Machine Intelligence Research
www.mi-research.net

20(6), December 2023, 884-896
DOI: 10.1007/s11633-022-1403-7

Multimodal Biometric Fusion Algorithm Based on Ranking Partition Collision Theory

Zhuorong Li Yunqi Tang

School of Criminal Investigation, People's Public Security University of China, Beijing 100038, China

Abstract: Score-based multimodal biometric fusion has been shown to be successful in addressing the problem of unimodal techniques' vulnerability to attack and poor performance in low-quality data. However, difficulties still exist in how to unify the meaning of heterogeneous scores more effectively. Aside from the matching scores themselves, the importance of the ranking information they include has been undervalued in previous studies. This study concentrates on matching scores and their ranking information and suggests the ranking partition collision (RPC) theory from the standpoint of the worth of scores. To meet both forensic and judicial needs, this paper proposes a method that employs a neural network to fuse biometrics at the score level. In addition, this paper constructs a virtual homologous dataset and conducts experiments on it. Experimental results demonstrate that the proposed method achieves an accuracy of 100% in both mAP and Rank1. To show the efficiency of the proposed method in practical applications, this work carries out more experiments utilizing real-world data. The results show that the proposed approach maintains a Rank1 accuracy of 99.2% on the million-scale database. It offers a novel approach to fusion at the score level.

Keywords: Image processing, convolutional neural network, multimodal, biometrics, fusion.

Citation: Z. Li, Y. Tang. Multimodal biometric fusion algorithm based on ranking partition collision theory. *Machine Intelligence Research*, vol.20, no.6, pp.884–896, 2023. http://doi.org/10.1007/s11633-022-1403-7

1 Introduction

Biometrics has rapidly evolved into a formidable instrument for public security authorities to uncover cases in recent years. Biometrics, which includes fingerprints, faces, and voiceprints, are natural identifiers. The facts of the case were established within 48 hours in the 615 bombing of a kindergarten in Xuzhou, Jiangsu Province, in 2017, using the face recognition platform of the Science and Information Bureau of Ministry of Public Security, China. However, the continuous growth of databases limits the performance of unimodal biometric algorithms. One of the key drawbacks of present unimodal biometric algorithms is that they place great demands on the quality of data received. Nevertheless, biometrics collected from crime scenes are generally of low quality, such as fragmented fingerprints, fuzzy facial images, and long-distance iris photographs, among other things. Existing approaches are unable to reconcile the imbalance between quality and accuracy. The majority of biometric recognition research has concentrated solely on how to exclude data of poor quality. There is an immediate need to solve the issue produced by low-quality biometrics. As a result, biometrics that integrate several modalities have emerged as a new development direction in the realm of public security.

Multimodal biometric recognition can be divided into multi-sensor biometric recognition, multi-algorithm biometric recognition, multi-instance biometric recognition, multi-native biometric recognition, and hybrid biometric recognition[1]. Currently, hybrid multimodal biometric recognition generally fuses unimodal biometrics from five levels: pixel, feature, score, ranking and decision. As the level of fusion increases, less original data is kept, and various fusion methods have advantages and disadvantages[2]. Due to its capacity to balance the amounts of information in many modalities and its processing simplicity, score-level fusion has been extensively studied. The heterogeneity of scores – that is, the fact that scores come from non-homogeneous modalities or distinct matching algorithms – is the main challenge in score fusion. The distribution of scores within the same interval exhibits considerable variances as a result, which has an impact on the significance of the meaning of scores.

The significance of the matching score for pattern recognition-based approaches relies on the algorithm employed to acquire the score and is often evaluated in terms of similarity or distance, which have no real-world applications. However, in ranking partition collision (RPC) theory, the virtual score meaning may be further

Research Article
Manuscript received on April 22, 2022; accepted on December 12, 2022; published online on April 13, 2023
Recommended by Associate Editor Ji-Liang Tang
Colored figures are available in the online version at https://link.springer.com/journal/11633

stated as a score value when paired with ranking information and given various weights. The approach applied RPC theory combines many score values to estimate the value of the evidence. It combines the many score meanings while transforming them into useful metrics for practical applications, which has a large influence on multimodal fusion and real-world applications in public security.

The matching scores and ranking ranges of unimodal and multimodal biometric retrieval have a significant impact on the detection of fake results in application scenarios. This work develops a method based on the Ranking Partition Collision theory to accomplish multimodal biometric fusion at the score level. Furthermore, this work is validated on a real-world biometric database. This paper makes the following contributions:

1) This work proposes RPC theory. The theory focuses on matching scores and ranking ranges, which is in accordance with the current demands of public security agencies. Score-level fusion is used to achieve fusion, which is governed by the ranking level.

2) The presented method is not intended for use with any specific modalities. The suggested approach is based on combining scores from distinct modalities, and the fused modalities are not restricted in any way. In this situation, modalities for fusion can be flexibly chosen based on the actual needs.

3) The suggested method is validated on a real-world biometric database to ensure that the RPC theory is effective and robust, as well as that it is reasonable in practice.

2 Related work

Researchers have been concentrating their efforts on biometric identification increasingly recently, leading to advancements in facial detection, palm print recognition and other perspectives[3–6].

Bruneli and Falavigna[7] were the first to apply fusion biometrics for authentication, and they used weighted averages of facial pictures and voiceprints to prove the viability of modal fusion. Bigün et al.[8] developed the term "multimodal," combining facial pictures and voiceprints using a Bayesian statistical framework. Verlinde and Chollet[9] created a simultaneous fusion of various outputs of classifiers. Multimodal fusion was improved into feature-level fusion, score-level fusion, and decision-level fusion by Ross and Jain[10], laying the groundwork for the fast development of multimodal biometric identification technology. Pixel-level fusion focused on fusing the original data[11–14]. Ning and Chen[13] used sensor fusion algorithms to achieve image alignment and complete 3D image fusion after removing image gaps. Yaman et al.[14] verified the effectiveness of pixel-level fusion. Feature-level fusion was oriented towards features extracted from multiple biometric features[15, 16], Kong[17] used a double layer feature for fusion, which is more stable than single-layer feature fusion. Score-level fusion concentrates on the matching scores provided by a specific comparison algorithm and merges them to produce a final score[18, 19]. Li et al.[20] fused static and dynamic features of gait by the fraction fusion method, which effectively improved the accuracy of gait recognition.

The matching scores provided by a certain comparison algorithm are studied in score-level fusion. Score-level fusion is an appropriate fusion method for multimodal fusion because it can easily balance raw information with data processing. Score level fusion approaches are now classified into three groups[19, 21]: 1) Transformation-based fractional fusion methods: The basic purpose is score normalization, in which various modal scores are normalized to the same interval and then synthesized into final scores; 2) Classification-based fractional fusion methods: Each matching score of candidates is considered as an element of eigenvectors, and the scores of various modalities are merged into new spaces, in which each matching score of candidates is considered as a feature vector[22]; 3) Probability density-based methods: The final result is obtained by converting the scores into category probabilities and performing the fusion calculation based on the likelihood ratio test and the multiplication principle, so it is critical to estimate the category probability density correctly.

Transformation-based fusion methods. Brunelli and Falavigna[7] normalized facial and voice information and applied weights to scores, achieving score level fusion by combining matched scores and ranking information. Hong and Jain[22] performed score modelling and normalization by ranking information. Jain et al.[23] investigated score normalization using geometric aspects of the face, fingerprint and hand information, highlighting the benefits and drawbacks of typical normalizing approaches. To achieve more robust user recognition, Alsaade et al.[24] added unconstrained cohort normalization (UCN) into the process of score-level fusion of face and speech recognition.

Classification-based fusion methods. Wang et al.[25] employed linear discriminant analysis (LDA) and radial basis functional neural network (RBFNN) as classifiers to combine the matching scores of face and iris recognition into a two-dimensional eigenvector. Tulyakov and Govindaraju[26] suggested a global algorithm to improve biometric performance by including interclassifier information and accounting for the interdependence of the output scores of a single classifier. For a face and iris biometric system, Wang and Han[27] implemented support vector machine (SVM) with a radial basis function as the kernel function. Eskandari and Toygar[28] fused the two modalities after tanh normalization using local and global feature extractors on face and iris pictures, respectively. By developing user-dependent and user-independent classifiers and merging them with simple linear combinations, Ylmaz and Yanıkoğlu[29] showed that score fusion is superior to feature fusion in signature recognition. Aravinth and Valarmathy[30] integrated the benefits of a rule-based classifier, a lazy classifier, and a learning classifier to achieve score level fusion for faces, fingerprints, and irises, resulting in improved accuracy. Madane and Thepade[18] combined iris, palmprint, and fingerprint scores based on similarity and distinct color spaces via categorical ternary block truncation coding. Supreetha et al.[31] investigated score-level fusion systems using basic transformation rules (min, max, and sum), dynamic weighting coefficients, and triangle rules (Frank, Hamacher, and Sugeno Weber).

Probability density-based fusion methods. Nandakumar et al.[32] proposed a combinatorial framework for matching scores, modelling the density distribution of matching scores between users and impostors using the Gaussian mixture model (FGMM), which outperformed weighted mean and SVM rule methods. Peng et al.[33] introduced a triangle criterion-based score fusion approach to better differentiate the score distributions of users and impostors and obtain a lower mistake rate.

Despite the fact that more studies have been performed on score-level fusion, there are still issues to be resolved. Ranking partition collision theory is developed in this study to steer score-level fusion based on ranking information to accomplish cross-level fusion.

3 Ranking partition collision theory

It is often impossible to judge whether the input characteristics of claimants (the biometrics to be identified) are in the candidate list (the existing database), so the task is an open-set retrieval event, which has different aims and difficulties than the verification task. In the practical application of public security organs, attention is also paid to the ranking information as well as the score information. Therefore, this paper proposed the RPC theory to guide score-level fusion based on ranking information to achieve multimodal biometric fusion.

3.1 Theoretical content

The algorithm in this paper first builds a modal proprietary network to train unimodal biometric features and obtains unimodal matching scores and ranking information; second, it divides the ranking information into partitions, attributing different assignments to sensitive and non-sensitive partitions; and finally, it integrates multimodal biometric fusion and retrieval according to the partition assignment. Fig. 1 depicts the algorithm framework proposed in this article.

The basic substance of ranking partition collision theory includes two parts: partition selection and modal matching score collision. The selection of partitions will affect the final fusion accuracy since it is based on the ranking information to split sensitive and non-sensitive partitions and assign values. We may direct the matching scores of multiple modes to merge by assigning partitions, which can effectively increase fusion accuracy in the situation of weak retrieval accuracy.

3.2 Formal description of the problem

We hypothesize multimodal biometric identification in

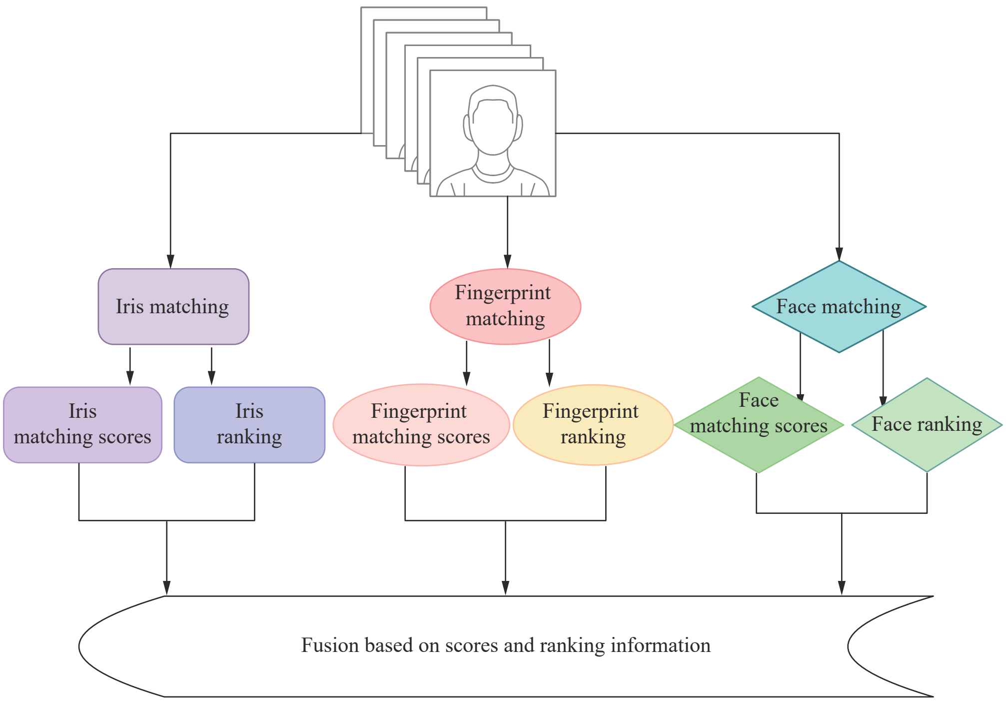

Fig. 1 Framework of ranking partition collision theory

a substrate with T candidates. Assume that candidates are $m\ (m = 1, 2, \cdots, T)$, that Y modal classifiers $I_y\ (y = 1, 2, \cdots, Y)$ are involved in the recognition process and that x denotes the biometric profile of an unknown claimant. Biometric recognition is the process of ranking claimant x versus the candidate list.

Each modal classifier generates a score by sorting all candidates and generates a $T \times 1$ dimensional vector$r_y = \{r_{y1}, r_{y2}, \cdots, r_{yT}\}$. The source identity of the unknown claimant x is expressed as $w_x \in \varphi = \{1, 2, \cdots, m\}$, and the identity of claimant w_x is m, which can be expressed as $w_x = m$. For an unknown claimant x, each modal classifier I_y generates a score ranking vector $r_y \in \mathbf{R}^{T\times 1}$, and all vectors form a $T \times Y$ dimensional matrix$r \in \mathbf{R}^{T\times Y}$, called the ranking score matrix. The row vectors correspond to all modal scores of a candidate, and the column vectors correspond to all candidate scores of a modality. For a ranking score matrix r, the number of elements in the corresponding value space is $(T!)^Y$. Denoting an element in this value space by n, then for a ranking score r, $r = n \in \{1, 2, \cdots, N\}$, $N = (T!)^Y$. Assuming that the proposed fusion decision function is f and that the scheme yields a decision based on the ranking score matrix r, the final decision is $f(x) \in \varphi = \{1, 2, \cdots, m\}$.

The recognition problem for x can be expressed as an optimization problem that maximizes the sum of the probabilities of identifying the unknown claimant w_x correctly over the full space with the ranking score matrix r and the fused decision scheme$f(x)$.

$$\max \sum_{m=1}^{T} \sum_{n=1}^{N} P\{f(x) = m, w_x = m, r = n\}. \tag{1}$$

As f only relies on the ranking score matrix r throughout the decision-making process, the decision is based only on r and is independent of w_x, the first term of (2) can be simplified. Decomposition and simplification of (1) using the Bayesian rule results in

$$\sum_{m=1}^{T} \sum_{n=1}^{N} P\{f = m|r = n\} P\{w_x = m, r = n\}. \tag{2}$$

In (2), the first term $P\{f = m|r = n\}$ depends on the fusion decision process, and the probability value can be expressed as $a_{mn} \in [0, 1]$; the second term $P\{w_x = m, r = n\}$ represents the ranking result of each modal classifier, depends on the performance of a classifier and is not related to fusion scheme f. This term can be further expressed as $P\{r = n|w_x = m\}P\{w_x = m\}$, and the joint probability can be determined by combining score matrices based on the different identities of the unknown claimant w_x.

By defining the fusion decision process $P\{f = m|r = n\}$ as taking values of $a_{mn} \in [0, 1]$, the objective function in (2) can be further expressed as

$$\sum_{m=1}^{T} \sum_{n=1}^{N} a_{mn} P\{f = m|r = n\}. \tag{3}$$

The final expression of the optimization problem is

Optimizing targets :

$$\max a_{mn\,q=1,2,\cdots,T,\ n=1,2,\cdots,N} \left\{ \sum_{m=1}^{T} \sum_{n=1}^{N} a_{mn} P\{f = m|r = n\} \right\}.$$

Qualifying conditions :

$$\sum_{m=1}^{T} a_{mn} = 1, \quad n = 1, 2, \cdots, N. \tag{4}$$

Since $P\{w_x = m, r = n\}$ is non-negative, the optimal solution to this optimization problem is

$$a^*_{mn} = \begin{cases} 1, & \text{if } m = \arg\max T\{w_x = k, r = n\} \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

3.3 Mathematical feasibility derivation

When the capacity of the known dataset is T and the number of modal classifiers is Y, to obtain the joint probability result of (7), it is necessary to consider all the possibilities of candidates and the ranking score matrix in a space ε, which has a total of potential scenarios. There are a total of $T \times (T!)^Y$ possible solutions in the event space. However, in this event space, most of the solutions are not optimal, and the full event calculation results in a waste of computational resources. To improve computational efficiency, the whole event space can be divided into non-intersecting subspaces, and then the sum of the probability of each event can be calculated on each subspace. Assuming that the mapping $\boldsymbol{Z}$ divides the full space ε into discrete event subspace partitions $M_{\boldsymbol{Z}}$, if the different partitions $\boldsymbol{Z}_w$ form an ordered set $D_{\boldsymbol{Z}} = \{\boldsymbol{Z}_1, \boldsymbol{Z}_2, \cdots, \boldsymbol{Z}_{M_{\boldsymbol{Z}}}\}$, then the partition M defines a new random variable:

$$d_{\boldsymbol{Z}} : \mathcal{I} \times \mathcal{S} \mapsto \{1, 2, \cdots, M_{\boldsymbol{Z}}\}. \tag{6}$$

In (8), $\mathcal{I}$ is the list of candidate identities, and $\mathcal{S}$ is the ordered set of score matrices. The random variable $d_{\boldsymbol{Z}}$ maps the full event space $\mathcal{I} \times \mathcal{S}$ to the set of sequences of an event partition so that the random variable can be expressed as $d_{\boldsymbol{Z}} = \boldsymbol{Z}(m, n)$. At this point, we can further expand the objective function in (2) by defining its summation over the partitioned subspace, which yields

$$\sum_{m=1}^{T}\sum_{n=1}^{N} P\ f(x)=m,\ \ w_x=m,\ \ r=n,\ \ d_{\boldsymbol{Z}}=\boldsymbol{Z}(m,n). \tag{7}$$

Using the Bayesian rules, we can obtain

$$\begin{aligned}&P\left\{f(x)=m, w_x=m, r=n, d_{\boldsymbol{Z}}=\boldsymbol{Z}(m,n)\right\}= \\ &\quad P\left\{f(x)=m \mid w_x=m, r=n, d_{\boldsymbol{Z}}=\boldsymbol{Z}(m,n)\right\}\times \\ &\quad P\left\{w_x=m, r=n \mid d_{\boldsymbol{Z}}=\boldsymbol{Z}(m,n)\right\}\times \\ &\quad P\left\{d_{\boldsymbol{Z}}=\boldsymbol{Z}(m,n)\right\}.\end{aligned} \tag{8}$$

Since the fusion decision f is based only on the score matrix, the combination (9) and (10) can obtain

$$\begin{aligned}&\sum_{m=1}^{T}\sum_{n=1}^{N} P\{f(x)=m|r=n\}\times \\ &\quad P\{w_x=m, r=n|d_{\boldsymbol{Z}}=\boldsymbol{Z}(m,n)\}\times \\ &\quad P\{d_{\boldsymbol{Z}}=\boldsymbol{Z}(m,n)\}.\end{aligned} \tag{9}$$

Equation (11) contains three probabilities. The first term is directly related to the fusion decision method f. The third term is related to the performance of the classifier and depends on the classification result of the classifier on the partition $\boldsymbol{Z}(m,n)$. The second term can be considered a transition term from partitions to sub-events. It defines the probabilistic relationship between the partitioning of coarse data and the source identity and ranking score matrix of individual data. In (11), the optimal solution of the fusion decision process depends on the partitioning of the data. Similarly, in terms of data partitioning, the objective function in (5) consists of two main components. The first is the space spanned by the free parameter a_{mn}, that is, the free parameter space. The second is to combine all the classifier result statistics across the space, that is, all the estimation parameters $P\{r=n, w_x=m\}$, which is also named partition.

When choosing a ranking-based multi-classifier fusion method, it is often impossible to obtain an optimal solution to a practical problem because the classifier behavior cannot be predicted. An accurate estimate of the full probability of the classifier cannot be made, so a reasonably specific partition needs to be selected. In this paper, the ranking information is partitioned, and the matching scores are fused with the ranking information. Usually, the ranking information is more reliable than the scores, especially when the unknown biometrics are of low quality; then, both the true identity and the impersonation scores will be low. At this point, when the scores of different modalities are fused, the ranking information is still a stable quantity, and the source identity may be identified in combination with other modal ranking information. Conversely, when the classifier yields incorrect results, the source identity usually lies in the top ranks. In this study, the ranking score matrix is divided into two parts, namely, a sensitive partition and a non-sensitive partition. The fusion method obtains the optimal solution when the probability of a correct decision is maximized, at which point the classifier observation statistic $P\{d_{\boldsymbol{Z}}=\boldsymbol{Z}(m,n)\}$ obtains a unique value that satisfies

$$\begin{aligned}&P\{r=n, w_x=k|d_{\boldsymbol{Z}}=\boldsymbol{Z}(k,n)\}\times \\ &\quad P\{d_{\boldsymbol{Z}}=\boldsymbol{Z}(k,n)\}\geq \\ &\quad P\{r=n, w_x=m|d_{\boldsymbol{Z}}=\boldsymbol{Z}(k,n)\}\times \\ &\quad P\{d_{\boldsymbol{Z}}=\boldsymbol{Z}(k,n)\},\ \ m=1,2,\cdots,T.\end{aligned} \tag{10}$$

The classifier at this point is close to the ideal classifier case.

3.4 Fusion method based on ranking partition collision theory

The selection of partitions is at the heart of the ranking partition collision theory-based fusion approach. Fig. 2 illustrates the three steps of the matching and identification process that take place when a trace is discovered: mass flow monitoring, refined investigation, and expert identification. The trace is compared to all trace databases during the mass flow monitoring phase to identify a "relevant trace database" that is connected to the trace, as shown in Outcome 1. Most biometric automated identification algorithms can provide the findings of this step, and investigators can complete it without having specialized skills. The top traces from Outcome 1 are further compared with the suspect traces during the refined investigation phase to provide a list of candidates with ranking information, as shown in Outcome 2. It should be emphasized that the ranking of the candidate list is only an algorithmic ranking, and different algorithms may provide results that are noticeably different from one another. A multimodal recognition algorithm in $1:N$ identification mode, with the investigators as the executive, can be used to produce the findings of this phase. When an identification document is issued, which is Outcome 3, the trace is recognized with each of the top-ranked candidates individually by an expert with specialized expertise, and it becomes evidence. Due to constraints on time, effort, and case efficiency, it is obvious that the higher up on the candidate list the source identification of the claimant appears, the more likely it is to be accepted as evidence – or what we refer to as having a greater evidential value.

Because the outcome of a distance classifier is determined by the distance between features in the feature space, the scores fall as the distances grow. In many cases, the feature distance at the top of the ranking is more sensitive than that at the bottom. In other words, a

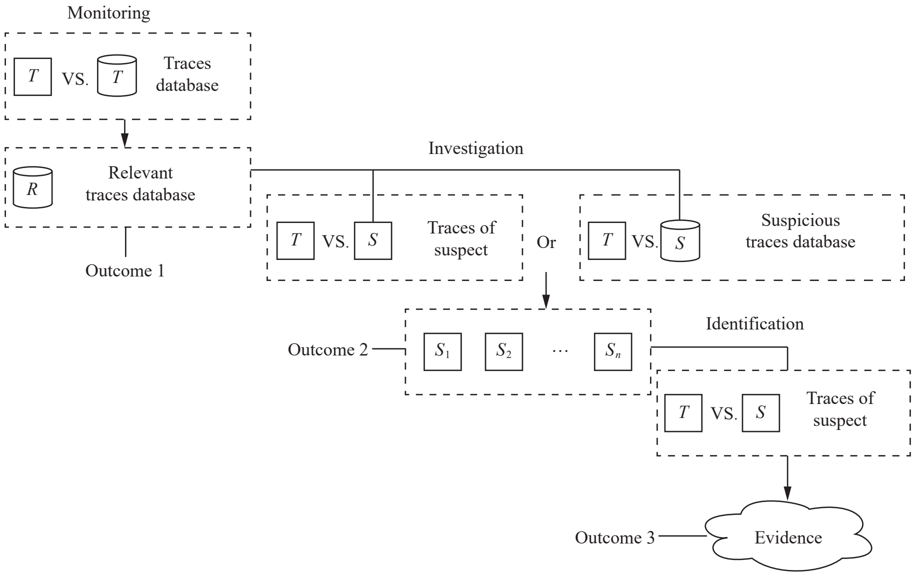

Fig. 2 Overview of the current forensic processing of a trace

smaller increase in the distance between features in the top ranking tends to result in a change in the ranking, whereas the distance between features in the bottom ranking is likely to result in a larger change in the ranking. As a consequence, higher-order information is more vulnerable. The divisions are divided as follows in this paper:

$$\hat{r}_{my} = \begin{cases} 1, & \text{if } r_{my} \geq \mu_y \\ 0, & \text{if } r_{my} < \mu_y \end{cases} \tag{11}$$

where μ is the algorithm parameter and r_{my} denotes the ranking of the m-th candidate's y-th modality. (13) indicates that among all the T rankings, this paper gives higher attention to the first μ_y rankings and gives equal attention to the latter $T - \mu_y$ candidates. That is, the current ranking has only $\mu_y + 1$ classes. The new ranking score matrix is defined as $\hat{R}$. Define $\beta(\hat{r}_{m_1}, \hat{r}_{m_2}, \cdots, \hat{r}_{m_Y})$ as the number of scores of zero in the ranking score matrix $\hat{R}$ for all Y modalities of the candidate m. Combined with (13), this is the number of modalities for which the ordering of all Y modalities is at post $T - \mu_y$. Then, for candidate m, the final fusion score can be defined as

$$\hat{s}_m = \sum_{y=1}^{Y} s_{my} \times c_{my} \tag{12}$$

$$c_{my} = \begin{cases} 1, & \text{if } \hat{r}_{qt} \neq 1 \\ T - k, & \text{if } \hat{r}_{qt} = 1 \text{ and } \beta(\hat{r}_{m_1}, \hat{r}_{m_2}, \cdots, \hat{r}_{m_Y}) = k \end{cases} \tag{13}$$

where $\hat{s}_m$ denotes the fusion score for all modalities for candidate m, and c_{my} represents the weight of each modality.

4 Experimental configuration

4.1 Dataset

Biometrics can be divided into physical characteristics (face, fingerprint, iris, etc.) and behavioral characteristics (gait, voiceprint, handwriting, etc.). Behavioral biometrics are easier to obtain than physiological biometrics, but they are less safe and reliable.

The proposed approach does not apply to specific modalities; rather, it merely integrates on the basis of modal proprietary network networks. It is possible to select any modality that may deliver equal scores in this circumstance. For the time being, public security agencies are mandated to keep track of suspects' fingerprints, faces, irises, and voiceprints and have built ministerial-scale databases. This study chooses three common physiological biometrics, namely, face, fingerprint, and iris, as the research subjects for the fusion of score level, based on the practical demands of public security organs.

Due to the unavailability of a large-scale opensource multimodal biometric dataset, this work uses unimodal public datasets from the Institute of Automation, Chinese Academy of Sciences, China, and Tsinghua University, China as virtual homologous multimodal biometric datasets. The treatment of virtual homology has no effect on the results since there is no association between homologous heterogeneous and heterogeneous biometrics[34, 35].

The raw data in this research are treated in particular to imitate low-quality data collected from cases. No additional processing is performed due to the poor quality and low resolution of the facial picture. Only 30% of the original fingerprint picture quality and 20% of the original iris image quality are kept when the iris and fingerprint images are shrunk to 1/4 of their original size. The composition of the dataset is given in Table 1.

Table 1 Composition of the dataset

	Face	Fingerprint	Iris
Training set	11 060	11 060	11 060
Gallery	1 449	1 449	1 449
Query	985	985	985

4.1.1 Iris dataset

The iris dataset in this paper consists of two parts: The first is the CASIA-IrisV4-Interval dataset[36], which contains 357 categories of data with a resolution of 320 × 280 and jpg format; the second is derived from data collected by public security organs using a homemade CASIA iris collection device, which contains 2 341 categories of data with a resolution of 640 × 320 and bmp format. The combined dataset is more closely matched to the practical application situation because the second portion of the data is obtained in a non-laboratory context.

We preserve more than five samples in each category to guarantee that the experiment is as accurate as possible. This article also preprocesses iris pictures to remove superfluous information such as corner, eyelash, and eyelid information. The Hough transform is used to locate the inner and outer circles of the iris, and the outer circle diameter is used as the edge length to crop the image to squares. Each of the 2 712 categories in the combined iris dataset has at least five items.

4.1.2 Fingerprint dataset

The fingerprint dataset utilized is the CASIA-Fingerprint V5 dataset[36]. The dataset, which has been collected using a URU4000 fingerprint sensor, has over 20 000 fingerprint scans from 500 people with a resolution of 328 × 356 pixels and bmp format. Each person collected data on eight fingers (thumb, index, middle, and ring) and five images of each finger, for a total of 40 images. Fingerprints from various fingers are considered separate categories in this work to expand the dataset. The therapy has little effect on the eventual outcome since homologous biometrics have no association[34, 35]. We only utilized the top 2 712 categories of the fingerprint dataset to match with the iris dataset.

4.1.3 Face dataset

The face dataset is the WebFace260M dataset[37]. The algorithm purifies the web face images after millions of alignments using an image resolution of 112 × 112 and jpg format. The challenge with this dataset is that it is based on online data that have not been cleaned, and the photos within the class span a wide time period with many fluctuations. Moreover, the resolution of the dataset is inadequate. This research only utilizes the top 2 712 categories of the face dataset, each having no less than five categories, to match with the iris and fingerprint databases.

4.2 Experimental settings

This research investigates the fusion method by processing the dataset and selecting the traditional common backbone network to mimic the low accuracy of the unimodal algorithm due to a lack of a large-scale and low-quality multimodal biometric dataset.

The backbone networks in this article are the standard convolutional neural networks VGG16, ResNet50, DenseNet169, and VGG16-BN with the addition of a batch normalization layer before the activation function. In this study, the same network structure is employed to train the various modal data, resulting in more generalized and generalized outputs. Additionally, in this study, the algorithm parameter μ is taken to be $\mu = 10$.

This study employs assessment measures such as mAP (mean average precision), Rank N (N takes 1, 5, 10) and performance indicators such as CMC (Cumulatively match charitable) curves.

The following experimental settings are employed in this study: Ubuntu 20.04 operating system, AMD EPYC 7 702 CPU running at 2.0 GHz, 128 GB of RAM, and an NVidia RTX3090 graphics card. PyTorch1.7 is the experimental framework, with an epoch of 120, stochastic gradient descent (SGD) optimizer, base learning rate of 0.004, weight decay of 0.000 1, learning rate descent technique of cosine, and input picture of 256 × 256.

5 Results and discussion

5.1 Unimodal biometric recognition

The main research subject of score-level fusion is the matching score given by the modal classifier. The outputs of the unimodal network need to be transformed first. In this paper, unimodal matching scores are obtained using both Euclidean distance and cosine similarity. The score transformed according to the Euclidean distance can be regarded as a distance score, which is transformed according to the distance between features. Suppose the features of the claimant are $X = (x_1, x_2, \cdots, x_n)$, one of the candidate features is $Y = (y_1, y_2, \cdots, y_n)$, and the 2-norm is $\|X\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} = \|X\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} = 10$; then, $d = \|X - Y\|_2 = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$, $d \in [0, 20]$. At this point, the distance score is $S_d(X, Y) = (20 - d)/20$. The unimodal accuracy using the Euclidean distance metric is shown in Table 2.

The scores transformed according to cosine similarity can be regarded as similarity scores, which are transformed according to the similarity between features. The formula is $d = \|X - Y\| = \frac{XY}{\|X\|_2 \|Y\|_2} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum x_i^2 \sum y_i^2}}, d \in$

$[-1, 1]$. In this case, the similarity score is $S_s(X, Y) = (d+1)/2$. The unimodal accuracy measured using the cosine similarity metric is shown in Table 3.

Table 2 Accuracy of unimodal biometrics using Euclidean measurements

Backbone	Modality	mAP (%)	Rank1 (%)	Rank5 (%)	Rank10 (%)
ResNet50	Iris	93.1	96.1	99.0	99.4
	Face	82.6	89.0	96.8	98.1
	Fingerprint	88.4	91.6	98.4	99.4
VGG16	Iris	91.6	94.8	98.3	99.3
	Face	81.9	88.1	95.6	97.7
	Fingerprint	86.1	89.7	98.3	99.4
VGG16_BN	Iris	92.6	95.8	99.2	99.6
	Face	84.0	89.5	97.4	98.4
	Fingerprint	88.5	92.1	98.6	99.4
DenseNet169	Iris	93.7	96.0	99.3	99.6
	Face	85.9	91.2	97.1	97.7
	Fingerprint	89.8	93.0	98.8	99.5

Table 3 Accuracy of unimodal biometrics using cosine measurements

Backbone	Modality	mAP (%)	Rank1 (%)	Rank5 (%)	Rank10 (%)
ResNet50	Iris	88.6	93.4	98.0	99.0
	Face	71.2	81.1	90.8	94.7
	Fingerprint	74.4	82.4	94.0	96.3
VGG16	Iris	86.4	92.3	97.1	98.5
	Face	68.6	79.1	90.4	94.1
	Fingerprint	67.7	74.7	92.0	95.9
VGG16_BN	Iris	86.8	92.4	97.5	98.6
	Face	72.9	81.7	92.1	94.5
	Fingerprint	71.3	78.3	93.7	96.9
DenseNet169	Iris	91.3	94.8	98.2	99.0
	Face	78.6	85.9	94.8	96.3
	Fingerprint	78.3	83.5	95.1	97.2

Fig. 3 illustrates the unimodal biometric accuracy based on Euclidean distance and cosine similarity. Fig. 3(a) presents the unimodal mAP line graph, whereas Fig. 3(b) depicts the unimodal Rank1 line graph. The results reveal that Euclidean distance has a much greater accuracy than cosine similarity and that Euclidean distance is a preferable distance description approach in biometric recognition.

Iris retrieval accuracy is substantially better than that of fingerprints and faces, indicating that irises are more distinctive and secure. Due to the difficulty of obtaining iris images, the problem can be mitigated to some extent by using commonly accessible biometrics such as faces and fingerprints as fusion possibilities. DenseNet169 is the most efficient of the four backbone networks for unimodal biometric retrieval, whereas VGG16 is the least effective.

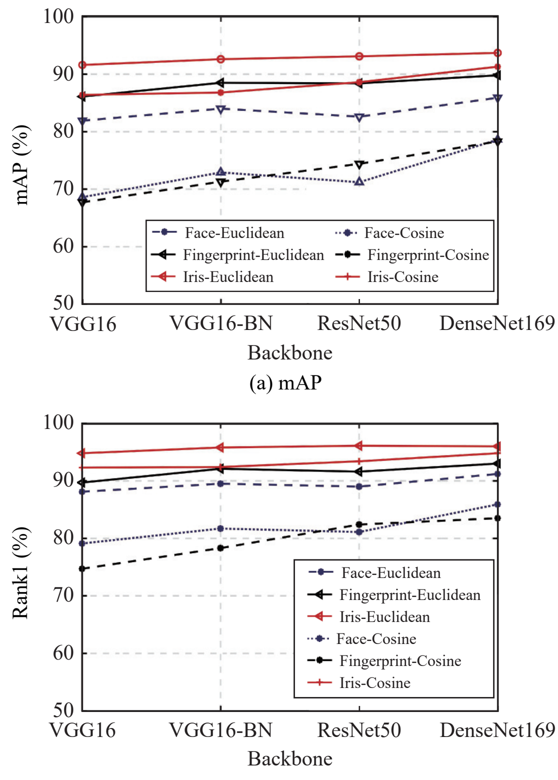

Fig. 3 Line graphs based on different distance description methods for mAP and Rank1 of unimodal biometrics

The CMC curves are used to assess the accuracy of unimodal biometric retrieval using various distance description approaches, as illustrated in Fig. 4. The performance of the Euclidean distance-based method is substantially superior to that of the cosine similarity method, as seen intuitively in Fig. 4. Likewise, DenseNet169 training is the most appropriate in unimodal retrieval.

5.2 Score level fusion based on ranking partition collision theory

The backbone network in the score level fusion strategy based on ranking partition collision theory is trained using Euclidean distance and cosine similarity. Table 4 shows the experimental results between our method and Yaman's[14], which serves as a baseline. The ranking partition collision theory introduced in this paper performed well on the test set, with 100% mAP and Rank1 in all four backbone networks, as shown in the table. The findings revealed that the proposed method performed better for both Rank1 and mAP. Rank1 and mAP are both improved by 2.9% and 4.5%, respectively, using the approach when assessed in terms of Euclidean distance. mAP and Rank1 are both improved by 6.3% and 2.7%, respectively, via the cosine distance approach.

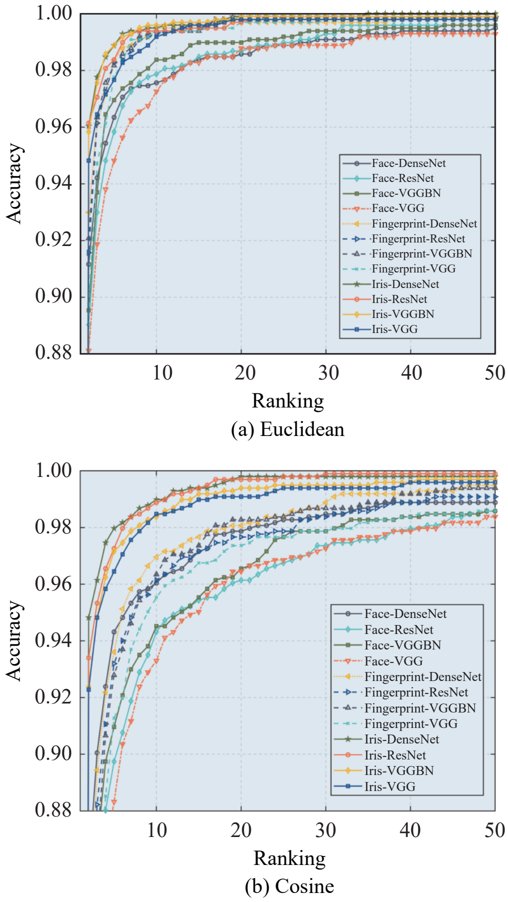

Fig. 4 CMC curves of unimodal biometric retrieval based on different distance description methods

The CMC curves of the proposed approach and other methods are compared in Fig. 5. The proposed scheme has a surprising impact, with the proposed fusion scheme mAP increasing to 100% in cases when the peak mAP of unimodal is less than 95% and the lowest mAP is less than 85%.

The proposed approach offers a very evident and considerable improvement, as shown by the CMC graph. The employment of the commonly available biometric features face and fingerprint, as well as the high security biometric feature iris, as fusion alternatives is fair and reasonable. The ranking partition collision theory has been adapted to the practical needs of public security organs, and the fusion system based on this theory achieves 100% mAP and Rank1 accuracy.

5.3 Discussion

According to the findings, the score-level fusion algorithm paired with RPC theory is efficient and in-

Table 4 Performance of the score fusion scheme based on ranking partition collision theory

Distance	Backbone	Method	mAP (%)	Rank1 (%)	Rank5 (%)	Rank10 (%)
Euclidean	VGG16	Our	100	100	100	100
		[14]	95.5	97.1	99.9	100
	VGG16-BN	Our	100	100	100	100
		[14]	96.9	98.6	100	100
	ResNet50	Our	100	100	100	100
		[14]	96.2	97.1	100	100
	DenseNet169	Our	100	100	100	100
		[14]	98.1	99.1	100	100
Cosine	VGG16	Our	99.5	99.9	100	100
		[14]	93.7	97.3	99.8	100
	VGG16-BN	Our	99.7	100	100	100
		[14]	95.0	97.9	100	100
	ResNet50	Our	99.5	99.8	100	100
		[14]	93.6	96.4	99.7	100
	DenseNet169	Our	99.9	100	100	100
		[14]	96.1	98.1	99.8	100

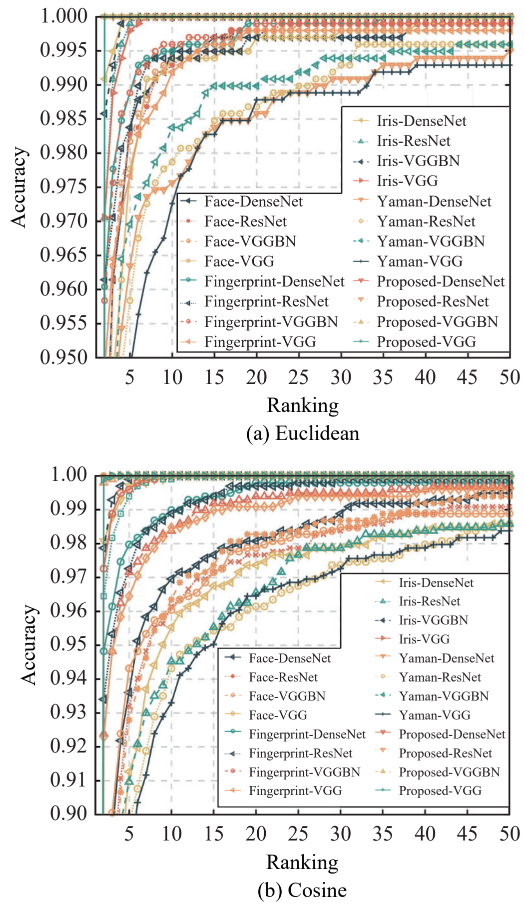

Fig. 5 CMC curve of the proposed scheme and other methods

creases the rate at which biometric data correctly match individuals and the effectiveness of the case solution. Suspicious traces for a case go through three steps before becoming admissible evidence in court: mass flow monitoring, refined investigation, and expert identification. Through successive technological advancements, we narrow the identity of a subject, and the standards for algorithm reliability are progressively raised.

Most automatic biometric identification algorithms may be used to derive the mass flow monitoring phase. In these algorithms, a wider number of candidates are circled, and the source identity of the suspicious trace is only anticipated to fall within the circled range. The source identification of the suspicious trace is likely to be at the top of the list of candidates during the refine investigation phase, which involves the fusion of many modalities and can be obtained through a multimodal fusion algorithm. The top of the candidate list must be manually matched by experts during the expert identification step; therefore, the more reliable the source identity of the suspect trace at the top of the list, the more likely it is to become evidence. Direct identification of experts based on algorithm ranking results may lead to a major loss of efficiency and effort. To give the scores meaningful worth, this study divides the ranking information into partitions and assigns different weights to each section. Few studies have evaluated the meaning of the scores from the standpoint of the score values since the majority of pattern recognition-based biometric approaches present ways from the perspective of mathematical statistics, rendering the given scores practically meaningless. The method combined with RPC theory fuses the matching scores of various modalities that contain the score value information so that the score value gradually approaches the evidence value. RPC theory, on the other hand, assigns different partition weights to the ranking so that the candidate list ranking contains the score value information. It can fully meet both forensic and judicial needs.

5.4 Additional experiment on real-world dataset

This research uses the real-world database collected by public security organs from real cases to validate the performance of the suggested algorithm and its rationale for practical applications. A database with a capacity of 1.1 million persons and 500 people registered to be retrieved is being built. For validation experiments, the top unimodal recognition software development kit (SDK) in China is employed as the modal proprietary network, and the matching scores provided by the unimodal SDK are used as the foundation for the fusion of the score level.

The fingerprint recognition algorithm adopts the fingerprint recognition SDK of Tongyuanwei Intelligent Technology. The face recognition algorithm adopts the open source offline face recognition SDK of ArcSoft. The iris recognition algorithm adopts the iris recognition SDK from IrisKing.

The received data are further processed to verify the robustness of the proposed methodology and to raise the complexity of the validation experiments. During the test, some data are exchanged for the same modal, as shown in Table 5. For claimants 1–50, the face uses information from 301–350, and the rest of the modalities are their own; for claimants 51–100, the fingerprint uses information from 351–400, and the rest of the modalities are their own; for claimants 101–150, the iris uses information from 401–450, and the rest of the modalities are their own. The facial information of claimants 1–50 comes from claimants 1–50 because the transferred information is included in the test set. The face information of claimants 1–50 comes from claimants 301–350, as the transferred information is included in the test set. This indicates that claimants 1–50′s face modal will provide high ratings, which is "false". As a result, it is possible that it will have an influence on the final recognition result, making the validation experiment incredibly tough right now.

Table 6 shows the accuracy of the unimodal and proposed approaches. The focus here is exclusively on the accuracy of the algorithm before Rank20 due to the great accuracy of the unimodal itself. In this validation experiment, there are two major challenges. First and foremost, the database is massive, with data derived from biometrics collected from genuine cases. Second, this artificial procedure reduces the accuracy of unimodal SDK. Simultaneously, the modal proprietary network may produce higher but "wrong" ratings, indicating that unimodal cannot be employed or that information is entered improperly, necessitating more resilience.

As seen in Table 6, adjusting the obtained data makes identification more difficult, and the accuracy of unimodal recognition declines dramatically, with Rank1 accuracy falling below 80% and Rank20 accuracy falling below

Table 5 Processing of test data

	Index	1–50	51–100	101–150	151–300	301–350	351–400	401–450	451–500
Test	Face	301–350	–	–	–	1–50	–	–	–
	Fingerprint	–	351–400	–	–	–	51–100	–	–
	Iris	–	–	401–450	–	–	–	101–150	–

90%. The proposed algorithm, however, is still quite accurate; Rank1 is 99.2% correct, and the CMC graph is given in Fig. 6. The validation experiments assess the proposed method in terms of data scale and algorithm robustness, confirming that it still works effectively in the presence of large datasets and unimodal data instability. It also shows that the presented method is suitable for real-world cases.

Table 6 Validation results of the proposed algorithm

Index	Face	Fingerprint	Iris	Proposed
1	0.796	0.794	0.798	0.992
5	0.798	0.820	0.798	0.994
10	0.798	0.846	0.802	0.996
20	0.800	0.876	0.808	0.996

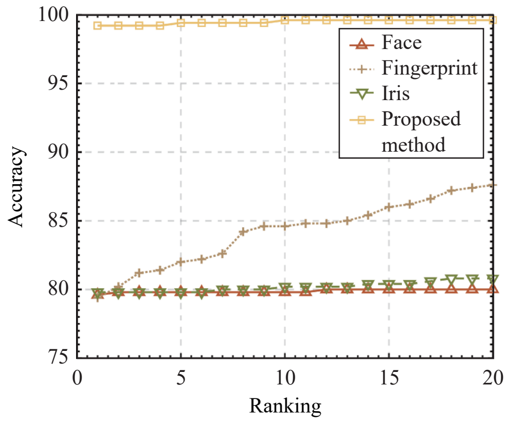

Fig. 6 Curve of CMC for the proposed method on the million-sized dataset

It is vital to note that the unimodal accuracy here relates solely to the accuracy of presentation after data adjustment and does not mean the full accuracy of the unimodal algorithms. The "wrong" discrimination seen here is due to data adjustment rather than false acceptance or false rejection in the unimodal identifications.

6 Conclusions

This paper introduces ranking partition collision theory and applies it to multimodal biometric fusion at the score level, delivering advanced results in both virtual homogeneous datasets and real-world applications. The proposed method combines ranking information with scores in practical applications and partitions them with different weights so that the virtual scores have realistic meanings and are eventually brought close to evidential values through multimodal fusion. It completes the transfer from the meaning of scores to the value of scores and, ultimately, to the value of evidence. This work achieves the unification of the meaning of heterogeneous scores. Existing approaches simply consider automated identification from a mathematical and statistical standpoint, but this paper applies RPC theory to address forensic and judicial demands, greatly assisting investigators and enhancing the effectiveness of case solving. This research makes significant advancements in the fields of multimodal biometric fusion and forensic science. However, this work still suffers from the problem of insufficient data due to the difficulty of acquiring biometrics. The next step is to increase the size of the dataset, and it could be possible to work together on further levels of fusion.

Acknowledgements

This work was supported by Double First-Class Innovation Research Project for People's Public Security University of China (No.2023SYL06).

Declarations of conflict of interest

The authors declared that they have no conflicts of interest to this work.

References

[1] J. A. Unar, W. C. Seng, A. Abbasi. A review of biometric technology along with trends and prospects. *Pattern Recognition*, vol. 47, no. 8, pp. 2673–2688, 2014. DOI: 10.1016/j.patcog.2014.01.016.

[2] Z. N. Sun, R. He, L. Wang, M. N. Kan, J. J. Feng, F. Zheng, W. S. Zheng, W. M. Zuo, W. X. Kang, W. H. Deng, J. Zhang, H. Han, S. G. Shan, Y. L. Wang, Y. W. Ru, Y. H. Zhu, Y. F. Liu, Y. He. Overview of biometrics research. *Journal of Image and Graphics*, vol. 26, no. 6, pp. 1254–1329, 2021. (in Chinese)

[3] L. Song, J. F. Yang, Q. Z. Shang, M. A. Li. Dense face network: A dense face detector based on global context and visual attention mechanism. *Machine Intelligence Research*, vol. 19, no. 3, pp. 247–256, 2022. DOI: 10.1007/s11633-022-1327-2.

[4] L. Y. Xu, Z. Gajic. Improved network for face recognition based on feature super resolution method. *International Journal of Automation and Computing*, vol. 18, no. 6, pp. 915–925, 2021. DOI: 10.1007/s11633-021-1309-9.

[5] M. Jacquet, C. Champod. Automated face recognition in forensic science: Review and perspectives. *Forensic Science International*, vol. 307, pp. 110–124, 2020. DOI: 10.1016/j.forsciint.2019.110124.

[6] L. Fei, B. Zhang, Y. Xu, C. Tian, D. Zhang. Jointly heterogeneous palmprint discriminant feature learning. *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 9, pp. 4979–4990, 2021. DOI: 10.1109/TNNLS.2021.3066381.

[7] R. Brunelli, D. Falavigna. Person identification using multiple cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 10, pp. 955–966, 1995. DOI: 10.1109/34.464560.

[8] E. S. Bigün, J. Bigün, B. Duc, S. Fischer. Expert conciliation for multi modal person authentication systems by Bayesian statistics. In *Proceedings of the 1st International Conferences on Audio-based and Video-based Biometric Person Authentication*, Springer, Crans-Montana, Switzerland, pp. 291–300, 1997. DOI: 10.1007/BFb0016008.

[9] P. Verlinde, G. Chollet. Combining vocal and visual cues in an identity verification system using K-NN based classifiers. In *Proceedings of the 2nd IEEE Workshop on Multimedia Signal Processing*, Redondo Beach, USA, pp. 59–64, 1998. DOI: 10.1109/MMSP.1998.738913.

[10] A. Ross, A. Jain. Information fusion in biometrics. *Pattern Recognition Letters*, vol. 24, no. 13, pp. 2115–2125, 2003. DOI: 10.1016/S0167-8655(03)00079-5.

[11] Y. L. Zhang, J. Yang, H. T. Wu, Y. F. Xue. Swipe fingerprint sequences mosaicking based on the extension of phase-correlation. *Journal of Shanghai Jiaotong University*, vol. 40, no. 3, pp. 471–475, 2006. DOI: 10.16183/j.cnki.jsjtu.2006.03.024. (in Chinese)

[12] J. G. Heo, S. G. Kong, B. R. Abidi, M. A. Abidi. Fusion of visual and thermal signatures with eyeglass removal for robust face recognition. In *Proceedings of Conference on Computer Vision and Pattern Recognition Workshop*, IEEE, Washington DC, USA, 2004. DOI: 10.1109/CVPR.2004.351.

[13] X. L. Ning, Z. X. Chen. Face panoramic image mosaic algorithm for face diagnosis. *Chinese Journal of Medical Physics*, vol. 37, no. 4, pp. 456–462, 2020. DOI: 10.3969/j.issn.1005-202X.2020.04.011. (in Chinese)

[14] D. Yaman, F. I. Eyiokur, H. K. Ekenel. Multimodal age and gender classification using ear and profile face images. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, IEEE, Long Beach, USA, pp. 2414–2421, 2019. DOI: 10.1109/CVPRW.2019.00296.

[15] A. Rattani, M. Tistarelli. Robust multi-modal and multi-unit feature level fusion of face and iris biometrics. In *Proceedings of the 3rd International Conferences on Advances in Biometrics*, Springer, Alghero, Italy, pp. 960–969, 2009. DOI: 10.1007/978-3-642-01793-3_97.

[16] Q. Hu, C. D. Wu, J. N. Chi, X. S. Yu, H. Wang. Multi-level feature fusion facial expression recognition network. In *Proceedings of Chinese Control and Decision Conference*, IEEE, Hefei, China, pp. 5267–5272, 2020. DOI: 10.1109/CCDC49329.2020.9164733.

[17] J. Kong. Biometric identification based on two-layer feature fusion. *Journal of Beihua University* (*Natural Science*), vol. 21, no. 1, pp. 110–117, 2020. DOI: 10.11713/j.issn.1009-4822.2020.01.024. (in Chinese)

[18] M. S. Madane, S. D. Thepade. Score level fusion based Multimodal biometric identification using Thepade's sorted ternary block truncation coding with variod proportion of iris, palmprint, left fingerprint & right fingerprint with asorted similarity measures & different colorspaces. In *Proceedings of International Conference on Automatic Control and Dynamic Optimization Techniques*, IEEE, Pune, India, pp. 824–828, 2016. DOI: 10.1109/ICACDOT.2016.7877702.

[19] K. Nandakumar, Y. Chen, S. C. Dass, A. Jain. Likelihood ratio-based biometric score fusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 342–347, 2008. DOI: 10.1109/TPAMI.2007.70796.

[20] H. A. Li, Z. M. Du, Z. L. Li, Q. J. Hui, J. H. Bai. Research on gait recognition algorithm based on double features using the layer matching fusion method. *Journal of Graphics*, vol. 40, no. 3, pp. 441–446, 2019. DOI: 10.11996/JG.j.2095-302X.2019030441. (in Chinese)

[21] M. Eskandari, Ö. Toygar, H. Demirel. A new approach for face-iris multimodal biometric recognition using score fusion. *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 27, no. 3, Article number 1356004, 2013. DOI: 10.1142/S0218001413560041.

[22] L. Hong, A. Jain. Integrating faces and fingerprints for personal identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1295–1307, 1998. DOI: 10.1109/34.735803.

[23] A. Jain, K. Nandakumar, A. Ross. Score normalization in multimodal biometric systems. *Pattern Recognition*, vol. 38, no. 12, pp. 2270–2285, 2005. DOI: 10.1016/j.patcog.2005.01.012.

[24] F. Alsaade, A. M. Ariyaeeinia, A. S. Malegaonkar, M. Pawlewski, S. G. Pillay. Enhancement of multimodal biometric segregation using unconstrained cohort normalisation. *Pattern Recognition*, vol. 41, no. 3, pp. 814–820, 2008. DOI: 10.1016/j.patcog.2007.06.028.

[25] Y. H. Wang, T. N. Tan, A. K. Jain. Combining face and iris biometrics for identity verification. In *Proceedings of the 4th International Conferences on Audio-based and Video-based Biometric Person Authentication*, Springer, Guildford, UK, pp. 805–813, 2003. DOI: 10.1007/3-540-44887-X_93.

[26] S. Tulyakov, V. Govindaraju. Use of identification trial statistics for the combination of biometric matchers. *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 4, pp. 719–733, 2008. DOI: 10.1109/TIFS.2008.2004287.

[27] F. Wang, J. Han. Multimodal biometric authentication based on score level fusion using support vector machine. *Opto-electronics Review*, vol. 17, no. 1, pp. 59–64, 2009. DOI: 10.2478/s11772-008-0054-8.

[28] M. Eskandari, Ö. Toygar. Fusion of face and iris biometrics using local and global feature extraction methods. *Signal, Image and Video Processing*, vol. 8, no. 6, pp. 995–1006, 2014. DOI: 10.1007/s11760-012-0411-4.

[29] M. B. Yılmaz, B. Yanıkoğlu. Score level fusion of classifiers in off-line signature verification. *Information Fusion*, vol. 32, pp. 109–119, 2016. DOI: 10.1016/j.inffus.2016.02.003.

[30] J. Aravinth, S. Valarmathy. Multi classifier-based score level fusion of multi-modal biometric recognition and its application to remote biometrics authentication. *The Imaging Science Journal*, vol. 64, no. 1, pp. 1–14, 2016. DOI: 10.1080/13682199.2015.1104067.

[31] G. H. Supreetha, H. G. Kumar, M. Imran. Multimodal biometric verification system: Evaluation of various score level fusion rules. In *Proceedings of IEEE International Conference on Electrical, Computer and Communication*

Technologies, Coimbatore, India, pp. 1–4, 2019. DOI: 10.1109/ICECCT.2019.8869429.

[32] K. Nandakumar, A. K. Jain, A. Ross. Fusion in multibiometric identification systems: What about the missing data? In *Proceedings of the 3rd International Conferences on Advances in Biometrics*, Springer, Alghero, Italy, pp. 743–752, 2009. DOI: 10.1007/978-3-642-01793-3_76.

[33] J. L. Peng, A. A. A. El-Latif, Q. Li, X. M. Niu. Multimodal biometric authentication based on score level fusion of finger biometrics. *Optik*, vol. 125, no. 23, pp. 6891–6897, 2014. DOI: 10.1016/j.ijleo.2014.07.027.

[34] K. Su, G. P. Yang, B. Wu, L. Yang, D. F. Li, P. Su, Y. L. Yin. Human identification using finger vein and ECG signals. *Neurocomputing*, vol. 332, pp. 111–118, 2019. DOI: 10.1016/j.neucom.2018.12.015.

[35] R. Snelick, U. Uludag, A. Mink, M. Indovina, 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. DOI: 10.1109/TPAMI.2005.57.

[36] CASIA database, 2016, [Online], Available: http://english.ia.cas.cn/db/201611/t20161101_169922.html, March 6, 2023.

[37] Z. Zhu, G. Huang, J. Deng, et al. WebFace260M: A benchmark unveiling the power of million-scale deep face recognition. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Nashville, USA, pp. 10487–10497, 2021. DOI: 10.1109/CVPR46437.2021.01035.

Zhuorong Li received B. Sc. degree in forensic science from People's Public Security University of China, China in 2020. She is currently a master student in forensic science at Department of Criminal Investigation, People's Public Security University of China, China.

Her research interests include forensic science and biometric recognition.

E-mail: lizhuorong@stu.ppsuc.edu.cn

Yunqi Tang received the B. Sc. degree in computer science from Civil Aviation University of China, China in 2005, the M. Sc. degree in computer science from Beihang University, China in 2008, and the Ph. D. degree in computer science and technology from National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China in 2013. He is currently an associate professor with Department of Criminal Investigation, People's Public Security University of China, China.

His research interests include pattern recognition and machine learning.

E-mail: tangyunqi@ppsuc.edu.cn (Corresponding author)