Contactless Palm Vein Identification using Multiple Representations

Yingbo Zhou, Ajay Kumar
Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
Email: yingbo.zhou@ieee.org, ajaykr@ieee.org

Abstract—This paper investigates some promising approaches for the automated personal identification using contactless palmvein imaging. We firstly present two new palmvein representations, using Hessian phase information from the enhanced vascular patterns in the normalized images and secondly from the orientation encoding of palmvein line-like patterns using localized Radon transform. The comparison and combination of these two palmvein feature representations, along with others in the palmvein literature, is presented for the contactless palmvein identification. We also evaluate the performance from various palmvein representations when the numbers of training samples are varied from minimum. Our experimental results suggest that the proposed representation using localized Radon transform achieves better or similar performance than other alternatives while offering significant computational advantage for online applications. The proposed approach is rigorously evaluated on the CASIA database (100 subjects) and achieves the best equal error rate of 0.28%. Finally, we propose a score level combination strategy to combine the multiple palmvein representations. We achieve consistent improvement in the performance, both from the authentication and recognition experiments, which illustrates the robustness of the proposed schemes.

I. INTRODUCTION

BIOMETRICS refers to the use of physiological and/or behavioral characteristics of humans for the personal identification and is being increasingly employed for forensic, civilian and variety of online business applications. Several biometrics technologies are prone to spoof attacks [1]-[3] in which covertly acquired biometric characteristics, e.g. face, iris, fingerprint, can be employed to develop spoof samples and compromise the deployed biometrics system. The main drawback of such biometrics systems is that they use the extrinsic features of humans which are easier to observe, but also, easier to acquire covertly from a distance and generate spoof samples. However, in contrast, the intrinsic or hidden physiological pattern, are naturally, harder to observe and therefore offer high degree of privacy and security for biometric identification. In this context, the handvein/palmvein identification has emerged as a promising component of biometrics study [20]. The complex/rich vascular structures residing inside the palm make it very difficult to forge, more convenient to present, can automatically and simultaneously ensures liveness in the presented biometric sample.

The palmvein imaging requires infrared illumination and is employed in several commercially available products [17]. The infrared illumination is one component of multispectral illumination and also employed to generate multispectral palm images as detailed in [5]-[6]. However, the multispectral images add significant computational requirements while only marginally improving the performance. Therefore Wang et al. [4] have investigated the bispectral imaging, i.e. palmprint and palmvein images, using ‘Laplacianpalm’ representation by using locality preserving projection (LPP). Hao et al. [5] combined multispectral palm at image level but using the information from feature level. Zhang et al. [6] have recently investigated the usage of bank of Gabor filters to extract the features from multispectral palm images and achieved promising results by using score fusion. The features from palm vein and palmprint creases textures have also been investigated in [7]. In [8] matched filter has been employed for palm vein feature extraction while a minutiae based matching approach is investigated in [9]. The summary of prior work suggests that there has not been any comparative study to ascertain which palmvein representation can be the best and for its suitability for nature of imaging (contactless [5], [17] versus constrained [6]-[9]).

A. Our Work

The key contributions from this paper can be summarized as follows: Firstly, we investigate on the problem of contactless palmvein identification and investigate a new approach using multiple palmvein representations for higher performance. Secondly, this paper investigates two new approaches, which extract two different kinds of palmvein features and illustrate promising performance. The localized Radon transform based approach achieves best performance and also offers computationally simpler alternative to existing palmvein identification approaches presented in the literature. Thirdly, there have been several efforts on palmvein authentication but none have comparatively evaluated them to ascertain the best and therefore this paper presents such comparison. The minimum number of training samples is desirable for biometrics systems and therefore we present rigorous performance evaluation with the variation in the number of training samples. Finally, the palmvein literature has so far examined on the authentication but not on recognition performance and therefore this paper also presents the palmvein recognition results.

II. PRE-PROCESSING

A. Region of Interest Segmentation

The acquired palmvein images are firstly normalized to minimize the rotational, translational and scale changes. Therefore, to make the identification process effective and efficient, it is necessary to construct a coordinate system that is invariant/robust to (or nearly) those variations. It is intuitive to associate the coordinate system with the palm itself, and thus the web between index finger and middle finger together with the web between ring finger and little finger were utilized as the reference points/line to build up the coordinate system (figure 1).

In order to account for the potential scale variations in the acquired contactless palmvein images, the location as well as
the size of region of interest (ROI) are selected based on the distance between the two webs ($L_W$) and is illustrated in the following equation:

$$\begin{align*}
L_D &= \alpha L_W \\
L_{ROI} &= \beta L_W
\end{align*}$$

(1)

where $L_{ROI}$ denotes the side length of ROI, $L_D$ denotes the distance between the ROI and the reference line, and $L_W$ represents the distance between the two webs, $\alpha$ and $\beta$ are the factors that control respectively the location and size of the ROI. The present approach is very similar to the method used in [5], however, our computations are simpler since no additional sampling is employed. After this segmentation, ROI is resized to a fixed size (128 × 128 pixels in our experiments) to facilitate the identification processes.

Figure 1: Robust segmentation of palmvein ROI from contactless images.

B. Image Enhancement

Since the palmvein images employed in our work were acquired under near infrared illumination (NIR), the image appears dark. Therefore to achieve better visualization of vascular and surface details, contrast enhancement is required. We firstly estimate the background intensity profiles by dividing the image into slightly overlapped 32 × 32 blocks (3 pixel overlapping between two blocks to address the ‘blocky effect’), and the average values of each block are calculated. Subsequently, the estimated background intensity profile is resized to the same size as the original image using bicubic interpolation and then is subtracted from the original ROI image. Finally, the histogram equalization was employed to obtain the enhanced image. The comparison between figure 3(a) and 3(b) suggests that the enhancement scheme has been quite successful in enhancing the local surface details.

Figure 2: Block diagram for the personal identification using multiple representations of palmvein images.

III. Feature Extraction and Representation

A. Multiscale Vessel Enhancement

The local palmvein image characteristics can be observed from the enhanced images using Taylor series expansion in the neighborhood of a point. This approach [10] allows us to observe the curvature of the enhanced palmvein images using second order (Hessian) representation of local image structures. Recall that the eigenvector of a matrix corresponds to the basis/principal directions of the matrix, and thus the corresponding eigenvalue of the Hessian matrix (second order derivative) will reflect the principal orientation of the curvature in the local image. Let $\lambda_m$ denote the eigenvalue with the m-th smallest magnitude, for an ideal vessel-like structure in a 2D image the eigenvalues should have two local characteristics as analyzed in the following. Firstly, the summation of the norm of the eigenvalues will be small at the location where no structure information is shown since the contrast variation is low, and it will become larger when the structure occupies higher contrast since at least one of the eigenvalue will be large. Secondly, the ratio between $|\lambda_2|$ and $|\lambda_1|$ will be large when the blob-like structure appears in the local area, and will be very close to zero when the structure shown is line-like. Mathematically the two measures is represented as following,

$$L_1 = \sqrt{\sum_{i=1}^{N} h^2}$$

(2)

$$L_l = \frac{|\lambda_1|}{|\lambda_2|}$$

(3)

where $L_1$, and $L_l$ are the measures of the local structure and local line characteristics with $N$ denoting the dimension of the image. Therefore, the local ‘vessleness’ of position $p$ in scale $s$ is assigned using the following equations:

$$V_p(s) = \begin{cases} 
0, & \text{if } \lambda_2 < 0 \\
\exp \left(-\frac{1}{2b^2} \left(1 - \exp \left(-\frac{s_1}{2b^2} \right) \right) \right), & \text{otherwise}
\end{cases}$$

(4)

$$I_g(p) = \max_{V_{p,s}} V_p(s,p)$$

(5)

where $a$ is fixed to 0.5 and $b$ equals to half of the maximum Hessian norm, $I_g$ represents for the enhanced image and $S$ is the set which contains all the defined scales.

The enhanced image (see figure 3(c)) is then binarized to form the template (Hessian phase, see figure 3 (d)), and a localized binarization scheme was investigated. This scheme essentially creates two local windows, i.e. inner and outer window, and the local threshold value is selected based on the following equation [11]:

$$T_{bin} = (1 - \omega_1)\mu + \omega_2 \left(\frac{\text{inner}}{\text{outer}}\right) (\mu - M) + \omega_3 M$$

(6)

$$\omega_2, \omega_3 = h_{1,2} \left(\frac{\text{inner}}{\text{outer}}\right)^{\tau}$$

(7)

where $\mu$, $\sigma_{\text{inner}}$, $M$ is the mean, standard deviation and the minimum intensity value of the inner window, $\sigma_{\text{outer}}$ represents for the standard deviation of the outer window, $\omega_i$, $h_i$, $h_2$, and $\tau$ are constants and fixed to value 0.1, 0.25, 0.04 and 3 respectively in this experiment.

B. The Localized Radon Transform

Radon transform is an effective tool in finding line structures in the image, the localized Radon transform (LRT)
demonstrated its effectiveness in extracting the line-like features in paper [12] and our previous work [13]. The vein/vessel patterns observed in the palm vein images also illustrate the curved and line-like features. Therefore this approach was also employed to extract/encode those features. The idea of this approach is that the curved/straight lines can be estimated by small piecewise joint line segments and it integrates the intensity value in a local region in all defined orientations; but instead of integrating all the pixel values inside the local region, only the pixels that fall into the confined line width area is integrated, and the orientation that gives the minimum (or maximum depending on the intensity of local features) integration value is selected as the dominant direction. This dominant direction index is encoded as a characteristic feature (figure 3(e)).

C. Ordinal Representation

The ordinal representation using orthogonal line ordinal features has illustrated promising performance on palmprint [14], multi-spectral palm [5] and for iris identification [15]. The absolute intensity information of an object may vary depending on different imaging conditions while the relative order of the intensity information is comparatively stable. Based on this observation, this approach measures the ordinal relationship among different areas and those relationships are regarded as ordinal features. In order to measure the line-like features two different 2-D Gaussian filters with an angle difference of π/2 is used to extract the line features, and is expressed as following [14]:

\[
OL(\theta) = g(x,y,\theta) - g(x,y,\theta + \pi/2)
\]

(8)

\[
g(x,y,\theta) = \exp \left\{ -\frac{\left(\cos \theta + \sin \theta \right)^2}{\delta_x^2} - \frac{\left(x \sin \theta + y \cos \theta\right)^2}{\delta_y^2} \right\}
\]

(9)

The ratio between \(\delta_x\) and \(\delta_y\) should be large so that the shape of the Gaussian lope is similar to a line, and in this experiment \(\delta_x\) and \(\delta_y\) was fixed to 10 and 3, and \(OL(\theta)\) \(OL(\pi/6)\) and \(OL(\pi/3)\) was employed to extract the ordinal code (see figure 3 (f)).

D. The Laplacianpalm

Principal component analysis (PCA) is one of the highly popular appearance based technique in pattern recognition for feature extraction and dimension reduction. The features extracted by PCA, however, are generally global features which may be quite limited in representing the discriminating/local characteristics of different classes. The Laplacianpalm utilizes the locality preserving projection (LPP) to map the palm images into subspaces, and the LPP preserves the local neighborhood structure of the data, i.e. maintains some local characteristics. This representation in palm vein identification shows promising results in [4], and thus was also investigated in our experiments.

IV. MATCHING SCORE GENERATION

In our work, the cosine similarity was employed to generate the matching score for the ‘Laplacianpalm’ representation. The matching scores from the other three palmvein representations, i.e., Hessian phase using multiscale vessel enhancement, LRT and ordinal code/representation, is generated as follows:

\[
M(F_n, F_q) = \frac{\text{hamdist}(F_n, F_q)}{\|F_n\| \|F_q\|}
\]

(10)

where \(F_n\) and \(F_q\) represent the registered and the test/probe feature template, \(\text{hamdist}\) is the Hamming distance between the two feature templates (i.e. number of bits that are different), and \(\|\ |\) is the cardinality operator. In addition, to take into account the potential translational variations, the test template is shifted from \([-c, c]\) in both horizontal and vertical direction while matching. In our experiments the \(c\) is empirically selected as one tenth of the template size.

V. SCORE COMBINATION

In this work, we investigated four different representations for the palmvein surface features. These representations can be typically categorized into three classes: Phase-based representation, Appearance-based representation and Orientation-based representation. The Hessian phase and the ordinal code represent a kind of phase features (quantized values according to some threshold or relationship); ‘Laplacianpalm’ belongs to the appearance/subspace-based features; LRT encodes the orientation features. It can be argues that the features from the subspace/appearance-based approach are expected to be highly independent from the other two, while there is no such clear separation between the phase- and the orientation-based representations. Since the phase-based features (see figure 3(d) and 3(f)) also occupies orientation characteristics, which means that the phase representation also occupies some down-graded orientation information. According to the above analysis, we can develop judicious combination of matching scores from the following heuristics [16].

1. Scores that have higher degree of correlation are preferred to be combined by average, since it is expected to reduce both the error and standard deviation of the scores.
2. Scores that are highly independent are combined using product as it is an approximation of the posterior probability.

The above two heuristics motivates the strategy for combining four groups of matching scores as follows:

\[ S_T(u) = \alpha (S_H(u) + S_D(u)) + (1 - \alpha)S_R(u) \quad (11) \]

where \( u \in U \), and \( U \) denotes the set that contains all the subjects/users in the employed database, \( S_T \) represent the final or consolidated matching score while \( S_H, S_D, S_L \), and \( S_R \) represents the normalized matching scores from the Hessian phase, ordinal code, ‘Laplacianpalm’ and LRT respectively and \( \alpha \) is a fixed constant. All the scores are normalized to the same range and using similarity measure. The matching score from the Hessian phase and ordinal code are firstly combined using their average/sum following the first heuristics (above), since they represent common category of features and are expected to be highly correlated [19]. The score from ‘Laplacianpalm’ are then combined with the scores from last step using product rule following the second heuristics above, as the two scores represents different categories of features (i.e. phase and subspace representations) hence are expected to be highly independent. Finally, the score from LRT is combined with the score from last step using sum rule.

VI. EXPERIMENTS AND RESULTS

A. Database

All the experiments reported in this paper for the palmvein identification CASIA Multi-Spectral Palmprint Image Database V1.0 (CASIA database) [18]. This CASIA database has been acquired using a contactless imaging device and have images from 100 users. Six images were acquired from each user and these images were acquired in two different data acquisition sessions (three images in each session) with a minimum interval of one month. Since our work is focused on palmvein identification and the vascular details are typically observed in the NIR illumination, only the images that were acquired under 850nm wavelength illumination from CASIA database were utilized in the following experiments.

B. Experiments and Results

We performed rigorous experiments to ascertain the performance for the authentication and recognition using the approaches detailed in section III and V. The comparative average performance in terms of equal error rate (EER) is illustrated in table 1-2 and the receiver operating characteristics (ROC) is illustrated in figure 4-5, while the ROC of the left palm from CASIA database is not reproduced simply to the conserve space. In this set of results, all the training samples were from the data acquired in the first session while the rest images were used as validation data. It is quite known in the biometric literature that the minimum numbers of training samples are always desirable. Therefore it is prudent to ascertain the actual performance from various palmvein representations and their combination, when the numbers of training samples are varied/minimum. Therefore in this study we also evaluate the performance of different approaches when the numbers of training samples are varied.

The ROC (see figure 4) suggests significant improvement from all the four approaches with the increase in the number of training samples for the CASIA database and this may be due to several reasons. This database was acquired under a less constrained environment which is expected to introduce higher intra-class variations of the acquired image/samples, and thus it gains benefits while trained with additional samples, since this increases the chance for the classifier to get the essential/variation information from individuals, in the meanwhile also increasing the probability for the matcher to get best possible matching scores.

We now combine all the scores using equation (11) and the achieved performance from the authentication experiments is illustrated in figure 5. Significant performance improvement is achieved from this combination (see figure 5), and we observed that fixing \( \alpha \) to 0.5 achieves a relatively good performance, but the value 0.6 (estimated from training samples) achieves the best performance in terms of EER.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Left Hand</th>
<th>Right Hand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hessian Phase</td>
<td>2.24%</td>
<td>2.00%</td>
</tr>
<tr>
<td>LRT</td>
<td>1.03%</td>
<td>0.78%</td>
</tr>
<tr>
<td>Ordinal Code</td>
<td>2.00%</td>
<td>1.00%</td>
</tr>
<tr>
<td>Laplacianpalm</td>
<td>5.00%</td>
<td>3.63%</td>
</tr>
</tbody>
</table>

It is prudent to examine on the advantages of selecting the value of \( \alpha \) as 0.5 with those when the value of \( \alpha \) as is
computed empirically (as 0.6). Firstly, the selection of $\alpha$ as 0.5 in equation (11) is equivalent to the combining the matching score from LRT with the combined score (i.e. fused score from Hessian phase, ordinal code and Laplacianpalm methods) using average rule, and therefore no training is necessary for our score combination scheme. This is a critical advantage, especially in the one-training cases, where it is very difficult to get the representative training scores, in particular, for the database that has samples with high intra-class variations. Secondly, the trained approaches may not necessarily outperform the fixed approaches [16] since the prior one heavily depend on the quality of training data, i.e. how representative the training dataset is for the actual test data. If the training data is biased, the classifier combination may be trained inappropriately, and finding a set of ‘good’ training samples is itself a complex task. Thirdly, the resultant trained score combination may not necessarily optimize the performance. For example, weighted linear combination is a commonly used scheme in score combination, and the weights to combine different classifier is selected based on the training data and certain criteria such as the weight that obtain the lowest EER or the lower false rejection rate at a certain false acceptance rate. The resultant combination may perform better in that operating point/region; however, the whole/average performance may not improve. Last but not least, our fixed approach saves the computation and makes the combination process simpler, especially when the dataset is large or only limited computational power is made available.

<table>
<thead>
<tr>
<th>Scores Combined</th>
<th>Left hand</th>
<th>Right hand</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_H$ and $S_0$</td>
<td>1.67%</td>
<td>0.67%</td>
</tr>
<tr>
<td>$S_H$, $S_0$ and $S_L$</td>
<td>0.49%</td>
<td>0.63%</td>
</tr>
<tr>
<td>All the scores</td>
<td>0.38%</td>
<td>0.28%</td>
</tr>
</tbody>
</table>

The performance from the score combination also improves as the training sample size is increased. This can be attributed to the improvement in the selection of best matching score as the size of training samples is increased. The proposed fusion strategy achieves significant performance improvement, especially in lower FAR region (see figure 5), and the plausible reason for such improvement lies in the large intra-class variations in the employed database largely resulting from the contactless imaging, e.g. scaling, orientation variations, motion blur (defocusing), projective distortions and translational changes. Therefore the information/features from multiple sources/representations is more likely to form complementary relationships and thus benefit for the combined performance. We also performed palmvein recognition experiments using different number of training samples and the results are summarized in table 3-4 and figure 6. The experimental results illustrate noticeable performance improvement using the combination strategy.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Left hand</th>
<th>Right hand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hessian Phase</td>
<td>91.00%</td>
<td>95.00%</td>
</tr>
<tr>
<td>LRT</td>
<td>91.33%</td>
<td>96.67%</td>
</tr>
<tr>
<td>Ordinal Code</td>
<td>91.00%</td>
<td>97.67%</td>
</tr>
<tr>
<td>Laplacianpalm</td>
<td>75.00%</td>
<td>77.67%</td>
</tr>
<tr>
<td>Combined Score</td>
<td>96.67%</td>
<td>98.00%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Approach</th>
<th>Left hand</th>
<th>Right hand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hessian Phase</td>
<td>97.33%</td>
<td>97.67%</td>
</tr>
<tr>
<td>LRT</td>
<td>98.33%</td>
<td>99.00%</td>
</tr>
<tr>
<td>Ordinal Code</td>
<td>98.67%</td>
<td>99.00%</td>
</tr>
<tr>
<td>Laplacianpalm</td>
<td>87.00%</td>
<td>91.67%</td>
</tr>
<tr>
<td>Combined Score</td>
<td>99.67%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

† The absolute mean rotational variation automatically estimated during the normalization of the entire left hand database is 21.0°

Table 2. The EER performance from the score combination using all/maximum the training samples in first session.

Table 3. Rank-1 recognition rate from different approaches using minimum (one) number of training samples.

Table 4. Rank-1 recognition rate from different approaches using maximum (three) number of training samples.

Figure 5. ROC from score combination of different training samples, (a) and (b) illustrates the 1-training and 3-training performance from the left palm respectively, (c) and (d) shows the 1-training and 3-training performance from the right palmvein respectively.

Figure 6. Cumulative match characteristics of left palm vein images from CASIA database using (a) one training, and (b) three training.
The performance from the recognition experiments follows the same trend as those from the verification experiments. It is notable that the performance of LRT and ordinal code are relatively close in both verification and identification scenario, and the similar observation holds good with the variations in the size of training samples. Regarding the computational complexity, however, LRT approach is more efficient, and is illustrated as following. Assume a $K \times K$ OL filter (see section III C) is applied on an $M \times M$ image to extract the ordinal code, and the computation required is $K^2 \times M^2$ multiplications and $(K^2 - 1) \times M^2$ additions; on the other hand, applying LRT to an $M \times M$ image to extract O orientation features with line width $w$ and length $K$ just requires $w \times K \times (M/w)^2 \times O$ additions. For simplicity let us assume (reasonable) that the computational complexity of multiplication and addition are equivalent in the sense of complexity, although the exact computational complexity of multiplication is greater than addition. According to this assumption, the number of operations required from LRT based feature extraction is still about $2Kw/O$ times less as compared with those from the spatial convolution based OL filter. Please note that in most of the case the value of $K$ is much larger than the value of $O$, which means LRT is more computationally effective, in general.

Partial palmvein images may be presented in the contactless systems. Such images can also be effectively utilized by introducing mask that can indicate valid palmvein region while matching. In more extreme cases, it is possible that only a small portion of palmvein images is acquired. One of the possible solutions for this circumstance is to divide the template into smaller pieces/blocks and search the best match among the blocks.

VII. CONCLUSIONS

This paper has presented a new approach for reliable personal identification using multiple palmvein representations. The two new palmvein feature representations, i.e. Hessian phase and LRT, were investigated and comparatively evaluated with the prior approaches (ordinal code and ‘Laplacianpalm’). The LRT based palmvein representation proved to be computationally simpler and better or similar in performance than best approach. Our experimental results achieved equal error rate of 0.38% and rank-one recognition accuracy of 99.67% from the left hand palmvein images when the acquired images from the second session where employed as the independent test data while those from the first session were employed as training set. The multiple palmvein representation can achieve noticeable performance improvement for the contactless imaging in which intra-class image variations are expected to be significant. This paper also presented performance evaluation on palmvein images for varying training samples and the achieved results suggest the robustness of the proposed scheme, both for the verification and recognition experiments.

VIII. ACKNOWLEDGEMENT

This work is supported by the internal competitive research grant from The Hong Kong Polytechnic University (2009-2010), grant no. PJ70. Authors thankfully acknowledge the Chinese Academy of Sciences for providing CASIA-MS-PalmprintV1 database used in this work.

IX. REFERENCES