International Journal of Computer Applications (0975 8887)
Volume 79 - No. 11, October 2013

Fuzzy Commitment Scheme for Masked Iris Codes

Osama Ouda
Information Technology Department
School of Computers and Information Sciences
Mansoura University, Mansoura 35516, Egypt

ABSTRACT

The fuzzy commitment scheme is one of most popular biometric cryptosystems that aim at securing cryptographic keys using biometrics. Because of the high recognition accuracy exhibited by the iris, iris-based fuzzy commitment schemes, among other modalities, provide the most practical performance rates. Unfortunately, existing iris-based fuzzy commitment schemes do not incorporate noise masks, generated along with iris-codes to highlight unwanted regions of the iris, because there is no way to know the mask of the decoding iris sample in advance. Therefore, the decoding accuracy of iris-based fuzzy commitment schemes is much less than the recognition accuracy of the underlying iris recognition system. This paper presents an iris-based fuzzy commitment scheme that uses the noise mask of the encoding iris sample at both encoding and decoding stages. Experimental results show that the proposed scheme provides a remarkable improvement in the decoding accuracy of iris-based fuzzy commitment schemes.

Keywords:

Biometric Cryptosyatems, Fuzzy Commitment Scheme, Iris Codes

1. INTRODUCTION

Recent advances in biometric technology paved the way for the emergence of several new applications other than the traditional application of personal authentication/identification. Providing secure management of cryptographic keys is one of such applications. Systems that employ biometric traits to secure cryptographic keys are often called biometric cryptosystems [13]. The Fuzzy Commitment Scheme (FCS) [8] is one of the most popular biometric cryptosystems that has been applied successfully to iris [2, 7, 9, 12, 15], fingerprint [10], and face [1] biometrics. FCS binds cryptographic keys, encoded using Error Correction Codes (ECC), to binary biometric feature vectors using simple XOR operation.

Due to the high recognition accuracy exhibited by iris biometric [5], iris-code based FCSs provide higher performance rates, compared to FCSs applied to other biometric characteristics. Besides, unlike other biometric traits from which real-valued feature vectors are extracted, iris-codes are binary feature vectors extracted from iris texture using the standard iris recognition algorithm of Daugman [3]. That is, the FCS can be applied directly to iris-codes whereas an extra binarization step, that may cause information loss, is required before it can be applied to other biometric traits.

In Daugman's approach to iris recognition, in addition to the iris-code, a corresponding noise mask of the same size is generated to designate unwanted regions of the iris such as eyelids, eyelashes, and specular highlights. Incorporating masks of enrolment and authentication iris samples in the matching process improves the recognition accuracy significantly [4]. However, incorporating those masks in iris-code based FCSs would introduce implementation difficulty because only the enrolment iris sample is available at the time of encoding (key binding) [6]. As a result, current iris-based FCSs [2, 7, 9, 12, 15] do not use masking information neither at encoding nor at decoding phases and hence the decoding accuracy of such schemes is much less than the recognition accuracy of the underlying iris recognition system.

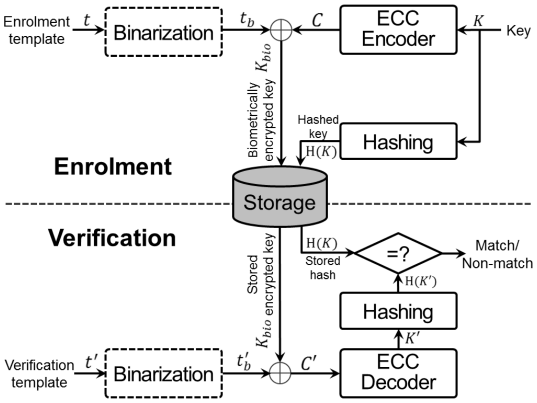

Fig. 1. Illustration of the Fuzzy Commitment Scheme.

This paper proposes an iris-code based FCS that benefits from noise masks and hence improves the decoding (key release) accuracy. In this scheme, the noise mask of the enrolment iris sample is employed at both encoding and decoding stages. Experimental results show that the decoding accuracy of the proposed scheme, which incorporates a single noise mask, outperforms the decoding accuracy of FCS implementation that does not employ any masks.

The rest of this paper is organized as follows: a brief review of the FCS is given in Sect. 2. The proposed scheme is explained in Sect. 3. Experimental results are presented in Sect. 4 and Sect.5 concludes the paper.

2. FUZZY COMMITMENT SCHEME

Figure 1 shows an illustration of the FCS. As illustrated in the figure, the FCS requires a biometric template, t, to be represented as a binary string. Hence, a non-binary template need to be binarized via an optional binarization module into a binary template, t_b, be-

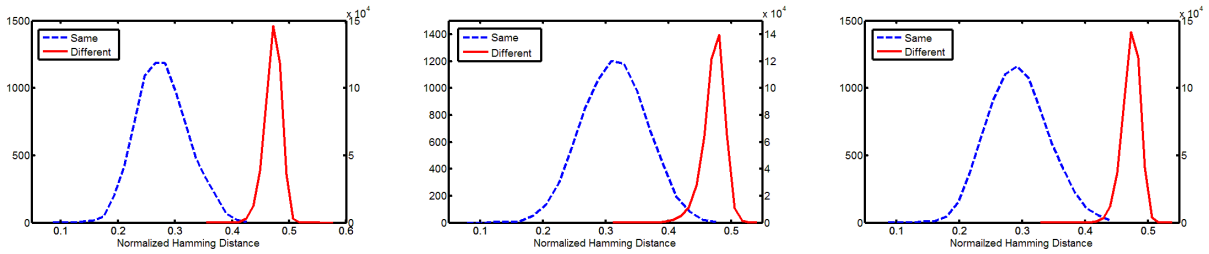

Fig. 2. Genuine and imposter distributions. (a) incorporating noise masks, (b) without noise masks (iris-codes only), and (c) incorporating noise masks of the enrolment samples only.

fore the FCS can be applied. For iris-codes, t_b is identical to t since iris-codes are binary by default. On enrolment, the cryptographic key K, that we want to secure using the biometric templates, is encoded using an appropriate ECC(s) into a codeword C of length $n = \|t_b\|$. Both the binary template and the encoded key are then XORed to produce a biometrically secured key K_{bio}, also called a biometric key, as follows:

$$K_{bio} = C \oplus t_b \tag{1}$$

Furthermore, the hash value of the random key $\mathrm{H}(K)$ is computed and stored along with the biometric key in a central storage or a user-specific token.

At the time of verification, a binary template $t_b^{'}$ is extracted from a live biometric sample captured from the person being verified and XORed with the stored biometric key to obtain a possibly corrupted codeword C':

$$C' = t_b' \oplus K_{bio} \tag{2}$$

The obtained codeword is decoded using the ECC(s) employed on enrolment to get the verification key K'. Finally, the hash value of the recovered key, $\mathrm{H}(K')$, is computed using the same hashing function employed on enrolment and compared to the stored hash value, $\mathrm{H}(K)$. Only if the two hash values are identical, the key is released; otherwise, the authentication process fails.

3. FCS FOR MASKED IRIS-CODES

Iriscode is a binary representation of discriminative features extracted from an iris image. Thus, the similarity between any two iriscodes, A and B, of length n can be measured simply using the normalized Hamming distance as follows:

$$d_H = \frac{\|(A \oplus B)\|}{n}, \tag{3}$$

where $\oplus$ denotes XOR and $\|\cdot\|$ the norm of the binary vector. Typically, iris images include unwanted regions such as eyelashes, eyelids, specular reflections, etc. Therefore, a corresponding binary noise mask, of the same size as the iris-code, is generated in addition to the iris-code to highlight those noisy regions. Hence, the similarity between iris-codes taking into account their corresponding noise masks can be measured as follows [4]:

$$d_H = \frac{\|(A \oplus B) \cap Mask_A \cap Mask_B\|}{\|Mask_A \cap Mask_B\|}, \tag{4}$$

where $\cap$ denotes the AND operation and $Mask_A$ and $Mask_B$ are the binary masks corresponding to iris-codes A and B, respectively.

Incorporating noise masks in the matching process improves the recognition accuracy significantly. We demonstrate this experimentally using CASIA-IrisV3-Interval iris database [14] and the open source iris recognition system described in [11]. This database contains 2639 8-bit grey scale images, with a resolution of 320×280 pixels, collected from 395 different classes (eyes) of 249 subjects. Because the adopted iris recognition system cannot derive iris-codes correctly from poor quality images in this database, we excluded erroneous iris samples and used a dataset of 2422 samples from 393 irises in our experiments. Figures 2(a) and 2(b) show the Hamming distance distributions for comparisons between iris-codes generated from same and different irises with and without noise masks respectively. These figures show clearly that employing masks provides much better separability between genuine and imposter distributions.

Because masking information of the decoding sample is not available a priori (at the encoding stage) iris-based FCSs do not incorporate noise masks. In this paper, we show that using the noise mask of only the enrolment sample at both encoding and decoding stages can improve the decoding accuracy of iris-based FCSs. The Hamming distance between a pair of iris-codes, A and B, incorporating only the noise mask of A is:

$$d_H = \frac{\|(A \oplus B) \cap Mask_A\|}{\|Mask_A\|}, \tag{5}$$

Figure 2(c) shows same and different Hamming distance distributions computed using Eq. (5). Although the separability between the two distributions is not as good as in the case of using masks of both the encoding and decoding iris codes (Fig. 2(a)), it is much better than the case where no masks are used at all (Fig. 2(b)). Obviously, this is due to the fact that masks generated from the same iris sample are more similar than masks generated from different irises. For CASIA-IrisV3-Interval iris database, we found that, using the above-mentioned setup, the average similarity between masks of same irises is 88.57% while the average similarity between masks of different irises is 35.14%. This implies that using the mask of the encoding sample at both encoding and decoding stages in iris-based FCSs, instead of doing without any masking information, would improve the decoding accuracy of such schemes significantly.

The proposed iris-based FCS is illustrated in Fig.3. Similar to existing iris-based FCS implementations, the proposed scheme is a

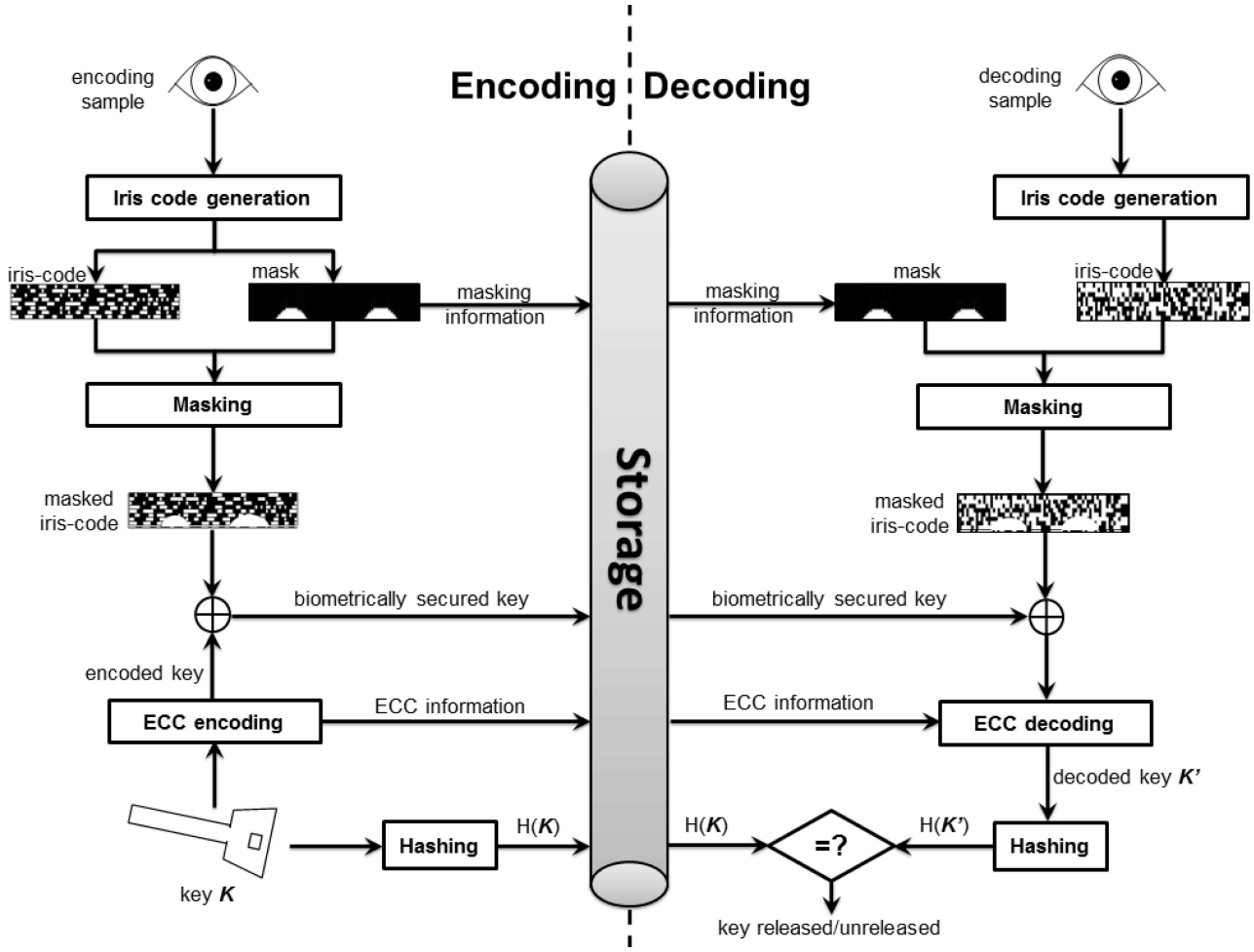

Fig. 3. Proposed iris-based FCS

two-factor scheme. At the encoding stage, an iris sample is captured from the user and both of the iris-code and noise mask are derived from that sample employing the standard iris recognition system proposed by Daugman [3]. Masking information is added to iris features information, represented by the generated iris-code, by ANDing it with the noise mask. Since the noise mask of the encoding sample will be employed at the decoding stage, it must be stored in a user-specific token or card as depicted in Fig. 3. At the same time, the cryptographic key K, to be committed by the masked iris-code, is prepared by the appropriate ECC(s) to obtain an encoded key of the same size as the iris-code. Finally, the commitment step is carried out by XORing the encoded key and the masked iris-code. The result of this step is a biometrically secured key which is stored along with the ECC(s) information and the hash value H(K) of the original key in the user-specific storage. At the decoding stage, an iris sample is acquired, from the user who asks for releasing the key, and iris-code is generated from that sample. The masked version of this sample is obtained by ANDing it with the noise mask of the encoding sample which can be retrieved from the user-specific storage. Once the masked iris-code is obtained, it is XORed with the biometrically secured key and the result is decoded using the stored ECC information to obtain the EC decoded key K'. At the last step of the decoding stage the hash value of the k' is computed, using the same hashing function employed at the encoding stage, and compared to the hash value of the original key. If both values are identical, the decoded key K' is released for further processing; otherwise, the key release process fails.

4. EXPERIMENTS AND DISCUSSION

To evaluate the improvement in decoding accuracy after masking out noisy bits of the enrolment iris sample, we follow the well-known iris-based FCS implementation of Hao et al. [7]. This system was tested on a small private set of ideal iris images (70 classes, with 10 samples from each class). Experiments in this paper are carried out on the publicly available CASIA-v3-Interval iris dataset that consists of 2639 8-bit gray scale images, with a resolution of 320×280 pixels, collected from 395 different classes (eyes) of 249 subjects. It is therefore not expected to match the performance reported by Hao et al. However, since our concern here is to com-

pare the decoding accuracy of FCS implementation for masked and unmasked iris-codes, it is not critical to match their reported performance for unmasked iris-codes. As mentioned in the previous section, iris-codes are generated using the open source MATLAB implementation for iris recognition provided in [11]. This implementation generates 9600-bit iris codes, together with their corresponding 9600-bit noise masks.

At enrolment, a masked iris-code is obtained from each iris-code via ANDing it with its corresponding noise mask and a cryptographic key K of length l is encoded using the two-layer ECC scheme, described in [7], into an $n_p(= 9600)$-bit codeword c. This concatenated ECC scheme combines Hadamard and Reed-Solomon ECCs to deal with background and burst errors in iris codes, respectively. In the first layer, bits in K are divided into k_{RS} blocks of k-bit each. This set of blocks is then represented as a message of k_{RS} symbols over F_{2^k} and encoded into a codeword of n_{RS} symbols using a (n_{RS}, k_{RS}, t_{RS}) Reed-Solomon(R-S) code with a correction capacity $t_{RS} = (n_{RS} - k_{RS})/2$. In the second layer, each of the resulting n_{RS} symbols is represented as a k-bit binary word and encoded into a 2^{k-1} bit codeword using a (2^{k-1}, k, 2^{k-2}) Hadamard code. Such Hadamard code is generated from a Hadamard matrix of order $k-1$ and can correct up to $2^{k-3}-1$ erroneous bits in each codeword. The following equation shows how the size n_p of the final encoded key and the parameters of both Hadamard and R-S codes are related:

$$n_p = n_{RS} \times 2^{k-1} \tag{6}$$

For instance, a 40-bit key can be encoded into a 9600-bit codeword via encoding it using the R-S(75,5), to get a $8 \times 75 = 600$-bit codeword and then encoding each 8-bit chunk in this codeword using the (128, 8, 64) Hadamard code to obtain the final $75 \times 128 = 9600$-bit codeword. In our experiments, the (128, 8, 64) Hadamard code is found to give the best correction results for background errors. Thus, the correction capacity of the implemented ECC scheme depends only on the R-S parameters, n_{RS} and k_{RS}. The lowest correction capacity is obtained when $n_{RS} = k_{RS}$ (i.e., only Hadamard encoding is employed). In other words, the correction capacity in this case would not exceed 25% of the encoded codeword, since the correction capability of sole Hadamard coeds is up to 25%. On the other hand, the correction capacity of the overall ECC scheme increases as the difference between n_{RS} and k_{RS} increases. Biometrically secured keys are then obtained by XORing the obtained encoded key with each masked iris-code.

At verification, the intra-user decoding accuracy is evaluated using leave-one-out cross-validation strategy. For each class, only one sample is considered as the enrolment sample and its noise mask is used to generate masked codes for the remaining iris samples for that class. Each verification iris-code is then XORed with the biometric key generated using the enrolment sample and the obtained result is decoded using ECCs employed at enrolment. Verification templates are shifted circularly up to 10 bits in both directions, to account for misalignment, and the above procedure is repeated after each shift. The overall process is repeated for other iris samples of the same class. On the other hand, to evaluate the inter-user decoding accuracy, each iris-sample for each class is treated as the enrolment sample and matched against all samples from all other classes.

Table 1 shows the experimental results obtained using different R-S codes. The results show that, for unmasked iris-codes based

Table 1. Experimental results

	Masked iris-codes		Unmasked iris-codes	
Key length	FRR(%)	FAR(%)	FRR(%)	FAR(%)
376 R-S (75, 47)	28.0103	0.0755	75.5297	0
360 R-S (75, 45)	25.0258	0.1149	72.9845	0
344 R-S (75, 43)	22.5323	0.1554	70.9302	0
328 R-S (75, 41)	19.8708	0.2110	68.3592	0
312 R-S (75, 39)	17.7132	0.2832	65.7235	0
296 R-S (75, 37)	15.8140	0.3849	60.5814	0
280 R-S (75, 35)	13.6176	0.5249	57.9845	0
264 R-S (75, 33)	11.7054	0.7898	55.2196	0
248 R-S (75, 31)	11.1628	0.9392	54.6512	0
232 R-S (75, 29)	9.8837	1.1026	49.9871	0
216 R-S (75, 27)	7.5581	1.5095	49.1990	0.0002
200 R-S (75, 25)	6.1886	1.8783	43.5142	0.0002
184 R-S (75, 23)	5.5297	2.2352	41.1499	0.0009
168 R-S (75, 21)	4.7416	2.7877	37.9716	0.0009
152 R-S (75, 19)	4.0310	3.3951	39.4961	0.0016
136 R-S (75, 17)	3.4109	4.1903	32.7778	0.0027
120 R-S (75, 15)	3.2946	4.9674	30.0388	0.0031
104 R-S (75, 13)	2.4806	6.4259	27.2093	0.0040
88 R-S (75, 11)	2.1964	6.5220	25.1034	0.0062
72 R-S(75,9)	1.8734	8.0896	21.6408	0.0147
56 R-S(75,7)	1.2274	11.4554	16.0207	0.0548
40 R-S(75,5)	1.0207	13.7854	12.5969	0.0914

FCS implementation, the false acceptance rates (FARs) are generally low ($< 1\%$) and they decrease as key length increases. On the other hand, the false rejection rates (FRRs) are high ($> 12\%$ for 40-bit key) and they increase as key length increases. These results are expected because all the imposter Hamming distances, in case of unmasked iris-codes, are higher than 30% which is beyond the correction capacity of the employed ECC scheme. Also, many genuine Hamming distances are high (4669 comparisons $> 30\%$) because masking information are not used. This makes correct key decoding for such genuine iris samples not possible even for short key sizes. For masked iris-codes based implementation, both genuine and imposter Hamming distance distributions are shifted to the left, as shown in Fig. 2(c). This resulted in an increase in FARs and a decrease in FRRs. As shown in Table.1, as key lengths increase, the correction capacity decreases and hence the FARs decrease and the FRRs increases.

Generally speaking, although the FARs in case of unmasked iris-codes is low, practical FRRs could not be obtained even for short keys. On the other hand, practical FRRs and FARs were obtained for long enough key lengths when masking information of the enrolment iris sample was employed. However, we should assume that the masking information are known to attackers and hence the key sizes shown in Table. 1, in case of masked iris-codes, do not reflect the actual security of the committed key. Therefore, the average number of noisy bits in an iris-code was computed for the adopted iris dataset. We found that approximately 20% of bits in an iris-code are noisy and should be masked out. That is, the actual key lengths that can be secured using the ECC parameters shown in Table.1, for the case of masked iris-codes only, are 0.8 of the lengths shown in the table.

It should also be noted that the disclosure of noise masks to attackers may affect users' privacy. However, this effect may not be critical because of two reasons. First, because these masks designate unwanted regions such as eyelid and eyelashes that are common in any iris image, they look similar in most of iris images. Second, since users enrol in different applications using different iris samples, iris masks cannot be identical for the same user. Rather, they differ due to the changing conditions of the acquisition process.

5. CONCLUSION

This paper presented an iris-based fuzzy commitment scheme which, unlike other existing iris-based fuzzy commitment schemes, incorporates noise information in the commitment process. Because the noise mask of the decoding iris sample cannot be available in advance, the noise mask of the encoding sample is employed at both encoding and decoding stages. Experimental results showed that the decoding accuracy of the proposed scheme outperforms existing iris-based fuzzy commitment schemes that do not use noise information.

6. REFERENCES

[1] Meng Ao and Stan Z. Li. Near infrared face based biometric key binding. In *Proc of the 3rd Int Conf on Biometrics (ICB'09)*, pages 376–385, 2009.

[2] Julien Bringer, Hervé Chabanne, Gérard D. Cohen, Bruno Kindarji, and Gilles Zémor. Theoretical and practical boundaries of binary secure sketches. *IEEE Transactions on Information Forensics and Security*, 3(4):673–683, 2008.

[3] John Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(11):1148–1161, 1993.

[4] John Daugman. How iris recognition works. *IEEE Trans. Circuits Syst. Video Techn.*, 14(1):21–30, 2004.

[5] John Daugman. Probing the uniqueness and randomness of iriscodes: Results from 200 billion iris pair comparisons. *Proceedings of the IEEE*, 94(11):1927–1935, 2011.

[6] Feng Hao. *On using fuzzy data in security mechanisms*. PhD thesis, University of Cambridge, April 2007.

[7] Feng Hao, Ross Anderson, and John Daugman. Combining crypto with biometrics effectively. *IEEE Trans. Computers*, 55(9):1081–1088, 2006.

[8] Ari Juels and Martin Wattenberg. A fuzzy commitment scheme. In *6th ACM Conf on Computer and Communications Security*, pages 28–36. ACM Press, 1999.

[9] Sanjay Ganesh Kanade, Dijana Petrovska-Delacrétaz, and Bernadette Dorizzi. Cancelable iris biometrics and using error correcting codes to reduce variability in biometric data. In *IEEE Int Conf on Computer Vision and Pattern Recognition (CVPR'09) Workshops*, pages 120–127, 2009.

[10] Peng Li, Xin Yang, Hua Qiao, Kai Cao, Eryun Liu, and Jie Tian. An effective biometric cryptosystem combining fingerprints with error correction codes. *Expert Syst. Appl.*, 39(7):6562–6574, 2012.

[11] Libor Masek and Peter Kovesi. Matlab source code for a biometric identification system based on iris patterns. *The School of Computer Science and Software Engineering, The University of Western Australia*, 2003.

[12] Christian Rathgeb and Andreas Uhl. Systematic construction of iris-based fuzzy commitment schemes. In *Proc of the 3rd Int Conf on Biometrics (ICB'09)*, pages 940–949, 2009.

[13] Christian Rathgeb and Andreas Uhl. A survey on biometric cryptosystems and cancelable biometrics. *EURASIP J. Information Security*, 2011:3, 2011.

[14] The Chinese Academy of Sciences, CASIA Iris Image Database. `http://www.cbsr.ia.ac.cn/IrisDatabase`.

[15] Sheikh Ziauddin and Matthew N. Dailey. Robust iris verification for key management. *Pattern Recognition Letters*, 31(9):926–935, 2010.