Journal of Database Management
Volume 34 • Issue 3

One-Factor Cancellable Fingerprint Template Protection Based on Index Self-Encoding

Yalan Feng, Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University, Hefei, China

Huabin Wang, Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University, Hefei, China*

Dailei Zhang, Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University, Hefei, China

Jiahao Li, Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University, Hefei, China

Liang Tao, Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University, Hefei, China

ABSTRACT

The existing one-factor cancellable biometrics algorithms generally require random sequences to reorder the biometrics, which reduces the discrimination of the transformed biometrics. Some algorithms hide and transmit the random sequence by XORing the random sequence with original biometrics, which may cause the leakage of the original biometrics. Therefore, this paper proposes a one-factor cancellable fingerprint template protection based on index self-encoding. First, the integer sequence generated by the hash function is used as the index. The random sequence is automatically encoded directly by the index value, and the generated binary sequence retains the original biological characteristics to the greatest extent. Second, self-encoding binary sequence and random binary sequence are XORed to obtain the encoded key without directly storing binary factor sequences. Experiments are implemented on the fingerprint database of FVC2002 and FVC2004, the results show that the recognition rate is enhanced; meanwhile, it fits the design criteria of cancellable biometrics.

KEYWORDS

Binary, Biometrics, Encoded Key, Hash Function, Performance, Privacy, Recognition, Revocability, Security, XOR

INTRODUCTION

Today, due to the progress of technology in daily life, a variety of data security issues emerge one after another (Dong et al.,2022). People begin to pay attention to data security and privacy protection (Turesson et al., 2021), and information security has become increasingly important in daily life (Bollle

DOI: 10.4018/JDM.321546 *Corresponding Author

et al., 2002; Lin et al., 2020). In identity management, people are also paying increasing attention to the identification and protection of biometrics. Common biometrics include fingerprint (Yang et al., 2022), iris (Lai et al., 2017), finger vein (Kirchgasser et al., 2019), and so on. Fingerprint identification technology (Wang & Hu, 2016) is the most convenient and widely used biometric technology with strong adaptability, easy operation, and high stability. At the same time, there are some problems with biometric templates that are worth noting: because biometric features are irrevocable, they cannot be reissued once they are damaged, and the authors need to ensure that the generated cancellable biological template is reproducible; in addition, biometrics are unique, and new biometrics cannot be generated if user characteristics are stolen (Ratha et al., 2001). Therefore, biometric template protection is a pivotal and urgent matter.

Biometric template protection falls into two types, including the two-factor cancellable method and the one-factor cancellable method. The two-factor cancellable biometric template protection algorithm needs an extra specific parameter from the user, which is a token or password, along with biometrics as input, to guarantee the unlinkability and revocability of the converted template. For example, reference (Teoh et al., 2004, 2006) generated a binary vector by the inner product of a feature vector and a user-specific nonsquare orthogonal random matrix and then performed threshold binarization to generate a scheme of cancellable biometric template, Biohash. Reference (Wang et al., 2017) proposed a cancellable fingerprint template using the local Hadamard transform method, which used a randomly generated token k to construct a submatrix of a Hadamard matrix for local Hadamard transformation to obtain a cancellable biometric template. An original approach is proposed in the reference (Wang & Hu, 2017) and used a randomly generated token matrix A to further hide the original biometric information. In a typical two-factor cancellable scheme, user-specific parameters (password or token) are important input factors, but there are also some problems caused by external factors (user-specific parameters): (i) it is necessary to keep the token or remember the password. (ii) External factors may be lost, forgotten, or stolen. (iii) The exposure of user-specific parameters may lead to the risk of conversion template intrusion, especially for some salt-based schemes.

The one-factor cancellable biometric template protection algorithm, as you can see from the name, does not require additional input factors and only requires biometrics as input; it effectively avoids the problems caused by external factors. Reference (Lee et al., 2018) put forward a one-factor cancellable template protection algorithm based on extended feature vector hashing (EFV). After copying and expanding the original biometrics, a hash function is used to generate a permutation factor to array the random sequence again to obtain a revocable template. Reference (Kong et al., 2021) proposed a one-factor sliding window algorithm based on fingerprints (WSE). After combining the extended binary biometric vector through the sliding window jumping value, biometric information is hidden by corresponding steps, such as the hash function. Reference (Zhang et al., 2021) proposed a one-factor cancellable fingerprint template protection algorithm called feature enhanced hashing (FE). After copying and expanding the original biometric features, the improved hash function is used to calculate the replacement factor and randomize it. The sequence is reordered, and a cancellable template is generated by shortening the random sequence before and after by the same length. Reference (Li & Wang, 2022) proposed a one-factor fingerprint feature template protection scheme based on the novel minimum hash signature (NMHS) and the secure extended feature vector (SEFV). NMHS has generated the fused hash code and SEFV is used to map. Generate a pseudo identifier for matching during registration and verification without additional storage of keys and biometrics. Finally, the pseudo identifier is matched and identified.

When the above one-factor cancellable biological template protection scheme reorders the random sequence after calculating the hash function, the hash function may generate different permutation factors, and the same cancellable biological template may be rearranged to generate the same cancellable biological template, resulting in information leakage. For the storage of random binary sequences, the existing algorithm is to perform a simple XOR operation with the original biometric vector to acquire the encoded key. If the database is stolen, some original feature vectors can be

restored. Therefore, the biometric template protection framework is based on the reference (Lee et al., 2018). This paper presents a one-factor cancellable fingerprint template protection algorithm based on index self-encoding. This algorithm primarily improves the encoding method of converting decimal to binary and the method of storing random binary sequences. First, this paper uses the integer sequence generated by the hash function as an index and marks the position corresponding to the index value as '1' and the other positions as '0' to obtain a new binary sequence, that is, the binary sequence generated by the index self-encoding. It realizes the difference and uniqueness after encoding different categories of features while realizing the same category of features after encoding. Second, when the random binary sequence is stored, the binary sequence generated by the index self-encoding is XORed with the random binary sequence to obtain the encoded key and stored in the database. Improved the security of random binary sequence storage.

The article refers to the fingerprint template in binary vector form (Jin et al., 2016), and the experiment uses fingerprint datasets FVC2002 (Maio et al., 2002) and FVC2004 (Maio et al., 2004). It is also demonstrated that the scheme defends against both attacks on safety and meets the design criteria of cancellability.

The main contributions are as follows:

1. An algorithm based on index self-encoding is proposed. The integer sequence generated by the hash function is used as the index, and the index value is directly used for automatic encoding without introducing additional random sequences. The generated binary sequence can preserve the original to the greatest extent. biological characteristics.
2. A new random sequence storage algorithm is designed. The two binary sequences obtained by index self-encoding and random binary sequences are respectively XOR operated, so as to obtain the revocable fingerprint template for matching and the encoding key for storing in the database, which can realize the revocability of biometrics and prevent attackers from recovering the original biometrics from the encoding key. So as to prevent information leakage.

RELATED WORK

Locality Sensitive Hashing

LSH (Datar et al., 2004) is a way to maximize similarity while dimensionality reduction processing on high-dimensional data by hashing so that the probability that similar features are mapped to similar positions after hashing is increased (Aydar & Ayvaz, 2019). LSH family H formula:

$$\begin{aligned} &P_{\mathrm{h}\in H}\left(h_i\left(M\right)=h_i\left(N\right)\right)\le P_1 \qquad if\ \ S\left(M,N\right)<R_1\\ &P_{h\in H}\left(h_i\left(M\right)=h_i\left(N\right)\right)\ge P_2 \qquad if\ \ S\left(M,N\right)>R_2 \end{aligned} \tag{1}$$

LSH is a probability distribution over hash functions family H, make $P_{\mathrm{h}\in H}\left(h_i\left(M\right)=h_i\left(N\right)\right)=S\left(M,N\right)$, where S is a similarity function (Charikar & Moses, 2002) that addresses the set of objects M and M. LSH handles sets M and M through some hash functions h_i. The pairwise distances of M and M are to be approximated in calculating the probability of collision because h_i.

Biometric Identification Scheme

Biometric template protection (Rathgeb & Uhl, 2011) is generally divided into cancellable biometrics (Patel et al., 2015) and biometric cryptosystems (Jain et al., 2008), of which cancellable biometrics

are the focus of this paper. A cancellable biometric is the generation of an irreversible biometric template from a common biometric template through a transfer function and user-unique parameters. A cancellable biometric template protection scheme needs to possess four properties (Patel et al., 2015): noninvertibility, unlinkability, revocability, and performance preservation. The cancellable biometric authentication methods include two-factor cancellable methods that combine external factors and biometrics and one-factor cancellable methods that only use biometrics. This subsection presents related work about these two methods.

The extended feature vector hashing (EFV) algorithm is a one-factor cancellable biometric template protection algorithm. Specifically, the enrollment process has two inputs: a biometric vector $\boldsymbol{x}$ and a random binary vector $\boldsymbol{r}$. The two vectors perform a series of operations and are then arranged as a "cancellable template". The random binary vector (original key) and fingerprint feature vector are XOR encoded to generate the encoded key, and the encoded key together with the revocable template are stored in the database. What is noteworthy is that the original key is deleted after enrollment. The verification process, a query biometric and the encoded key are XORed to generate the decoded key, and the query template is generated by the decoded key. Since two permutation factors are biometrics independent of enrollment and query, this scheme does not require a second input of permutation factors like typical permutation-based methods.

Biohash is a two-factor cancellable biometric template protection algorithm. Specifically, generating a randomization matrix on the basis of each user's unique token, and inner product the matrix with the biometric vector, and then the inner product vector is subjected to a threshold algorithm to obtain a binarized hash code. The Biohash algorithm (Teoh et al., 2004, 2006) was first applied in the field of fingerprint biometrics. Because the Biohash method has noninvertibility and unlinkability, the dataset test achieves good results with a recognition rate of 100% in an ideal state. The algorithm satisfies the revocability, and the user token can be replaced to cancel the original template and replace the new encrypted template when the template information leaks. The Biohash algorithm requires biometric vectors $\boldsymbol{x} \in \mathbb{R}^n$ and orthogonal random matrices $R \in \mathbb{R}^{n \times q}$ together as input, where $q \leq n$. Cancellable biological template generation steps in the Biohash algorithm are as follows:

Obtain the inner product vector $\boldsymbol{y}$ by calculating $\boldsymbol{y} = R^T \boldsymbol{x}$; (ii) Binarize $\boldsymbol{y}$ based on a predefined threshold $\boldsymbol{\tau}$, and bioCode $\boldsymbol{b} \in \left[0,1\right]^q$ is generated by formula (2).

$$b_i = \begin{cases} 0, & if\ \mathrm{y}_i \geq \tau \\ 1, & otherwise \end{cases} \tag{2}$$

where $i = 1...q$. The Biohash algorithm is also applied to other biometric modalities, not limited to fingerprints, but it must combine biometrics and user tokens to assure the accuracy of the algorithm. If the user token is leaked, it will seriously affect the algorithm's accuracy. From this, it can be seen that external factors play a decisive role. Some methods, such as compromised key algorithms and orthogonal matrices, have also been found to recover the original biometrics.

The two-factor cancellable biological template protection algorithm represented by Biohash requires external factors as input, so in the process of enrollment and authentication, the attacker is likely to steal the token or password; in addition, the storage of the token or password is a problem for the user. Therefore, this paper proposes the one-factor cancellable biometric template protection algorithm, which improves the recognition rate and security.

METHOD

This section mainly elaborates on four aspects: framework, algorithm steps, index self-encoding process, and cancellable template generation.

System Framework

From Figure 1, the framework of the one-factor cancellable template authentication system based on index self-encoding is proposed. This method only uses biometric features $\boldsymbol{x}$. $\boldsymbol{r}$ is the ancillary data, which is a random binary vector. Both only need to perform certain actions to receive the cancellable biometric template.

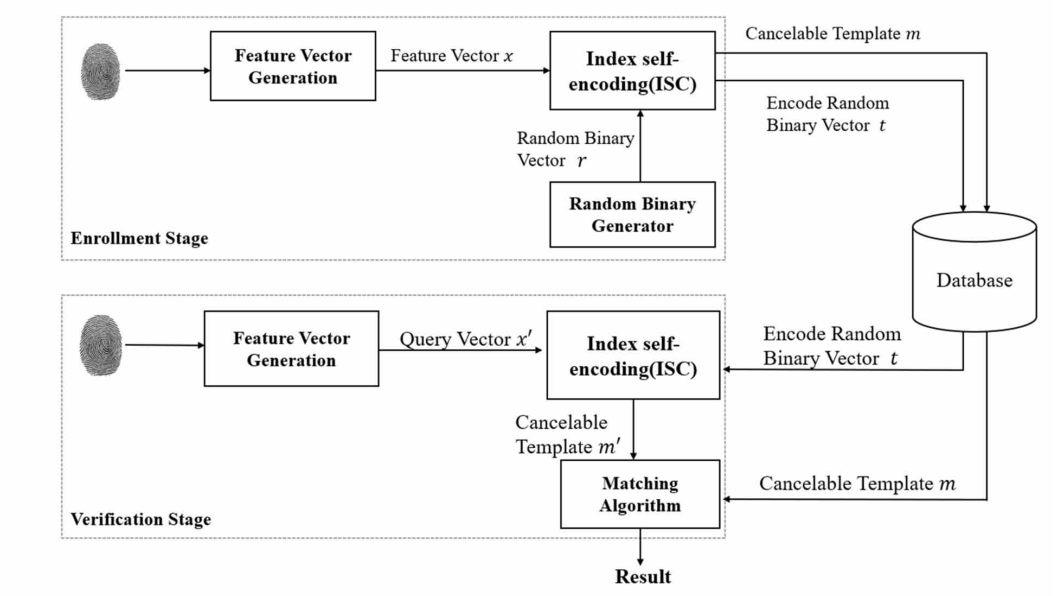
The main process of the authentication system is to apply it to the fingerprint matching scenario. The authentication system is applied in the scene of fingerprint matching in the enrollment stage. The input binary biometric vector $\boldsymbol{x}$ is copied and expanded to obtain the sequence $\overline{\boldsymbol{x}}$. The binary vector is combined by the sliding window jumping value (Kong et al., 2021) and then converted into a decimal number sequence $\hat{\boldsymbol{x}}$. Next, the authors use the index value $\boldsymbol{y}_1$ obtained by the hash function calculation. A new binary sequence $\boldsymbol{w}_1$ is obtained by self-encoding the index value $\boldsymbol{y}_1$. Finally, the sequence $\boldsymbol{w}_1$ and the key $\boldsymbol{r}$ are XORed to obtain the final cancellable template $\boldsymbol{m}$. Meanwhile, it uses another hash function to generate a hash code $\boldsymbol{y}_2$ for self-encoding to obtain a binary sequence $\boldsymbol{w}_2$ and uses the generated $\boldsymbol{w}_2$ and the key $\boldsymbol{r}$ to the XOR operation to obtain a random binary vector t. After the enrollment phase, the authors need to memorize the data in the database, including the encoded random binary vector t and cancellable template $\boldsymbol{m}$. In the verification phase, the authors need to decode the generated key $\boldsymbol{r}'$ from the ciphertext t, which is used to generate a cancellable template $\boldsymbol{m}'$. Finally, $\boldsymbol{m}$ and $\boldsymbol{m}'$ are matched to estimate whether the match is ture or not. In this paper, $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are irreversible binary sequences obtained by self-encoding after hashing. Irreversible binary sequences are obtained by XOR of $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ and random binary sequences $\boldsymbol{r}$. Compared with the irreversible binary sequences are obtained by XOR of $\boldsymbol{x}$ and $\boldsymbol{r}$ proposed in the previous paper, the method in this paper is more difficult than the original method to derive the original biological characteristics, which improves the irreversibility and further ensures the security of data. The improved self-encoding method and the storage of random binary vectors in this paper have increased noninvertibility, and the security and recognition rate have also been obviously improved.

Hash Algorithm Based on Index Self-Encoding

Figure 2 shows the specific flow chart of the index self-encoding hash algorithm. The detailed steps of the index self-encoding algorithm are as follows:

1. Let $\boldsymbol{x} \in \left[0,1\right]^{h}$ be a binary biometric vector, copy $\boldsymbol{x}$ to compose an extended binary biometric vector $\overline{\boldsymbol{x}} \in \left[0,1\right]^{hn}$, in which h is the length of the vector and n is the system parameter. This step allows for better protection of biological characteristics by replication.
2. Each $\overline{x}_i \in \overline{\boldsymbol{x}}$ appends the corresponding $2(k-1)$ elements that come from $\overline{\boldsymbol{x}} \in \left[0,1\right]^{hn}$, where k is the system parameter and is called the window size. This generates a subbit block $\overline{x}_{b_i} = [\overline{x}_i \mid \overline{x}_{i+2} \mid \overline{x}_{i+4} \mid \cdots \mid \overline{x}_{i+2(k-1)}]$ via a sliding and extracting window, where | represents the connection operation. According to step 2, $\overline{\boldsymbol{x}}$ can be transformed into subbit blocks $\overline{\boldsymbol{x}}_b$ so that the vector $\boldsymbol{x}$ can be protected. For example, if $\overline{\boldsymbol{x}} = [\overline{x}_1, \overline{x}_2, \cdots \overline{x}_{hn}]$, $k = 2$, then each $\overline{x}_i$ appends

Figure 1.
Framework of authentication system



$2*(2-1)$ elements from $\bar{\boldsymbol{x}}$. If $\bar{x}_i$ is the last element $(\bar{x}_{hn})$ of $\bar{\boldsymbol{x}}$, the second element of $\bar{\boldsymbol{x}}$ will be appended. That is, $\bar{\boldsymbol{x}}_b = [\bar{x}_1 \mid \bar{x}_3, \bar{x}_2 \mid \bar{x}_4, \cdots \bar{x}_{hn-1} \mid \bar{x}_1, \bar{x}_{hn} \mid \bar{x}_2]$.

3. Each subbit block $\bar{x}_{b_i}$ of $\bar{\boldsymbol{x}}_b$ is transformed into $\hat{x}_i \in Z$, and each binary subbit block is transformed to decimal. The resulting decimal number is processed below, the transformation of the two hash functions $y_i = [i^{(\hat{x}_i+1)} + (\hat{x}_i + 1)] mod(hn+1)$, $y_j = [(\hat{x}_j + 1) * j] mod(hn+1)$ to generate a real-valued vector. For both hash functions, all elements are incremented by '1' to avoid the situation of $\hat{x}_i = 0, \hat{x}_j = 0$, because the index value obtained by adding '1' and not adding '1' will be very different, which may adversely affect the recognition rate; $mod(hn+1)$ to ensure that $\bar{x}_i$ is transformed to obtain the maximum value of $\boldsymbol{y}$ is hn. If the result of the modulo operation is 0, set $\boldsymbol{y} = 1$. This procedure generates $\boldsymbol{y_1} = \left[1, hn\right]^{hn}$ as a vector of integers, and $\boldsymbol{y_2} = \left[1, hn\right]^{hn}$ that helps to store random binary sequences.
4. Generating an all-zero vector of the same length as $\boldsymbol{y_1}$, use $\boldsymbol{y_1}$ as an index, mark the position where the index occurs as '1', and mark the rest as '0'. For example, $\boldsymbol{y_1} = 300$, and the 300th position of the all-zero vector is marked as '1'. By checking the set $\left(1, hn\right)$ of all position indices, a new binary sequence $\boldsymbol{w_1}$ is generated. In the same way, $\boldsymbol{y_2}$ also generates a new binary sequence $\boldsymbol{w_2}$.
5. Let $\boldsymbol{r} \in \left[0,1\right]^{hn}$ be a binary vector as the ancillary data of the algorithm. The cancellable template $\boldsymbol{m}$ is obtained by XORing the new binary sequence $\boldsymbol{w_1}$ obtained from index self-encoding with $\boldsymbol{r}$.

The following is the pseudocode of hash algorithm for index self-encoding.

Input: binary biometric vector $\boldsymbol{x} \in [0,1]^h$, multiple of expansion n, window size k.

Output: Cancelable fingerprint template $\boldsymbol{m} \in [0,1]^{hn}$

Step 1: Extended binary biometric vector

For $i = 1:n$

Copy $\boldsymbol{x}$ extension n times and assign it to $\bar{\boldsymbol{x}}$

End for

Step 2: Generating subbits blocks

For $i = 1:hn$

$\bar{x}_{b_i} = [\bar{x}_i \mid \bar{x}_{i+2} \mid \bar{x}_{i+4} \mid \cdots \mid \bar{x}_{i+2(k-1)}]$, where $\mid$ is the connection operation

End for

Step 3: Each subbit block is converted and the hash function is computed

For $i = 1:hn$

Convert $[\bar{x}_i \mid \bar{x}_{i+2} \mid \bar{x}_{i+4} \mid \cdots \mid \bar{x}_{i+2(k-1)}]$ to $\hat{x}_i$

Then

For $i = 1:hn$

Replace $\hat{x}_i$ into $y_i = [i^{(\hat{x}_i+1)} + (\hat{x}_i + 1)] mod(hn+1)$

If $y_i == 0$

Set y_i to 1

End if

End for

Step 4: Index self-encoding generates the binary sequence

For $i = 1:hn$

$\boldsymbol{w_1} = \boldsymbol{P(y_1)}$, Index self-encoding for $\boldsymbol{y_1}$

End for

Step 5: XOR generates a revocable template $\boldsymbol{m}$

$\boldsymbol{m} = \boldsymbol{w_1} \oplus \boldsymbol{r}$ Generating vector $\boldsymbol{r} \in [0,1]^{hn}$ of pseudorandom binary number generator

The Process of Index Self-Encoding

As shown in Figure 3, the combination of binary vectors is converted into decimal numbers through the sliding window hopping value, and the hash function is used to compute the hash value. Then, the authors generate an all-zero vector and find the corresponding position of the all-zero vector for each hash value. The position corresponding to the vector is marked as '1'. If the position is repeated, the position remains '1'. Index positions that do not appear are marked with '0'. In addition, so on, traverse all the hash values. Finally, a binary sequence based on index self-encoding is obtained.

Cancellable Template Generation

If the cancellable template is destroyed, the cancellable template $\boldsymbol{m}$ and its encoded key t can be readily replaced via a new random binary vector $\boldsymbol{r} \in [0,1]^{hn}$ to receive a new cancellable template $\boldsymbol{m} \in [0,1]^{hn}$. This step fully reflects the revocability and replacement in this scheme.

Figure 2.
Hash algorithm for index self-encoding $(h = 6, n = 2, k = 2)$

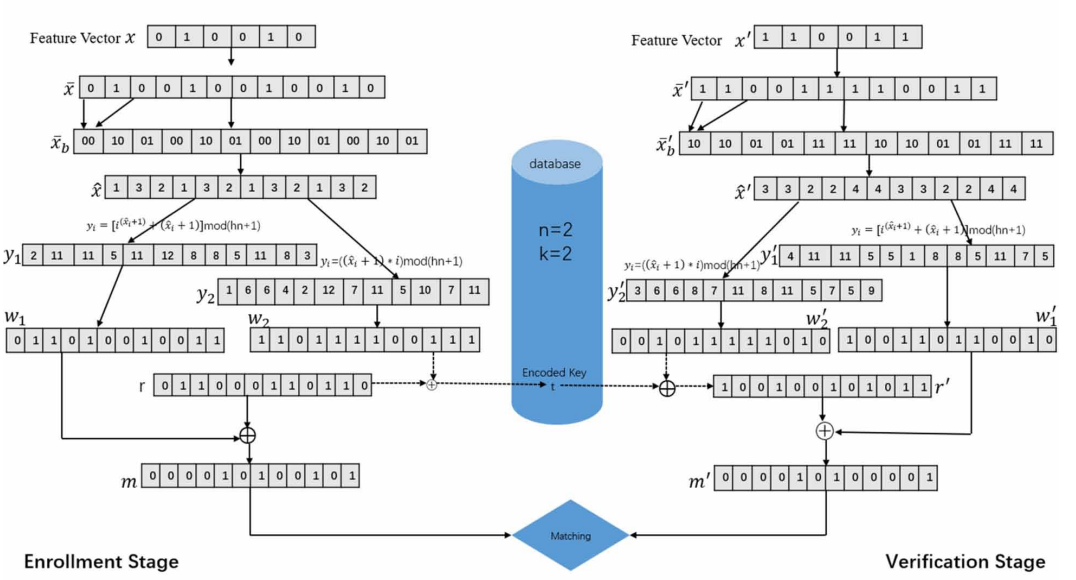
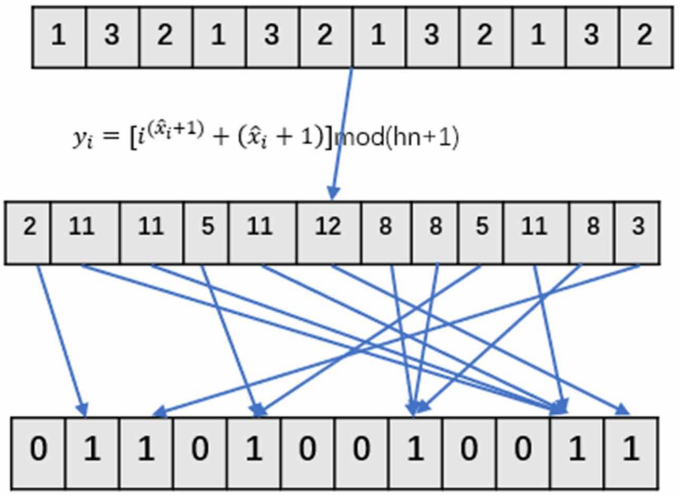


Figure 3.
Index self-encoding process



EXPERIMENTS AND DISCUSSIONS

This paper uses a 256-bit binary fingerprint vector, and the generation steps of the vector are: 1) tiny descriptor extraction (Cappelli et al., 2010), 2) transformation based on kernel learning, and 3) feature vector binarization (Lim et al., 2012). To prove that this method has better recognition performance, this paper conducts experiments on four publicly available fingerprint datasets: FVC2002 (DB1, DB2) (Maio et al., 2002) and FVC2004 (DB1, DB2) (Maio et al., 2004). Each database involves 100 users,

and each user has 8 samples, adding up to 800 fingerprint image samples. Here, 5 out of 8 samples of each user are chosen for testing, and the rest are used for training. The matching result of the two binary vectors is obtained by comparing the Hamming distance of the enrollment and query biometrics.

The assessment criteria of this experiment are based on the reference (Cappelli et al., 2006), which assesses the accuracy of the authentication system according to the Genuine/Imposter matching score and the Equal Error Rate. For each database, five samples per class generate cancellable templates, which can generate $1000\left(100\times C_5^2\right)$ true match scores and $4950\left(C_{100}^2\right)$ false match scores. Due to the application of random binary sequences, to estimate the scheme by rule and line, this paper uses five different ancillary data $\boldsymbol{r}$ for testing. Finally, the obtained equal error rate (EER(%)) is averaged as the final result.

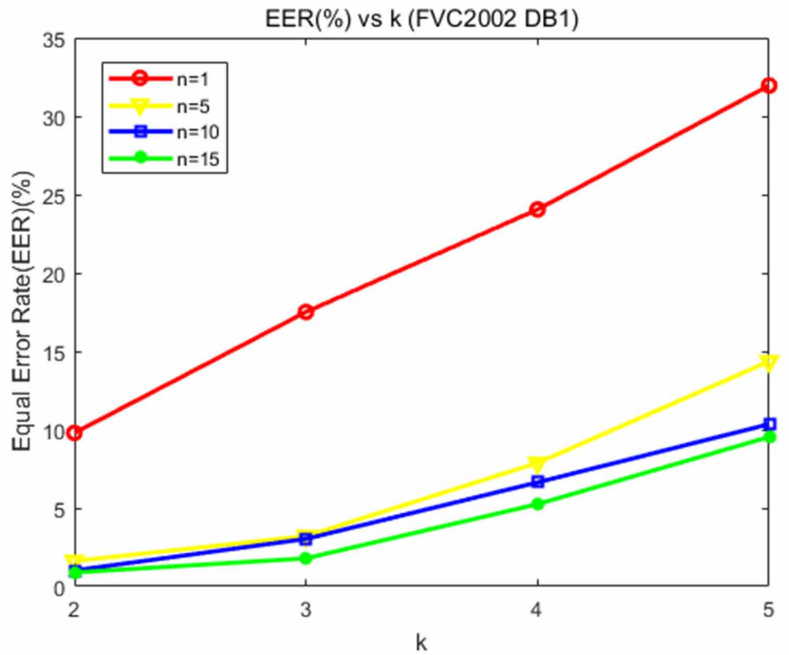
Parameters for Index Self-Encoding

In this section, the authors investigate the effect of parameters on the authentication performance of this method in the light of EER (%). The EER is related to recognition performance. The two system parameters in this scheme are:(i) Window size $k(k\geq 2)$.(ii) Number of repetitive concatenations $n(n\geq 1)$.

Influence of Parameter k

For the parameter $k(k\geq 2)$, the authors take the values of 2, 3, 4, and 5 in turn and conduct experiments on magnifications of 1, 5, 10, and 15, respectively. In Figure 4 "EER(%)-vs-k" (FVC2002 DB1), when k increases, EER(%) increases. As described in Algorithm 1, k is a subbit block, so an increasing number of bits will be added as k increases, which increases the probability that the subbit block is influenced by noise bits, and thus the EER(%) increases. It can also be surveyed that the EER(%) decreases when k does not change and n increases.

Figure 4.
"EER(%)-vs-k" (FVC2002 DB1)



Influence of Parameter n

For the parameter $n(n \geq 1)$, on the basis of $k = 2$, the expansion multiplier n is set to be 1, 5, 10, 15, 20, 50, 100, 200, 500, 800, 1000, and 1200 to conduct experiments. Figure 5 "EER(%)-vs-n" (FVC2004) shows that with the increase in the expansion multiple, the EER(%) decreases dramatically when the multiple does not exceed 100. The rate at which the expansion factor increases EER(%) decreases slowly when the multiple is greater than 100. To reduce the gap between the two random sequences generated from enrollment and query biometrics, n should be large. However, n cannot be extended discretionarily because a template that is too long will result in waste and problems that attackers are easily steal.

Performance Evaluation

In this section, the various parameters that identify the best performance in the previous section are selected, and experiments are performed on the databases FVC2002 (DB1, DB2) and FVC2004 (DB1, DB2) when $k = 2$, $n = 1200$. Table 1 lists the results of several methods for comparative analysis. After a series of experimental comparisons, the authors find that the one-factor cancellable fingerprint template protection scheme based on index self-encoding has significantly lower equal error rates (EER) on the four databases compared with the same type of one-factor cancellable biometric template protection algorithm. Compared with the two-factor cancellable biometric template protection algorithm, the EER of the ISC algorithms is basically lower than those algorithms. Only the fourth database of GRP-IOM Hashing has a moderate error rate smaller than the ISC algorithm. Since the error rate is lower, the recognition rate is higher. The index self-encoding algorithm enhances the recognition rate. because more information of the original features is retained when the hashed integer sequence is indexed and self-encoded so that the similarity of the biometric template matching is higher. On the whole, the recognition rate of the index self-coding (ISC) algorithm has been improved and solves the problem of external factors, which improves security. Therefore, the index self-coding algorithm is better.

Figure 5.
"EER(%)-vs-n" (FVC2004)

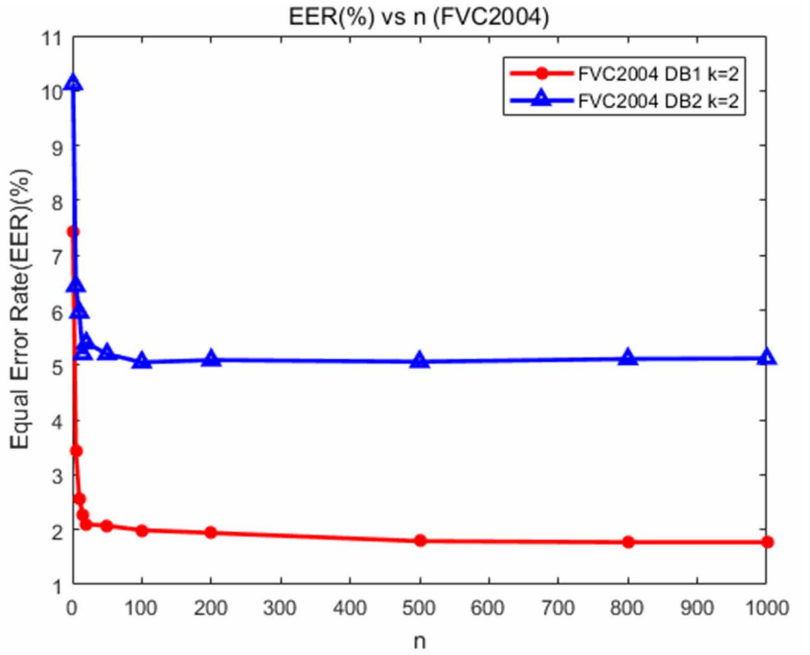


Table 1.
Performance accuracy and comparisons

Methods	EER(%) for FVC2002		EER(%) for FVC2004	
	DB1	DB2	DB1	DB2
EFV Hashing (Lee et al., 2018)	0.30	0.56	2.42	6.27
WSE Hashing (Kong et al., 2021)	0.20	0.62	2.60	7.13
FE Hashing (Zhang et al., 2021)	0.20	0.22	2.00	4.41
NMHS and SEFV (Li & Wang, 2022)	0.43	0.32	1.92	4.97
GRP-IOM Hashing (Jin et al., 2018)	0.22	0.47	4.74	**4.10**
URP-IOM Hashing (Jin et al., 2018)	0.46	2.10	4.51	8.02
Bloom Filter (Abe et al., 2015)	2.30	1.80	13.40	8.10
Proposed Scheme	**0.19**	**0. 10**	**1.92**	4.21

Average Processing Time

Table 2 shows the ISC processing time of $n = 1200$ and $k = 2$ in MATLAB R2017a. The processing time includes the total time of the enrollment stage and verification stage. Table 2 shows that the average processing time for both stages = 0.029 (seconds).

SECURITY AND PRIVACY ANALYSIS

Analysis of Noninvertibility

For noninvertibility, this paper generates a new binary sequence y_2 and a random binary sequence $\boldsymbol{r}$ by XORing with a self-encoding method to obtain t and stores it in the database. If the encoding key t and the revocable template $\boldsymbol{m}$ in the database are stolen by the attacker and can deduce $\boldsymbol{x}$ or $\overline{\boldsymbol{x}}$. Conversely, this means that the noninvertibility is not satisfied. The previous algorithm may restore part of an original biometric template by the postencoded key t and template $\boldsymbol{m}$. In this article, $\boldsymbol{t} = \boldsymbol{y_2} \oplus \boldsymbol{r}$ and the database, the authors can only steal t from the database, and two parameters y_2 and $\boldsymbol{r}$ are unknown, so the thief cannot accurately know the information of one of the parameters; thus, another parameter information cannot be inferred. Therefore, $\boldsymbol{r}$ or y_2 cannot be deduced inversely. If the brute force method is used to crack, it requires $2^{hn} = 2^{256\times1200} = 2^{307200}$ guesses, and the authors will be able to know that the actual calculation is infeasible, so recovering the biometric vector by guessing is hard. This enhances the security of the template.

Table 2.
ISC processing efficiency $\left(n = 1200, k = 2\right)$

Databases	Enrolment Stage	Verification Stage
Average Time(s) for FVC2002 DB1	0.02869	0.02908
Average Time(s) for FVC2002 DB2	0.02968	0.03024
Average Time(s) for FVC2004 DB1	0.02875	0.02915
Average Time(s) for FVC2004 DB2	0.02889	0.02984

Revocability Analysis

For revocability, it means that a new template should be generated to substitute the damaged template when a template is destroyed. Genuine match score, Imposter match score, and Mated-Genuine match score distributions were computed and evaluated. Figure 6 shows the revocability analysis on the four databases, from which it can be observed that they have considerable overlap in the Mated-Genuine and Imposter score distributions. This indicates that for the same user, templates generated with different keys $\boldsymbol{r}$ are indistinguishable from each other, which satisfies the revocability.

Unlinkability Analysis

According to the requirement of unlinkability, the keys $\boldsymbol{r}$ are unlinkable. Different cancellable biometric templates $\boldsymbol{m}$ are generated by XORing the same biometrics with different keys $\boldsymbol{r}$. This paper verifies the unlinkability of ISC by following the benchmark framework in reference (Gomez-Barrero et al., 2018). Cross-matching the cancellable biometric templates generated by ISC with the mated/nonmated sample fraction distribution model. The unlinkability of templates is calculated according to two different measurement methods, local measures and global measures, proposed in the reference (Gomez-Barrero et al., 2018). Two measurement methods are computed from the mated/nonmated sample distribution. The local measure $D_{\hat{s}} \in [0,1]$ is a local score measure, which represents the link degree of the cancellable template on a scoring basis. The global measure $D_{\overrightarrow{sys}}$ evaluates the unlinkability of the total system, the value of $D_{\overrightarrow{sys}}$ from 0 to 1, and it can compare more fairly with other cancellable schemes at the level of unlinkability. As the global measure $D_{\overrightarrow{sys}}$ decreases, the unlinkability of the cancellable template increases.

Figure 6.
Revocability analysis

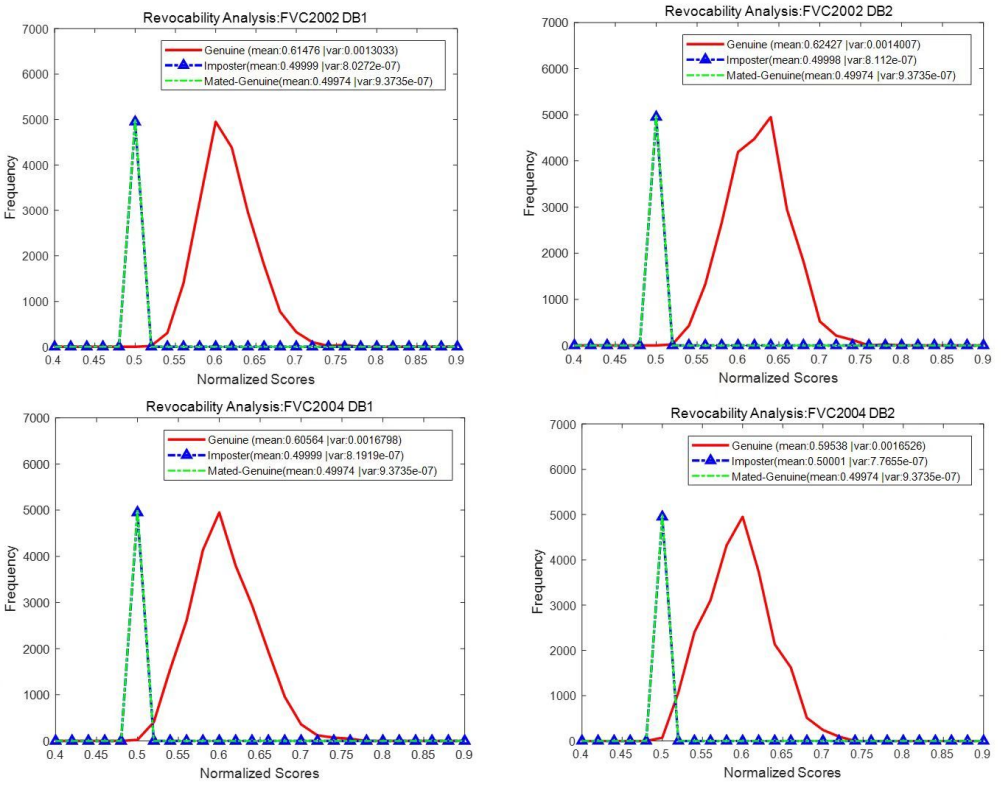
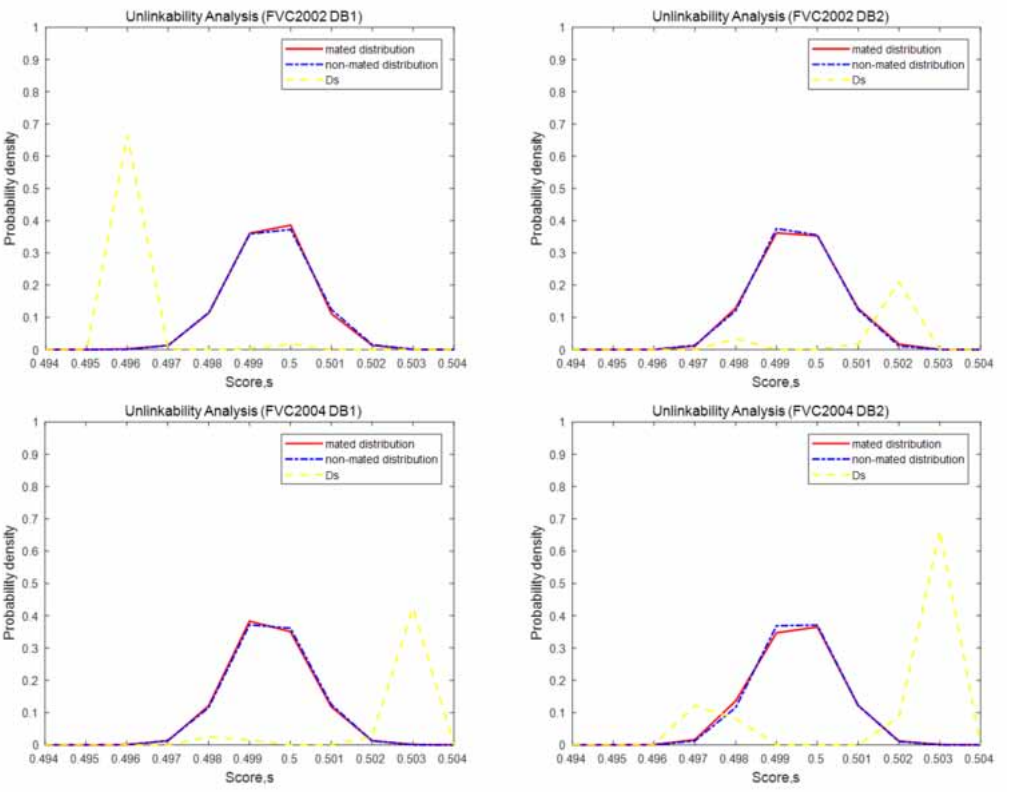


Figure 7.
Unlinkability analysis



The unlinkability analysis of the four databases from Figure 7 clearly shows that the mated/nonmated sample score distribution curves overlap, indicating that cancellable templates are indistinguishable from the same user or not. Therefore, the indexed self-coding (ISC) algorithm satisfies the unlinkability criterion. Table 3 lists the detailed values of $D_{\overrightarrow{sys}}$ for all test datasets of ISC and other one-factor schemes for comparison. It serves to show that the $D_{\overrightarrow{sys}}$ value of the ISC algorithm on the database is smaller, indicating that the ISC algorithm has higher unlinkability and better security.

Security Analysis

Brute force attack (Najafabadi et al., 2014) is an attack method about safety, which is to brute force the query cancellable template used by a user to match by enumeration. For the ISC algorithm, the

Table 3.
Unlinkability based on global variable $D_{\overrightarrow{sys}}$

Methods	FVC2002DB1	FVC2002DB2	FVC2004DB1	FVC2004DB2
EFV Hashing (Lee et al., 2018)	0.0404	0.0473	0.0465	0.0459
WSE Hashing (Kong et al., 2021)	0.0257	0.0235	0.0271	0.0250
FE Hashing (Zhang et al., 2021)	0.0226	0.0288	0.0266	0.0215
ISC	0.0146	0.0142	0.0142	0.0182

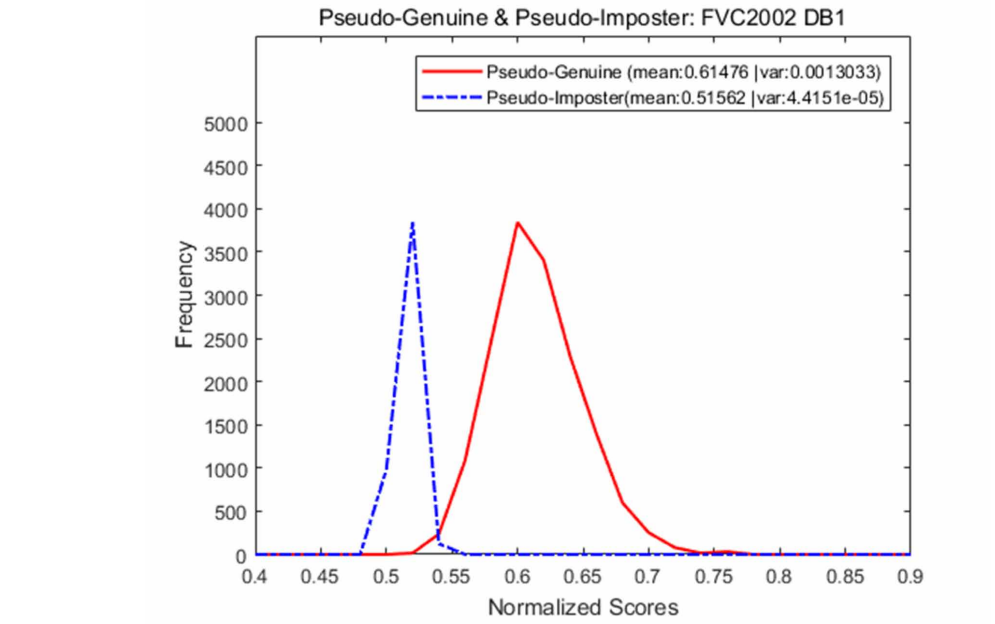
length of the cancellable template is $hn = 256 \times 1200 = 307200$ bits, and it takes $2^{hn} = 2^{307200}$ enumerations to crack the cancellable template. Therefore, the calculation result of 307200 is obviously impracticable.

A falsely accepted attack is more critical and practical than a brute force attack. It is to illegally access the system through fewer attempts. Here, the authors employ an authentication system; if the matching score exceeds the threshold τ set by the system, the authentication system will allow access. Taking the fingerprint database FVC2002 DB1 as a verification example, the parameter values are $n = 1200$ and $k = 2$ for the experiment. Figure 8 illustrates that the threshold $\tau = 0.54$, which reveals that the minimum number of matches for falsely accepted attack attempts is $hn\tau = 256 \times 1200 \times 0.54 = 165888$. Therefore, the complex rate of the falsely accepted attack is calculated as $2^{hn\tau} = 2^{165888}$. Although the complexity is nearly half that of a brute force attack, this number is still difficult to achieve in practice.

CONCLUSION

This paper proposes a one-factor cancellable biometric template protection scheme. In this scheme, binary biometrics are used as the only input, avoiding some safety issues attributed to the introduction of external factors in the two-factor cancellable biometric template protection algorithm. The one-factor cancellable biometric template protection algorithm based on index self-encoding (ISC) is proposed in this paper. The theory and experiments reveal that the method of index self-encoding can assure the accuracy of biological characteristics and further improve the safety of the program while comparing other scheme identification rates. The ISC algorithm also satisfies noninvertibility, revocability, and unlinkability. For future work, the current one-factor scheme only involves fingerprint biometrics, and the authors can try to protect the biometric template after fusion while ensuring that the security and recognition rate of the algorithm are guaranteed.

Figure 8.
Genuine-imposter matching



ACKNOWLEDGMENT

Funding Agency

This research was supported by the National Natural Science Foundation of China under Grant 61372137, and the Natural Science Foundation for the Higher Education Institutions of Anhui Province under Grant No. 2022AH050091. Huabin Wang is the corresponding author. E-mail address: wanghuabin@ahu.edu.cn.

REFERENCES

Abe, N., Yamada, S., & Shinzaki, T. (2015). Irreversible fingerprint template using Minutiae Relation Code with Bloom Filter. *IEEE International Conference on Biometrics Theory* (pp.1-7). IEEE. doi:10.1109/BTAS.2015.7358770

Aydar, M., & Ayvaz, S. (2019). An improved method of locality-sensitive hashing for scalable instance matching. *Knowledge and Information Systems*, *58*(5), 275–294. doi:10.1007/s10115-018-1199-5

Bolle, R. M., Connell, J. H., & Ratha, N. K. (2002). Biometrics perils and patches. *Pattern Recognition*, *35*(12), 2727–2738. doi:10.1016/S0031-3203(01)00247-3

Cappelli, R., Ferrara, M., & Maltoni, D. (2010). Minutia cylinder-code: A new representation and matching technique for fingerprint recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(12), 2128–2141. doi:10.1109/TPAMI.2010.52 PMID:20975113

Cappelli, R., Maio, D., Maltoni, D., Wayman, J. L., & Jain, A. K. (2006). Performance evaluation of fingerprint verification systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(1), 3–18. doi:10.1109/TPAMI.2006.20 PMID:16402615

Charikar & Moses, S. (2002). Similarity estimation techniques from rounding algorithms. *Thiry-fourth Acm Symposium on Theory of Computing* (pp. 380-388). ACM.

Datar, M., Immorlica, N., Indyk, P., & Mirrokni, V. (2004). Locality sensitive hashing scheme based on p-stable distributions. *Scg '04 Proceedings of the Twentieth Annual Symposium on Computational Geometry, 34*(2), 253-262. Princeton University.

Dong, Y., Zhang, S., Xu, J., Wang, H., & Liu, J. (2022). Random Forest Algorithm Based on Linear Privacy Budget Allocation. [JDM]. *Journal of Database Management*, *33*(2), 1–19. doi:10.4018/JDM.309413

Gomez-Barrero, M., Galbally, J., Rathgeb, C., & Busch, C. (2017). General framework to evaluate unlinkability in biometric template protection systems. *IEEE Transactions on Information Forensics and Security*, *13*(6), 1406–1420. doi:10.1109/TIFS.2017.2788000

Jain, A. K., Nandakumar, K., & Nagar, A. (2008). Biometric template security. *EURASIP Journal on Advances in Signal Processing*, 1–17.

Jin, Z., Hwang, J. Y., Lai, Y. L., Kim, S., & Teoh, A. B. J. (2018). Ranking based locality sensitive hashing enabled cancelable biometrics: Index-of-max hashing. *IEEE Transactions on Information Forensics and Security*, *13*(2), 393–407. doi:10.1109/TIFS.2017.2753172

Jin, Z., Lim, M. H., Teoh, A. B. J., Goi, B. M., & Yong, H. T. (2017). Generating fixed-length representation from minutiae using kernel methods for fingerprint authentication. *IEEE Transactions on Systems, Man, and Cybernetics. Systems*, *46*(10), 1–14.

Kirchgasser, S., Uhl, A., Lai, Y. L., & Jin, Z. (2019). Finger-Vein Template Protection based on Alignment-Free Hashing. *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)* (pp.1-10). IEEE.

Kong, X. J., Li, X. J., & Jin, Z. (2021). One-factor Cancellable Biometrics Verifification Scheme. *Acta Automatica Sinica*, *47*(5), 1159–1170.

Lai, Y. L., Jin, Z., Teoh, A., Goi, B. M., Yap, W. S., & Chai, T. Y. (2017). Cancellable iris template generation based on indexing-first-one hashing. *Pattern Recognition*, *64*(10), 105–117. doi:10.1016/j.patcog.2016.10.035

Lee, M. J., Jin, Z., & Teoh, A. B. J. (2018). One-factor Cancellable Scheme for Fingerprint Template Protection: Extended Feature Vector (EFV) Hashing. *IEEE International Workshop on Information Forensics and Security* (pp. 1-7). IEEE. doi:10.1109/WIFS.2018.8630782

Li, H. J., & Wang, X. Y. (2022). One factor cancellable fingerprint scheme based on novel minimum hash signature and secure extended feature vector. *Multimedia Tools and Applications*, *81*(9), 13087–13113. doi:10.1007/s11042-022-12424-y

Lim, M. H., Teoh, A. B. J., & Toh, K. A. (2012). An efficient dynamic reliability-dependent bit allocation for biometric discretization. *Pattern Recognition*, *45*(5), 1960–1971. doi:10.1016/j.patcog.2011.11.011

Lin, C., Kunnathur, A. S., & Li, L. (2020). The Cultural Foundation of Information Security Behavior: Developing a Cultural Fit Framework for Information Security Behavior Control. [JDM]. *Journal of Database Management*, *31*(2), 21–41. doi:10.4018/JDM.2020040102

Maio, D., Maltoni, D., Cappelli, R., Wayman, J. L., & Jain, A. K. (2002). FVC2002: Second Fingerprint Verification Competition. *Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02).* IEEE Computer Society. doi:10.1109/ICPR.2002.1048144

Maio, D., Maltoni, D., Cappelli, R., Wayman, J. L., & Jain, A. K. (2004). *Fvc2004: third fingerprint verification competition*. Lecture Notes in Computer Science.

Najafabadi, M. M., Khoshgoftaar, T. M., Kemp, C., Seliya, N., & Zuech, R. (2015). Machine Learning for Detecting Brute Force Attacks at the Network Level. *IEEE International Conference on Bioinformatics & Bioengineering*. IEEE.

Patel, V. M., Ratha, N. K., & Chellappa, R. (2015). Cancelable biometrics: A review. *IEEE Signal Processing Magazine*, *32*(5), 54–65. doi:10.1109/MSP.2015.2434151

Ratha, N. K., Connell, J. H., & Bolle, R. M. (2001). Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal*, *40*(3), 614–634. doi:10.1147/sj.403.0614

Rathgeb, C., & Uhl, A. (2011). A survey on biometric cryptosystems and cancelable biometrics. *Eurasip Journal on Information Security*, *2011*(1), 3. doi:10.1186/1687-417X-2011-3

Teoh, A. B. J., Goh, A., & Ngo, D. C. L. (2006). Random multispace quantization as an analytic mechanism for biohashing of biometric and random identity inputs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(12), 1892–1901. doi:10.1109/TPAMI.2006.250 PMID:17108365

Teoh, A. B. J., Ngo, D. C. L., & Goh, A. (2004). Biohashing: Two factor authentication featuring fingerprint data and tokenised random number. *Pattern Recognition*, *37*(11), 2245–2255. doi:10.1016/j.patcog.2004.04.011

Turesson, H. K., Kim, H., Laskowski, M., & Roatis, A. (2021). Privacy Preserving Data Mining as Proof of Useful Work: Exploring an AI/Blockchain Design. [JDM]. *Journal of Database Management*, *32*(1), 69–85. doi:10.4018/JDM.2021010104

Wang, S., Deng, G., & Hu, J. K. (2017). A partial hadamard transform approach to the design of cancelable fingerprint templates containing binary biometric representations. *Pattern Recognition*, *61*(17), 447–458. doi:10.1016/j.patcog.2016.08.017

Wang, S., & Hu, J. K. (2016). A blind system identification approach to cancelable fingerprint templates. *Pattern Recognition*, *54*(1), 14–22. doi:10.1016/j.patcog.2016.01.001

Wang, S., & Hu, J. K. (2017). Alignment-free cancelable fingerprint template design: A densely infinite-to-one mapping (DITOM) approach. *Pattern Recognition*, *45*(12), 4129–4137. doi:10.1016/j.patcog.2012.05.004

Yang, W., Wang, S., Kang, J. J., Johnstone, M. N., & Bedari, A. (2022). A linear convolution-based cancelable fingerprint biometric authentication system. *Computers & Security*, *114*, 102583–102596. doi:10.1016/j.cose.2021.102583

Zhang, L. P., Wang, H. B., & Tao, L. (2021). One-factor cancelable fingerprint template protection based on feature enhanced hashing. *Twelfth International Conference on Graphics and Image Processing, 11720*(17), 1-10. doi:10.1117/12.2589436

Yalan Feng received the B.S. degree in information security from Huaibei Normal University,in 2020. She is currently pursuing the M.S. degree in computer science and technology with Anhui University, Hefei, China. Her research interest includes biometric template protection.

Huabin Wang received the B.S. degree in computer science and technology from the Anhui University of Finance and Economics, China, in 2005, and the M.S. degree in signal and information processing and the Ph.D. degree in computer application technology from Anhui University, Hefei, China, in 2008 and 2011, respectively. He is currently an Associate Professor and the Director of the School of Computer Science and Technology, Anhui University. His research interests include biometric recognition, template protection, and medical image processing.

Dailei Zhang received the B.S. degree in communication engineering from Hefei University, Hefei, China, in 2016, and the the M.S. degree in computer science and technology from Anhui University, Hefei, China, in 2022. His research interests include signal and image processing.

Jiahao Li is currently pursuing the B.S. degree in computer science and technology with Anhui University, Hefei, China. His research interest includes biometric template protection.

Liang Tao received the B.S. degree in radio technology and the M.S. degree in circuit and system from Anhui University, Hefei, China, in 1985 and 1988, respectively, and the Ph.D. degree in information and communication engineering from the University of Science and Technology of China, Hefei, in 2013. He is currently a Professor with the School of Computer Science and Technology, Anhui University. His research interests include pattern recognition, intelligent information processing, and multimedia signal processing. He was the Chair for a number of National Natural Science Foundation projects, the Anhui Provincial Natural Science Foundation Project, and the key research projects of Natural Science in the Anhui Provincial Department of Education. He is the Academic and Technical Leader of Anhui Province.