Journal of Information Hiding and Multimedia Signal Processing ©2019 ISSN 2073-4212
Ubiquitous International Volume **10**, Number **1**, January **2019**

Template Protection based on Chaotic Map and DNA Encoding for Multimodal Biometrics at Feature Level Fusion

Zhifang Wang, Jinjin Dong, Jiaqi Zhen and Fuzhen Zhu

Department of Electronic Engineering
Heilongjiang University
No.74, Xufu Road,Nangang District, Harbin
Corresponding author: xiaofang_hq@126.com

Received October 2017; revised September 2018

Abstract. *Multimodal biometrics becomes the current development trend of the biometric recognition technology because of its better applicability, higher security and better performance. Compared with unimodal biometric templates, multimodal biometric template contains more sensitive information. So the safety of the multimodal biometric template needs to be more concerned. This paper proposes a novel template protection algorithm based on chaotic map and DNA encoding for multimodal biometric at feature level fusion. Each original multimodal biometrics template is permuted by different chaotic sequence produced by the master key and user identification number. Then DNA encoding and complementary substitutions are imported to further confuse the information of the template. Experiment results show that the proposed algorithm significantly improves the original system performance of the classic feature level fusion methods and ensures the security of multimodal biometric template.*

Keywords: Multimodal biometrics, Template protection, Chaotic map, DNA coding, Feature level

1. **Introduction.** Biometric recognition has been developed as a reliable and efficient technology for person authentication, which relies on person physiological or behavioral characteristics such as face, iris, voice and signature. However, with the wide-spread use of biometric recognition, some problems of the traditional unimodal biometric have been appeared such as large intra-class variations, non-universality and spoofing attacks. The multimodal biometrics, which can overcome most of these above defects, has been developed rapidly as a new development trend of biometric recognition technology.

Because of the uniqueness of the biometrics features, the biometric template should not been stored directly. If the attackers steal the biometrics template, they can use it directly to other authentication systems or create a spoofing sample to access the biometric system. Comparing with unimodal biometrics, multimodal biometric template is more sensitive. Once the multimodal biometric template is leaked, it will bring more security problems. So it is more crucial to protect the multimodal biometric template. On the base of unimodal biometrics, multimode biometric template algorithms are also proposed [1, 2, 3, 4, 5]. Fuzzy vault and fuzzy commitment are two well-known methods. Rathgeb [1] proposed a reliability-balanced feature level fusion for fuzzy commitment scheme which binds cryptographic keys with iris feature vector to protect multimodal biometric template. Wang [2] presented a multimodal biometric template protection scheme based on fuzzy commitment and chaotic system. Nagar [3] implemented the feature-level fusion framework using fuzzy vault and fuzzy commitment. However, fuzzy vault is usually applicable to the biometric features in point set format, while the biometric features are mostly vector format. Fuzzy commitment usually combines with error correcting codes to solve the contradiction between the ambiguity of biometric features and the accuracy of cryptography. Each matching process requires encoding or decoding, so the computation and the efficiency need to be considered.

Here, this paper proposes a novel multimodal biometrics template protection approach based on chaotic map and DNA encoding at feature level fusion. Each original multimodal biometric template is permuted by different chaotic sequence produced by the master key and user identification number(ID). Then DNA encoding is imported to confuse the information of the template. Finally the decimal vector is restored as the final template in the database. The key characteristics of our algorithm are as follows: (1) The proposed algorithm is applicable to the classic feature level fusion multimodal biometrics such as series rule, weighted sum rule and combined Fisherface; (2) The template is cancelable and renewable; (3) Our algorithm significantly improves the performance of multimodal biometric system.

The remainder of the paper is organized as follows: Section 2 introduces the proposed algorithm in preprocessing on logistic map and DNA encoding. The proposed template protection algorithm is described in section 3. Section 4 shows the experiments and security analysis. At last, Section 5 concludes this paper.

2. **The preprocessing.**

2.1. **Generalized logistic map.**

The logistic map as a typical chaotic map attracts much attention in image encryption. It produces fundamental results on non-linear dynamics. A logistic map is defined as follows:

$$x_{n+1} = \mu x_n(1 - x_n) \tag{1}$$

Where n is a non-negative integer and represents the dimension of the logistic map, $x_0 \in [0, 1]$ is the initial value of the logistic map, and μ is a control parameter. If μ is different, the logistic sequence x_n is different. Fig.1 shows the bifurcation diagram of the logistic map while $\mu \in (0, 4]$. We can see that the logistic sequence displays chaotic phenomenon when $3.57 < \mu \leq 4$.

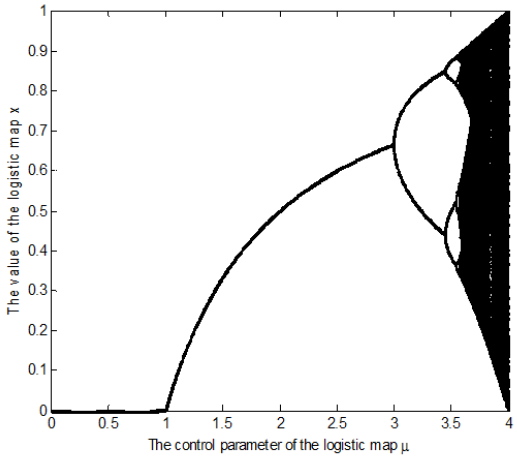
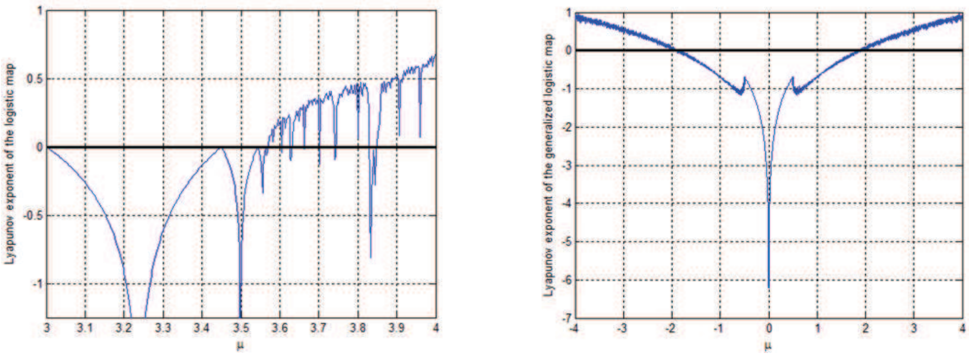


FIGURE 1. The bifurcation diagram of the logistic map

For more clarity, we give the Lyapunov exponent curve in fig.2. When the value of Lyapunov exponent is positive number, the system is chaotic. It is noted that the logistic map has some drawbacks such as non-uniform behavior and blank windows in the chaotic region displayed. In fig.2 (a), there are some areas where the Lyapunov exponent is either zero or negative. To overcome the issue of the logistic map, we introduce a generalized logistic map [6] defined as follows:

$$x_{n+1} = \frac{4\mu^2 x_n(1-x_n)}{1+4(\mu^2-1)x_n(1-x_n)} \tag{2}$$

where $-4 \leq \mu \leq 4$. Fig. 2(b) shows the Lyapunov exponent of the generalized logistic map generated by eq.(2). It can be sure that the logistic sequence is chaotic if we just choose the values of μ from these two intervals: $-4 \leq \mu \leq -2$ and $2 \leq \mu \leq 4$.



(a) Logistic map

(b) Generalized logistic map

FIGURE 2. Lyapunov exponent curves of logistic map and generalized logistic map

2.2. **DNA encoding and complementary substitution.** DNA encoding is commonly used in image encryption. This paper uses it to further confuse the original template information. The full name of DNA is deoxynucleotide, which is the main chemical component of chromosomes. Each DNA sequence contains four bases: adenine (A), thymine (T), guanine (G), and cytosine (C). Here, A and T, C and G are two complementary pairs. Each pixel value of gray digital image can be represented by a binary sequence of 8 bits [7], where 0 and 1 are complementary, 00 and 11, 01 and 10 are also complementary. If we use four bases A, T, G and C to represent the binary numbers 00, 11, 01 and 10 respectively, then each pixel can be encoded into a string of nucleotides. For example, the binary sequence of the gray value 225 is 11100001, the corresponding nucleotides string is TCGA. The above procedure is performed according to rule 1 in table 1. There are 24 types of combinations for the four nucleotides. However, only 8 coding combinations are suitable for the principle of complementarity. These rules are summarized in table 1.

The biometric features are usually not integers, not even positive numbers, which do not belong to the range $[0, 255]$. So we should preprocess the biometric feature. Firstly, each element in the biometric feature vector should be normalized to the range $[0, 1]$ and expanded to the range $[0, 255]$. Eq.(3) converts x to $y \in [0, 1]$, and eq.(5) expands y to the integer $z \in [0, 255]$.

$$[y, PS] = mapminmax(x, 0, 1) \tag{3}$$

$$z = round(y \times 255) \tag{4}$$

Table 1. The rules of DNA encoding

base	rule 1	rule 2	rule 3	rule 4	rule 5	rule 6	rule 7	rule 8
A	00	00	01	01	10	10	11	11
T	11	11	10	10	01	01	00	00
C	01	10	00	11	00	11	01	10
G	10	01	11	00	11	00	10	01

where *mapminmax*() is a normalized function of MATLAB. Then the biometric feature can be encoded according to DNA coding rules. It is easy to know that the inverse transformation of eq.(3) and eq.(4) are as follows:

$$y = round(z \div 255) \tag{5}$$

$$x = mapminmax('reverse', y, PS) \tag{6}$$

Because normalization changes the value of biometric features, we take face recognition and palm recognition based on principal component analysis(PCA) as the examples to verify whether the above process affects the recognition performance. Fig.3 shows the DET curve comparison with or without DNA encoding. It can be found that two curves are overlapped. So this step does not affect the recognition performance of biometric system though it changes the value of biometric template.

Meanwhile, in order to improve the security of biometric template, we further perform complementary substitution. Suppose x_i is the ith users template after DNA encoding, and x_i^j is the jth dimension of x_i. So x_i^j is a sequence constructed by four nucleotide bases, then in accordance with the principle of the complementary base, we set the nucleotide string x_i of the encoding nucleotides as follows:

$$\begin{cases} x_i^j \neq D(x_i^j) \neq D(D(x_i^j)) \neq D(D(D(x_i^j))) \\ x_i^j = D(D(D(D(x_i^j)))) \end{cases} \tag{7}$$

where $D(x_i^j)$ is the complementary of x_i^j. These base pairs need to meet the conditions of injective mapping. There are 6 types of rational complementary combinations of base pairs according to eq.(7) in table 2 as follows:

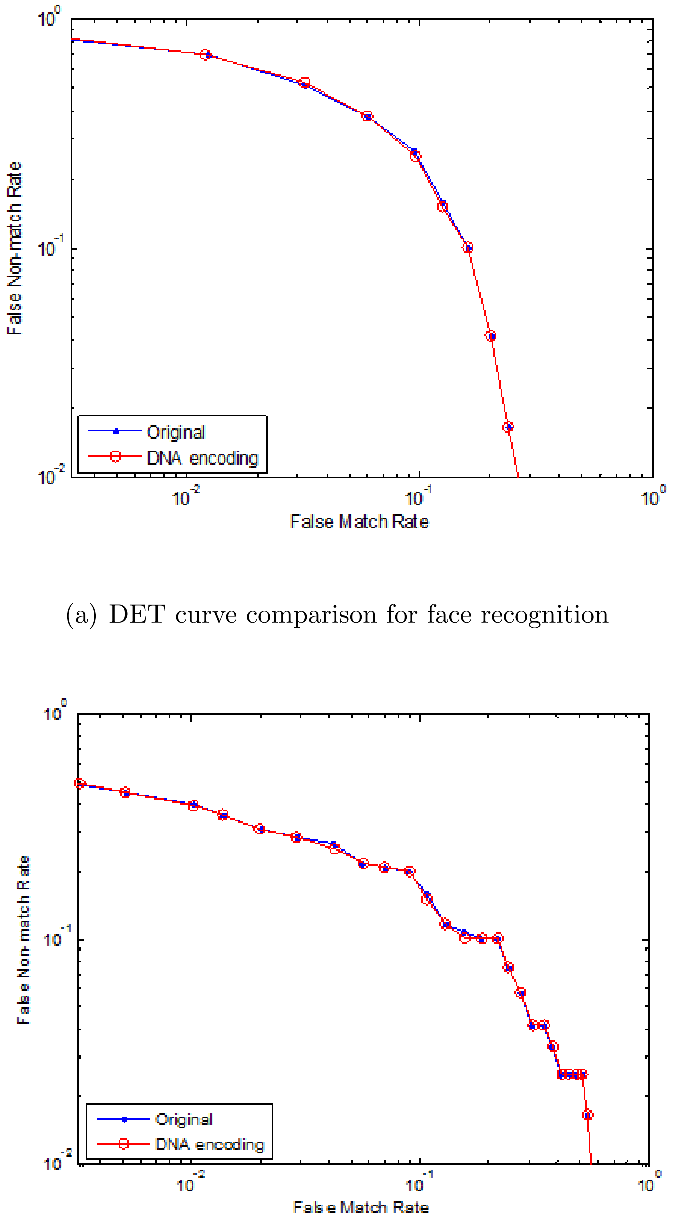
Table 2. The types of complementary substitutions of base pairs

type 1	(AT)(TC)(CG)(GA)
type 2	(AT)(TG)(GC)(CA)
type 3	(AC)(CT)(TG)(GA)
type 4	(AC)(CG)(GT)(TA)
type 5	(AG)(GT)(TC)(CA)
type 6	(AG)(GC)(CT)(TA)

The complementary substitution is cyclic. For example, in type 1, (AT) means T is the complementary of A. the final pair (GA) means A is the complementary of G. This type is a closed ring.

3. **Proposed algorithm.**

In this section, we present our template protection algorithm for multimodal biometrics recognition using chaotic map and DNA encoding. In this procedure, generalized logistic map is used to permute the original fusion template. Then DNA encoding further diffuses the information of permuted biometric template. Fig.4 shows the flow of template generation. The detail steps are described as follows:



(a) DET curve comparison for face recognition

(b) DET curve comparison for palm recognition

Figure 3. Performance comparison with or without DNA encoding

(1) The parameter initialization: our algorithm needs five parameters: logistic mapping requires two, and DNA encoding requires three. Section 2.1 shows that the initial value $x_0 \in [0,1]$ and the control parameter $\mu(-4 \leq \mu \leq -2$ or $2 \leq \mu \leq 4)$ are necessary to generate the logistic sequence. For DNA encoding, two different coding rules $r_1(1 \leq r_1 \leq 8)$ and $r_3(1 \leq r_3 \leq 8)$ are requested in table 1. Meanwhile, we also need to choose a complementary type $r_2(1 \leq r_1 \leq 6)$ from table 2. In order to generate the above five parameters at a time, our algorithm encodes the master key and user ID into a vector satisfied the input request of hash function. We can take five parameters from the output of hash function according to their respective values.

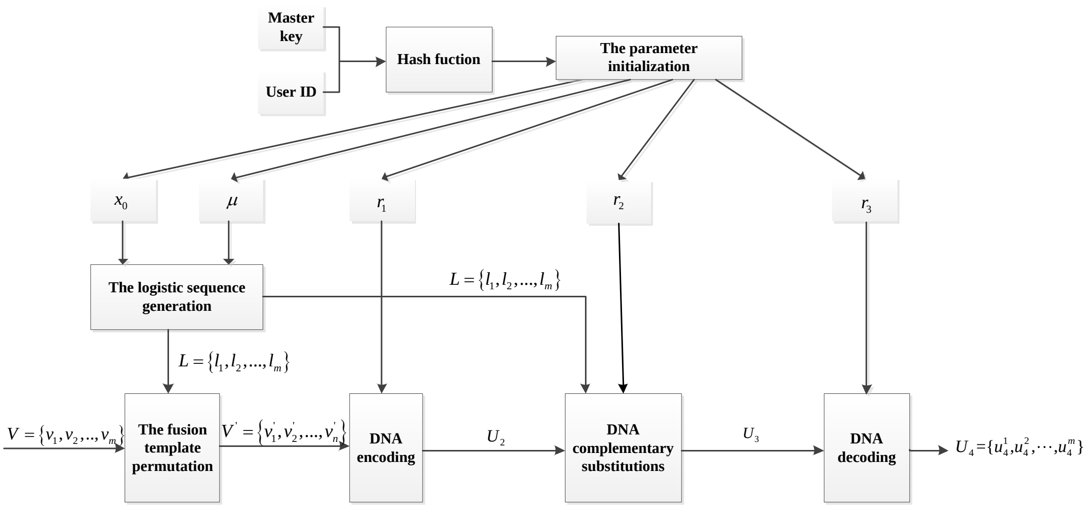

Master key
User ID
Hash fuction
The parameter initialization
x_0
μ
r_1
r_2
r_3
The logistic sequence generation
$L=\{l_1,l_2,...,l_m\}$
$L=\{l_1,l_2,...,l_m\}$
$V=\{v_1,v_2,..,v_m\}$
The fusion template permutation
$V^{'}=\{v_1^{'},v_2^{'},...,v_n^{'}\}$
DNA encoding
U_2
DNA complementary substitutions
U_3
DNA decoding
$U_4=\{u_4^1,u_4^2,\cdots,u_4^m\}$

Figure 4. The flow of template generation

(2) The logistic sequence generation: With the initial value x_0 and the control parameter μ, the logistic sequence can be generated according to eq.(2). Because different user has different ID, the initial conditions of the logistic map are not the same, then the logistic sequences are different. Certainly, this step can only use master key to generate the same logistic map for all users' templates. We also give corresponding experimental results in subsequent section. It can increase the difficulty of deciphering that different user use different logistic sequence.
(3) The fusion feature permutation $(V \to V')$**:** suppose $V = \{v_1, v_2, \cdots, v_m\}$ be a multimodal biometric fusion feature where m is the dimension, the dimension of logistic sequence is often greater than m. We select the former m dimension of the logistic sequence gained by step (2) as the final used chaotic sequence $L = \{l_1, l_2, \cdots, l_m\}$. Then sort $L = \{l_1, l_2, \cdots, l_m\}$ in ascending order of size and obtain the new sequence $L' = \{l'_1, l'_2, \cdots, l'_m\}$ where $l'_1 < l'_2 < \cdots < l'_m$. Meanwhile, the substitution index $S = \{s_1, s_2, \cdots, s_m\}$ is produced where s_i is the position in L of l'_i. Then $V = \{v_1, v_2, \cdots, v_m\}$ is replaced by $V' = \{v'_1, v'_2, \cdots, v'_m\}$ according with S. So $V' = \{v'_1, v'_2, \cdots, v'_m\}$ is one users permuted biometric feature.
(4) DNA encoding$(V' \to U_1 \to U_2)$**:** before DNA encoding, the permuted biometric feature should be firstly normalized to satisfy the demand of DNA encoding. We take $V' = \{v'_1, v'_2, \cdots, v'_m\}$ as an example. For each dimension of V' , eq.(3) and eq.(4) in section 2.2 are performed. Let $U_1 = \{u_1^1, u_1^2, \cdots, u_1^m\}$ be the normalized result of V' in which u_1^i is corresponding to v'_i, we convert U_1 into the DNA sequence U_2 according to the coding rule r_1 . So each dimension of U_2 is presented by four nucleotide bases.
(5) DNA complementary substitutions$(U_2 \to U_3)$**:** the logistic sequence $L = \{l_1, l_2, \cdots, l_m\}$ generated in step (2) is used to form the iteration number of DNA complementary substitutions. Suppose $C = \{c_1, c_2, \cdots, c_m\}$ represent the final number of iterations, the ith dimension c_i of C is computed as shown in eq.(8).

$$c_i = fix(mod(l_i \times 10, 4)) \qquad (8)$$

where $fix(x)$ is a function which rounds the elements of x to the nearest integer towards zero, and $mod()$ is the module operation. Then r_2 generated in step(1) decides the type of complementary base pairs shown in table 2. Finally, according to

c_i , u_3^i in the complementary sequence U_3 of the DNA sequence U_2 can be obtained by u_2^i as follows:

$$\begin{cases} u_3^i = u_2^i, & if \quad c_i = 0; \\ u_3^i = D(u_2^i), & if \quad c_i = 1; \\ u_3^i = D(D(u_2^i)), & if \quad c_i = 2; \\ u_3^i = D(D(D(u_2^i))), & if \quad c_i = 3. \end{cases} \tag{9}$$

(6) DNA decoding ($U_3 \to U_4$)**:** because the final template is stored in a decimal form in the database, we need to perform DNA decoding. r_3 obtained in step(1) decides the rule of DNA decoding, we convert u_3^i into a binary sequence according to the rule r_3 and further transform into a decimal number u_4^i . So $U_4 = \{u_4^1, u_4^2, \cdots, u_4^m\}$ is the final template of the original biometric fusion feature $V = \{v_1, v_2, \cdots, v_m\}$. The above six steps are the total procedure of template generation and performed on the fusion features to obtain the final templates of all users.

Before matching, the final templates should be transformed as the inverse process of template generation. The users template U_4 is converted into a binary sequence and further transformed into a DNA coding sequence U_3' according to the coding rule r_3. DNA complementary substitution is a cyclic operation. According to the complementary type r_2 , U_2' is obtained from U_3' . After DNA decoding according to r_1, U_1' is the corresponding decimal number of U_2' . Then we perform eq.(5) and eq. (6) which are the inverse transformation of eq. (4) and eq. (3) and obtain the real number sequence V" . According to the substitution index $S = \{s_1, s_2, \cdots, s_m\}$, we perform inverse permutation on V" and get T to match with the testing fusion features. All of the above steps are for one users template. Each user's template needs to be executed.

4. Experimental Results And Analysis.

4.1. Feature extraction and fusion method.

This paper selects face and palm as two distinct biometric characteristics to test our algorithm because they are easy to capture and register comparing with other biometrics. The experiments were performed on ORL face database and PolyU palm database. ORL face database includes 40 people, 10 different images with pose and expression variation per person. PolyU multispectral palm images were collected from 250 volunteers, 6 different images for each palm. In order to fuse two feature sets, the number of samples should be coordinated. Our solution is to select 40 classes in which three samples per class were selected as the training sets, three samples for testing. As one of the classical methods of subspace learning, principal component analysis (PCA) is used in our algorithm to extract features of face and palm and unify the dimension of fusion features. To further eliminate the differences in the order of magnitude and the distribution between two distinct feature sets, we use the z-score model to normalize two feature sets before fusion.

Multimodal biometric technology is divided into four levels: pixel level, feature level, score level and decision level. Comparing with other three fusion levels, feature level can reduce the redundant information to avoid calculation con-sumption, and simultaneously acquire the discriminative information to improve the system performance. In general, there are two basic modes for feature level fusion: serial rule and weighted sum rule [8]. The former connects two feature vectors into a longer fusion vector. This rule consumes large computational resources. For weighted sum rule, the fusion feature is the sum of the two unimodal features multiplied by the respective weighted value. In our experiments, we set three weighted ratios for face feature and palm feature: 7 : 3, 3 : 7 and 5 : 5. Besides, Yang [9] proposed the combined Fisherface method which takes two feature vectors as real part and imaginary part of a complex vector. Suppose a and b be two feature vectors derived from two biometric modes respectively, the fusion feature is obtained as $f = a+ib$ where i denotes imaginary unit. So our algorithm performed on five fusion methods: series rule, three weighted sum rules and combined Fisherface.

4.2. **Experimental results and analysis.** Although there is no uniform standard to measure the performance of template protection algorithm, an ideal template protection algorithm should satisfies four conditions [10]: performance, diversity, restorability and security. This section introduces our experimental results from the above four aspects.

(1) Performance: The template protection algorithm cannot reduce the performance of the original multimode biometric system. Our algorithm performed the experiments on five fusion methods: series rule, three weighted sum rule and combined Fisherface. Section 2.1 describes two cases: different users use different logistic maps, and all users use the same logistic map. So, for each fusion methods, we compare three cases: the original system without template protection (the original system), the system with template protection using same logistic map (same logistic), and the system with template protection using different logistic map (the proposed algorithm). Fig.5 shows the DET curves of five fusion methods respectively. From fig.5, we find that the two DET curves of the original system and same logistic basically overlap. And the curve of the proposed algorithm is the lowest of three curves in five fusion methods. It displays that the method of same logistic basically does not affect the performance of the original system, and the method of the proposed algorithm even improve the original system performance. Meanwhile, table 3 shows the comparison of EER in three cases. It is easier to see from the exact data of table 3 that the proposed algorithm significantly improves the performance of the original system.

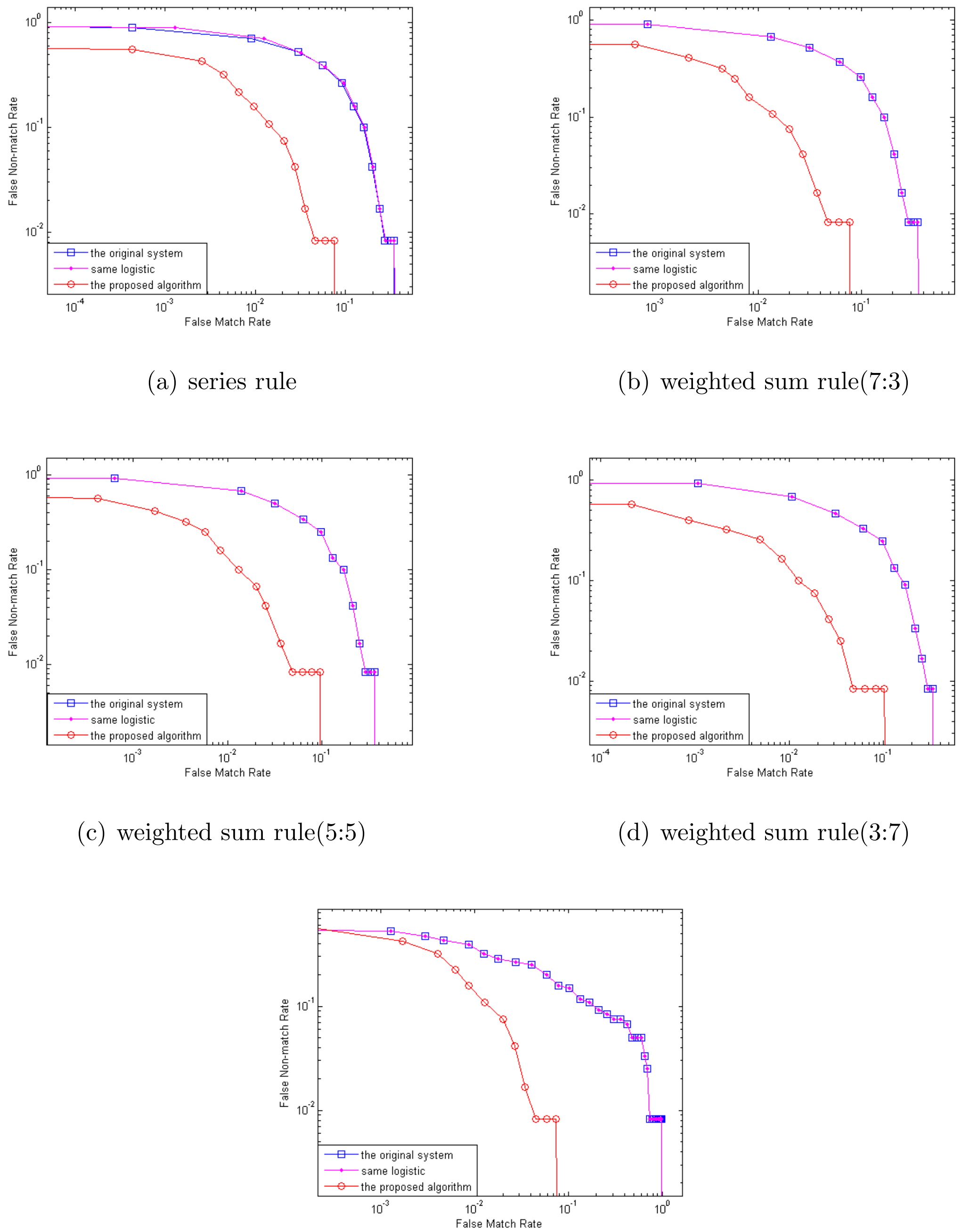
Table 3. EER comparison of five fusion methods(%)

Methods	the original system	same logistic	the proposed algorithm
Series rule	13.64	13.80	3.10
Weighted sum rule(7:3)	13.93	13.93	3.12
Weighted sum rule(5:5)	13.15	13.16	3.04
Weighed sum rule(3:7)	13.20	13.18	3.14
Complex Fisherface	12.53	12.51	3.00

(2) Diversity: a security template do not allow cross-matching between multiple databases. For this requirement, we can change the five initial variables(x_0, μ, r_1, r_2, r_3 to obtain different final template. It can be ensure that the different database use the different template.

(3) Restorability: the system can easily cancel the template and regenerate a new template once the template is leaked. As long as we change the five initial variables, we can generate a new template to replace the leaked template.

(4) Security: A security template must ensure that the original biometric information cannot be leaked from itself. That is to say, an attacker is prevented from being able to reverse the biometric feature even if he gets a biometric template, and cannot forge a user's biometric sample. The use of the hash function in the proposed algorithm ensures that the master key will not be compromised. In addition, the security of the five initial parameters should be considered. We analyze the key space imitating the image encryption algorithm. The greater the key space, the algorithm will be more secure.



(a) series rule

(b) weighted sum rule(7:3)

(c) weighted sum rule(5:5)

(d) weighted sum rule(3:7)

(e) complex Fisherface

FIGURE 5. DET curves of five fusion methods

In our experiment, there are 40 users, and the fusion feature of each person has been permuted by different logistic map. Meanwhile the logistic map is also used to form the number of iterations of DNA complementary substitutions. So, the logistic map has been used 41 times in total in our algorithm. The sensitivities to the initial values and parameters of the logistic map are both considered to be 10^{-16} [11]. Therefore, we can define the key space of the initial values of the logistic map: $S_{x_0} = 10^{16}$. For the variation of the parameters, $\mu \in (3.6, 4]$ and $S_\mu = 0.5 \times 10^{16}$. There are only 8 kinds of DNA coding rules, and there are 6 kinds of complementary types. Therefore, the key space of the random integers is $S_{r_1} = S_{r_3} = 8$ and $S_{r_2} = 6$. The total key space is:

$$S = (10^{16} \times 0.5 \times 10^{16})^{41} \times 8 \times 6 \times 8 \approx 1.746 \times 10^{1302} \tag{10}$$

Therefore, the encryption algorithm has a sufficiently large key space to repel all kinds of brute-force attacks.

5. **Conclusion.** In this paper, we propose a novel multimodal biometric template protection approach based on chaotic map and DNA coding at feature level fusion. The original multimodal biometric fusion features is permuted by different chaotic sequence produced by the master key and user identification number. Then DNA encoding and complementary substitutions are imported to further confuse the information of the permuted fusion features. We take face and palm as two different experimental objections to construct the fusion features through three classic fusion methods: series rule, weighted sum rule and combined Fisherface. The final template can be diverse and renewable by adjusting the initial parameters. Meanwhile, experimental results show our algorithm significantly improves the performance of the original multimodal biometric system and ensures the security of multimodal template.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (No.61201399,61501176,61601174) and Startup Fund for Doctor of Heilongjiang University.

REFERENCES

[1] C. Rathgeb, A. Uhl, P. Wild. Reliability-balanced feature level fusion for fuzzy commitment scheme, *Proc. of International Joint Conference on Biometrics*, pp. 1-7, 2011.

[2] N. Wang, Q. Li, A. A. Abd El-Latif. A novel template protection scheme for multibiometrics based on fuzzy commitment and chaotic system, *Signal image and Video processing*, vol. 9, vol. 1, pp. 99-109, 2015.

[3] A. Nagar, K. Nandakumar, A. Jain. Multibiometric cryptosystems based on feature-level fusion, *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 1 , pp. 255-268, 2012.

[4] O. Nafea, S. Ghouzali, W. Abdul, E. U. Qazi. Hybrid Multi-Biometric Template Protection Using Watermarking, *Computer Journal*, vol. 59, no. 9, pp. 1392-1407, 2016.

[5] M. Gomez-Barrero, E. Maiorana, J. Galbally, P. Campisi, J.Fierrez, Multi-biometric template protection based on Homomorphic Encryption, *Pattern Recognition*, vol. 67, no. C, pp. 149-163, 2017.

[6] X. Song, S. Wang, A. Ahmed A. El-Latif, X. Niu. Quantum image encryption based on restricted geometric and color transformations, *Quantum Information Processing*, vol. 13, no. 8, pp. 1765-1787, 2014.

[7] Q. Zhang and X. Wei. A novel couple images encryption algorithm based on DNA subsequence operation and chaotic system, *International Journal for Light and Electron Optics*, vol. 124, no. 23, pp. 6276-6281, 2013.

[8] Q. Zhang, Y. L. Yin, D. C. Zhan, J. L. Peng. A novel serial multimodal biometrics framework based on semisupervised learning techniques, *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 10, pp. 1681-1694, 2014.

[9] J. Yang, J. Y. Yang, A. F. Frangi. Combined Fisherfaces framework, *Image and Vision Computing*, vol. 21, no. 12, pp.1037-1044, 2003.

[10] D. Maltoni, D. Maio, A. K. Jain. Hand Book of Fingerprint Recognition, *New York: Springer Verlag*, 2003.

[11] Y. Liu, J. Tang, and T. Xie. Cryptanalyzing a RGB image encryption algorithmbased on DNA encoding and chaosmap, *Optics and Laser Technology*, vol. 60, no. 5, pp. 111-115, 2014.