Abstract

Finger vein is a new member in the biometrics family. Current finger vein recognition algorithms provide high accuracy in good quality images. However, previous researches have shown that the poor quality samples lead to a significant reduction in the accuracy of a single biometric trait system. For non-ideal finger vein images, this paper proposes a method based on multi-instance matching score fusion. First, the features of finger vein images from two or more fingers of a user are extracted by local binary pattern (LBP) algorithm and the matching scores are delivered by Hamming distance. Second, the matching scores are fused using combination-based approach and the decision is made comparing with the threshold. The experimental results show that the proposed method could significantly improve the finger vein recognition performance and the more fingers combined the higher accuracy obtained.

Keywords: Finger Vein Recognition, Multiple Instances, Score Level Fusion, Local Binary Pattern

1. Introduction

Finger vein recognition is a biometric identification technology that uses the universality and uniqueness of finger vein patterns. Comparing with traditional biometrics (e.g., fingerprint, face, etc.), this new generation of biometrics has the following advantages [1]: (1) As a part of internal physiological characteristics, finger veins are hidden underneath the skin's surface so it avoids forgery; (2) Vein patterns are stable and clearly defined, allowing the use of low-resolution cameras to capture vein images for small-size, simple data image processing; (3) The use of near-infrared light (wavelengths between 700 and 1000 nanometers) allows for non-invasive, contactless imaging that ensures both convenience and cleanliness for the user.

Now, finger vein recognition about the non-ideal low-quality has gained more and more attention on specific applications. The low quality finger vein images are caused by the followings [2]: (1) There are some kinds of noise (e.g., sensor noise, finger noise, etc.); (2) The slight movement of the finger and the light scattering through the skin layer can make the image blurred; (3) The impact of visible light produces the interference patterns; (4) The uneven illumination causes large dark or bright areas on the image that makes it difficult to detect vein patterns; (5) The strong reflection from the skin's surface and the shallow penetration of light under the skin make the contrast very small. There are several typical low quality finger vein images that are used in this paper in Figure1.

![Figure 1. Non-ideal Finger Vein Images](a) large dark (b) large bright (c) blurriness (d) noise (e) low contrast

For poor quality finger vein images, the current researches focus on image enhancement. Typical filter-based image enhancement methods are adopted in [3, 4]. J. Wang [5] enhanced the region of
interest of finger vein images using genetic programming (GP) algorithm and obtained more accurate features. In order to deal with skin scattering and optical blurring of finger vein image, E.C. Lee and K.R. Park [6] proposed a new finger vein image restoration method based on the skin scattering model using an estimated point spread function (PSF) and a constrained least square (CLS) filter, so, the finger vein patterns were greatly emphasized and the recognition rate was improved. Among other possible ways, multi-biometric can be used to mitigate the effect of non-ideal conditions. Multiple biometrics consist of different sources [7]: (a) These sources may come from multiple biometrics traits, such as iris and palmprint [8], finger vein, fingerprint and face [9], and so on. (b) Other sources may be multiple sensors, multiple instances (i.e., multiple units), multiple snapshots, or multiple matching algorithms of a single biometric trait. For example, the fusion of fingerprints from the solid-state sensor and optical sensor was studied in [10]. The multiple matching algorithms of the finger vein were researched in [11], [12] and [13]. B. Ulery et al. [14] fused the fingerprints of all ten fingers of a user with logistic regression and the accuracy was up to 99.95%. Zhao Rui et al. [15] proposed the fusion of multi-finger knuckleprints, and the experimental results showed that it was more than 6 percentages higher than the best result of using a single finger knuckleprint. Fusion of multiple instances of a single biometric trait is an inexpensive way of improving system performance since this does not entail deploying multiple sensors nor incorporating additional feature extraction and (or) matching modules [16]. However, as far as we know, there was no previous research about multiple instances fusion in finger vein recognition. Based on this, a finger vein recognition method based on multi-instance matching score fusion is proposed in this paper. The score level fusion is implemented using combination-based method.

This work has two contributions: (1) In order to improve the recognition accuracy of the non-ideal finger vein images, this paper proposes the method of matching score fusion based on multiple instances; (2) Considering the uncertainty relationship between vein images from different fingers of a same user, the triangular norms (t-norms) operations are adopted to integrate matching scores.

The rest of this paper is organized as follows: Section 2 presents the proposed method. Section 3 introduces experiments and discusses the results. Finally, Section 4 concludes this paper.

2. Finger vein recognition based on multiple instances

The framework of the proposed method is shown in Figure 2.

![Figure 2. The Schematic Diagram of Fusing multi-instance](image)

The method of matching score fusion based on multiple instances consists of two phases. The first phase is to obtain the matching scores. The second phase is to fuse the matching scores for making decision. The fusion of vein images of right index finger and right middle finger is taken as an example in Figure 2. Firstly, in feature extraction module, binary code features of the two fingers vein images of
a user are extracted by local binary pattern (LBP) algorithm, respectively. Secondly, the Hamming distance (HD) is invoked for delivering the score to match features of the user’s right index finger with features of the right index finger in the template database. In the same way, the matching score is obtained from the user’s right middle finger vein. In the next step, the two matching scores are fused at score level using the combination-based method, and then the final score is gotten. Finally, the final score is compared with the threshold to answer the question of “whether the person is the one that he/she claims to be” in decision module.

2.1. LBP algorithm

A LBP algorithm is a typical texture image analysis method. The method extracts finger vein codes in the whole region of interest without accurate detection of vein region. It implements simply and it is robust against high saturation and irregular shadings.

The LBP operator as a nonparametric 3×3 kernel for texture classification was proposed. The LBP can be defined as an ordered set of binary values determined by comparing the gray values of a center pixel with its eight neighboring pixels, as shown in Figure 3. The binary values can be expressed in decimal form as shown in Equation (1) [17].

\[
LBP(x_c, y_c) = \sum_{n=0}^{8} s(p_n - p_c) \cdot 2^n
\]

\[
s(x) = \begin{cases} 
1 & x \geq 0 \\
0 & x < 0 
\end{cases}
\]

where \(p_c\) and \(p_n\) represent the gray value of the center pixel \((x_c, y_c)\) and the gray values of the eight surrounding pixels, respectively. The function \(s(x)\) is defined as Equation (2). If the size of preprocessing finger vein image is 96×64, feature code of 46624 (94×62×8) bits is extracted.

![](image)

**Figure 3.** The LBP Operator

The feature code of a finger vein image of a user A is denoted by \(\text{code}_A\), and the corresponding template B is denoted by \(\text{code}_B\). The HD represented by Equation (3) is generally adapted to measure dissimilarities between two binary patterns [18].

\[
HD(A, B) = \frac{(\text{code}_A \oplus \text{code}_B)}{\text{CodeLength}},
\]

where \(\oplus\) is a Boolean exclusive-OR operator between corresponding pairs of bits. Therefore, the HD can be calculated by dividing the number of binary codes (CodeLength = 46624 in our experiment). The range of the HD is [0, 1], and the smaller value of the HD, the more similarity of two images.

2.2. Matching score fusion

In a multi-biometric system, fusion can be performed at four different levels, i.e., sensor level, feature level, matching score level, and decision level, corresponding to sensor module, feature extraction module, matching module, and decision-making module[7]. Sensor level fusion refers that the raw data from the sensors are combined [10]. Fusion at feature level refers to combining different feature vectors that are obtained by either using multiple biometrics or employing multiple feature extraction algorithms on the same biometric trait. Although fusion of feature vectors uses rich original information, features from different representation algorithms or multimodal biometrics may not be compatible. Moreover, concatenating two or more feature vectors may result in the feature vector with
very large dimensionality, which leads to the “curse of dimensionality” [16]. Score level fusion is to combine the matching scores of the individual matchers. Fusion at the score level is fairly popular and simple because the scores have easy availability and contain ample information about the input pattern next to the feature vectors [7]. Fusion at decision level refers to fuse matching results from different matchers to make final decision, and it is considered to be the rigid due to the lack of available information. Based on the analysis, score level fusion is a reasonable choice in multiple instances method.

Matching score fusion focuses on different integrated approaches. Matching score fusion techniques can be divided into three categories as follows [7, 19]: (a) density-based matching score fusion, (b) classifier-based matching score fusion, and (c) combination-based or transformation-based matching score fusion. The density approach is based on the likelihood ratio test and it requires explicit estimation of matching score density function by the parametric or non-parametric methods. Gaussian mixture model (GMM) has been successfully used to estimate arbitrary densities [19]. In the classifier-based approach, a feature vector is constructed using the matching score by the individual matcher, and this feature vector is then classified into one of two classes, i.e., genuine user or impostor user. The classifier used for this purpose is capable of learning the decision boundary irrespective of how the feature vector is generated. Generally, neural network (NN), decision tree (DT), k nearest neighborhood (k-NN) and support vector machine (SVM) have been used to arrive at a decision. In the combination or transformation approach, the matching scores are combined to generate a single scalar score which is then used to make the final decision based on the threshold. Since the scores by different matching algorithms can be heterogeneous (e.g., distance or similarity metric, different numerical ranges, etc.), normalization is required to transform these scores into a common domain before combination. Product rule, sum rule, max rule and min rule are often used to determine the final score.

In this paper, the combination-based method is used to fuse the matching scores derived from multiple fingers vein images of a user. Ross and Jain have shown [20] that the sum rule is sufficient to obtain a significant improvement in the matching performance of a multimodal biometric system. For the consideration of the uncertainty and imperfection exist in the different modalities, M. Hanmandlu et al. [21] confirm t-norms are more effective to fuse the matching scores for multimodal biometrics. In view of these, the sum rule and the t-norms method are adopted to fuse the matching scores in our experiments.

Assuming use \( n \) fingers of a user for finger vein recognition, the \( s_i \) represents the matching score of the first \( i \) finger, the \( s_f \) represents the final score fusing \( n \)-fingers scores, and then the weighted sum rule can be expressed in Equation (4):

\[
s_f = w_1s_1 + w_2s_2 + \cdots + w_ns_n.
\]  

(4)

The notation \( w_i \) stands for the weight which is assigned to the \( i \)-finger, for \( i = 1, 2, \cdots, n \). The weights are varied over the range \( [0, 1] \), such that the constraint \( w_1 + w_2 + \cdots + w_n = 1 \) is satisfied.

A t-norm is a commutative, associative, monotonous operation \( T: [0, 1] \times [0, 1] \rightarrow [0, 1] \). It has been widely used in fuzzy logic systems, fuzzy control systems and artificial intelligence, etc. In this paper, the three t-norms which we use can be expressed by Equations (5), (6) and (7). Equations (5) and (6) are non-parameter [21]. Equation (7) has the parameter \( p(p \in [0, \infty]) \); one can get different values by adjusting the parameter.

Einstein product: 

\[
T(s_1, s_2) = \frac{s_1s_2}{2 - (s_1 + s_2 - s_1s_2)},
\]  

(5)

Hamacher: 

\[
T(s_1, s_2) = \frac{s_1s_2}{s_1 + s_2 - s_1s_2},
\]  

(6)
In t-norms, any two scores (e.g., $s_i$ and $s_{i1}$) are first combined to yield $T(s_{i1}, s_i)$, which is in turn combined further with $s_{i2}$ to yield $T(s_{i1}, T(s_{i2}, s_i))$. The order of combination is immaterial due to the associative and commutative properties of these norms. The final score $s_f$ is obtained by Equation (8).

$$s_f = T(s_{i1}, T(s_{i2}, s_i)).$$

In the next step, the fused score $s_f$ will be compared to a pre-specified threshold $t$. We declare a user to be genuine if $s_f \leq t$, otherwise, we declare him (or her) an impostor.

### 3. Experimental results and analysis

#### 3.1. Database and experimental design

The dataset which is used in our experiment is a subset of the SDUMLA-HMT finger vein database. There are 80 subjects were asked to offer vein images of the right index finger, the right middle finger, and the right ring finger, and each finger provides 6 images. Therefore, 18 images from 3 fingers were collected from each subject. In total, the dataset contains 1440 ($80 \times 3 \times 6$) images from 240 fingers. The dataset include lots of non-ideal images (as Figure 1), and there are obvious varieties of translation, shift, and rotation between the samples from the same finger. The original image captured by the device is a 24-bit color image with a size of 320×240. In order to reduce the noise and computational complexity, the preprocessing is done including image gray processing (an 8-bit gray image), region of interest cropping, gray normalization, and size normalization with a size of 96×64.

In this paper, three experiments are designed to evaluate the proposed method. The first experiment evaluates matching performance by the LBP algorithm based on a single finger, i.e., the right ring finger (rr), the right middle finger (rm), and the right index finger (ri). The second experiment evaluates the recognition performance of the multi-instance matching score fusion using the sum rule, i.e., fusion of the ring finger and middle finger (denoted by rr+rm), fusion of the ring finger and index finger (denoted by rr+ri), and fusion of the three fingers (denoted by rr+rm+ri). The third experiment evaluates the recognition performance of the multi-instance matching score fusion by the t-norms method. The experiments are implemented in MATLAB 2011a.

**Experiment 1.** In the first experiment, take the right index finger as an example, there are 80 classes and 480 different images. To obtain statistical results, the different samples of the same index finger for each subject are matched, and the matching between them is counted as a genuine matching, so 1200($80 \times C_2^6$) genuine scores have been generated all together. Matching a index finger sample of one subject with a index finger sample of another subject is counted as an imposter matching, and then in total 3160($C_8^{60}$) imposter scores have been gotten. Similarly, this work is done on the right middle finger and ring finger of a subject. So there are three group matching scores including 1200 genuine scores and 3160 imposter scores per group. Figure 4(a)-(c) show the score distribution under a single finger, and there are obvious overlaps between the distributions of genuine and imposter scores.
To evaluate the performances of the three experiments, the false acceptance rate (FAR) and false rejection rate (FRR) are reported. The FAR is the fraction of the number of falsely accepted imposter scores divided by the total number of imposter scores. The FRR is the fraction of the number of falsely rejected genuine scores divided by the total number of genuine scores. The FAR and FRR will change correspondingly as the threshold changes, that is, if the FAR increases, the FRR would tend to reduce. According to the FAR and the FRR, the receiver operating characteristics (ROC) which depicts the overall performance of a biometric system is plotted. The ROC curves shown in Figure 5 directly reflect the difference of the performances by single finger and fused multiple fingers (in experiment 2 and experiment 3). The equal error rate (EER) of a system is the value that the FAR and FRR value are to be equal. Table 1 shows the EERs in the first experiment.

<table>
<thead>
<tr>
<th>Finger</th>
<th>rr</th>
<th>rm</th>
<th>ri</th>
</tr>
</thead>
<tbody>
<tr>
<td>EER (%)</td>
<td>4.71</td>
<td>3.82</td>
<td>5.41</td>
</tr>
</tbody>
</table>

Experiment 2. The goal of this experiment is to investigate the accuracy based on fusing vein images from two or more fingers of a user. The combination approach is the sum rule. Suppose that we fuse the information of vein images from the three fingers of a user, each template in the enrolment
database will be composed by features of vein images from the three fingers. When matching, a user’s vein features of the three fingers will be matched to features of the corresponding finger of a template, respectively. Then three HD values will be gotten, i.e., matching scores. The three matching scores will be fused according to the sum rule to obtain the final score, by which the client’s identity can be verified. Similarly, this work is done on fusion of two different fingers.

In the sum rule, considering that there is no enough evidence to determine which finger is more helpful to the identity, the equal weight is adopted. The recognition results are presented in Table 2.

<table>
<thead>
<tr>
<th>Table 2. EERs Obtained in Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fingers in fusion</td>
</tr>
<tr>
<td>EER (%)</td>
</tr>
</tbody>
</table>

Experiment 3. This experiment adopts the t-norms method to fuse the matching scores. The processing of the experiment is consistent with experiment 2. The fusion scores are gotten by Equation (5)-(8), where the parameter of Equation (7) is obtained by experience. Figure 4(d) shows the score distribution where scores of the three fingers are fused using Frank t-norm with p=0.2, and there is lower overlap between genuine and imposter score distributions. The recognition results are presented in Table 3. Compared with Table 1, the recognition performance by integrating two or three fingers vein images is significantly improved.

<table>
<thead>
<tr>
<th>Table 3. EERs (%) Obtained in Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fingers in fusion</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>rr+rm</td>
</tr>
<tr>
<td>rr+ri</td>
</tr>
<tr>
<td>rm+ri</td>
</tr>
<tr>
<td>rr+rm+ri</td>
</tr>
</tbody>
</table>

Figure 5. ROC curves

3.2. Experiment results analysis

According to the three experiments above, key results and analysis are summarized as follows:

- For non-ideal finger vein image, the proposed method based on multi-instance can effectively improve the recognition accuracy.
Compared with two fingers, three fingers fusion has better performance.

Two fusion approaches at score level are effective in this paper, but the t-norms method is better than the sum rule on the whole. The assumption of statistical independence of a multimodal biometric system is used in the sum rule. However, the t-norms method is based on the uncertainty relationship between different modalities. At present, for the relationships between different fingers vein images from the same individual, there is no enough evidence to prove they are statistically independent. Consequently, it is more reasonable to treat the relationship between different fingers as uncertainty by experiment.

In our dataset, besides the poor quality images, the samples of the same finger have obvious shift and rotation. The LBP algorithm is sensitive to the shift and rotation of images, which lead to genuine matching score much higher.

4. Conclusion

Finger vein recognition technology for low-quality image has got more and more attention. Apart from concerning to enhance finger vein image, integration of multiple fingers vein images of the same individual provides an optional approach. Fusion based on multi-instance is to expect using diverse information carried by the different fingers to improve recognition performance. The proposed method is simple and has low computational complexity. Firstly, the feature extraction by the LBP and matching by the HD are simple; secondly, integration of multiple instances costs less; thirdly, the combination-based matching score fusion is simple and does not require any learning and training. The experiments prove that the proposed method can significantly improve the recognition accuracy. Although this method has a positive impact on the verification of low-quality finger vein dataset, we do not consider the factors of image quality measure for integration. Further research will focus on other popular matching algorithms and how to incorporate image quality information into the integration approach.

5. Acknowledgement

This work is supported by National Natural Science Foundations of China under Grants No. 61173069 and 61070097, and Natural Science Foundation of Anhui Provincial Universities under Grant No. KJ2012A214. The authors would like to thank Ying Li and Yanbin Ning for their helpful comments and constructive advices on structuring the paper. In addition, the authors would particularly like to thank the reviewers for their helpful suggestions.

6. References