2018 24th International Conference on Pattern Recognition (ICPR)
Beijing, China, August 20-24, 2018

Cancelable Biometrics Using Noise Embedding

Dae-Hyun Lee
Institute of New Media and
Communications (INMC)
Seoul National University
Seoul, South Korea
Email: dhlee@ispl.snu.ac.kr

Sang Hwa Lee
Institute of New Media and
Communications (INMC)
Seoul National University
Seoul, South Korea
Email: lsh529@snu.ac.kr

Nam Ik Cho
Institute of New Media and
Communications (INMC)
Seoul National University
Seoul, South Korea
Email: nicho@snu.ac.kr

***Abstract*—This paper presents a cancelable biometric (CB) scheme for iris recognition system. The CB approaches are roughly classified into two categories depending on whether the method stresses more on non-invertibility or on discriminability. The former is to use non-invertibly transformed data for the recognition instead of the original, so that the impostors cannot retrieve the original biometric information from the stolen data. The latter is to use a salting method that mixes random signals generated by user-specific keys so that the imposters cannot retrieve the original data without the key. The proposed CB can be considered a combination of these methods, which applies a non-invertible transform to the salted data for binary biocode input. We use the reduced random permutation and binary salting (RRP-BS) method as the biometric salting, and use the Hadamard product for enhancing the non-invertibility of salted data. Moreover, we generate several templates for an input, and define non-coherent and coherent matching regions among these templates. We show that salting the non-coherent matching regions is less influential on the overall performance. Specifically, embedding the noise in this region does not affect the performance, while making the data difficult to be inverted to the original.**

I. Introduction

If biometric data are stolen by impostors, then it would raise many serious problems because we cannot change our biometric information. Thus it is desirable that a biometric authentication system does not keep the original data, and moreover does not use the original data for authentication. Instead, the system keeps only non-invertibly transformed or salted data, and uses these data for the recognition. When it is found that these data are stolen, then we can replace the data by using another transform or salting methods, though we need to ask the users to capture the data again. Also, it is desired that the impostors cannot recover the original information from these transformed and/or salted data. This scheme is called cancelable biometric (CB) system.

The biometric template protection (BTP) systems have attracted attention for its efficiency [2], [3], [4], which define some important conditions that CB should satisfy:

- Unlinkability/Diversity: The same cancelable template should not be used in two different applications.
- Reusability/Renewability: If a template or key is compromised, it should be revoked and reproduced readily to generate a new template with a replaced key.
- Irreversibility/Non-invertibility: Complete biometric data should not be reconstructed from the current template, or the reconstruction process should be computationally infeasible.
- Accuracy/Performance: matching score between the templates should not be severely degraded even if compromise event occurs.

The conventional CB approaches attempt to satisfy above stated conditions by designing a non-invertible transform and using the transformed data, or by salting the data. The principle of transform design or salting method is to keep the discriminability of templates while minimizing the possibility of recovering the original template from the transformed and/or salted data. In this paper, we propose a new CB method that combines the advantages of biometric salting with non-invertible transform methods. For biometric salting, we use performance-oriented reduced random permutation and binary salting (RRP-BS) technique, which satisfies unlinkability. Also, Hadamard product and noise embedding method are introduced, which are designed to satisfy the non-invertibility. The noise embedding is applied to non-coherent matching regions obtained from some enrollment templates, so that it is robust against known threats.

II. Related Works

There are two main approaches to enabling the non-invertibility. One is to apply a non-invertible transform to the original data and the other is to insert random data, i.e., biometric salting.

A. Non-Invertible Transform Approaches

The main objective of these approaches is to design a non-invertible transform that ensures that the original biometric cannot be reconstructed from the transformed data. These methods usually make the template data partially unavailable or make it computationally difficult to retrieve the original data. Since most of these algorithms are focused on the non-invertibility, the discriminating capability of the transformed template becomes lower than the original.

For some specific examples of non-invertible transform approaches, Ratha *et al.* proposed three non-invertible transforms for the cancelable fingerprint authentication such as polar, Cartesian and surface folding [5]. Zuo *et al.* proposed a

978-1-5386-3788-3/18/$31.00 ©2018 IEEE

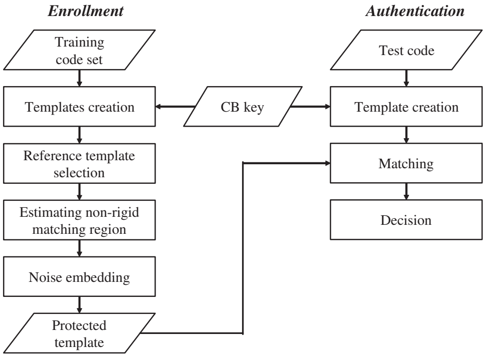


Fig. 1. Proposed CB system.

method based on the row shift and combinations to generate iris templates [6]. Hämmerle-Uhl *et al.* proposed a block remapping and image warping method for CB [7], where the re-mapping process eliminates some blocks by overlapping. The bloom filter was also introduced to the CB systems by Gomez-Barrero *et al.* [8]. It creates the filter using decimal number indexing. The indexing-first-one (IFO) hashing was proposed by Lai *et al.* [9]. This algorithm is based on min-hashing which estimates how similar the two sets are. if the hash value is repeatedly obtained.

B. Biometric Salting Approaches

Biometric salting means that the biometric data are blended with auxiliary data such as user specific random keys. The independence of keys ensures the discriminability and also satisfies the unlinkability. These approaches use user-specific keys which are revocable and regenerated when compromised.

The main research in biometric salting is focused on biohashing and its variation. Jin *et al.* introduced biohashing to generate secure biocode template [10]. Pillai *et al.* introduced a sectored random projection (SRP) method, where they handled different qualities in each partial iris region [11]. This method divides the iris into sectored regions and applies random projections separately to each sector followed by concatenating Zuo *et al.* proposed basic salting algorithms such as GRAY-SALT and BIN-SALT (BS) that blend random matrix with biocodes [6]. More detailed and various reviews of CB are introduced in many literatures [12], [13], [14].

III. Proposed CB System

The overall schematic diagram of the proposed algorithm is shown in Figure 1. We explain the main steps of the algorithm in this section.

A. Template Creation

The authentication system takes several images (training data), and generates a set of codes $\mathcal{X} = \{X_1, X_2, \cdots, X_{N_E}\}$ of $M \times N$ matrix $X_i \in \{0,1\}^{M \times N}$ where M and N are respectively the number of rows and columns and N_E be the number of training images per class. Then each code X_i is converted to a template Z_i through the three processes: RRP-BS, Hadamard product, and decimal encoding as described in the figure. The RRP-BS is adopted for discriminability, Hadamard product is used for non-invertibility, and the decimal encoding is performed for generating a row template and block matching of binary codes. We repeat three processes r times to vertically aggregate row templates (W_{ir}). Since the discriminability is proportional to the size of template, larger r better ensures the discriminability. Also, we use several training codes to generate the non-coherent matching map, which will be explained in the following subsection.

To be more precise with the first process, i.e., RRP-BS, we have a set of binary codes $\mathcal{X}$ for a class. For each X_i, we repeat the RRP-BS code generation $r \times q$ times, i.e., we obtain the RRP-BS codes

$$Y_{ijk} = S_{ijk} \oplus P_{ijk} X_i, \tag{1}$$

for $j = 1, \cdots, r$, and $k = 1, \cdots, q$, where r is the height of a template, q is the order of the Hadamard product, S is an $h \times N$ random key matrix with binary elements where h is a row control parameter that decides to keep some rows and to drop the others, P is an $h \times M$ permutation matrix with $h < M$, and the operator $\oplus$ is the elementwise logical exclusive-OR. In this process, notice that S_{ijk} and P_{ijk} act as CB keys and are generated by a personal identification number (PIN) of each class as shown in Figure 1. Since the randomly generated keys are not correlated with each other, the diversity is satisfied *i.e.,* we can use different keys for other devices or systems. In addition, the keys can be easily replaced, satisfying the renewability condition. And RRP-BS also improves the security by changing the order of rows of X and removing some of them. Specifically, only partial original information can be recovered even when the impostors stole the salting parameters.

The second process is the Hadamard product of order q, which is the logical AND operation of RRP-BS codes as

$$\tilde{Y}_{ij} = \prod_{k=1}^{q} Y_{ijk}. \tag{2}$$

In general, each element of X_i is a binary random variable that is 0 or 1 with equal probability, and so is the element of Y_{ijk}. Applying the Hadamard product reduces the number of 1's to $hN(1/2)^q$ and thus enhances non-invertibility.

The third process is to convert the binary vector to a decimal number. The result of (2), can also be expressed as

$$\tilde{Y}_{ij} = [\tilde{y}_{ij1} \cdots \tilde{y}_{ijN}] \tag{3}$$

where $\tilde{y}_{ijk}$ is an $h \times 1$ binary vector. We define a mapping function $f : \{0,1\}^{h \times 1} \to \mathbb{N}_0$. Then a $h \times N$ matrix $\tilde{Y}_{ij}$ converts into a $1 \times N$ row vector W_{ij} with decimal elements expressed as

$$W_{ij} = [f(\tilde{y}_{ij1}) \cdots f(\tilde{y}_{ijN})]. \tag{4}$$

The above steps are repeated r times and we obtain a final template Z_i of $r \times N$ matrix for X_i as follows

$$Z_i = [W_{i1}^T \cdots W_{ir}^T]^T. \tag{5}$$

Then, for each of the enrolled codes $\mathcal{X}$, we have the protected template set $\mathcal{Z} = \{Z_1, \ldots, Z_{N_E}\}$.

B. Reference Template Selection

From the set of templates $\mathcal{Z}$ that represents a class, we find a region that all the templates have similar properties, which is called coherent region. To be precise, the templates are from several images of a biometric of a person (class) which are individually different. But these also have some common regions which correspond to the coherent region of the templates for the robust biometric recognition. For obtaining the coherent region, we first define a reference template $\tilde{Z}$ among the elements of $\mathcal{Z}$, which has the least sum of distances to other templates as

$$\hat{Z} = \underset{Z_a \in \mathcal{Z}}{\operatorname{argmin}} \sum_{Z_b \in \mathcal{Z} \setminus Z_a} dist(Z_a, Z_b) \tag{6}$$

where the distance $dist(Z_i, Z_j)$ is actually a dissimilarity measure defined by

$$dist(Z_a, Z_b) = 1 - \frac{\|B_{Z_a,Z_b} \wedge B_{Z_a} \wedge B_{Z_b}\|_0}{\|B_{Z_a} \wedge B_{Z_b}\|_0} \tag{7}$$

where B_{Z_a} is a binary matrix whose entry is 1 if Z_a is a positive or 0 otherwise, B_{Z_a,Z_b} is a binary matrix with the value of 1 if the entry of Z_a is equal to Z_b, or 0 otherwise, $\|\cdot\|_0$ is the number of non-zero entry in a matrix and $\wedge$ is the entrywise logical AND operation. The dissimilarity measure is a variation of the similarity measure used in [9] and is related to Jaccard similarity.

C. Finding Coherent and Non-Coherent Matching Region

We examine the same value among the positive values between the reference template $\tilde{Z}$ and the others Z_j. A binary map that shows whether the values are equal is called a coherent matching map $B_{coherent}$ defined as

$$B_{\text{coherent}} = \bigvee_{Z_a \in \mathcal{Z} \setminus \hat{Z}} (B_{\hat{Z},Z_a} \wedge B_{\hat{Z}} \wedge B_{Z_a}) \tag{8}$$

where $\bigvee$ is the entrywise logical OR operation. Since the coherent matching region is the best matching area for the training data in the same class, we can infer that the region will also be well matched with the test data for the same class. The non-coherent matching map $B_{\text{non-coherent}}$ is expressed as $B_{\text{non-coherent}} = \neg B_{\text{coherent}}$ where $\neg$ is the entrywise logical negation operator.

D. Noise Embedding

Since the non-coherent matching region has quite different code values for each code in a class, adding an arbitrary value in this region does not greatly affect the matching result. So we assign an arbitrary value that follows the distribution of the coherent domain values into the non-coherent matching region, making it impossible for the intruder to distinguish the coherent domain and maintain the matching performance. Since the entries in Z_i are the decimal numbers converted from the h bit vector $\tilde{y}_{ijk}$ in (3) and the number of 1's of Y_{ij} is $hN(1/2)^q$, the probability of 1 per bit in coherent region is $(1/2)^q$. Let G be an $h \times 1$ binary vector whose entry g_i is a random variable with Bernoulli distribution pmf of $\Pr(g_i = 1) = (1/2)^q$ and $\Pr(g_i = 0) = 1 - (1/2)^q$. Then, the noise decimal $f(G)$ is injected into the non-coherent region and the noise-embedded protection template Z^* is expressed as

$$Z^*(m,n) = \begin{cases} f(G) & \text{if } B_{\text{non-coherent}}(m,n) = 1 \\ \hat{Z}(m,n) & \text{otherwise.} \end{cases} \tag{9}$$

Since Z^* is an $r \times N$ decimal matrix and each entry uses h bits, the size of $h \times r \times N$ bits is needed to store one template. For example, let the original code be an $M \times N = 20 \times 512$ matrix. If $h = 4$ and $r = 12$, then $24,576$ bits $= 24$ kbits are required, which is 2.4 times the original size. Although the final template Z^* is larger than the original, it is reasonable to embed the noise so as to maintain good performance and to prevent inversion.

E. Modifications for Alignment

Biocode inputs are not aligned in general, so they must be circular shiftable by left-right or up-down movement in such cases. It means that when measuring the distance between the templates Z_a and Z_b in (6), one of them is circularly shifted to select the correct B_{coherent}. Using this fact, (6) is changed as

$$\hat{Z} = \underset{Z_a \in \mathcal{Z}}{\operatorname{argmin}} \sum_{Z_b \in \mathcal{Z} \setminus Z_a} \left[\min_{U_b \in H_{Z_b}} dist(Z_a, U_b) \right] \tag{10}$$

where H_{Z_b} is a set of circularly shifted templates of Z_b from left to right, and (8) is reformulated as

$$B_{\text{coherent}} = \bigvee_{Z_b \in \mathcal{Z} \setminus \hat{Z}} (B_{\hat{Z},K_{\hat{Z},Z_b}} \wedge B_{\hat{Z}} \wedge B_{K_{\hat{Z},Z_b}}) \tag{11}$$

where $K_{\hat{Z},Z_b}$ is the aligned template of Z_b based on $\hat{Z}$ so that the distance between $\hat{Z}$ and Z_b is the minimum score as follows,

$$K_{\hat{Z},Z_b} = \underset{U_b \in H_{Z_b}}{\operatorname{argmin}}\, dist(\hat{Z}, U_b). \tag{12}$$

F. Authentication

The authentication step is similar to the template creation process of the enrollment step. Given a test code X_t, the test template Z_t is created by (5), and it is compared with the protected template Z^* of the database by (7). If the score is smaller than the criterion, it is authenticated, otherwise rejected.

IV. Experiments and Discussion

A. Experimental Databases

The proposed algorithm can be used for all biometric. It requires binary biocodes input as mentioned in Section III-A. Other forms of biocode should be changed to binary to use our algorithms. We tested our algorithm with popular iris database ND-iris-0405[15]. This dataset was acquired in 2004-2005 at Notre Dame with LG 2200 iris image camera. It has about 65,000 iris images in a diverse and challenging environment

TABLE I
EER Performance According to h and q with $r = 12$.

q	EER (%)					
	$h = 6$	$h = 8$	$h = 10$	$h = 12$	$h = 14$	$h = 16$
2	0.00	0.00	0.20	0.00	0.02	0.39
3	0.00	0.00	0.00	0.00	0.00	0.25
4	0.02	0.01	0.04	0.00	0.05	0.00
5	0.25	0.25	0.06	0.25	0.25	0.05

TABLE II
EER Performance According to r with some h values and $q = 3$.

h	EER (%)				
	$r = 4$	$r = 8$	$r = 12$	$r = 16$	$r = 20$
4	0.62	0.50	0.25	0.22	0.00
8	0.50	0.00	0.00	0.00	0.00
12	0.55	0.00	0.00	0.00	0.00
16	0.65	0.25	0.25	0.00	0.00

of size 640x480 from 712 classes obtained from 356 people. We randomly select 15 samples in 80 classes to create an experimental environment similar to [11] using ND-iris-0405. Among the 15 samples, 10 are used for enrollment images and 5 as authentication images. To divide and encode the iris region, the Masek algorithm was used to create a 240×20 binary iris code [16].

B. Scores for evaluation

In order to evaluate the performance change in various situations, four scores are measured such as genuine (GE), imposter (IM), pseudo genuine 1 (PG1), and pseudo genuine 2 (PG2). We first create the templates using different key for each class, and the GE is referred to as the score between the templates of the same class, and the IM is referred to as the score between the templates of different classes. Next, we create templates that use the same key for all classes, and the PG1 is referred to as the score between distinct classes. Finally, we refer to PG 2 as the score obtained by comparing one template with 50 templates after generating templates using 51 different keys for an iris code of each class.

GE is the benchmark for comparing three different score distributions, and we produce three performance results such as GE-IM, GE-PG1, and GE-PG2. The GE-IM refers to the performance of the algorithm in normal situations, the GE-PG1 shows how much the performance is degraded compared to the GE-IM when a user's keys are stolen, and the GE-PG2 shows how different templates are generated in the situation when one iris is used in multiple devices using different keys. All the performances are expressed in terms of the equal error rate (EER), which means that the false accept rate (FAR) and the false reject rate (FRR) are equal.

C. Effect of parameters

In order to satisfy the conditions for the BTP system, the proposed algorithm uses three tuning parameters such as h, r and q. We fix one parameter and evaluate the performance change according to the other two parameters.

Table I shows the performance according to h and q for fixed $r = 12$ with a reasonable template size, showing good EER distribution trends. As q increases by 1, the number of valid 1's for matching is reduced by the power of $1/2$, so we experiment with the algorithm in the range of q from 2 to 5. Since the height M of all iris codes used in this experiment is fixed at 20, h values are chosen to be less than 20. In Table I, it can be seen that the EER increases with the increase of h as q approaches 2. The EER usually depends on the size of the coherent matching region, because the non-coherent matching area, where the noise is embedded, does not contribute to lowering the EER. If h increases, the number of bits to make an entry of Z_i in (5) increases, so that the matching performance becomes poor and the coherent matching region is reduced, resulting in the increase of EER.

On the other hand, the situation is different with large q. When q approaches 5, the EER decreases as h increases, which is opposite to the case of $q = 2$. Although the coherent matching region decreases by half when q increases by 1, the active range of entry values in Z_i decreases more quickly where the "active range" means the range of numbers that are more frequently selected from $[0, 2^h - 1]$. In the case of large q, the range of active values has a greater effect because the size of the coherent matching region is very small. For example, given $q = 2$, the entries of Z_i are appropriately distributed in the interval $[0, 15]$ when $h = 4$. However, as q is changed to 5, they usually remain at a limited number such as $0, 1, 2, 4, 8$ and the others are rarely selected. This increases the ambiguity between the classes, which results in the increase of EER. However, when q is kept to 5 and h is increased to 18, the entries of Z_i have active range of $0, 1, 2, ..., 2^{17}$. The longer the range of active values, the less the ambiguity and EER. In summary, as h increases for a fixed r, the performance decreases when q is small, and increases when q is large. This tendency is similar for all databases.

Table II shows the EER performance of h and r for fixed $q = 3$ with the best performance.. As expected, the EER decreases with the increase of r regardless of h. It means that increasing r also increases the size of the coherent matching region to enhance the discriminability of Z_i.

D. Comparison with other algorithms

In this section we compare our method with the state-of-the-art algorithms on cancelable biometrics. First, the parameters of $q = 3$, $r = 12$, and $h = 10$ are used for the proposed algorithm. We compare our method with five algorithms, such as biohasing [10], SRP [17], block remapping [7], bloom filter [8], and IFO hashing [9]. Their parameters were manipulated to have the best performance. The biohashing in [10] was applied to the fingerprint, but it is compared here as it is a representative and universal algorithm in the cancelable biometrics field. The bloom filter of [8] was also applied to the face database, but we also include it for comparison because it was adopted for enhancing the unlinkability of previous iris

TABLE III
PERFORMANCE COMPARISON BETWEEN THE PROPOSED METHOD AND OTHER ALGORITHMS IN TERMS OF EER IN DATABASES.

Method	EER (%)		
	GE-IM	GE-PG1	GE-PG2
Unprotected	1.28	N/A	N/A
Biohashing [10]	**0.00**	16.20	**0.00**
SRP [11]	**0.00**	5.16	**0.00**
Block Remapping [7]	1.10	12.58	4.11
Bloom Filter [8]	1.09	5.32	1.25
IFO hashing [9]	0.82	2.84	1.10
Proposed	**0.00**	**0.88**	**0.00**

TABLE IV
EXCUTION TIME IN SECONDS FOR THE PROPOSED SYSTEM WITH OTHER ALGORITHMS

Method	Enrollment time		Authentication time	
	Key/Templ. generation	Total	Key/Templ. gen + Comparison	Total
Biohashing	0.5999/0.0063	0.6062	0.5875/0.0753	0.6628
SRP	0.0990/0.0018	0.1008	0.0942/0.0097	0.1039
Block Remapping	0.0010/0.0007	0.0017	0.0011/0.0065	0.0076
Bloom Filte	0.0013/0.0017	0.0030	0.0009/0.0017	0.0027
IFO hashing	0.0049/0.0030	0.0080	0.0049/0.0444	0.0493
Proposed	0.0046/0.0272	0.0319	0.0046/0.0407	0.0453

recognition method [18]. We test the algorithms on the same environment by implementing them in MATLAB. Since there are very few comparative experiments with various databases in a common environment in the CIB field, we hope that our experiment would help to investigate the properties of the compared algorithms.

Table III shows the EER for the above stated algorithms. As mentioned in Section IV-B, GE-IM means the discriminability of an algorithm in a normal situation, and GE-PG1 shows the ability to withstand the key-dependent attacks when the keys are stolen. GE-PG2 is the authentication performance of generated templates when one class is used on multiple devices with different keys and affects the unlinkability. The proposed algorithm shows good performance for all three categories in terms of EER. In GE-IM and GE-PG2, two biometric salting algorithms [10],[17] take advantage of user-specific projection and show the excellent efficiency of zero EER. In GE-PG1, however, biohashing shows poor results because it allows similar projection between classes with the stolen key. Three non-invertible transform-based algorithms [7],[8],[9] are slightly worse than salting methods on GE-IM and GE-PG2 on average, but better in GE-PG1. Among them, IFO hashing shows overall stable performance for all three cases in all datasets due to the discriminability of min-hash and repeated hash values. Our algorithm shows better performance than others in most scores because the robustness is enhanced and the accuracy is improved by composing a plurality of templates. Also, the templates are created by using repetitive hash values based on RRP-BS, which makes them coherent matching areas. This result is also demonstrated in the receiver operating characteristics (ROC) curves of all the databases as shown in Figures 2.

Table IV shows the running times of the compared algorithms. The key generation process in biohashing takes much time because it needs large amount of computations in the orthogonalization process. The block re-mapping method in enrollment stage is very fast as the random shuffling of the duplicating blocks is done only. The bloom filter method needs the least time because of its alignment-free property during the authentication. Since the proposed algorithm uses several images when enrollment, it takes about 27.2 msec to generate the template. However, when the authentication is performed, the template is created with a single iris code and it needs similar time to IFO hashing. The number of registered images used in Table IV is four, but if the number of training images increases, the registration time may become longer. However, the execution time of the proposed algorithm is modest because a small number of registered images can achieve sufficient accuracy.

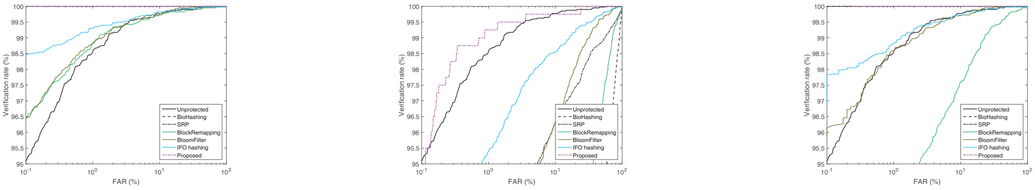
E. Unlinkability

The unlinkability of the proposed algorithm depends on the key matrices P and S. They are generated by different pseudo random numbers for each device and have different values. We have seen that the performance of our method in terms of GE-PG2 is good enough in Section IV-C. It means that the iris templates among other devices have little similarity and the unlinkability is satisfied.

In this subsection, we also test whether the distribution of PG2 is ultimately ambiguous compared to that of IM. As mentioned in Section IV-B, the score of IM is obtained from the comparison of the other classes, and the score of PG2 is derived from the comparison within the class. Therefore, the distribution of PG2 is always statistically closer to the distribution of GE than that of IM. When the distribution of PG2 is nearly identical to that of IM, and if the adversary gets the templates of two devices derived from that class, he/she would not be able to distinguish whether they are from the same class or not. Speficically, Figure 3 shows that the distributions of PG2 and IM are almost overlapping. In other words, the proposed algorithm can create completely different templates if the keys are different.

F. Non-invertibility Analysis

The proposed algorithm uses Hadamard product and noise embedding method for non-invertibility in the enrollment step. Suppose an attacker has acquired all the possible parameters, key matrices, and the template Z^*. To retrieve the original biometric, the attacker needs to fill in Z^* with the original entries instead of noise in the non-coherent matching region and perform an inverse operation of the Hadamard product. However, since neither B_{coherent} nor $B_{\text{non-coherent}}$ is stored, the attacker has no way of knowing it. Estimating the coherent region is very difficult due to the number of cases of $\binom{n(Z^*)}{n(B_{c1})}$ where $n(Z^*)$ is the number of elements of Z^* and $n(B_{c1})$



(a) (b) (c)

Fig. 2. ROC curves of the proposed system with the other algorithms for the best EER performance (a) GE-IM (b) GE-PG1 (C) GE-PG2.

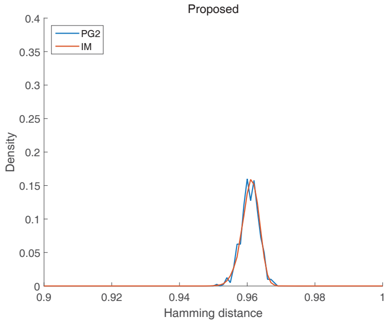


Fig. 3. PG2 and IM distributions of the proposed algorithm.

is the number of elements with 1 of B_{coherent} and the size of coherent matching region is usually different for each device because of using different PIN. And because the hadamard product increases the number of inverse operation candidates, it is almost impossible to estimate the original biometric which is effectively protected from brute-force, hill climbing, multiplicity or pre-image attacks through our algorithm.

V. Conclusion

The proposed CB algorithm adopts Hadamard product and noise embedding method based on RRP-BS. The unlinkability is statisfied by using the RRP-BS, and the Hadamard product and noise embedding enable the non-invertibility. To use the noise embedding method effectively, we defined a coherent matching region among several enrollment templates and embed the noise into the non-coherent region. The area information cannot be stolen because it is not stored in the authentication system. Without precise information on the coherent regions, an attacker would have a wrong estimate even if he/she performs an inverse operation on the Hadamard product. This makes the proposed algorithm robust to brute-force, hill climbing, multiplicity or pre-image attacks. The proposed method does not only guarantee the non-invertibility but also preserves the low error rates in terms of GE-IM, GE-PG1, and GE-PG2.

Acknowledgement

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2017-2016-0-00288) supervised by the IITP (Institute for Information & communications Technology Promotion).

References

[1] N. K. Ratha, J. H. Connell, and R. M. Bolle, “Enhancing security and privacy in biometrics-based authentication systems,” *IBM systems Journal*, vol. 40, no. 3, pp. 614–634, 2001.

[2] A. B. Teoh, A. Goh, and D. C. Ngo, “Random multispace quantization as an analytic mechanism for biohashing of biometric and random identity inputs,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 1892–1901, 2006.

[3] D. Maltoni, D. Maio, A. Jain, and S. Prabhakar, *Handbook of fingerprint recognition*. Springer Science & Business Media, 2009.

[4] Y. C. Feng, P. C. Yuen, and A. K. Jain, “A hybrid approach for generating secure and discriminating face template,” *IEEE transactions on information forensics and security*, vol. 5, no. 1, pp. 103–117, 2010.

[5] N. K. Ratha, S. Chikkerur, J. H. Connell, and R. M. Bolle, “Generating cancelable fingerprint templates,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 29, no. 4, pp. 561–572, 2007.

[6] J. Zuo, N. K. Ratha, and J. H. Connell, “Cancelable iris biometric,” in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. IEEE, 2008, pp. 1–4.

[7] J. Hämmerle-Uhl, E. Pschernig, and A. Uhl, “Cancelable iris biometrics using block re-mapping and image warping,” in *International Conference on Information Security*. Springer, 2009, pp. 135–142.

[8] M. Gomez-Barrero, C. Rathgeb, J. Galbally, C. Busch, and J. Fierrez, “Unlinkable and irreversible biometric template protection based on bloom filters,” *Information Sciences*, vol. 370, pp. 18–32, 2016.

[9] Y.-L. Lai, Z. Jin, A. B. J. Teoh, B.-M. Goi, W.-S. Yap, T.-Y. Chai, and C. Rathgeb, “Cancellable iris template generation based on indexing-first-one hashing,” *Pattern Recognition*, vol. 64, pp. 105–117, 2017.

[10] A. T. B. Jin, D. N. C. Ling, and A. Goh, “Biohashing: two factor authentication featuring fingerprint data and tokenised random number,” *Pattern recognition*, vol. 37, no. 11, pp. 2245–2255, 2004.

[11] J. K. Pillai, V. M. Patel, R. Chellappa, and N. K. Ratha, “Secure and robust iris recognition using random projections and sparse representations,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 9, pp. 1877–1893, 2011.

[12] C. Rathgeb and A. Uhl, “A survey on biometric cryptosystems and cancelable biometrics,” *EURASIP Journal on Information Security*, vol. 2011, no. 1, p. 1, 2011.

[13] V. M. Patel, N. K. Ratha, and R. Chellappa, “Cancelable biometrics: A review,” *IEEE Signal Processing Magazine*, vol. 32, no. 5, pp. 54–65, 2015.

[14] H. Kaur and P. Khanna, “Biometric template protection using cancelable biometrics and visual cryptography techniques,” *Multimedia Tools and Applications*, pp. 1–29, 2015.

[15] P. J. Phillips, W. T. Scruggs, A. J. O’Toole, P. J. Flynn, K. W. Bowyer, C. L. Schott, and M. Sharpe, “Frvt 2006 and ice 2006 large-scale experimental results,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 5, pp. 831–846, 2010.

[16] L. Masek, “Recognition of human iris patterns for biometric identification,” B.S. dissertation, The School of Computer Science and Software Engineering, The University of Western Australia, Crawley WA, Perth, Australia, 2003.

[17] J. K. Pillai, V. M. Patel, R. Chellappa, and N. K. Ratha, “Sectored random projections for cancelable iris biometrics.” in *ICASSP*, 2010, pp. 1838–1841.

[18] C. Rathgeb, F. Breitinger, and C. Busch, “Alignment-free cancelable iris biometric templates based on adaptive bloom filters,” in *2013 International Conference on Biometrics (ICB)*. IEEE, 2013, pp. 1–8.