Two-Factor Authentication or How to Potentially Counterfeit Experimental Results in Biometric Systems[⋆]

Christian Rathgeb and Andreas Uhl

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

Abstract. Two-factor authentication has been introduced in order to enhance security in authentication systems. Different factors have been introduced, which are combined for means of controlling access. The increasing demand for high security applications has led to a growing interest in biometrics. As a result several two-factor authentication systems are designed to include biometric authentication.

In this work the risk of result distortion during performance evaluations of two-factor authentication systems including biometrics is pointed out. Existing approaches to two-factor biometric authentication systems are analyzed. Based on iris biometrics a case study is presented, which demonstrates the trap of untruly increasing recognition rates by introducing a second authentication factor to a biometric authentication system. Consequently, several requirements for performance evaluations of two-factor biometric authentication systems are stated.

1 Introduction

Reliable personal recognition is required by a wide variety of access control systems. Examples of these systems include ATMs, laptops and cellular phones [1]. If these systems fail to meet the demands of reliable and robust authentication potential imposters may gain access to these systems. In order to enhance the security of access control systems two factor authentication (T-FA) has been introduced, wherein two factors are combined in order to authenticate a user. The key idea of T-FA is to sum up the security of two factors. These factors include, passwords, representing "something you know", or physical tokens, such as smart-cards, representing "something you have". Additionally, biometric traits are applied, respresenting "something you are".

However, several problems may occur when introducing biometric authentication to T-FA systems. Performance gain with respect to recognition rates is often achieved due to the assumption of unrealistic preconditions. Resulting performance distortions may not be recognized at first sight, yet, these could lead to serious security vulnerabilities. In order to shed light on the use of biometrics as additional factors in T-FA schemes we demonstrate a way of how to untruly

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

A. Campilho and M. Kamel (Eds.): ICIAR 2010, Part II, LNCS 6112, pp. 296–305, 2010.

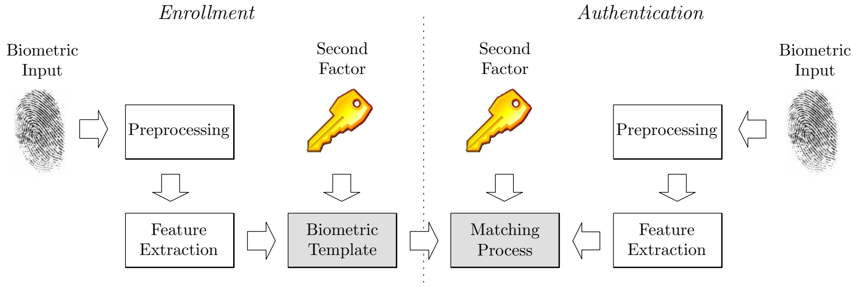

Fig. 1. The basic operation mode of a two-factor biometric authentication system. During enrollment and authentication additional factors are presented to the system.

improve recognition rates in an biometric recognition system by introducing a token-based T-FA scheme. As a consequence, we manifest requirements with respect to performance measurements in T-FA systems including biometrics.

The remainder of this paper is organized as follows: in Sect. 2 we first summarize the fundamentals of T-FA and biometric systems. Performance evaluation in biometric verification systems is described and existing T-FA systems involving biometric authentication are analyzed. In Sect. 3 we propose a case study in which we apply T-FA to an iris-biometric verification system. Experimental results are analyzed and a summary of T-FA and biometric verification systems is given in Sect. 4. In conclusion, requirements for performance evaluations in such systems are stated.

2 T-FA and Biometric Systems

An authentication factor is a piece of information used to authenticate or verify the identity of a user. In a T-FA system two different factors are combined in order to authenticate a user. It is claimed that T-FA generally delivers a higher level of authentication assurance compared to using just one factor. Three basic classes of factors can be distinguished: personal factors, such as user-defined passwords, physical factors, such as smart-cards [2] or human factors, such as biometric traits [3]. Combining two factors from two different classes yields T-FA where each factor is applied independently.

A common example of T-FA systems are ATMs, where physical factors are combined with personal factors. T-FA only applies to systems which use factors of different classes – authentication schemes based on, for example, two biometric modalities are referred to as multi-modal authentication [4]. Besides known vulnerabilities of T-FA schemes [5], such as “man-in-the-middle” attacks, where an imposter does not need to be in possession of any physical factor, we will focus on the risk of false performance evaluation of T-FA schemes involving biometrics as human factor. Fig. 1 shows the basic operation mode of a T-FA system including biometric authentication. At the time of enrollment biometric traits and a second personal or physical factor are presented. During authentication this factor is presented again in order to achieve successful authentication.

With respect to biometric authentication two different modes are distinguished: verification and identification [1]. Since we aim at analyzing T-FA schemes including biometrics we will only focus on verification, since the presentation of an additional personal or physical token represents an identity claim per se. Due to the variance in biometric measurements biometric systems do not provide perfect matching, as it is easily implemented in password or PIN-based authentication systems [3]. Thus, a fuzzy matching is performed where decision-thresholds are set up in order to distinguish between genuine and non-genuine users, respectively. Hence, several magnitudes define the performance of a biometric system. Widely used measures include False Rejection Rate (FRR), False Acceptance Rate (FAR) and Equal Error Rate (EER).

2.1 Biometric T-FA Systems

Obviously, T-FA increases the security of biometric authentication systems since potential imposters have to compromise the second factors in order to gain access to the system as a first attack stage. However, if a T-FA scheme is constructed where one factor is represented by a biometric trait additional factors are either personal or physical tokens. Both of these factors require a perfect matching. This means a wrong PIN or a wrong smart-card would result in a rejection of the imposters. If each imposter would be in possesion of the correct second factor of the account he wants to gain access to (i.e. the second factor has been compromised), the overall recognition rate of system is expected to remain the same. Applying a sequential check of both factors (regardless of the order), the recognition rate is equal to a system which only performs biometric authentication. That is, additional factors become meaningless with respect to recognition performance since these are potentially compromised by all imposters. By analogy, if no imposter would be in possesion of the correct second factor of the account he wants to gain access to, the overall recognition rate is of course expected to increase. This is because imposters which may have tricked the biometric authentication system are rejected at the time the additional factor is checked. Throughout literature several approaches have been proposed where T-FA is introduced to biometric systems. In any case, authors claim to introduce personal or physical tokens in order to enhance the security of the system. However, in some cases it is doubtful if the proposed system can maintain recognition rates without the use of a second factor (i.e. in case the second factor has been compromised). In order to underline the problem of evaluating the performance of biometric T-FA schemes in terms of recognition performance, we will discuss several approaches which we found questionable regarding reported performance results. Hence, we do not cover all approaches to biometric T-FA schemes, but only a small selection to emphasize that potential incorrect performance evaluations should be considered an important issue.

The introduction of biometrics to generic cryptographic systems resulted in a new field of research, named biometric cryptosystems [6]. Most approaches which can be subsumed under the category of biometric cryptosystems aim at extracting cryptographic keys out of biometric data. Performance evaluations are adopted such that correctly generated keys are equivalent to successful authentication and vice versa. Due to biometric variance a widespread usage of helper data, for example error correction codes, has proven to be worthwhile. However, in several approaches the application of helper data conceals the actual performance rates. In the following, we provide three examples for that.

Teoh *et al.* [7,8] introduced a technique to generate cryptographic hashes out of face biometrics which they refer to as "BioHashing". Like in generic T-FA schemes, in the BioHashing approach random numbers are associated with each user. These user-specific random numbers are used as seed in order to generate biometric hashes. These random numbers, which represent the second factor, have to be presented to the system in addition to biometric data at authentication. The authors report almost perfect performance rates for the generation of biometric hashes. In order to expose the true performance of BioHashing, Kong *et al.* [9] presented an implementation of FaceHashing. It was found that the reported performance was achieved under the hidden assumption that random numbers are never lost, stolen, or compromised. This assumption does not hold in general. Physical tokens can be stolen or duplicated while personal tokens can be easily guessed or broken, for example by dictionary attacks [10]. In order to associate cryptographic with biometric data Reddy *et al.* [11] proposed a so-called "fuzzy vault scheme" [12] based on iris biometrics. The authors achieve T-FA by embedding an additional layer of security, namely a password. With this password the generated vault as well as the secret key is encrypted. In experiments the security of a fuzzy vault scheme which exhibits a FRR of 8% and a FAR of 0.03% is increased, where a total number of 100 templates are used. As result of the hardening scheme the FRR increases to 9.8% due to misclassification of a few minutiae. At the same time the FAR decreases to 0.0%. It is claimed that this is due to the fact that minutiae are distributed more randomly. If this was the case for the use of one single password (identical for all users), this could be integrated into the original algorithm to increase performance. However, if passwords are compromised the systems' security decreases to that of an ordinary fuzzy vault scheme which indicates that the FAR of 0.0% was calculated under unrealistic preconditions. In recent work Jassim *et al.* [13] proposed a method of improving the performance of PCA based face recognition system. The authors introduce random projections based on the Gram-Schmidt process which are used to map biometric features onto secret personalized domains. For this purpose a secret permutation matrix is assigned to each user, which represents a second factor. This means, each user is in possession of a unique matrix. In experimental results accumulations of genuine users remain the same while inter-class distances increase. While the original PCA based system reveals an EER of 17% an EER of 0.2% is reported if random projections are applied for each user. The authors do not consider the case where the same permutation matrix is assigned to each user. Therefore, again the results are achieved under the assumption, that the second factor has not been compromised.

Ratha *et al.* [14] introduced the concept of "cancellable biometrics". Biometric data can be compromised and therefore become useless because it can not be modified ex post. The idea of cancellable biometrics consists of intentional, repeatable distortion of a biometric signal based on transforms where matching is performed in the transformed space. Thus, if a potential imposter is able to steal the stored template recovering of original biometric data becomes infeasible. In contrast to biometric recognition systems or biometric cryptosystems the scope of cancellable biometrics is template security. By definition, a T-FA system is constructed since user-specific transformations are applied. That is, different transformations represent second factors which are used to secure biometric templates. All of the approaches presented in Section 2.1 can be seen as systems which provide cancellable biometrics, since random numbers, passwords or transformations which are applied to biometric features can be updated easily. With respect to recognition performance approaches to cancellable biometrics aim at maintaining the performance of the original underlying systems (in general loss of performance is expected). This means, approaches to cancellable biometrics in which better performance as compared to the original algorithm is reported should be examined carefully. For example, the BioHashing approach of Teoh *et al.* [7,8] was extended to be used as cancellable biometrics. As pointed out by Cheung *et al.* [15], experimental results were obtained under the unpractical assumptions stated above.

3 T-FA and Iris Recognition: A Case Study

In this section we describe an existing iris recognition system which we apply to construct a T-FA system. First we will consider the performance of the biometric system. Subsequently, we will construct a generic T-FA scheme by introducing user-specific bit streams as second factor.

3.1 Iris Recognition System

In order to apply biometric authentication we use our own implementation of the algorithm of Ma *et al.* [16]. In their approach the iris texture is treated as a kind of transient signal which is processed using a 1-D wavelet transform. The local sharp variation points, which denote important properties of transient signals, are recorded as features. We always extract an iris texture from eye images as a first step. We assume the texture to be the area between the two almost concentric circles of the pupil and the outer iris. These two circles are found by contrast adjustment, followed by Canny edge detection and Hough transformation. After the circles are detected, unwrapping along polar coordinates is done to obtain a rectangular texture of the iris. In our case, we always resample the texture to a size of 512x64 pixels.

The texture is subsequently divided into N stripes to obtain N one-dimensional signals, each one averaged from the pixels of M adjacent rows. We used $N = 10$ and $M = 5$ for our 512x64 pixel textures (only the 50 rows close to the pupil are used from the 64 rows, as suggested in [16]). A dyadic wavelet transform

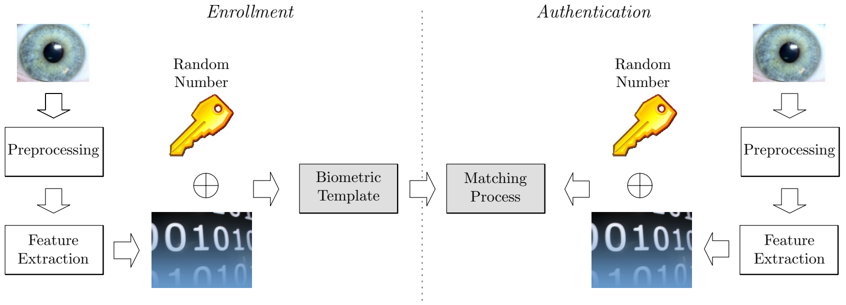

Fig. 2. T-FA scheme: random numbers are introduced with which iris codes are sequentially XORed during enrollment. At authentication the biometric template is XORed with another random number and the result is matched against an extracted iris code.

is then performed on each of the resulting 10 signals, and two fixed subbands are selected from each transform. This leads to a total of 20 subbands. In each subband we then locate all local minima and maxima above some threshold, and write a bitcode alternating between 0 and 1 at each extreme point. Using 512 bits per signal, the final code is then 512x20 bit.

Once bitcodes are obtained, matching can be performed on them and Hamming distance lends itself as a very simple distance measure. For matching to work well, we compensate for eye tilt by shifting the bit-masks during matching by three pixels in each direction.

3.2 Two-Factor Iris Recognition System

In algorithm described above, a users iris serves as the only authentication factor. In order to construct a T-FA system a second factor has to be introduced. Therefore we simply apply random numbers which are associated with specific users. These random numbers can be stored on a smart-card, representing a physical factor. Additionally, we choose rather short random bit streams, hence, these are easily remembered representing personal factors as well. At this point it is important that the application of random bit streams yields a generic approach to T-FA, since these just represent a user-specific secret.

During enrollment for each user a randomly generated bitstream is generated. The iris code of a user, which is extracted during enrollment is sequentially XORed with the random number in order to generate a secure template. That is, the stored iris code is protected by the random bit stream, similar to the approach presented by Zuo *et al.* [17]. If a user wants to authenticate with the system an appropriate random number has to be presented firstly. Subsequently, the stored template is sequentially XORed with this random number and the resulting iris code is matched against the one extracted from the presented iris image. Hence, a T-FA system is realized by simply introducing random numbers which are associated with users by sequentially XORing these with iris codes. The operation mode of the whole system is illustrated in Fig. 2.

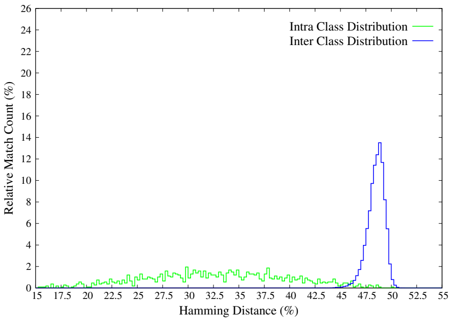

Fig. 3. Intra-class and inter-class distribution of the algorithm of Ma

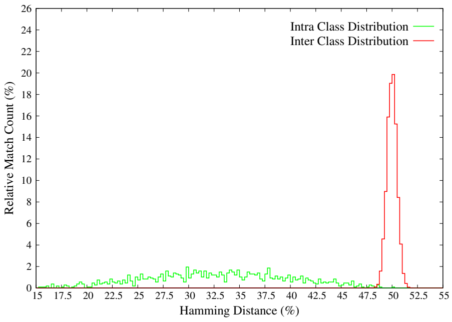

Fig. 4. Intra-class and inter-class distribution of the algorithm of Ma using T-FA

3.3 Experimental Results

Experiments are carried out using the CASIAv3-Interval iris database[1], a widely used test set of iris images of over two hundred persons. The database comprises iris images of size 320 × 280 pixels out of which normalized iris textures of 512 × 64 pixels are extracted in the preprocessing step as described earlier. Applying our implementation of the feature extraction of Ma *et al.* to preprocessed iris textures, a total number of 10240 bits are extracted per iris image. Matching is performed by calculating the Hamming distance between extracted iris codes, where a circular shift of three pixels to the left and right is implemented in order to provide some degree of rotation invariance. In Fig. 3 the distribution of the intra-class distance and the inter-class distance are plotted. Fig. 5 shows the FRR and the FAR resulting in an EER of 1.76%. For a threshold of 42% (in terms of correct bits), a FRR of 5.61% and zero FAR is achieved. In other words, the system will in general reject 5.61% of all genuine users while no imposters are untruly accepted. We are aware that these results are worse than those reported by Ma.*et al.*, however, the absolute performance of the algorithm is not the topic of this work. Thus, our implementation serves its purpose.

For the construction of a T-FA system we introduce random numbers consiting of 8 bits. Performance is measured in the same way as in the iris recognition system applying the same test set, however, now users have to present biometric data which has to pass the fuzzy match of the recognition system as well as a random number which is sequentially XORed with stored templates. Since genuine users are in possession of correct random numbers the construction of the T-FA system does not effect the intra-class distribution. Therefore, calculated Hamming distances between genuine iris codes remain the same as can be seen in Fig. 4. If we make the assumption that imposters are not in possesion of valid random numbers the performance of the whole system is increased. This means we calculate the inter-class distribution applying the random numbers to users which were assigned to them during enrollment. In other words, users claim the

[1] The Center of Biometrics and Security Research: CASIA Iris Image Database, `http://www.sinobiometrics.com`

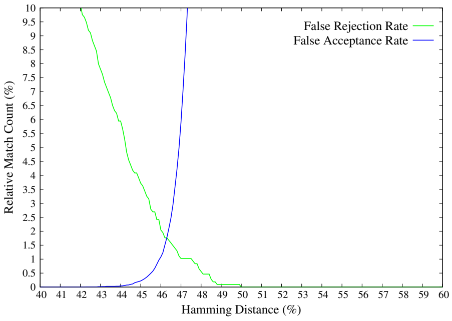

Fig. 5. FRR and FAR of the algorithm of Ma

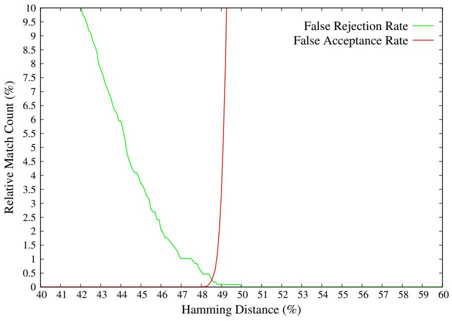

Fig. 6. FRR and FAR of the algorithm of Ma using T-FA

identity of other users but present their own biometric data as well as their own random number (a randomly chosen number could be used as well). The distribution of the resulting inter-class distances is also plotted in Fig. 4 and the FRR and the FAR are plotted in Fig. 6 resulting in a EER of 0.25%. It becomes clear that now inter-class distances accumulates around 50% instead of around 47% in the original system. This is because iris codes are now XORed with potentially different random numbers resulting in almost random bitstreams. Since binary iris codes are extracted, the Hamming distance between random bit streams is expected to be 0.5. This means we can now increase the threshold and result in overall (virtual) performance gain of the system. For an increased threshold of 48% correct bits we now achieve a FRR of 0.43% and zero FAR resulting in an EER of 0.25% which is about 1.5% better than the performance of the original iris recognition system (see Fig. 6).

3.4 Analysis

In the above presented biometric system two authentication factors are combined by sequentially XORing iris codes with 8-bit random numbers. Based on the assumption that random numbers are never compromised we increase inter-class distances and are able to gain performance by increasing the threshold which is used to separate genuine and non-genuine users.

The problem of the presented scenario is that additional factors are considered to never be stolen, lost, shared or duplicated where in practice the opposite is true. The assumption that imposters would try to infiltrate the system by presenting some random personal or physical factor is rather absurd. Additional factors such as passwords or PIN must not be considered secure since these are easily compromised [3]. As we demonstrated inter-class distributions increase. In case the decision-threshold is increased according to the new inter-class distribution the biometric system becomes more tolerant. This is because access to the system is even eased if potential imposters are in possession of a valid second factor which is a realistic scenario. In case a potential imposter is in possesion of a valid random number the recognition rate degrades to that of the original biometric system. In this case, a threshold of 48% correct bits would yield a FAR of 24.21%, that is, the system becomes highly vulnerable. Thus, performance only holds if random numbers are never compromised. Having adjusted the system decision threshold to the virtual performance as determined under unrealistic preconditions, the accuracy of the system is actually severely degraded, in case the second factor is compromised.

4 Summary and Conclusion

In all of the above discussed T-FA schemes additional factors (random numbers, passwords or permutation matrices) are considered to never be stolen, lost, shared or duplicated. If this would be the case the introduction of biometrics becomes meaningless since the system could rely on these random numbers or passwords without any risk. That is, authentication could be performed just by presenting appropriate random numbers or passwords. In case inter-class distributions are calculated under these assumptions for performance evaluations, the FAR of the system is kept artificially low. Hence, thresholds can be adapted to generate better results like in our presented scheme. That is, the biometric system is set to be more tolerant since inter-class distances become larger. However, if imposters are in possession of valid random numbers or passwords, T-FA systems become highly vulnerable as has been shown.

We conclude that it is incorrect and also severely misleading to claim that T-FA does increase the recognition performance of an biometric authentication system. In practice, security may be enhanced since two factors are necessary to achieve successful authentication, yet it is essential that the recognition performance and the corresponding decision parameters remain the same as compared to the employment of the "pure" biometric system. In all of the presented systems claimed performance is achieved through the unpractical assumption that non-genuine users are not in possession of valid second factors. If performance evaluations are carried out like in the above presented systems the true performance of the underlying biometric system is concealed.

4.1 Requirements for Performance Evaluations of T-FA Systems

From the analysis of all the above presented approaches and our case study several requirements to performance evaluations regarding T-FA schemes including biometric authentication can be derived:

1. It is required that in experiments, especially when calculating inter-class distances, any type of personal or physical token has to be considered compromised. Focusing on biometric systems T-FA must not be interpreted as a way to increase the recognition performance of a system.
2. If any sort of helper data is introduced, especially in biometric cryptosystems, this helper data must be considered compromised during experimental results if this helper data is not dependent on biometric data only.
3. The scenario where potential imposters are in possession of additional second factors must not be ignored since physical or personal factors are easily

compromised in general (decision thresholds have to be set up according to this scenario).

4. The security provided by introducing second factors to biometric systems must not and cannot be measured in terms of FRR or FAR.

References

1. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Trans. on Circuits and Systems for Video Technology 14, 4–20 (2004)
2. Chien, H., Jan, J., Tseng, Y.: An efficient and practical solution to remote authentication: Smart card. Computer and Security 21, 372–375 (2002)
3. Jain, A.K., Flynn, P.J., Ross, A.A.: Handbook of Biometrics. Springer, Heidelberg (2008)
4. Jain, A.K., Nandakumar, K., Ross, A.A.: Handbook of Multibiometrics. Springer, Heidelberg (2006)
5. Schneier, B.: Two-factor authentication: too little, too late. ACM Commun. 48(4), 136 (2005)
6. Uludag, U., Pankanti, S., Prabhakar, S., Jain, A.K.: Biometric cryptosystems: issues and challenges. Proceedings of the IEEE 92(6), 948–960 (2004)
7. Teoh, A.B.J., Ngo, D.C.L., Goh, A.: Biohashing: two factor authentication featuring fingerprint data and tokenised random number. Pattern Recognition 37, 2245–2255 (2004)
8. Teoh, A.B.J., Ngo, D.C.L., Goh, A.: Biometric Hash: High-Confidence Face Recognition. IEEE Transactions on Circuits and Systems for Video Technology 16(6), 771–775 (2006)
9. Kong, A., Cheunga, K.H., Zhanga, D., Kamelb, M., Youa, J.: An analysis of BioHashing and its variants. Pattern Recognition 39, 1359–1368 (2006)
10. Klein, D.V.: Foiling the cracker: a survey of, and improvements to, password security. In: Proceedings of the 2nd USENIX Workshop Security, pp. 5–14 (1990)
11. Reddy, E., Babu, I.: Performance of Iris Based Hard Fuzzy Vault. IJCSNS International Journal of Computer Science and Network Security 8(1), 297–304 (2008)
12. Juels, A., Sudan, M.: A fuzzy vault scheme. In: Proc. 2002 IEEE International Symp. on Information Theory, p. 408 (2002)
13. Jassim, S., Al-Assam, H., Sellahewa, H.: Improving performance and security of biometrics using efficient and stable random projection techniques. In: Proceedings of the 6th International Symposium on Image and Signal Processing and Analysis, ISPA '09, pp. 556–561 (2009)
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. Cheung, K.H., Kong, A.Z.D., Kamel, M., You, J., Lam, T.H.W.: An Analysis on Accuracy of Cancelable Biometrics based on BioHashing. In: Khosla, R., Howlett, R.J., Jain, L.C. (eds.) KES 2005. LNCS (LNAI), vol. 3683, pp. 1168–1172. Springer, Heidelberg (2005)
16. 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)
17. Zuo, J., Ratha, N.K., Connel, J.H.: Cancelable iris biometric. In: Proceedings of the 19th International Conference on Pattern Recognition 2008 (ICPR'08), pp. 1–4 (2008)