Cancellability for Local Binary Pattern Biometric Authentication

Munalih Ahmad Syarif*, Leslie Ching Ow Tiong*, Alwyn Goh*, Latifah Mat Nen* and Kay Win Lee*
*MIMOS Berhad, Technology Park Malaysia, 57000 Kuala Lumpur, Malaysia
E-mail: {ahmad.syarif, tco.leslie, alwyn.goh, latifah.matnen, kw.lee}@mimos.my

***Abstract*—Biometric cancellability enables protection and revocation of sensitive biometric data as used for authentication. This paper describes a cancellable biometric implementation by means of randomised Local Binary Pattern (LBP) feature vector, with user-specific secret-keys used to generate a biometric template. We also present experimental results to establish that authentication undertaken by our methodology is reliable, and also that the objectives of revocability and diversity are accomplished.**

I. Introduction

Biometric authentication is applicable in security systems and processes supportive of physical access control, and also transactions of both over-the-counter and electronic variety. The single most important consideration against large-scale use of biometric authentication has always been in relation to the privacy and protection of user biometric data; as arising from the permanence (or at least long-term immutability) of such data, and consequent to that the impossibility of biometric revocation. This characteristic allows for malicious interception and fraudulent replay of user biometric data [1].

Biometric cancellability [2] addresses this structural vulnerability of biometric authentication. The basic concept is to generate a biometric template which can be revoked and/or replaced, per the equivalent actions in response to password compromise or token loss. Cancellable biometric schemes can be broadly classified as being based on biometric-salting or non-invertible transformations. Biometric-salting utilises random auxiliary data as an additional input, equivalent to the input of salt into password hashing. This enables the computation of different biometric templates from singular biometric datum, as would be necessary for a particular user to satisfy a plurality of authentication requirements.

Biometric-salting requires that auxiliary data be handled as secret and exclusive to the user of interest, as equivalent to passwords. The security of this auxiliary data would not necessarily need to be handled separately from that of the constituent biometric data, per the biometric-hashing in Goh and Ngo [3], which blends auxiliary and biometric data in the construction of a secure biometric template. This generic approach allows for security of biometric data in the output transform domain without the need for server-side storage of the constituent input data. Use of this auxiliary data by an imposter would however result in the loss of any advantage in biometric-hashing, with recognition performance reverting to that of the original feature vector. Such compromise of the auxiliary data is referred to as the stolen-token scenario.

Non-invertible transformation, on the other hand, maps the input biometric data into a context-sensitive feature representation, which would not need secrecy for security. Ratha et al. [4] described non-invertible transforms constructed from surface foldings in Cartesian and polar coordinates, resulting in an output representation which cannot be inverted for recovery of the original input.

This paper proposes a new approach for biometric-salting by means of randomised Local Binary Pattern (LBP) blocks, as arising from user-specific key input. Our approach is based on input of a set of random keys with which to paarameterise permutation and transformation of the input biometric data into an output template. The outcome of this process results in template cancellability, and also renders infeasible recovery of the original biometric data.

This paper is structured in sections as follows: (II) review of related works in cancellable biometrics; (III) presentation of our approach of using random LBP blocks from user-specific keys for generation of biometric templates; (IV) presentation of experimental data and analysis; and (V) conclusions arising and outline of future work.

II. Related Work

The major challenge in designing biometric template protection that fulfils the following properties:

- Diversity: which prevents cross-matching of the same template across databases.
- Revocability: which enables cancellation of existing templates, and replacement thereof with new templates based on the same biometric data.
- Security: which stipulates that recovery of the original biometric data from biometric template is computationally infeasible.
- Performance: which stipulates retention of recognition performance after application of protection measures.

as set out in Teoh and Lim [5]. The basic concept is that cancellable biometric resembles passwords or user-specific random information, in that secret-keys are blended with biometric data resulting in biometric templates satisfying the above properties.

Several methods have been used to effect of cancellable templates broadly similar to biometric-hashing. Teoh et al. [6] described Random Multispace Quantisation (RMQ) of feature vectors, so as to generate biometric-hash outputs from inputs

of secret passwords and/or unique physical tokens. The results therein indicate that RMQ-derived biometric-hashes can be used in a manner equivalent to cryptographic keys, and that such use would not degrade recognition performance. Lumini and Nanni [7] proposed a modification of basic biometric-hashing via employment of threshold variations on the bit-extraction process, and also expansion the vector space by use of multiple random sequences and feature-vector permutations; resulting in more superior recognition performance under stolen-token conditions.

Chang et al. [8] introduced an alternative approach of stable key generation, resulting in stabilised outputs of user-specific keys. Their framework utilises user-dependent transforms to generate a larger set of distinguishable features, resulting in a longer and more stable bitstream.

Chikkerur et al. [9] proposed a representation based on localised patches with which to encode spatial-domain fingerprint data by means of a key-tuple with two independent elements, the matrix product of which results in the biometric-salting transformation. We use this notion of a key-tuple with multiple elements in our presentation, to the same net effect of better security against key compromise.

Nandakumar and Jain [10] studied the protection of individual templates with multiple passwords, and also the resultant security vulnerabilities. Their work introduced the use of a cryptographic fuzzy-vault to encode a biometric template with a high degree of stability, and thereafter for this encoded secret to be used in biometric-salting. The presentation therein was of iris data as the relatively stable vault-encoded salt, and fingerprint data as the typically less stable biometric input into subsequent transformation. The security of their approach stems from the irreversibility of the fuzzy-vault, the computational infeasibility of which can be quantified by means of a cryptographic analysis.

Bai and Hatzinakos [11] introduced a LBP-based biometric-hash scheme that generates discretisation outputs from an inner product sequence of random matrices and LBP feature vectors. Our work is also based on LBP description, but with the discretisation process also ”naturally” extracted out of the internal workings of the LBP process, as opposed an external biometric-hash stage.

Our previous work [12] proposed feature-level discretisation based on Most Intensive Block Locations (MIBL) from the biometric data. A MIBL would contain information from the location index of the image blocks, which is then amenable to filtering and discretisation by means of key-specific biometric-hashing. This present work reuse the concept of block-based feature extraction. We also managed to do zero knowledge (ZK) encoding on biometric data in our previous work [13]. The encoding method discretise biometric data into vector representation. Moreover, the previous work also included client-side masking of biometric data, as protective measure against leakage of biometric data on server-side storage, and additionally client-side encoding and corresponding serverside decoding, as protective measure against interception and/or leakage of biometric data in transit from client-to-server [13].

III. Proposed Method

In general, most of the existing cancellability approaches require server-side storage of the user-specific secret-keys, resulting in imposition of potentially burdensome secrecy and security requirements. The proposed method obviates these practical complexities and potential difficulties by eliminating the necessity for key storage. This is accomplished via the binding of particular keys to the user of interest, which only requires user presentation of keys (k_1, k_2) during enrolment and authentication, as shown in Fig. 1. The design objective here is to prevent attacker from obtaining the original biometric data via ”back-door” key recovery.

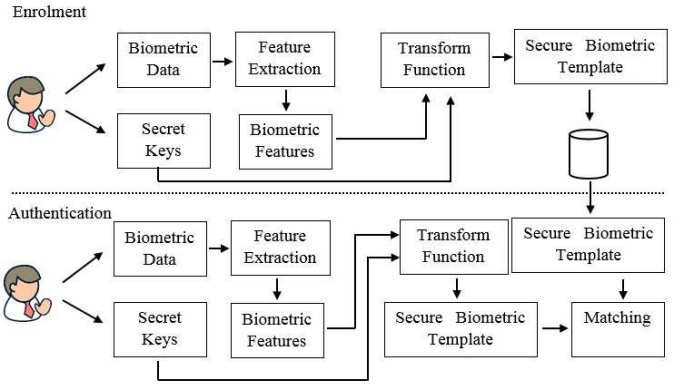

Fig. 1. Proposed Approach

Fig. 2 illustrates the proposed cancellability method, as based on LBP randomisation via secret-key input. Our method requires the use of key-parameters k_1 and k_2; with the former for randomisation of histogram bin establishment and assignment within a particular image block, and the latter for randomisation of the image block sequence. The outcome is a biometric template based on a randomised LBP descriptor, which can be revoked and replaced when necessary.

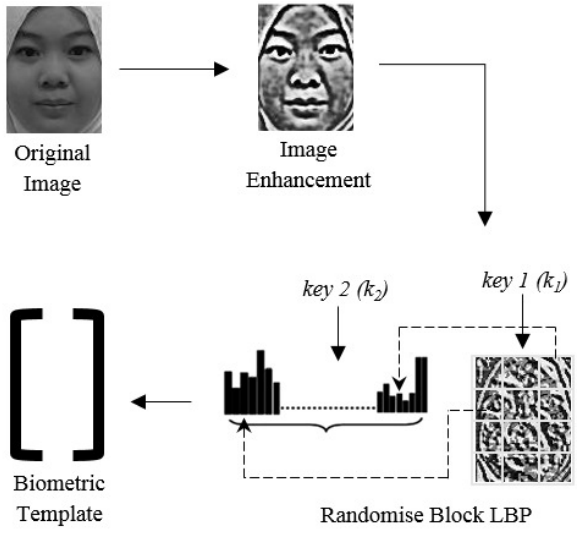

Fig. 2. Internal Operations of Proposed Approach

Application of our method is preceded by the enhancement method proposed by Tan and Triggs [14]. This method presented therein is a strong solution for various problems arising from illumination, as demonstrated by its effect on analysis of the Yale-B database [15].

LBP analysis is then undertaken on the enhanced image. The standard prescription is to compute pixel-level corrections based on each local cluster of 3×3 neighbourhood pixels by thresholding 8 surrounding pixels with the respect to the centre pixel, and then representing the result in binary number. Fig. 3 illustrates the LBP process as reference [16]. In the normal LBP process, the image is divided into *m*-number of blocks, and then computed into a set of histogram bins independently for each sub region of the image. Then, the histogram of each block is concatenated sequentially in order to form a feature vector.

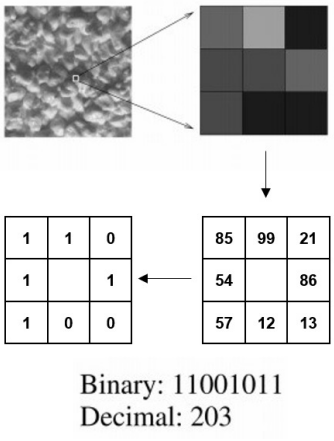

Fig. 3. Illustration of LBP Method [16]

However in this proposed method, the number of histogram bins generated from each image block are different based on k_1. Where k_1 is a set of different number which shows number of histogram bin generated from each image block. A variation of *n*-number of histogram bin can be set manually as long as it satisfies the following requirements:

1) *n* is positive integer
2) *m* mod *n* = 0 and
3) *n* >0

n is used to generate a population (*P*) which consists of a different number of histogram bins in the range of 2^2, 2^3, ..., 2^{n+1}, can be formulated as:

$$P = [2^2, 2^3, ..., 2^{n+1}, 2^2, 2^3, 2^{n+1}, ..., (r\ times)] \quad (1)$$

where *r* = *m* / *n*. At last k_1 is generated based on random permutation of *P*.

After k_1 is generated, the histogram bins for each block are computed and normalized respectively based on k_1. The normalised histogram bins from each block will be concatenated into single dimension of feature vector randomly based on k_2, where k_2 is a set of random numbers which is generated based on random permutation algorithm. The algorithm shuffles the sequence of the blocks *1, 2, ... , n* randomly and uniquely for each user. The biometric template is formed by concatenating histogram bins from each block randomly. With this proposed method, more than one biometric template is able to generate from same biometric data by providing different k_1 and/or k_2.

IV. Experimental Results

We characterise the proposed method using the following databases: 'DB-1' Extended Yale-B Cropped Images [15], and 'DB-2' images from our corporate database. DB-1 consists of 38 different persons; with each person having 60 sample images of 192×168 pixel dimension, as captured under different illumination. DB-2 consists of 12 different persons, with each person having 20 sample images of 96×72 pixel dimension.

Sample images of DB-1 and DB-2 are shown in Figs. 4 and 5 respectively. DB-1 is illustrative of face recognition undertaken under difficult operational conditions; while DB-2 is representative of much more amenable conditions.

Fig. 4. Sample Images from DB-1

Fig. 5. Sample Images from DB-2

Our experiments are organised in three parts, looking in detail at: (I) performance comparison between original LBP and the proposed approach; (II) analysis on revocability and diversity of the proposed approach; and (III) analysis on security.

A. Performance Comparison between LBP and Proposed Method

In this experiment, both LBP and proposed method are applied on DB-1 and DB-2. Recognition performance is then compared in terms of Equal Error Rate (EER). The number of blocks (*m*) is set to 64 for both LBP and our method. The number of histogram bins in each blocks is set to 256 for LBP; and four different quantities of histogram bins are randomly set for the proposed approach. Verification is then undertaken by means of Euclidean distance.

Table I shows that proposed method of user-specific LBP is able to outperform basic LBP for both DB-1 and DB-2, with the former attaining near-zero EERs. Basic LBP, in contrast, obtains an unsatisfactory 47.2% EER for the challenging DB-1, and 1.76% for the much more amenable DB-2.

TABLE I
PERFORMANCE COMPARISON BETWEEN LBP AND PROPOSED METHOD

Database	Extraction Method	EER (%)
DB-1	LBP	47.2
	Proposed Approach	1.38×10^{-4}
DB-2	LBP	1.76
	Proposed Approach	3.89×10^{-12}

Figs. 6 and 7 show the distribution of genuine and imposter distributions arising from basic and personalised LBP on both databases. It is quite clear that LBP in of itself does not result in reasonable separation of genuine and imposter classifications, resulting in poor recognition performance. LBP personalisation, in contrast, enables clear separation of genuine and imposter classifications, with a near-zero EER.

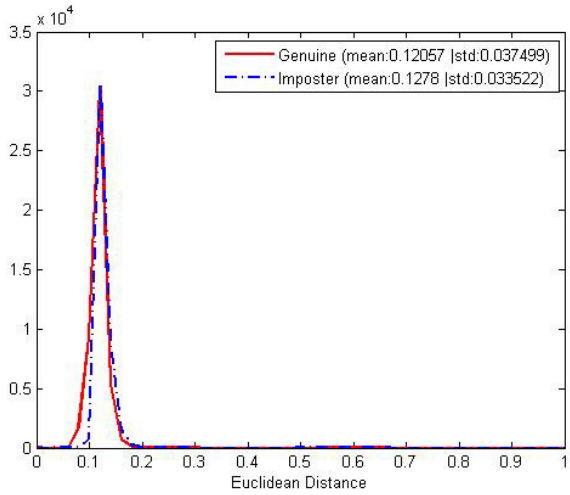

Fig. 6. Genuine and Impostor Score Distribution generated by LBP (DB-1)

Figs. 8 and 9 show the outcomes of the same comparative analysis undertaken on DB-2. Textbook LBP able to perform better in this experiment on DB-2, as indicated by the relatively small classification overlap between the genuine and imposter distributions. This better performance on the more verification-friendly database is likewise illustrated by the Receiver Operating Characteristics (ROC), per Figs. 10 and 11, of both recognition methods applied on DB-1 and DB-2. We are able to conclude that LBP personalisation improves recognition performance.

B. Revocability and Diversity Analysis

Biometric revocability and diversity requires the capability to generate different biometric templates from the same user

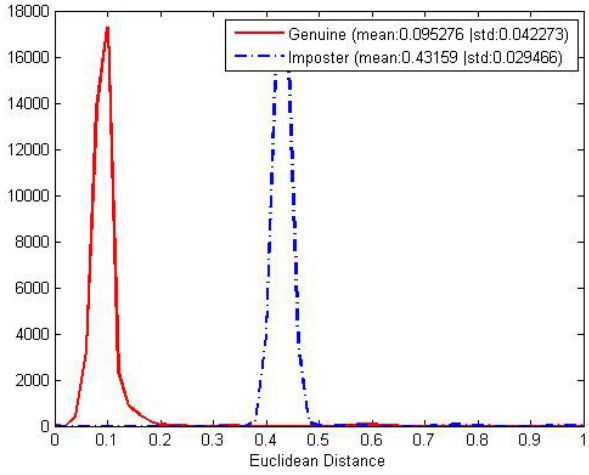

Fig. 7. Genuine and Impostor Score Distribution generated by proposed approach (DB-1)

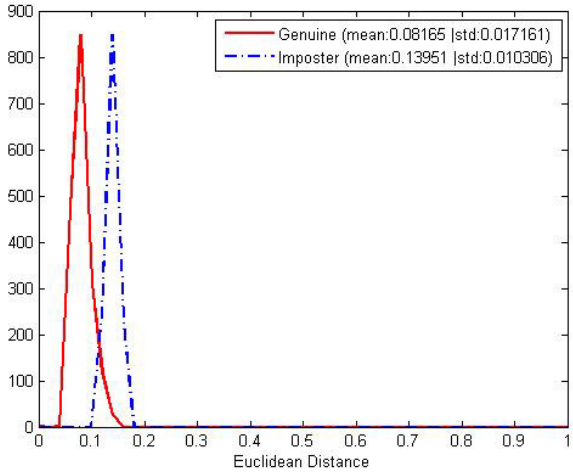

Fig. 8. Genuine and Impostor Score Distribution generated by LBP (DB-2)

biometric data. Biometric templates can then be revoked and re-registered in the event of compromise. The other important specification is for a plurality of biometric templates, each one specific to a particular service provider undertaking verification, such that cross-matching is infeasible.

Our method addresses revocability and diversity via use of different secret-keys to generate multiple LBP personalisations from the same user biometric (face) data. This allows for generation of a pseudo-impostor distribution, from distance measurements between the genuine (applicable) and pseudo-imposter (inapplicable) personalisations. Revocability and diversity can therefore be assessed by means of EER, in which a pseudo-imposter is to be regarded as equivalent to a ”regular” imposter in terms of desirable classification outcomes. This would be indicated by separation of the genuine and pseudo-imposter distributions, as equivalent to separation of the gen-

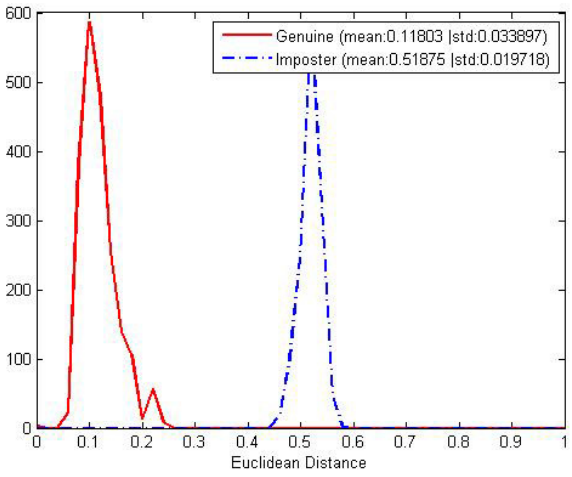

Fig. 9. Genuine and Impostor Score Distribution generated by proposed approach (DB-2)

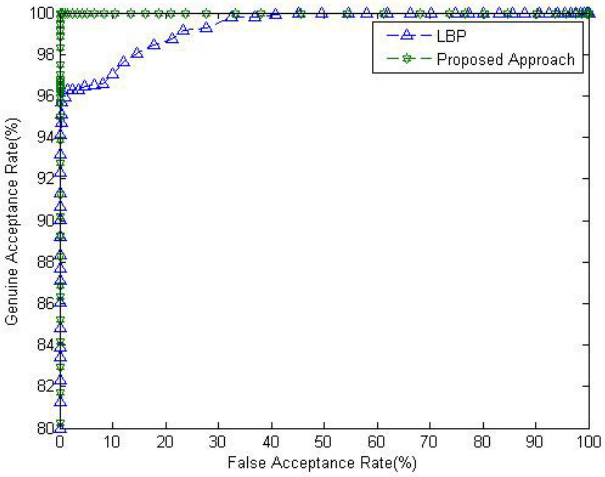

Fig. 11. ROC Curve of Experiment for DB-2.

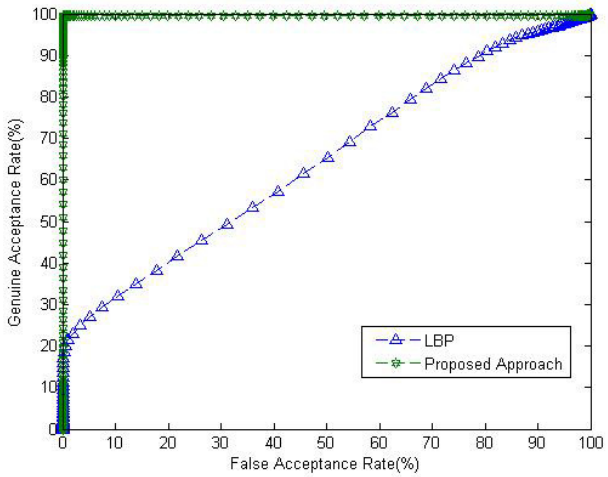

Fig. 10. ROC Curve of Experiment for DB-1.

uine and imposter distributions.

In this experiment, we generate 100 different secret-keys per user, resulting in that number of personalisations. This produces 100 pseudo-imposter measurements per user. Over the entire database, this yields 228,000 ($38\times60\times100$) measurements for DB-1, and 24,000 ($12\times20\times100$) measurements for DB-2.

Table II presents the EERs obtained from the pseudo-impostor verification process. The results confirm recognition performance comparable to "normal" genuine vs imposter verification. Table III also demonstrates the basic similarity of pseudo-imposter and imposter distributions, as characterised by mean (μ) and standard deviation (σ), for both DB-1 and DB-2. Figs. 12 and 13 is further illustrative of pseudo-imposter and imposter equivalence in terms of the high degree, to the point to near co-incidence, of the distribution overlaps.

In conclusion, use of different secret-keys results in biometric template personalisations that are qualitatively different with respect to one another, even if generated from the same biometric data. Different personalisations arising from the same user is indistinguishable from personalisations arising from different users. We can therefore conclude that our approach makes possible template revocability and diversity.

TABLE II
COMPARISON OF NORMAL AND PSEUDO-IMPOSTER VERIFICATIONS

Database	EER for Normal Verification (%)	EER for Pseudo Imposter Verification (%)
DB-1	1.38×10^{-4}	1.02×10^{-7}
DB-2	3.89×10^{-12}	4.7×10^{-11}

TABLE III
COMPARISON OF CHARACTERISTICS OF DIFFERENT DISTRIBUTIONS

Database	Genuine Distribution		Imposter Distribution		Pseudo-imposter Distribution	
	μ[a]	σ[b]	μ[a]	σ[b]	μ[a]	σ[b]
DB-1	9.5×10^{-4}	4.2×10^{-4}	0.0043	2.9×10^{-4}	0.0045	1.8×10^{-4}
DB-2	0.0012	0.0003	0.0052	0.0002	0.0053	0.0002

[a] μ: Mean
[b] σ: Standard Deviation

C. Security Analysis

We present experiments to analyse the following stolen-key scenarios, in which k_1 and k_2 are individually compromised. The objective here is to demonstrate that our approach is able to retain good recognition characteristics even if one of the personalisation keys is compromised. Table IV shows the

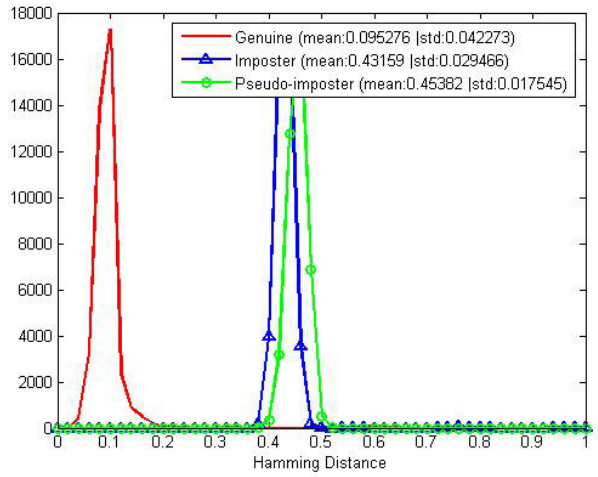

Fig. 12. Genuine, Impostor and Pseudo-impostor Distributions arising from Proposed Approach (DB-1)

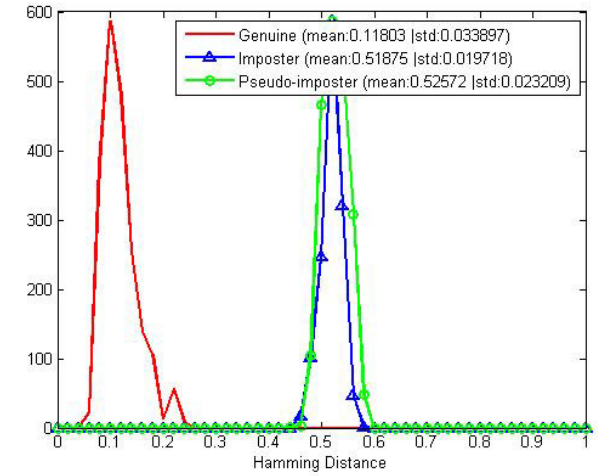

Fig. 13. Genuine, Impostor and Pseudo-impostor Distributions arising from Proposed Approach (DB-2)

performance under the single stolen key scenario for both DB-1 and DB-2. It is hence demonstrated that the proposed scheme continues to undertake recognition at near-zero EER.

This can also be seen for the genuine and impostor distributions under these conditions for DB-1, as in Figs. 14 and 15; and for DB-2, as in Figs. 16 and 17. The distributions here are qualitatively similar to those obtained under the normal genuine vs imposter analysis without consideration of key compromise. We can therefore conclude that the proposed method is secure against compromise of any single personalisation key.

The proposed scheme would not retain these advantages if both personalisation keys k_1 and k_2 are compromised. The consequence of that theft would be a major degradation in the recognition performance. Our experiments indicate a reduction in EER to 45.4% for DB1, and 1.26% for DB2; which can be

TABLE IV
PERFORMANCE UNDER SINGLE STOLEN KEY SCENARIO

Database	Condition	EER (%)
DB-1	k_1 Stolen	1.16×10^{-4}
	k_2 Stolen	2.09×10^{-4}
DB-2	k_1 Stolen	2.16×10^{-13}
	k_2 Stolen	5.34×10^{-12}

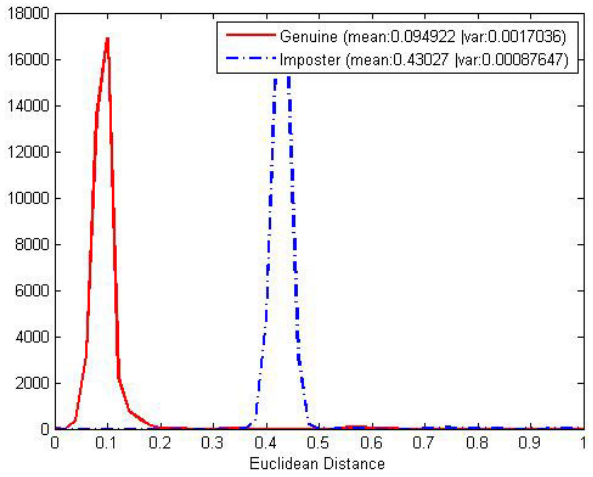

Fig. 14. Genuine, Impostor and Pseudo-Impostor Distribution of Proposed Approach under Stolen k_1 Scenario for DB-1

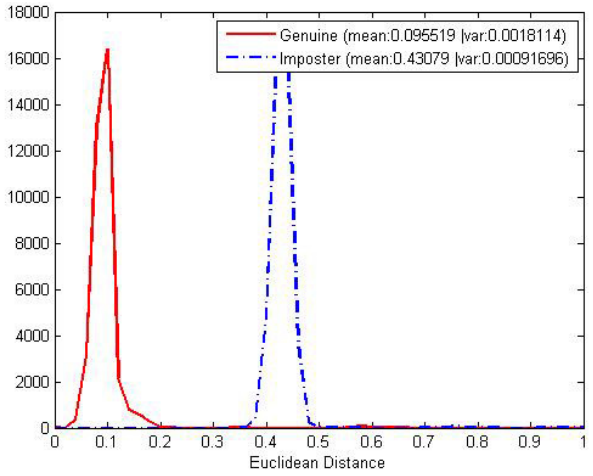

Fig. 15. Genuine, Impostor and Pseudo-Impostor Distribution of Proposed Approach under Stolen k_2 Scenario for DB-1

interpreted as a reversion to the performance of basic LBP.

Dual key compromise is avoidable in practical operational scenarios under specification that k_1 and k_2 are computed independently of one another. For instance, one element of the key-tuple could be generated from password inputs, as presumed secret and exclusive to particular user; and the other

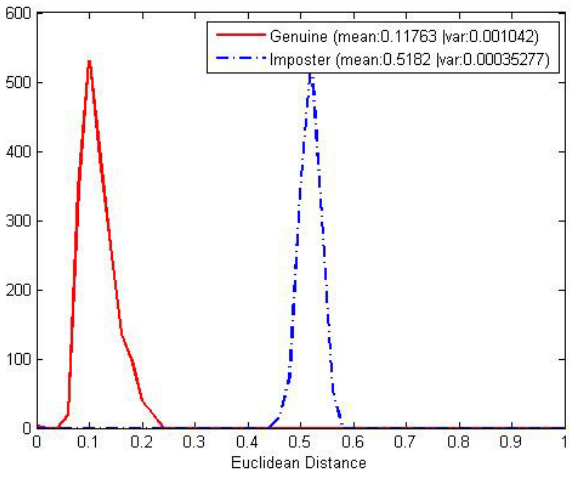

Fig. 16. Genuine, Impostor and Pseudo-Impostor Distribution of Proposed Approach under Stolen k_1 Scenario for DB-2

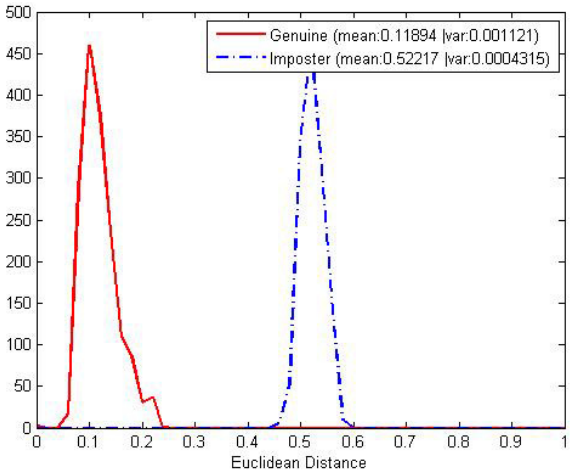

Fig. 17. Genuine, Impostor and Pseudo-Impostor Distribution of Proposed Approach under Stolen k_2 Scenario for DB-2

from hardware (or software) tokens, as presumed unique and in exclusive possession of that same user.

V. Conclusions

This paper demonstrates biometric cancellability by means of LBP personalisation via user-specific secret-keys. This is accomplished by means of key-specific permutation of the LBP internal structure of constituent blocks, and additionally the key-specific permutation of the histogram bins constituent to each block. Our prescription results in near-zero EER recognition performance on face image datasets illustrative of a broad range of environmental conditions.

LBP personalisation is also capable of generating multiple biometric templates from the same biometric data associated with a single user. This capability enables straightforward template cancellation and replacement, which is furthermore secure against cross-matching against templates originating from the same user.

In the future we will be investigating the key presentation within the context of an interactive challenge-response sequence. This interaction might also serve as an indication of user liveness.

References

[1] AK Jain, K Nandakumar, and A Nagar, "Biometric template security," *EURASIP Journal on Advances in Signal Processing,* pp 1-17, 2008.

[2] C Rathged and A Uhi, "A survey on biometric crytosystems and cancelable biometrics," *EURASIP Journal on Information Security,* no. 3, pp 1-25, 2011.

[3] A Goh and DCL Ngo, "Computation of Cryptographic Keys from Face Biometrics," *International Federation for Information Processing (IFIP) Technical Communications and Multimedia Security (CMS),* pp 1-13, 2003.

[4] NK Ratha, S Chickkerur, CH Jonathan and RM Bolle, "Biohashing: two factor authentication featuring fingerprint data and tokenised random number," *Pattern Recognition,* vol 37, no 11, pp 2245-2255, 2004.

[5] ABJ Teoh and MH Lim, "Cancelable Biometrics," *Scholarpedia,* vol. 5, no. 1, pp. 9201, 2010.

[6] ABJ Teoh, A Goh and DCL Ngo, "Random multispace quantization as an analytic mechanism for biohashing of biometric and random identity inputs," *IEEE Trans Pattern Anal Mach Intell,* vol 28, no 12, pp 1892-1901, 2006.

[7] A Lumini and L Nanni, "An improved biohashing for human authentication," *Pattern Recognition,* vol 40, no 3, pp 1057-1065, 2007.

[8] YJ Chang, W Zhang and T Chen, "Biometrics-based cryptographic key generation," in *IEEE Int Conf on Multimedia and Expo,* Taipei, Taiwan, 2004.

[9] S Chikkerur, N Ratha, J Connell and R Bolle, "Generating registration-free cancelable fingerprint templates," in *IEEE Int Conf on Biometrics: Theory, Applications and Systems,* Arlington, Virginia, 2008.

[10] K Nandakumar and A Jain, "Multibiometric template security using fuzzy vault," in *IEEE Int Conf on Biometrics: Theory, Applications and Systems,* Arlington, Virginia, 2008.

[11] Z Bai and D Hatzinakos, "LBP-based biometric hashing scheme for human authentication," in *Int Conf Control, Automation, Robotics and Vision,* Singapore, 2010.

[12] MA Syarif, TS Ong, ABJ Teoh and C Tee, "Improved biohashing method based on most intensive histogram block location," in *Neural Information Processing,* Springer, pp 664-652, 2014.

[13] A Goh, DCL Ngo, KS Ng, KW Lee and L Mat Nen, "Zero knowledge processing on biometric data in discretised vector representation," Malaysia Patent 2014/2934, 2014.

[14] X Tan and B Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Transactions on Image Processing,* vol 19, pp 1635-1650, 2010.

[15] KC Lee, J Ho and DJ Kriegman, "Acquiring linear Subspaces for face recognition under variable lighting," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol 27, pp 684-698, 2005.

[16] T Ahonen, A Hadid and M Pietikäinen, "Face description with local binary patterns: application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol 28, pp 2037-2041, 2006.