Hindawi
Security and Communication Networks
Volume 2018, Article ID 5783976, 12 pages
https://doi.org/10.1155/2018/5783976

Research Article

Multibiometric Fusion Authentication in Wireless Multimedia Environment Using Dynamic Bayesian Method

Zhendong Wu,[1] **Jiajia Yang**,[2] **Jianwu Zhang**,[2] **and Hengli Yue**[1]

[1]*School of Cyberspace, Hangzhou Dianzi University, Hangzhou, Zhejiang, China*
[2]*School of Communication Engineering, Hangzhou Dianzi University, Hangzhou, Zhejiang, China*

Correspondence should be addressed to Zhendong Wu; wzd@hdu.edu.cn

Received 10 May 2018; Revised 28 August 2018; Accepted 10 September 2018; Published 18 November 2018

Guest Editor: Zhaoqing Pan

Single biometric method has been widely used in the field of wireless multimedia authentication. However, it is vulnerable to spoofing and limited accuracy. To tackle this challenge, in this paper, we propose a multimodal fusion method for fingerprint and voiceprint by using a dynamic Bayesian method, which takes full advantage of the feature specificity extracted by a single biometrics project and authenticates users at the decision-making level. We demonstrate that this method can be extended to more modal biometric authentication and can achieve flexible accuracy of the authentication. The experiment of the method shows that the recognition rate and stability have been greatly improved, which achieves 4.46% and 5.94%, respectively, compared to the unimodal. Furthermore, it also increases 1.94% when compared with general multimodal methods for the biometric fusion recognition.

1. Introduction

Biometric feature analysis has been widely studied for decades as it is a vital way for authentication and safeguard in computer vision. However, traditional biometrics, such as fingerprinting and vein recognition, gradually reveals some of its drawbacks that it can already be assigned and mimicked by forging fingerprints or faces [1]. As fusing features such as facial features, fingerprints, palm prints, sounds, and irises improves the stability, accuracy, and unforgeability of biometrics, multimodal biometric systems could help relieve the problem brought by the single-modal biometric systems and provide tremendous help for more secure authentication and identification.

There have been some researches about multimodal biometrics. Conti et al. [2] fused fingerprint and iris using homogeneous biometric vector through Log-Gabor filtering. Nagar et al. [3] studied the fusion of three biological feature (iris, fingerprint, and face) by using fuzzy vault and fuzzy commitment model to form a biometric encryption system framework. Snelick et al. [4] used a new method of normalization and fusion strategies to fuse and identify the biometrics of fingerprints and face at the score level. Muthukumar et al. [5] fused iris and fingerprint at the score level based on an evolutionary algorithm, Particle Swarm Optimization, which can help the authentication system adapt to different security needs. Shekhar et al. [6] used sparse matrices fusing the same three characteristics (iris, fingerprint, and face). Sparse matrix method has good recognition robustness. The above improvement in biometric identification demonstrates that there are many advantages of multimodal biometric identification.

On the other hand, we find some limitations about the existing researches, they are as follows: (1) the above articles all chose to integrate multibiometrics at a certain level but did not take into account the fact that multiple biometric features may interfere with each other, thereby reducing the recognition effect. (2) Most of the fusing at the decision layer always takes fixed weights. This is based on the overall average quality, but it is not the best solution for every decision. For example, if the fingerprint recognition rate is higher than the voiceprint overall, then it will be given a higher weight in the fusion recognition; however, the fingerprints are not always better than the voiceprint quality.

Inspired by these ideas, we propose a multibiometric fusion authentication solution by using dynamic Bayesian decision method, which is named MFDB-decision (Multibiometric Fusion using Dynamic Bayesian decision). The key idea of this work is that using matching layer score assists decision layer in fusing fingerprints and voiceprints aiming at recovering identity information lost in decision layer and, besides, overcoming the above problems caused by fixed weights. This paper uses fingerprint and voiceprint in multimodal fusion, because of the stability of the fingerprint and the high user acceptance of the voiceprint. The method proposed in this paper can be extended to more dimensional biometric fusion authentication, instead of being limited to the voiceprint and fingerprint.

The outline of this paper is organized as follows. Section 2 presents some related work. The preliminary research about fingerprint feature extraction and voiceprint feature extraction is introduced in Section 3. The multibiometric authentication fusion algorithm MFDB-decision is described in Section 4. The analysis of the experiments and results is given in Section 5. Finally, the paper would be concluded in Section 6.

2. Related Work

There has been a great deal of research on the application based on single biometric identification, especially in the fields of biological key [7–9], cloud computing data security [10–13], blockchain [14], privacy preserving [15–18], and biological template protection [19–21]. However, there are still not so many studies on multibiometrics. The study by Windsor Holden [22] further increased the application of multibiometric methods in the fields of common life other than criminal investigation.

Multimodal biometrics research attempts to overcome the shortcomings of single-modal biometrics in recognition accuracy, robustness, and flexibility and provides richer and more reliable biometrics applications. At present, the multimodal biometric system mainly focuses on the fusion extraction of multimode features at different levels to provide a unified data manipulation interface at the application layer [23]. Mehrotra et al. [24] proposed a class of multimodal classification for relevance vector classifier, which combined incremental and granular learning, which could handle large-scale unbalanced datasets and achieve better performance in multimodal biometrics classification and evaluation. Abdolahi et al. [25] proposed a multimodal fusion system using fingerprint and iris with fuzzy logic, and the obvious improvement in recognition rate was achieved. However, this method does not give a quantitative analysis of the effectiveness of the fusion process, and the obtained effect is poorly generalized. Miao et al. [26] proposed a framework of bin-based classifier method for the fusion of multibiometrics, which embedded matching scores into a new image pixel space, and obtained richer feature information when performing image-based biometrics. Chen et al. [27] proposed a framework for face and fingerprint images fusion using a type of middle-layer semantic features extracted from local feature-image matrix. However, it is still not clear whether this feature has good feature expression for all kinds of biometrics. Khellat et al. [28] proposed a feature level fusion method for three biological traits, which mainly used the Fisher dimensionality reduction technique, which caused the occurrence of the feature fusion in the dimensionality reduction space. Mai et al. [29] proposed a binary feature fusion method, which was generated from the sequence of feature bits using a machine learning algorithm that minimizes intraclass differences by minimizing interclass differences. The above proposed feature fusion algorithms still lack effective theoretical proof. Some work has been done on the application of multimodal biometrics. Liu et al. [30] applied the multimodal biometrics authentication method to single difficult biometrics, fused the different feature modalities to recognize the short utterance speaker, and achieved remarkable performance improvement. Gomez et al. [31] studied the protection of multibiometric template and proposed a multibiometric template protection technology based on homomorphic probability encryption. Gurusamy et al. [32] studied the biometric characteristics of MRI, and it was found that wavelet transform could better highlight the features of MRI images. Meng et al. [33] proposed a method for effectively detecting image hidden information by combining various image features through a fast R-CNN network. At present, the interpretability of the R-CNN network is insufficient.

To the best of our knowledge, although there are quite a few studies on multimodal biometrics, the definitive demonstration of multimode biometrics fusion has not been discussed yet. In this paper, we conducted a study on the deterministic effect of multimodal biometrics, using fingerprints and voiceprint as a template. At present, the research on the fusion of these two types of biometrics is still very limited.

3. Preliminary Research

Fingerprints, irises, human faces, voiceprints, and finger veins are the most commonly used biological features in human biometrics. The samples collection of fingerprint, voiceprint, and face is more convenient, and the application rate is higher. However, the face needs a large number of face sample banks, and the training and operation cost is high. In this paper, we use the low-cost fingerprint and voiceprint characteristics as the research objects.

3.1. Fingerprint Authentication Technology. The fingerprint authentication process is mainly divided into 4 steps, as shown in Figure 1: (1) acquisition of fingerprint images, usually using optical instruments and other equipment; (2) fingerprint image preprocessing, which finally gets the fingerprint thinning map; (3) fingerprint feature extraction, which extracts the fingerprint feature point information which is serialized as fingerprint feature vector and stored as fingerprint feature template; (4) fingerprint matching, which matches the extracted fingerprint feature vector with the feature template in the fingerprint database to confirm the authenticator.

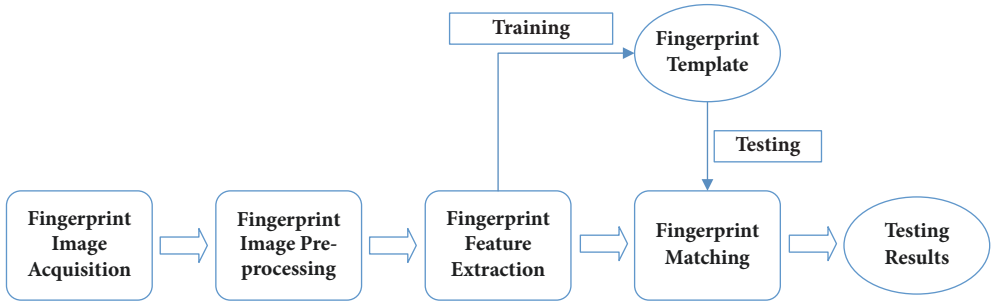


Figure 1: Fingerprint authentication process flowchart.

In general, fingerprint image authentication mainly depends on the uniqueness of individual fingerprints in the texture. Fingerprint image preprocessing is the process of removing noise and highlighting and clarifying fingerprint texture. The general process is shown in Figure 2. Fingerprint image preprocessing directly affects the performance of the entire fingerprint authentication system, and its mainly steps include the following: (1) fingerprint image enhancement: in this step, a specific algorithm such as frequency domain transform, filtering, denoising, and splicing of small block fingerprints is used to improve the quality of the image so that the fingerprint lines can have better connectivity and clearness, avoid false feature points, and improve fingerprinting characteristics accuracy of extraction. (2) Fingerprint image binarization: in this step, the fingerprint image is converted into a black-and-white binary image by the method of deleting the local image pixel point while maintaining the connectivity. As a result, the adhesion between the lines is removed, the complexity of the fingerprint feature extraction is reduced, and the subsequent image thinning operation is facilitated. (3) Fingerprint image refinement: in this step, the binarized fingerprint texture is refined into single-pixel lines, preserving the trend of the fingerprint lines without regard to the thickness of the lines. The refined fingerprint can extract details such as feature points more conveniently, so as to improve the accuracy of fingerprint matching.

Fingerprint feature extraction is a key step in the fingerprint template generation. Its main task is to obtain fingerprint feature point information. In the fingerprint identification process, the fingerprint feature point information is generally used as its main feature information, including the attribute point, position, and direction field value of the feature point. The comparison process determines whether the two feature points are the same according to the feature information. When two fingerprints have a certain number of the same feature points, the two fingerprints can be considered as one. The extracted feature point information is serialized to obtain a biometric template.

3.2. Voiceprint Authentication Technology. A complete speaker recognition system consists of acoustic feature extraction, voiceprint models establishing, and voiceprint matching calculations, as shown in Figure 3. The process of feature extraction is to extract the acoustic features of speech, such as Mel-scale Frequency Cepstral Coefficients (MFCC) from the original waveform signal, and to obtain a voiceprint model, such as Gaussian Mixture Model (GMM), which is used as a template to identify personal speech features. By calculating the voiceprint matching score, the system outputs the speaker authentication result.

The GMM is the most common voiceprint recognition model of the existing voiceprint models, as shown in Figure 4. The basic process is to extract the speech MFCC feature sequence and use the training data to calculate the model parameters and obtain the individual GMM template. The specific process is as follows.

For any D-dimensional vector x_t, $t = 1, 2, \ldots, M$, the Gaussian mixture probability density function used to calculate the likelihood is as follows:

$$p(x_t \mid \lambda) = \sum_{i=1}^{M} \omega_i b_i(x_t), \tag{1}$$

where ω_i is the $i-$th Gaussian component weight, satisfying $\sum_{i=1}^{M} \omega_i = 1$. Speaker model $\lambda = \{\omega_i \ \mu_i \ \Sigma_i\}$, where Σ_i is usually a diagonal matrix. And M is the number of Gaussian components; that is, the $i-$th mixed Gaussian probability density $b_i(x_t)$ is defined as follows:

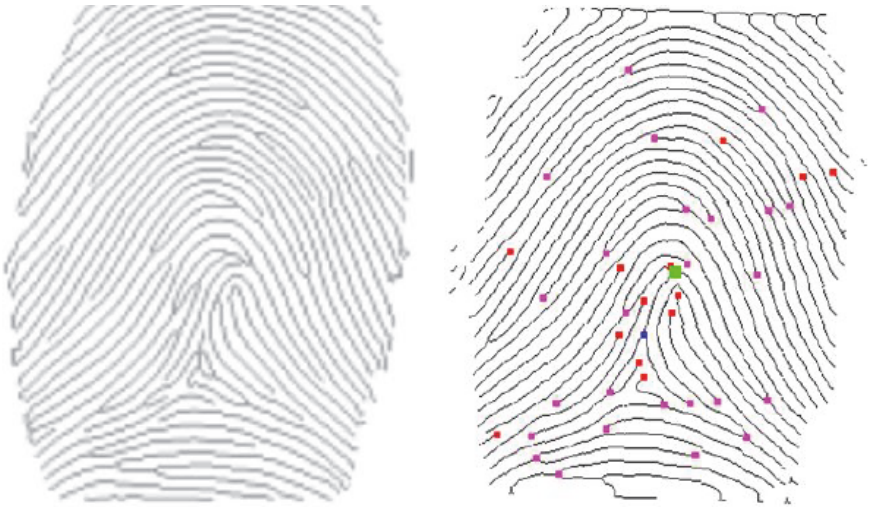
$$b_i(x_t) = \frac{1}{(2\pi)^{D/2} \left|\sum_i\right|^{1/2}} \cdot \exp\left\{-\frac{1}{2}(x_t - u_i)' \left(\sum_i\right)^{-1} (x_t - u_i)\right\}. \tag{2}$$

If $X = \{x_1, x_2, \ldots, x_m\}$ is the acoustic feature vector set of speaker I training and x_t is a D-dimensional vector, then the whole process of parameter estimation can be described as updating the model parameter λ^* satisfying $p(X \mid \lambda^*) \geq p(X \mid \lambda)$ iteratively until convergence. Given the trained feature vector set of speaker I, the model parameters are usually obtained by the EM iterative algorithm, and the iteration process is as follows:

(1) Weight iterative formula is as follows:

$$\omega_i = \frac{1}{T} \sum_{k=1}^{N} P(i \mid X_k, \lambda). \tag{3}$$

(a) Primitive image (b) Enhanced image (c) Binary image



(d) Thinning image (e) Feature point image

Figure 2: Fingerprint authentication process flowchart.

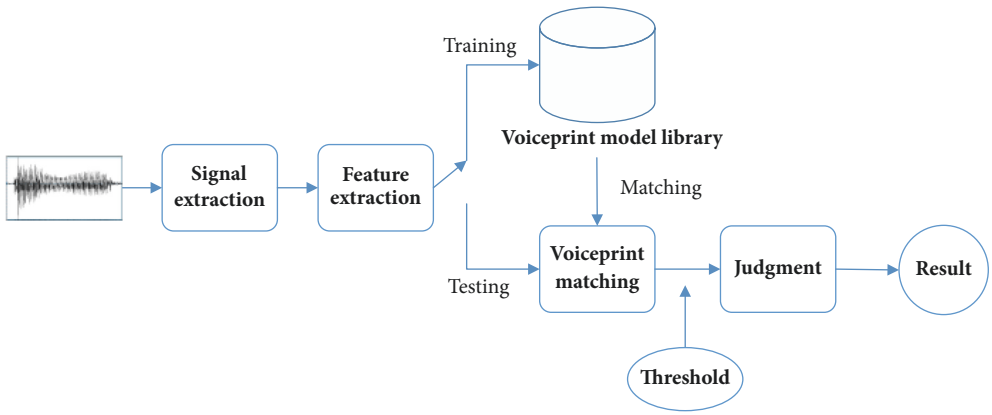


Figure 3: Speaker recognition system structure.

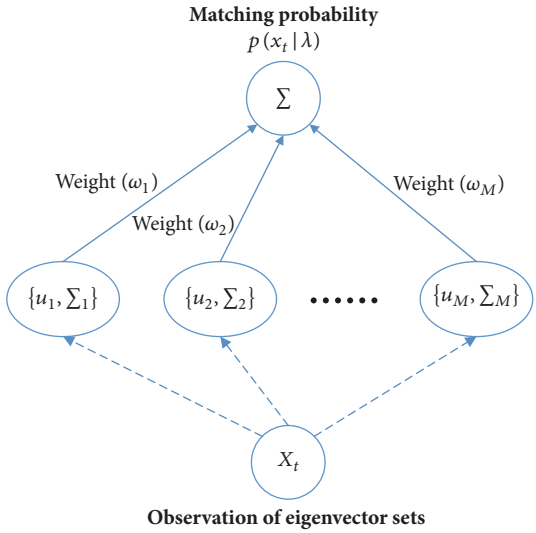


FIGURE 4: Fingerprint authentication process flowchart.

(2) Mean iterative formula is as follows:

$$u_i = \frac{\sum_{k=1}^{N} P(i \mid X_k, \lambda) \cdot X_k}{\sum_{k=1}^{N} P(i \mid X_k, \lambda)} \tag{4}$$

(3) Variance iterative formula is as follows:

$$\sum_i = \frac{\sum_{k=1}^{N} P(i \mid X_k, \lambda) \cdot X_k^2}{\sum_{k=1}^{N} P(i \mid X_k, \lambda)} \tag{5}$$

In the above formula, the posterior probability of component i is as follows:

$$P(i \mid X_k, \lambda) = \frac{\omega_i \cdot b_i(X_k)}{\sum_{t=1}^{M} \omega_t \cdot b_t(X_k)}. \tag{6}$$

The GMM parameters ω_i, u_i, $\sum_i$, etc. constitute a voiceprint biometric template.

Fingerprint and voiceprint features have their own characteristics. The fingerprint features are presented in the form of images. The features are hidden in the image texture. The recognition process requires fine processing of the image texture, which is susceptible to contamination and other kinds of interference and reduces the recognition accuracy. The high recognition accuracy of general fingerprints requires the assistance of a high-quality fingerprint collector. Speech frequency spectrum is the main analysis object of voiceprint feature. Fine processing of frequency domain features is needed in the process of recognition. The specificity of the features is not intuitive, but the anti-interference ability is slightly stronger than the fingerprint. The effective fusion of the two types of features can complement each other and enhance the anti-interference ability of feature recognition.

4. Multibiometric Fusion Authentication Algorithm MFDB-Decision

Multimodal biometric authentication based on image feature fusion generally achieves the ideal recognition accuracy in the limited sample test. However, the validity of the algorithm often lacks theoretical proof, making the generalization of the algorithm questionable. In this chapter, a demonstrable multimode fusion algorithm is derived from the combination of matching level and decision level making.

4.1. Strategy for Multimodal Fusion Optimizing. Multimodal biometric system uses various levels of fusion to combine two or more modalities [23], according to the different levels of integration. From low to high, it can be divided into the following:

(1) At the sensor layer, the captured images are pixel-level fused. It is worth paying attention to that retaining as much as information as possible is inefficient and has poor real-time performance due to the large amount of sensor data processed. Furthermore, considering the differences in signal acquisition equipment, sensor layer fusion is not feasible in most cases.

(2) At the feature extraction level, two or more modalities in the form of feature vectors are concatenated. Such a fusion often leads to very high dimensional vectors. But at present, the selection of characteristics is more random, and the specificity of the selected features generally lacks large-scale test.

(3) At the matching score level, it mainly combines the matching scores from different modalities. But for the two modes of pattern with different calculation methods, fusion will be difficult.

(4) At the decision level, the judgments of multiple verdicts are consolidated, and it has little requirement on the data relevance.

Through the abovementioned analysis, at present, although feature layer fusion may produce new effective features, it is difficult to guarantee the stability and reliability of new features. Considering that single-mode biometrics can easily excavate stable and specific features, we use the dynamic Bayesian method to combine feature recognition in the score layer and that in the decision layer based on fingerprint and voiceprint single-mode feature extraction, and the higher recognition accuracy and stability are obtained.

4.2. Dynamic Bayesian Decision for Minimizing the Error Risk. Because Bayesian judgments can achieve high judgment accuracy in mutually independent biometric modalities, we use Bayesian decision theory [34, 35] as the underlying mechanism. In the ideal case where all the relevant probabilities are known, Bayesian decision theory considers how to choose the optimal category marker based on these probabilities. Firstly, we suppose there are N possible category collections that can be shown like $Y = \{c_1, c_2, \ldots, c_n\}$, and λ_{ij} is the loss of classifying a sample of true labeled c_j as c_i. Based on the posterior probability $P(c_i \mid x)$, the expected loss produced by

classifying the sample x as c_i ("conditional error risk" on the sample x) can be expressed as follows:

$$R(c_i \mid x) = \sum_{j=1}^{N} \lambda_{ij} P(c_j \mid x). \tag{7}$$

Our task is to find a decision criterion $h : X \longrightarrow Y$ to minimize the cost of the error risk:

$$R(h) = E_x [R(h(x) \mid x)]. \tag{8}$$

Obviously, for each sample x, if h can minimize the conditional risk $R(h(x) \mid x)$, the overall cost of the error risk $R(h)$ will also be minimized. This produces dynamic Bayesian decision rule: To minimize the overall risk, it is needed to choose on each sample a category marker that minimizes the conditional risk $R(c \mid x)$; it can be written as follows:

$$h^* = \arg\min_{c \in Y} R(c \mid x), \tag{9}$$

where h^* means Bayesian optimal classifier, corresponding to the overall risk $R(h^*)$ called Bayesian risk, and $1 - R(h^*)$ reflects the notion that the classifier can achieve the best performance. When it comes to classification issues, λ_{ij} can be expressed as

$$\lambda_{ij} = \begin{cases} 0 & if\ i = j \\ 1 & otherwise. \end{cases} \tag{10}$$

Thus, the Bayesian optimal classifier that minimizes the classification error rate is

$$h^*(x) = \arg\max_{c \in Y} P(c \mid x). \tag{11}$$

It is obviously observed that the significance of maximizing posterior probability is to minimize the expected risk.

4.3. Multibiometric Fusion Authentication using Dynamic Bayesian Decision (MFDB-Decision). In order to fuse the fingerprints and voiceprint recognition systems together, a score vector $X = (X_1, X_2)$ containing multiple recognition system is constructed, where X_1 and X_2, respectively, represent the scores obtained from the fingerprint and voiceprint recognition system. Then, the question of identity conversion translates into the problem of classifying the two-dimensional fraction vector $X = (X_1, X_2)$ as accepting (genuine) or rejecting (imposter).

We know each modal classifier should have a different weight in multimodal classifier; in this paper, we should not fix the weight of biometrics because the dominant biometric is not always the same one. So we propose an algorithm called dynamic Bayesian decision (MFDB-decision) to get the best fusion recognition accuracy. Algorithm 1 is described in detail as follows.

The output of the algorithm is the category to which this feature belongs. The matching scores from each of the two classifiers $X = (X_1, X_2)$, $X_1 = P(x_{ij} \mid R_1)$, $X_2 = P(y_{ij} \mid R_2)$, $i = 1, 2, \ldots, N$, $j = 1, \ldots, k$; N means a total of N categories,

Input: $X = (X_1, X_2)$, $X_1 = P(x_{ij} \mid R_1)$,
$X_2 = P(y_{ij} \mid R_2)$, $i = 1, 2, \ldots, N$, $j = 1, \ldots, k$
(1) **repeat**
$fs \longleftarrow \max_{i=1,\ldots,N} P(x_{ij} \mid R_1)$
$fn \longleftarrow \arg\max_{i=1,2,\ldots N} P(x_{ij} \mid R_1)$
$vs \longleftarrow \max_{i=1,\ldots,N} P(y_{ij} \mid R_2)$
$vn \longleftarrow \arg\max_{i=1,2,\ldots N} P(y_{ij} \mid R_2)$
(2) **if** $fs < \sigma_1$ && $vs < \sigma_2$ **then**
(3) $n = false$
(4) **else if** $fs \geq \sigma_1$ && $vs < \sigma_2$ **then**
(5) $n = \arg_i \left(\max_{i=1,2,\ldots N} \sum_{m=1}^{2} \alpha_m vote_{im} \right), \quad \alpha_1 > \alpha_2$
(6) **else if** $fs < \sigma_1$ && $vs \geq \sigma_2$ **then**
(7) $n = \arg_i \left(\max_{i=1,2,\ldots N} \sum_{m=1}^{2} \alpha_m vote_{im} \right), \quad \alpha_2 \geq \alpha_1$
(8) **end if** {n is category judgment result}
(9) **until** the (k+1)-th test sample

Algorithm 1: The MFDB-decision algorithm.

and k represents a total of k test samples. α_1 is the weight of fingerprint recognition system; α_2 is the weight of voiceprint recognition system; σ_m is the quality failure threshold. Here, we set σ_2 to be the average score value of current person with N template and cycling σ_1 from 0.3 to 0.6, step 0.01. In step 5, the algorithm uses (10) to calculate $vote_{im}$. In step 7, because the voiceprint quality is high, the algorithm sets the weight of voiceprint α_2 to be greater than α_1.

4.3.1. Fingerprint Matching Score Calculation. Fingerprint matching algorithm is mainly divided into three kinds of schemes based on correlation, minutiae, and nonminutiae matching. Either of the schemes will firstly form a fingerprint feature vector template, which is authenticated by a template. After extracting the feature vector for two fingerprint images, we can express it as

$$\begin{aligned} A &= \left\{ m_{A_1}, m_{A_2}, \ldots, m_{A_p} \right\}, \\ &\text{where } m_{A_i} = \left\{ x_{A_i}, y_{A_i}, cn, \theta_{A_i} \right\},\ 1 \leq i \leq p \\ B &= \left\{ \mathrm{m}_{B_1}, \mathrm{m}_{B_2}, \ldots, \mathrm{m}_{B_p} \right\}, \\ &\text{where } m_{B_1} = \left\{ x_{B_j}, y_{B_j}, cn, \theta_{B_j} \right\},\ 1 \leq j \leq q. \end{aligned} \tag{12}$$

There are two finite sets of points in space: $A = \{x_1, y_1, cn, \theta_1\}$ and $B = \{x_2, y_2, cn, \theta_2\}$, where x and y represent the coordinates of the detail points, respectively, where cn denotes the type of the detail point, for example, $cn = 3$ for a fork, $cn = 1$ for an endpoint, and so on. θ denotes the direction along the main ridge. As fingerprints are collected during being pressed, it is easy for the collected ones to be offset. Therefore, in the authentication process, the geometric constraints on the details of the matching point are

proposed, including the geometric distance and the angle of detail deviation from the limit as follows:

$$dist_r\left(m_{A_i}, m_{B_j}\right) = \sqrt{\left(x_{A_i} - x_{B_j}\right)^2 + \left(y_{A_i} - y_{B_j}\right)^2} < r_\delta \tag{13}$$

$$dist_\theta\left(m_{A_i}, m_{B_j}\right) = \min\left(\left|\theta_{A_i} - \theta_{B_j}\right|, 360 - \left|\theta_{A_i} - \theta_{B_j}\right|\right) < r_\theta. \tag{14}$$

Following global registration, a local search can be performed [36], where r_δ means a reasonable distance threshold for the offset of the minutia and r_θ is a permissible deviation from the distortion estimate obtained from the ridge pattern. At the same time, two characteristic points satisfying formula (13) and (14) are considered as matching feature points. Two fingerprints with enough matching feature points are considered to be matched fingerprints. The specific score of fingerprint matching can be calculated as follows:

$$sim(A, B) = \frac{n_{match}^2}{n_A n_B} \tag{15}$$

where n_{match}^2 means the number of feature points that match within the threshold in both graphs. n_A and n_B are the number of feature points, respectively, owned by the template vector and the test vector.

4.3.2. Voiceprint Matching Score Calculation. For D-dimensional acoustic feature vector x_k, $k = 1, 2, \ldots, M$, the Gaussian mixture probability density function used to calculate the likelihood is as shown in (1). The whole process of likelihood parameter estimation can be described as updating the model parameter λ^* satisfying $p(X \mid \lambda^*) \geq p(X \mid \lambda)$ iteratively until convergence. According to Jensen inequality, the problem of parameter solving can be transformed into the problem of maximized $Q(\lambda, \lambda^*)$, and $Q(\lambda, \lambda^*)$ can be solved as follows due to $p(x_k, i \mid \lambda) = \omega_i p(x_k \mid \lambda, i)$:

$$Q(\lambda, \lambda^*) = \sum_{k=1}^{m}\sum_{i=1}^{M}\frac{\omega_i p(x_k \mid \lambda, i)}{p(x_k \mid \lambda)}\left[\lg \omega_i^* + \lg p(x_k \mid \lambda^*, i)\right] \tag{16}$$

Calculate the partial derivatives of mean value, weight, and covariance, respectively, and let the result be zero, and then an updated formula of the model parameters ω_i, μ_i, and Σ_i will be obtained. Given the trained feature vectors set of speaker I, the model parameters are usually obtained by the EM iterative algorithm, but the computational complexity is large. Therefore, this paper adopts the adaptive method proposed by Reynolds to solve the model parameters. Using the set of observed feature vectors of speaker I to fit the predictive speaker model through the maximum posteriori probability (MAP), the problem is actually transformed into an optimization problem. Similar to the E step in the EM algorithm, the sufficient statistics n_i, $E_i(x)$, $E_i(xx')$, and $p(i \mid x_k)$ of each Gaussian component of the UBM are first calculated, the difference being that the weight of the speaker model, the mean value, and covariance of the update process at the M step are as follows:

$$\omega_i^* = \left[\frac{\alpha_i^\omega n_i}{m} + \left(1 - \alpha_i^\omega\right)\omega_i\right]\zeta \tag{17}$$

$$\mu_i^* = \alpha_i^\omega E_i(x) + \left(1 - \alpha_i^\omega\right)\mu_i \tag{18}$$

$$\left(\sigma_i^*\right)^2 = \alpha_i^\upsilon E_i\left(xx'\right) + \left(1 - \alpha_i^\upsilon\right)\left(\sigma_i^2 + \mu_i\mu_i'\right) - \left(\mu_i^*\right)^2. \tag{19}$$

In the formula, $\alpha_i^\rho = n_i/(n_i + r^\rho)$, $\rho \in \{\omega, m, \upsilon\}$ is an adaptive parameter that controls the change between the old and new coefficients, where $r^\rho = 16$ is a fixed correlation factor. ζ makes sure the weight rollup is always 1. Usually, only update the mean value, that is, $\alpha_i^\omega = \alpha_i^\upsilon = 0$. If the test speech feature vectors set of speaker J (declared as the I-th speaker λ_1, abbreviated as λ) is χ, the common background model is λ_u, and the system logarithmic likelihood score is

$$\Lambda = \lg p(\chi \mid \lambda) - \lg p(\chi \mid \lambda_u). \tag{20}$$

4.4. Theoretical Support for the MFDB-Decision Algorithm. Algorithm 1 leads to the following lemma.

Lemma 1. *The MFDB-decision algorithm can be generalized to* L *classifiers when each one is independent.*

Proof. Assuming there are x_i $(i = 1, \ldots, m)$ samples to be identified and entered into L different classifiers, the output of a certain sample to be recognized after passing L classifiers is R_j $(j = 1, 2, \ldots, L)$, and x_i must be able to be identified as one of the N classes. According to the Bayes decision theory for minimizing the risk of loss, the fusing sample to be identified will be recognized as the highest posterior probability in N modal classes, and (11) can be written as

$$n = \arg_i\left(\max_{i=1,2,\ldots,N} P(x_i \mid R_1, R_2, \ldots, R_L)\right). \tag{21}$$

We assume that L classifier is independent. So (6) can be analyzed as follows:

$$P(x_i \mid R_1, R_2, \ldots, R_L) = \prod_{l=1}^{L} P(x_i \mid R_l). \tag{22}$$

We derive the following from bringing (22) into (21):

$$n = \arg_i\left(\max_{i=1,2,\ldots,N}\prod_{l=1}^{L} P(x_i \mid R_l)\right). \tag{23}$$

If it is assumed that the posterior probability $P(x_i \mid R_l)$ fluctuates above and below the prior probability $P(x_i)$ and is not large, just shown as

$$P(x_i \mid R_l) = P(x_i)(1 + \delta_{il}) \quad \delta_{il} \ll 1, \tag{24}$$

take formula (24) into (22) as follows:

$$\begin{aligned}\prod_{l=1}^{L} P(x_i \mid R_l) &= P(x_i) \prod_{l=1}^{L} (1 + \delta_{il}) \\ &\approx P(x_i) + P(x_i) \sum_{l=1}^{L} \delta_{il} \\ &= P(x_i) + P(x_i) \sum_{l=1}^{L} (P(x_i \mid R_l) - P(x_i)) \\ &= P(x_i) \left(1 - LP(x_i) + \sum_{l=1}^{L} P(x_i \mid R_l)\right).\end{aligned} \tag{25}$$

Approximately equal sign uses the Taylor series expansion, and we get a general formalized multiclassifier fusion strategy which holds for each independent feature classifier as follows:

$$n = \arg_i \left(\max_{i=1,2,\ldots,N} \sum_{l=1}^{L} P(x_i \mid R_l) \right). \tag{26}$$

Consider fusing the concept of minimum loss expressed by (11) and (26) and if each classifier can find the maximum posterior probability, we can find the best posterior probability comprehensively:

$$\begin{aligned} vote_{Kl} &= \begin{cases} 1, & K = \arg_i \left(\max_{i=1,2,\ldots,N} (P(x_i \mid R_l)) \right) \\ 0, & otherwise \end{cases} \\ n &= \arg_k \left(\max_{k=1,2,\ldots,N} \sum_{l=1}^{L} vote_{kl} \right). \end{aligned} \tag{27}$$

Under the premise that all L classifiers are correct, the fusion classification result of (27) is guaranteed. However, the abovementioned formula does not consider the influence of the classifier posterior probability $P(x_i \mid R_l)$ on the classification result. If some of the preceding classifications are wrong, the error probability will be accumulated and transmitted backwards, resulting in low robustness. In order to reduce the influence of $P(x_i \mid R_l)$ on fusion classification results, fractional layer information was added to assist in judgment which means the weight α_m needs to be dynamically adjusted in each fusing round and mainly rely on scores that the current user obtained at the matching layer. When the maximum posteriori is within the qualified threshold, then the quality is judged to be good and a larger weight a is assigned. If it is outside the qualified threshold, then the quality is judged to be poor and a smaller weight b is assigned. When a suitable threshold is found as the decision key, our algorithm using matching score to assist in making decision can be modified from formula (27) as

$$\begin{aligned} \alpha_l &= \begin{cases} a & \max_{k=1,2,\ldots,N} (P(x_k \mid R_l)) > \sigma_i \\ b & otherwise,\ b < a; \end{cases} \\ n &= \arg_k \left(\max_{k=1,2,\ldots,N} \sum_{l=1}^{L} \alpha_l vote_{kl} \right), \end{aligned} \tag{28}$$

where α_m is a dynamic weight based on the overall performance of each single-mode which can be judged from the matching layer. a and b are positive numbers that satisfy $a + b = 1$ and $b < a$. σ_i is the quality threshold in which we usually set the average score value of current person with N templates. The setting of details in this paper can be seen in Section 4.3. In this way, when an error is accumulated, the quality of the poor quality feature can be reduced to influence the final result of the voting, thereby improving the robustness of the algorithm. According to formula (28), the lemma is proved. □

Corollary 2. *If the L classifiers are independent and the number of categories is fixed, then the error rate of the MFDB-decision algorithm will be infinitely tending to zero when the number of classifiers tends to infinity.*

Proof. Assuming that L classifiers are independent and our MFDB-decision algorithm achieved the minimum error rate in Lemma 1, then the probability of classification error for each classifier is less than the randomly selected error rate for N categories, i.e., $(N-1)/N$. So it can be expressed as follows:

$$\min er \le \left(\frac{N-1}{N} \right), \tag{29}$$

where er represents the error rate of each classifier. When the number of categories N is fixed and the number of classifiers L increases to infinity, the classification error rate can be expressed as

$$\begin{aligned} 0 &\le \lim_{L \longrightarrow +\infty} (\min er)^L \le \lim_{L \longrightarrow +\infty} \left(\frac{N-1}{N} \right)^L \\ &= \lim_{L \longrightarrow +\infty} \left(1 - \frac{1}{N} \right)^L = 0. \end{aligned} \tag{30}$$

Since N is a fixed value, $1 - 1/N$ is a constant less than 1, so the abovementioned equation $\lim_{L \longrightarrow +\infty} (\min er)^L$ equals 0 according to Sandwich Theorem. Therefore, as the number of classifiers L increases, the classification error will be reduced to 0. □

This section not only proves the feasibility of MFDB-decision theoretically in this paper but also lists a situation showing that the classification error rate will decrease with the increase of classifiers. Therefore, the proposed algorithm is suitable for generalization of multimodal recognition beyond bimodality.

5. Experimental Results and Analysis

In order to test the effectiveness and practicability of the MFDB-decision algorithm, we used 3 common databases and 1 self-extracting database to conduct experimental tests. The three common databases were FVC2002 DB1 database (fingerprint), MIT Media lab Speech Dataset (speech), and TIMIT corpus (speech). The self-extracting database was hdu2016_40 database (short speech). FVC2002 DB1 database was a standard difficult fingerprint dataset with 100 fingers

and eight samples for each finger, which was provided by the National Institute of Standards and Technology (NIST). The recognition rate of difficult fingerprints by general algorithms was not high. Generally, if there was no optimization, the recognition rate would be lower than 95%. In the fingerprint recognition competition, the participants could increase the recognition rate to over 99% by specific optimization, but the versatility of such algorithms was not strong. Since this test mainly investigates the effects of the two types of biometric fusion, the high accuracy of single-mode feature recognition was not conducive to the test results. Therefore, all the following test procedures used the universal fingerprint recognition algorithm and did not specialize optimization for the FVC database. The MIT Media lab Speech Dataset consisted of 48 registrars (22 females and 26 males) and 40 attackers (17 females and 23 males). Recorded separately in the handheld microphone and external headphones, the recording environment included quiet indoor, slightly interfering laboratories and noisy intersections; each tester randomly read 108 words or sentences in 6 environments. The TIMIT corpus was designed by the Defense Advanced Research Projects Agency. The number of registrations in TIMIT was 630, and each person read 10 sentences with 6300 sentences. 630 people were made up of 8 regions, including 438 men and 192 women. Each person read two designated text phonetics (SA) in dialect, five phonetically compact sentences (SX), and three phonetically diverse sentences (SI). The hdu2016_40 was a corpus including 40 people, each person had 25 paragraphs of different short utterance, each paragraph would be recorded ten times, and each record lasted 2~3 seconds. The MIT Speech Dataset and the TIMIT corpus were both English datasets, and the hdu2016_40 database was short speech Chinese datasets. Although, in the field of long speech voiceprint recognition, the recognition accuracy rate had reached 98% in low noise environment, in the short speech voiceprint recognition environment (voice length less than 5s), even if the ambient noise was low, the recognition accuracy was still not ideal, less than 95%. This article tested for short speech voiceprint recognition. Because the common fingerprint and voiceprint from the same tester database were relatively rare, the experiment used the abovementioned four groups of database combination for testing. In order to made the experimental results more reliable, the samples in the different databases were randomly selected during the experiment for combination testing, and multiple random sampling was performed under the condition that the combinations were not repeated. To evaluate the performance, false rejection rate (FRR) and false acceptance rate (FAR) are used as the main indicators.

5.1. Single Modal and Multimodal Comparison. In order to investigate the effectiveness of the MFDB-decision algorithm, we firstly examine the improvement of the accuracy of the MFDB-decision algorithm when the single-mode algorithm has difficulty in achieving high accuracy. The abovementioned four databases constitute two sets of datasets, with one being the combination of FVC2002 DB1 database, the MIT speech database, and the TIMIT corpus and the other being the combination of the FVC2002 DB1 database and the hdu2016_40 short speech database. The TIMIT corpus was used to train the English speech Universal Background Model (UBM), and the Chinese UBM was trained with the network random grasp of Chinese speech set. In the experiment, the fingerprint recognition algorithm used a general feature-based matching algorithm. The voiceprint recognition algorithm used the GMM model. Fingerprint and voiceprint recognition algorithms had no additional targeted optimization measures, so the difficulties of fingerprint recognition rate and the short speech voiceprint recognition rate were not high, as shown in Table 1.

Table 1: Difference between single biometric and multimodal biometric (%).

Trait	FAR	FRR	Accuracy
Fingerprint	4.8	5.0	94.60
Voiceprint(#1)	7.0	7.0	93.12
Voiceprint(#2)	11.5	11.7	88.38
MFDB-decision(#1)	1.0	1.2	99.06
MFDB-decision(#2)	1.6	1.8	98.35

Table 1 shows the experimental differences between single-mode and multimodal states. The '#1' and '#2', respectively, represent the collection of two types of voiceprint test databases in Chinese and English. Since the Chinese database is a self-acquisition voice database, the voice quality is better, so the recognition accuracy is relatively high. The result shows the superiority of our algorithm compared to single-mode recognition. We use the method proposed in Section 4.3.1 in the process of the fingerprint matching, using formulae (13) and (14) calculating the geometric distance and the angle of detail deviation; here, we set $r_\delta = 10$ and $r_\theta = 9$ and get similarity score calculated by (15). We performed the voiceprint process using the method proposed in Section 4.3.2, and 128 mixtures are used by GMM model, and the likelihood score is calculated by using formula (20). From the results, the MFDB-decision algorithm merged two single modes and achieved more stable and accurate results.

5.2. Robustness of the MFDB-Decision Algorithm. The fingerprints of 100 people (FVC2002 DB1, totaling 100 training pieces and 700 testing pieces) were divided into 60 groups, with each group consisting of 1 40 people, 2~41 people, 3~42 people, etc. There were 40 individuals in each group, each with 1 training fingerprint, 7 testing fingerprints, 25×5 training voices, and 25×5 test voices, for a total of 40×7×25×5=35,000 tests. Each fingerprint and each voiceprint were paired and the average of 35,000 tests was taken as the test result of this group. We plotted the recognition rate (Genuine Acceptance Rate (GAR)) of each group after using the MFDB-decision algorithm in Figure 5, where "1~10" in the legend indicated the recognition rate obtained after the first to tenth groups of voiceprints and fingerprints were fused.

It can be seen from Figures 5 and 6 that the fusion model shows high stability as the recognition rate concentrated in 97% to 100%. And the fusion recognition rate is increased by

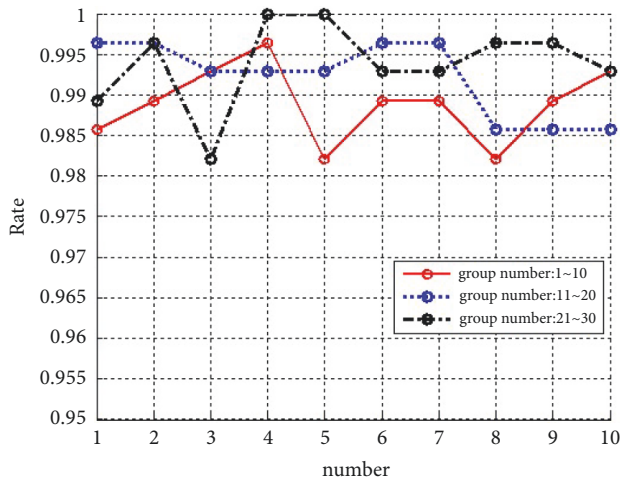


Figure 5: The recognition rate (GAR) of voiceprint fusing with fingerprint Group (a).

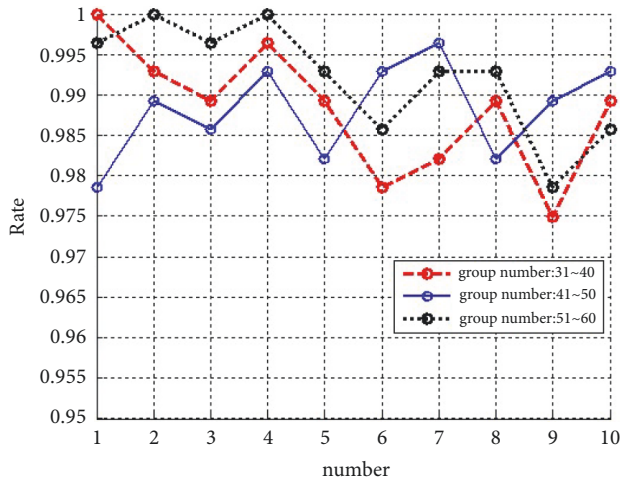


Figure 6: The recognition rate (GAR) of voiceprint fusing with fingerprint Group (b).

4.46% and 5.96% compared with single-model of fingerprint and voiceprint, respectively.

5.3. Effectiveness of the MFDB-Decision Algorithm. We test the effectiveness of the MFDB-decision algorithm by comparing the recognition rate of the MFDB-decision algorithm with several general fusion algorithms. The experimental process used the same grouping method in Section 5.2. The fusion recognition rate of each group was calculated by different algorithms, and the average recognition rate of each group was used as the final recognition rate of the fusion algorithm. The averaging in the abovementioned process was advantageous to avoid contingency. We randomly plotted 10 sets of recognition rates and compared them with the other two methods (AND as well as fixed weight voting method [37]). As shown in Figure 7, it can be seen that the MFDB-decision algorithm not only achieves a high recognition rate but also obtains good stability. The recognition rate of the

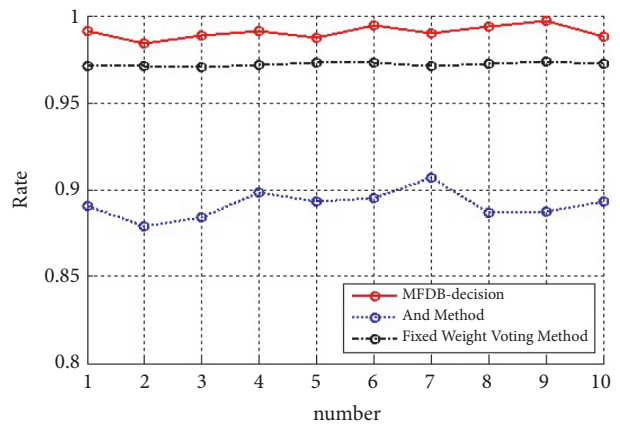


Figure 7: The recognition rates (GAR) of the three methods.

Table 2: Recognition rate of methods (%).

Method	Accuracy
And	89.10
Fixed weight Voting Method	97.21
MFDB-decision	99.06

MFDB-decision algorithm was higher than other algorithms. The fixed weight method was more stable than the MFDB-decision algorithm, but its recognition rate was not high enough. In many cases, it is worthwhile to sacrifice some stability and get a better recognition rate. Table 2 lists the average accuracy of the three multimodal methods.

MFDB-decision-making algorithm uses the score information of matching layer to assist decision-making recognition, which is helpful to recover the lost data in the decision-making process. Figure 8 shows the DET curves for various fusion methods. The PCA method uses the principal component analysis method mentioned in [38]. Since the unprocessed voiceprint MFCC sequence was not specific in the voiceprint recognition process, in this experiment, the accuracy of the PCA method was not high. The fuzzy rule method used fuzzy logic in the decision-making layer in [25] and achieved significant performance improvement. The trend of all curves is similar and decreases with the increase of FRR. The results show that all kinds of fusion methods are effective in the fusion of fingerprint and voiceprint, but our algorithm has achieved better results than other algorithms. We find that each curve intersects with the diagonal, indicating FRR=FAR, which is the equal error rate point EER. Generally, the lower the EER, the better the performance of the algorithm.

6. Conclusions

In this paper, we proposed a multimodal biometric recognition algorithm (named MFDB-decision) and demonstrated its effectiveness. We solved the problem that the fixed weight value could not be adaptively assigned in multimodal recognition and it would result in poor fusion performance. We compared the result of fusion with the result of single-modal

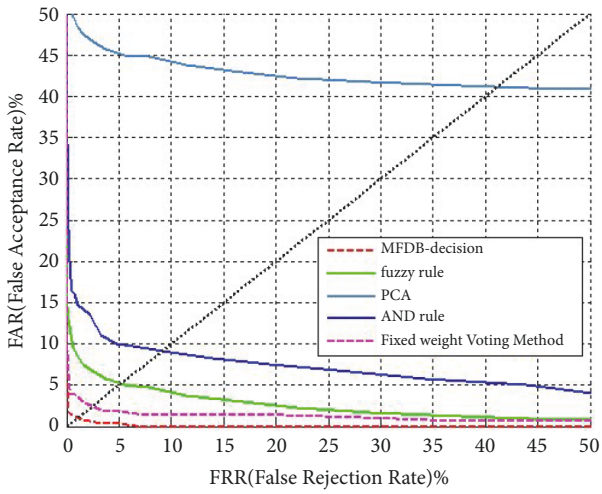


Figure 8: The DET curve for the methods.

recognition as well as the other methods and found that the method improved the recognition rate by an average of 5.0% or more. The multimodal fusion methods we developed are also greatly useful in the fusion recognition of other patterns. Future work will focus on multimodal biometric key extraction, ubiquitous identity authentication, and encryption technologies.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research is supported by National Key R&D Program of China (no. 2016YFB0800201), National Natural Science Foundation of China (no. 61772162), joint fund of National Natural Science Fund of China (no. U1709220), and Zhejiang Natural Science Foundation of China (no. LY16F020016).

References

[1] P. Wild, P. Radu, L. Chen, and J. Ferryman, "Robust multimodal face and fingerprint fusion in the presence of spoofing attacks," *Pattern Recognition*, vol. 50, pp. 17–25, 2016.

[2] V. Conti, C. Militello, F. Sorbello, and S. Vitabile, "A frequency-based approach for features fusion in fingerprint and iris multimodal biometric identification systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 40, no. 4, pp. 384–395, 2010.

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

[4] R. Snelick, U. Uludag, A. Mink, M. Indovina, and A. Jain, "Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 450–455, 2005.

[5] A. Muthukumar, C. Kasthuri, and S. Kannan, "Multimodal biometric authentication using particle swarm optimization algorithm with fingerprint and iris," *ICTACT Journal on Image and Video Processing*, vol. 02, no. 03, pp. 369–374, 2012.

[6] S. Shekhar, V. M. Patel, N. M. Nasrabadi, and R. Chellappa, "Joint sparse representation for robust multimodal biometrics recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pp. 113–126, 2014.

[7] Z. Wu, B. Liang, L. You, Z. Jian, and J. Li, "High-dimension space projection-based biometric encryption for fingerprint with fuzzy minutia," *Soft Computing*, vol. 20, no. 12, pp. 4907–4918, 2016.

[8] Z. Wu, L. Tian, P. Li, T. Wu, M. Jiang, and C. Wu, "Generating stable biometric keys for flexible cloud computing authentication using finger vein," *Information Sciences*, vol. 433-434, pp. 1339–1351, 2018.

[9] G. Bajwa and R. Dantu, "Neurokey: Towards a new paradigm of cancelable biometrics-based key generation using electroencephalograms," *Computers & Security*, vol. 62, pp. 95–113, 2016.

[10] L. Jin, L. Sun, Q. Yan et al., "Significant permission identification for machine learning based android malware detection," *IEEE Transactions on Industrial Informatics*, p. 12, 2018.

[11] J. Shen, T. Zhou, X. Chen, J. Li, and W. Susilo, "Anonymous and Traceable Group Data Sharing in Cloud Computing," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 4, pp. 912–925, 2018.

[12] J. Shen, Z. Gui, S. Ji, J. Shen, H. Tan, and Y. Tang, "Cloud-aided lightweight certificateless authentication protocol with anonymity for wireless body area networks," *Journal of Network and Computer Applications*, vol. 106, pp. 117–123, 2018.

[13] J. Shen, C. Wang, T. Li, X. Chen, X. Huang, and Z.-H. Zhan, "Secure data uploading scheme for a smart home system," *Information Sciences*, vol. 453, pp. 186–197, 2018.

[14] Q. Lin, H. Yan, Z. Huang, W. Chen, J. Shen, and Y. Tang, "An ID-based linearly homomorphic signature scheme and its application in blockchain," *IEEE Access*, 2018.

[15] C. Gao, Q. Cheng, P. He, W. Susilo, and J. Li, "Privacy-preserving Naive Bayes classifiers secure against the substitution-then-comparison attack," *Information Sciences*, vol. 444, pp. 72–88, 2018.

[16] Y. Li, G. Wang, L. Nie, Q. Wang, and W. Tan, "Distance metric optimization driven convolutional neural network for age invariant face recognition," *Pattern Recognition*, vol. 75, pp. 51–62, 2018.

[17] X. Zhang, Y. Tan, C. Liang, Y. Li, and J. Li, "A Covert Channel Over VoLTE via Adjusting Silence Periods," *IEEE Access*, vol. 6, pp. 9292–9302, 2018.

[18] J. Li, Q. Lin, C. Yu, X. Ren, and P. Li, "A QDCT- and SVD-based color image watermarking scheme using an optimized encrypted binary computer-generated hologram," *Soft Computing*, pp. 1–19, 2016.

[19] 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. 370/371, pp. 18–32, 2016.

[20] Y. Liu, J. Ling, Z. Liu, J. Shen, and C. Gao, "Finger vein secure biometric template generation based on deep learning," *Soft Computing*, vol. 21, no. 1, pp. 1–9, 2017.

[21] C. Yuan, X. Li, Q. M. J. Wu, J. Li, and X. Sun, "Fingerprint liveness detection from different fingerprint materials using convolutional neural network and principal component analysis," *Computers, Materials and Continua*, vol. 53, no. 4, pp. 357–371, 2017.

[22] W. Holden, "Securing public faith in biometrics," *Biometric Technology Today*, vol. 2016, no. 9, pp. 7–9, 2016.

[23] M. He, S.-J. Horng, P. Fan et al., "Performance evaluation of score level fusion in multimodal biometric systems," *Pattern Recognition*, vol. 43, no. 5, pp. 1789–1800, 2010.

[24] H. Mehrotra, R. Singh, M. Vatsa, and B. Majhi, "Incremental granular relevance vector machine: A case study in multimodal biometrics," *Pattern Recognition*, vol. 56, pp. 63–76, 2016.

[25] M. Abdolahi, M. Mohamadi, and M. Jafari, "Multimodal biometric system fusion using fingerprint and iris with fuzzy logic," in *International Journal of Soft Computing and Engineering*, vol. 2, pp. 504–510, 2013.

[26] D. Miao, M. Zhang, Z. Sun, T. Tan, and Z. He, "Bin-based classifier fusion of iris and face biometrics," *Neurocomputing*, vol. 224, pp. 105–118, 2017.

[27] Y. Chen, J. Yang, C. Wang, and N. Liu, "Multimodal biometrics recognition based on local fusion visual features and variational Bayesian extreme learning machine," *Expert Systems with Applications*, vol. 64, pp. 93–103, 2016.

[28] S. Khellat-Kihel, R. Abrishambaf, J. L. Monteiro, and M. Benyettou, "Multimodal fusion of the finger vein, fingerprint and the finger-knuckle-print using Kernel Fisher analysis," *Applied Soft Computing*, vol. 42, pp. 439–447, 2016.

[29] G. Mai, M.-H. Lim, and P. C. Yuen, "Binary feature fusion for discriminative and secure multi-biometric cryptosystems," *Image and Vision Computing*, vol. 58, pp. 254–265, 2017.

[30] Z. Liu, Z. Wu, T. Li, J. Li, and C. Shen, "GMM and CNN Hybrid Method for Short Utterance Speaker Recognition," *IEEE Transactions on Industrial Informatics*, vol. PP, no. 99, pp. 1-1, 2018.

[31] M. Gomez-Barrero, E. Maiorana, J. Galbally, P. Campisi, and J. Fierrez, "Multi-biometric template protection based on Homomorphic Encryption," *Pattern Recognition*, vol. 67, pp. 149–163, 2017.

[32] R. Gurusamy and V. Subramaniam, "A machine learning approach for MRI brain tumor classification," *Computers, Materials and Continua*, vol. 53, no. 2, pp. 91–109, 2017.

[33] R. Meng, G. S, J. Wang, and X. Sun, "A machine learning approach for mri brain tumor classification," *Computers Materials and Continua*, vol. 55, no. 1, p. 16, 2018.

[34] N. Ueffing and H. Ney, "Bayes decision rules and confidence measures for statistical machine translation," in *Advances in Natural Language Processing*, vol. 3230 of *Lecture Notes in Computer Science*, pp. 70–81, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[35] J.-T. Chien, C.-H. Huang, K. Shinoda, and S. Furui, "Towards optimal bayes decision for speech recognition," in *Proceedings of the 2006 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2006*, pp. I45–I48, France, May 2006.

[36] J. Abraham, P. Kwan, and G. Junbin, *Fingerprint Matching using A Hybrid Shape and Orientation Descriptor*, 2011.

[37] S. Lei and M. Qi, "Multimodal Recognition Method based on Ear and Profile Face Feature Fusion," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 9, no. 1, pp. 33–42, 2016.

[38] C. Chibelushi, "Feature-level data fusion for bimodal person recognition," in *Proceedings of the 6th International Conference on Image Processing and its Applications*, pp. 399–403, Dublin, Ireland.