamp, E.: Factoring Polynomials Over Finite Fields. Bell Systems Technical Journal 46, 1853–1859 (1967)
20. The Center of Biometrics and Security Research, CASIA Iris Image Database, http://www.sinobiometrics.com