BioHashing with Fingerprint Spectral Minutiae

Berkay Topcu[1,2], Hakan Erdogan[2], Cagatay Karabat[1], Berrin Yanikoglu[2]

[1] eID Program
TUBITAK - BILGEM - UEKAE
Gebze, Kocaeli, TURKEY 41470
{berkay.topcu, cagatay.karabat}@tubitak.gov.tr

[2] Sabancı University
Tuzla, Istanbul, TURKEY 34956
{haerdogan, berrin@sabanciuniv.edu}

Abstract: In recent years, the interest in human authentication has been increasing. Biometrics are one of the easy authentication schemes, however, security and privacy problems limit their widespread usage. Following the interest in privacy protecting biometric authentication, template protection schemes for biometric modalities has increased significantly in order to cope with security and privacy issues. BioHashing, which is based on transforming the biometric template using pseudo-random projections that are generated using a user-specified key or token, has received much attention as it improves verification accuracies over using only the biometric data, allows template revocation and preserves privacy. In our work, we develop a new BioHashing scheme for fingerprints. A fixed-length feature vector is required in order to design a BioHashing scheme. In the literature, most of the studies on fingerprint BioHashing uses features extracted from fingerprint texture. On the other hand, our new BioHashing scheme is based on minutia based feature vectors. We use the spectral minutiae representation for obtaining a fixed-length feature vector for a fingerprint sample. Then, we use a random projection matrix, which is generated from user's key/token, in order to generate a BioHash vector. We propose to randomly project each column of the spectral minutiae feature matrix via a single matrix which allows fast bit string extraction and adaptive quantization. Experiments on FVC2002 databases show the promise of the proposed system for fast and secure verification.

1 Introduction

Personal authentication using biometric systems that use physiological or behavioral characteristic are becoming widespread. Among various biometric modalities, fingerprint is preferred in many settings, due to its distinctiveness and performance, as well as the practicality and low cost of the fingerprint readers. Most fingerprint recognition systems are based on matching fingerprint minutiae which are the endpoints and bifurcations of fingerprint ridges. They are known to remain unchanged over an individual's lifetime and allow a very discriminative classification of fingerprints [MMJP09].

Increasing widespread use of fingerprint identification, as well as other biometric modal-

ities, raises privacy concerns significantly [JNN08] and protecting biometric fingerprint templates (or mostly minutiae templates) becomes a requirement. Combining fingerprint recognition with template protection puts two important constraints to a fingerprint recognition system [XV10]. No relative pose-alignment of two fingerprints is possible due to the encrypted storage and a fixed-length feature vector is required as input of template protection schemes such as fuzzy commitment and helper data schemes [TAK+05, Jue07].

Extracting fixed-length feature vector from fingerprints has been an interesting research topic in the last decade and several texture-based methods are proposed. In [JPHP00], Jain et al. presented FingerCode that is based on the fingerprint texture. Using a bank of Gabor filters to capture both local and global information, a compact fixed-length feature vector (FingerCode) is formed. The author concluded that FingerCodes are not as distinctive as minutiae. Similar to FingerCode, [TAK+05] suggested a quantization algorithm based on local orientation of ridges. Some other studies that follow the same idea (texture patterns of fingerprints) are [BCK08] and [CV11], and they require several samples per user at the enrollment to extract stable feature vectors.

Apart from texture-based approaches, several methods for fixed-length feature extraction from minutiae have also been investigated. Park et al. used histogram of the quantized distances between all possible minutiae in the ROI determined by the core point as a reference [PSBL05]. The method is very sensitive to minutiae insertions and deletions and minutiae orientations are not taken into account; therefore, the performance is not satisfying. [DKM+07] presented a statistical model of the relationship between the enrollment biometric and noisy biometric measurement taken during authentication and designed specific encoding and decoding algorithms to deal with displacement, erasure and insertion of minutiae using some stored public information about the biometric template (helper data). In [NRV10], numbers of minutiae in local cuboids are used for binary representation. However, in addition to minutiae information, ridge orientation map and ridge frequency is also used in this study and a preliminary registration step before comparison is required. In [BD10], a fingerprint is characterized by its similarity with a fixed number set of representative local minutiae vicinities. This approach by representative leads to a fixed length binary representation, and, as the approach is local, it enables to deal with local distortions that may occur between two acquisitions. Capelli et al. recently presented the Minutiae Cylinder-Code [CFM10] where a minutiae cylinder record the neighborhood information of a minutiae as a 3D function. A cylinder contains several layers and each layer represents the density of neighboring minutiae along the corresponding direction. Another study represents minutiae in spectral domain (Spectral Minutiae Representation [XV10]) and creates a fixed-length feature vector using log-polar transform.

In this work, we use the spectral minutiae representation to create a fixed-length feature vector, using minutiae location and direction information (x, y and θ) in ISO 19794-2 standard [ISO05]. The obtained fixed-length template for a fingerprint sample is then combined with pseudo-random data, generated from a user specified key or token used as a seed, to generate a unique code per person, using the BioHash scheme [JLG04]. Mixing of pseudo-random number and biometric data - BioHashing - leads to protection of the biometric template against biometric fabrication without possession of the corresponding token or knowledge of the randomization. Token-based randomization also enables

revocation of one's biometric template via token replacement.

This paper is organized as follows. First we review the spectral minutiae representation and give details of spectral minutiae representation in Section 2. Section 3 presents the basic BioHashing idea and our approach for spectral minutiae representation. In Section 4, we discuss the experimental results. Finally we draw our conclusion in Section 5.

2 Spectral Minutiae Representation

The spectral minutiae representation of a minutiae set is a fixed-length feature vector that is invariant to translation, rotation and scaling [XV10]. These characteristics enable the combination of fingerprint recognition systems with template protection schemes and allow for fast minutiae-based matching. The spectral minutiae representation can be applied on minutiae sets without any other requirement, therefore it is compatible with most of the existing fingerprint databases and minutiae-based fingerprint verification systems.

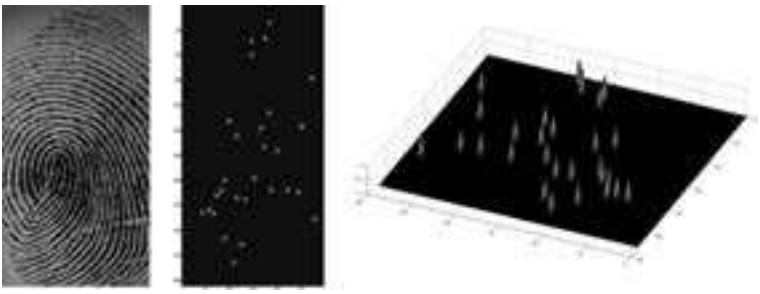

Figure 1: Minutiae locations and set of minutiae represented by Gaussian functions.

Complex spectral minutiae (SMC) is one of the three possible spectral minutiae representations, proposed by Xu et al. in order to obtain a fixed-length feature vector using minutiae location and orientation [XV10]. The other two alternatives are location based spectral minutiae (SML) and orientation based spectral minutiae (SMO). In SMC, each minutiae is represented by a Dirac pulse and in order to reduce the sensitivity to small variations in minutiae locations in the spatial domain, a Gaussian low-pass filter is used to attenuate the higher frequencies. This corresponds to a convolution in the spatial domain where every minutia is now represented by an isotropic two-dimensional Gaussian function with standard deviation σ_C. Minutiae locations on a fingerprint image together with the Gaussian functions are illustrated in Figure 1. The minutiae orientation is incorporated into this representation by assigning each Gaussian a complex amplitude $e^{j\theta_i}$, where θ_i is the orientation of the corresponding minutiae. For a set of Z minutiae with locations $(x_i, y_i)|_{i=1}^{Z}$, by evaluating the magnitude of the Fourier spectrum (1) on a polar-logarithmic grid, we

obtain a complex spectral representation $\mathbf{M_C}(w_x, w_y; \sigma_C^2)$:

$$\mathbf{M_C}(w_x, w_y; \sigma_C^2) = \left| \exp\left(-\frac{w_x^2 + w_y^2}{2\sigma_C^{-2}} \right) \sum_{i=1}^{Z} \exp(-j(w_x x_i + w_y y_i) + j\theta_i) \right| \quad (1)$$

where w_x and w_y are the spatial frequencies in the x and y directions.

The Fourier spectral magnitude is mapped onto a polar-logarithmic coordinate system as $\lambda = \sqrt{w_x^2 + w_y^2}$ and $\beta = \arctan(w_y/w_x)$ where λ corresponds to the radial direction and β corresponds to the angular direction. In the radial direction $M = 128$ samples are used between $\lambda_l = 0.05$ and $\lambda_h = 0.63$. In the angular direction $N = 256$ samples are used between $\beta = 0$ and $\beta = 2\pi$. The resulting complex spectral representation of a minutiae set is a 128×256 matrix.

3 BioHash for Protecting the SMC Template

BioHashing, applied to fingerprint biometric by Jin et al. [JLG04], is a two factor authentication approach that combines fingerprint feature with a user specified key/token and generates a unique compact code per person. A bit string from biometric data is created by inner product between the pseudo-random number sequence generated using the key as the seed and fixed-length fingerprint feature vector and deciding each bit on the sign of the result after subtracting a threshold.

A fixed-length biometric feature vector, $\mathbf{f} \in \mathbb{R}^d$ with length d, is reduced down to a bit string $\mathbf{b} \in \{0,1\}^p$, with p the length of the bit string ($p < d$), via a pseudo random pattern, $\mathbf{r} \in \mathbb{R}^p$ whose entries are uniformly distributed between -1 and 1. Details of this operation can be found in [JLG04].

In our study, to apply the BioHash scheme to complex spectral minutiae features, we reduce an $M \times N$ spectral fingerprint feature ($\mathbf{M_C}$) down to a bit string, $\mathbf{b} \in \{0,1\}^p$. Each column of $\mathbf{M_C}$ is a M-dimensional column vector. Randomly projecting each column of $\mathbf{M_C}$ to k dimensions and then thresholding the resulting vector, we obtain a k-length bit string. Mean value of the k-dimensional feature vector is used as the threshold for the quantization. We apply the same procedure to each column of $\mathbf{M_C}$ and concatenate the bit strings to create p-length bit string, where $p = k \times N$.

Each column $\mathbf{f_n}$ of this matrix is a 128-dimensional column vector and it is reduced to k dimensions by calculating its multiplication ($\mathbf{R} \cdot \mathbf{f_n}$) with the random projection matrix, $\mathbf{R}$ (which is a $k \times 128$ matrix). Thresholding the resulting k-dimensional feature vector by using its mean value as the threshold, we obtain a k-length bit string. The outputs of each column of $\mathbf{M_C}$ are then concatenated in order to create a bit string of length $k \times 256$. In this work, we evaluated different values of k to obtain a high verification accuracy with the smallest feature vector and used $k = 4$ resulting in a 1024-bit final feature vector (Figure 2).

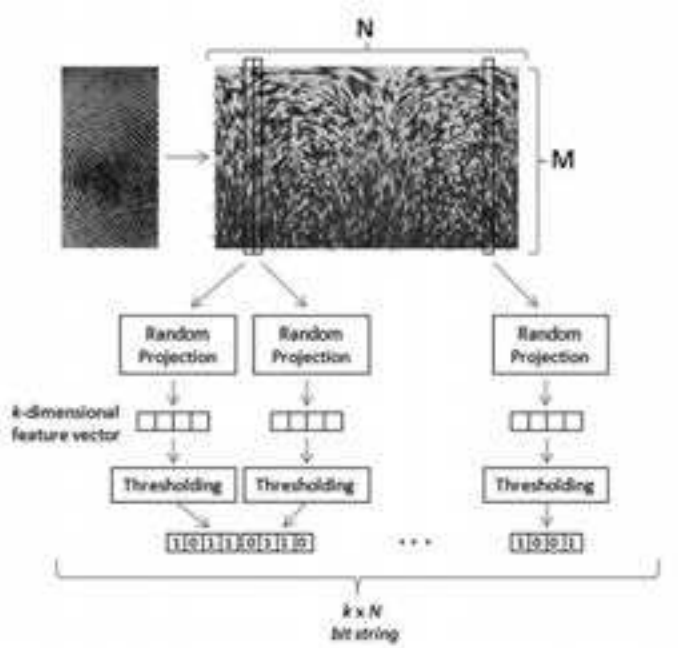

Figure 2: BioHashing procedure - from spectral representation to bit string.

4 Experimental Results

4.1 Experimental Settings

The proposed algorithm has been evaluated on publicly available FVC2002 fingerprint databases, namely DB1A, DB2A and DB3A [MMC+02]. DB1 and DB2 consist of fingerprint images captured with optical sensors and images in DB3 are captured with a capacitive fingerprint sensor. We have selected these databases in order to evaluate the performance of the proposed method for different image capturing technologies and left DB4 out in our experiments because it is a synthetic fingerprint database.

For the performance evaluation we adopt the equal error rate (EER). The minutiae sets are obtained by a commercial automatic minutiae extractor (Verifinger 4.4 SDK). We propose to use our algorithm in a high security scenario. In FVC2002 databases, some of the samples were obtained by requesting the users to provide fingerprint with exaggerated displacement and rotation. In a high security scenario where the user is aware that cooperation is crucial for security reasons, (s)he will be cooperative. Therefore, only four out of eight samples are chosen for each subject (1-2-6-7 for DB1 and DB3, 1-2-7-8 for DB2). Following the verification setting described in FVC competitions, we used all possible combinations for matching genuine pairs and the first sample of each subject is chosen for imposter matches (a total of 600 genuine and 4950 imposter matches for 100 subjects).

4.2 Results

We tested the BioHash of spectral minutiae representation on three databases. For comparison, we also included results from two other matching methods: i) matching two fingerprints based on the correlation of their complex spectral minutiae (SMC-Correlation) and ii) the minutiae-based commercial matcher which is also used for minutiae extraction. Equal error rates (EER) for both methods together with minutiae-based matching results on three databases are given in Table 1. As can be seen in this table, we obtain 0% EER for all the databases, when we apply BioHashing over the spectral minutiae features.

Table 1: EER on FVC2002 databases

	SMC-Correlation	**Minutiae Matching**	**SMC-BioHash**	**Stolen Key**
DB1	6.50%	0.50%	0.00%	14.77%
DB2	6.47%	0.83%	0.00%	13.10%
DB3	11.68%	2.50%	0.00%	26.46%

In addition, we evaluated the performance of the proposed scheme on a stolen key scenario, where an unauthorized imposter acquires the secret key/token of a genuine user but does not have the claimed person's fingerprint information. In this case, the imposter sends his/her fingerprint template and the secret key/token of the genuine user in order to be authenticated as the genuine user. This is a serious threat to the system as the pseudo-random vectors generated using the secret key has a considerable influence on the generated bit string, therefore on the matching score.

Assuming the key to be unknown at all times (never stolen) makes using the biometric unnecessary for real authentication scenarios. In order to analyze the effect of the key/token on the resulting bit strings, we have conducted experiments with stolen key scenario where an imposter attempt has the same secret key with the user that (s)he is intended to authenticate as (Table 1). While the error rates are considerably high in this case, they are in the same range as other results obtained with fingerprint BioHash implementations. For instance the straightforward BioHash implementation using FingerCode ([JPHP00]) reported in [LN06] achieves 15%, 15% and 27% EER on FVC2002 DB1-DB3 databases respectively for the stolen key scenario (see BASE row in Table 5 of the reference). Our error rates for this case are slightly better in the same scenario. The same authors report improved results (7%, 6.8% and 22% on FVC2002 DB1-DB3 respectively) with a classifier combination approach that aims to reduce the stability issues of BioHash, presumably with a system significantly slower and larger than ours [LN06].

5 Conclusion

In this study, we proposed a BioHashing approach for fingerprint identification based on minutiae information. Using the spectral minutiae representation of a fingerprint minutiae

set, we create a fixed-length bit string by randomly projecting spectral minutiae feature vectors. With this approach, one can obtain perfect separation between genuine and imposter population and the system provides 0% equal error rate, which is desired for all identity verification systems. In addition, in case the secret key of a valid user is stolen, the system allows acceptable error rates for imposter authentication attempts with a valid secret key. Also, biometric revocation becomes feasible through secret key (token) replacement, which addresses the cancellability issue.

Our main contribution is providing the first implementation of the BioHash scheme with the spectral minutiae representation. The proposed scheme is computationally fast as it only uses column-wise random projection of the spectral minutiae matrix, while achieving the 0% EER in the verification scenario. The original spectral minutiae features are 8096-dimensional (128×256) and in order to create a 1024-bits string, one needs to generate a random projection matrix of size 1024×8096. Instead, we propose to use a single 4×128 random projection matrix for multiplying with each column of SMC (which are 128-dimensional column vectors). This results in a computationally low random projection operation as well as adaptive thresholding for each column of SMC, instead of generating a larger projection matrix (which takes much time to generate as orthonormalization of vectors is required for higher number of vectors - 1024 instead of 4) and using a single threshold for quantization.

6 Acknowledgments

This work has been performed by the BEAT project 7th Framework Research Programme of the European Union (EU), grant agreement number: 284989. The authors would like to thank the EU for the financial support and the partners within the consortium for a fruitful collaboration. For more information about the BEAT consortium please visit http://www.beat-eu.org.

References

[BCK08] Julien Bringer, Hervé Chabanne, and Bruno Kindarji. The best of both worlds: Applying secure sketches to cancelable biometrics. *Sci. Comput. Program.*, 74(1-2):43–51, December 2008.

[BD10] J. Bringer and V. Despiegel. Binary feature vector fingerprint representation from minutiae vicinities. In *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS'2010)*, pages 1–6, 2010.

[CFM10] Raffaele Cappelli, Matteo Ferrara, and Davide Maltoni. Minutia Cylinder-Code: A New Representation and Matching Technique for Fingerprint Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(12):2128–2141, 2010.

[CV11] C. Chen and R. N. J. Veldhuis. Binary Biometric Representation through Pairwise Adaptive Phase Quantization. *EURASIP Journal on Information Security*, 2011:543106, March 2011.

[DKM+07] S.C. Draper, A. Khisti, E. Martinian, A. Vetro, and Jonathan S. Yedidia. Using Distributed Source Coding to Secure Fingerprint Biometrics. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, volume 2, pages II–129–II–132, 2007.

[ISO05] ISO/IEC 19794-2:2005. Biometric Data Interchange Formats - Part 2: Finger Minutiae Data, 2005.

[JLG04] Andrew Teoh Beng Jin, David Ngo Chek Ling, and Alwyn Goh. Biohashing: two factor authentication featuring fingerprint data and tokenised random number. *Pattern Recognition*, 37(11):2245–2255, 2004.

[JNN08] Anil K. Jain, Karthik Nandakumar, and Abhishek Nagar. Biometric template security. *EURASIP J. Adv. Signal Process*, 2008:113:1–113:17, January 2008.

[JPHP00] A. K. Jain, S. Prabhakar, L. Hong, and S. Pankanti. Filterbank-based fingerprint matching. *Trans. Img. Proc.*, 9(5):846–859, May 2000.

[Jue07] Ari Juels. Fuzzy Commitment. In Pim Tuyls, Boris Skoric, and Tom Kevenaar, editors, *Security with Noisy Data*, pages 45–56. Springer London, 2007.

[LN06] Ra Lumini and Loris Nanni. An improved BioHashing for human authentication. *Pattern Recognition*, 40:1057–1065, 2006.

[MMC+02] D. Maio, D. Maltoni, R. Cappelli, J.L. Wayman, and A. K. Jain. FVC2002: Second Fingerprint Verification Competition. In *In Proceedings of 16th International Conference on Pattern Recognition (ICPR2002), Quebec City*, pages 811–814, 2002.

[MMJP09] Davide Maltoni, Dario Maio, Anil K. Jain, and Salil Prabhakar. *Handbook of Fingerprint Recognition*. Springer Publishing Company, Incorporated, 2nd edition, 2009.

[NRV10] Abhishek Nagar, Shantanu Rane, and Anthony Vetro. Alignment and bit extraction for secure fingerprint biometrics. In Nasir D. Memon, Jana Dittmann, Adnan M. Alattar, and Edward J. Delp, editors, *Media Forensics and Security II, part of the IS&T-SPIE Electronic Imaging Symposium, San Jose, CA, USA, January 18-20, 2010, Proceedings*, volume 7541 of *SPIE Proceedings*, page 75410. SPIE, 2010.

[PSBL05] Chul-Hyun Park, Mark J. T. Smith, Mireille Boutin, and Joon-Jae Lee. Fingerprint matching using the distribution of the pairwise distances between minutiae. In *Proceedings of the 5th international conference on Audio- and Video-Based Biometric Person Authentication*, AVBPA'05, pages 693–701, Berlin, Heidelberg, 2005. Springer-Verlag.

[TAK+05] P. Tuyls, A. H. M. Akkermans, T. A. M. Kevenaar, G-J. Schrijen, A. M. Bazen, and R. N. J. Veldhuis. Practical Biometric Authentication with Template Protection. In T. Kanade, A. K. Jain, and N. K. Ratha, editors, *5th Int. Conf. on Audio- and Video-Based Personal Authentication (AVBPA), Rye Brook, New York*, volume LNCS 3546, pages 436–446, Heidelberg, July 2005. Springer-Verlag Berlin.

[XV10] Haiyun Xu and Raymond N.J. Veldhuis. Complex Spectral Minutiae Representation For Fingerprint Recognition. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPRW*, pages 1–8. IEEE Computer Society Press, June 2010.