Fusion of Palmprint and Palm Vein Images for Person Recognition Based on “Laplacianpalm” Feature

Jian-Gang Wang, Wei-Yun Yau, Andy Suwandy
Institute for Infocomm Research
21 Heng Mui Keng Terrace, Singapore 119613
{jgwang, wyyau, suwandy}@i2r.a-star.edu.sg

Eric Sung
Nanyang Technological University
Singapore 639798
eeric sung@ntu.edu.sg

Abstract

Unimodal analysis of palmprint and palm vein has been investigated for person recognition. However, they are not robust to noise and spoof attacks. In this paper, we present a multimodal personal identification system using palmprint and palm vein images with fusion applied at the image level. The palmprint and palm vein images are fused by a novel integrated line-preserving and contrast-enhancing fusion method. Based on our proposed fusion rule, the modified multiscale edges of palmprint and palm vein images are combined as well as the image contrast and the interaction points (IPs) of the palmprints and vein lines are enhanced. The IPs are novel features obtained in our fused images. A novel palm representation, called “Laplacianpalm” feature, is extracted from the fused images by Locality Preserving Projections (LPP). We compare the recognition performance using the unimodal and the proposed fused images. We also compared the proposed “Laplacianpalm” approach with the Fisherpalm and Eigenpalm on a large dataset. Experimental results show that the proposed multimodal approach provides a better representation and achieves lower error rates in palm recognition.

1. Introduction

Currently, much of the research efforts has been devoted to unimodal analysis of palmprint [18, 28, 15, 14, 27] or palm vein [9, 5] identification. However, most of the unimodal methods do not perform sufficiently well for identification. They have to contend with a variety of problems such as noisy data, intra-class variations, restricted degrees of freedom, non-universality, spoof attacks, and unacceptable error rates. A robust identification system may require fusion of several modalities. Ambiguities in one modality such as caused by illumination problem of palmprints may be compensated by another modality like vein features. Multi-modal identification system hence promises to perform better than any one of its individual components. Another advantage of fusion is that it gives better protection against spoof attacks because both palmprint and palm vein are required simultaneously by the system.

The fusion can be done at the feature level, matching score level or decision level with different fusion models [16]. Most of the existing studies on multi-modal focused on combining the outputs of multiple classifiers at the decision level. In a recent work, the eigenpalm and eigenfinger are fused at the matching score level for user identification [4]. The eigenpalm and eigenfinger features are extracted based on the K-L transform; the system was tested on a database of 237 people (with 1,820 hand images). The experimental results showed the effectiveness of the system in terms of the recognition rate (100 percent), the equal error rate (EER = 0.58 percent), and the total error rate (TER = 0.72 percent).

In this paper, we proposed a multimodal personal identification system where palmprint and palm vein images are combined in a single image at the image level. The motivation is that there are unique features in a palmprint and palm vein image that can be used for personal identification. The principal lines, wrinkles, ridges, minutiate points, singular points, and texture are regarded as useful features for palmprint and palm vein representation. Lines and texture features can still be clearly observed in low-resolution palmprint and palm vein images. However, lines are more appealing than texture for the human vision. In this paper, an integrated line-preserving and contrast-enhancing image fusion is proposed to fuse palmprint and palm vein images. Multiscale edge-based image fusion [1] and Multiscale edge-based image enhancement [31] have been investigated separately. They are more interested in highlighting the features of interest and for reconstructing the image for human visual interpretation. Also, using wavelet transforms, they could delete a lot of unwanted codes to achieve data compression. Different from the existing wavelet based image fusion [1, 32], we integrate both multiscale based image fusion and enhancement in a new fusion rule. In the new fusion rule, the modified...
multiscale edges instead of the edges themselves are used. By doing so, not only the line feature of the palmprint and palm vein are retained but also the contrast between the line and the surrounding area are enhanced.

The issue of how to represent the palm features for effective classification is an open problem. Various palmprint representations have been proposed for palmprint recognition, including line [11], points [21], Fourier spectrum [22], Mophological [24], Texture [25], Wavelet signatures [26] Gabor [15], Fusion code [27], and competitive code [28]. Recently, Sun et al [18] proposed to unify several state-of-the-art palmprint representations, such as competitive code [28] and fusion code [27], by using ordinal features. They claimed their algorithm is a general framework for the most current palmprint representations, for example Zhang et al.’s Gabor based representations [15, 27, 28], which reported the best recognition performance in the literature. However, the theoretical foundation of the operator has not been formulated. Recently hand vein has attracted increasing interest from both research communities [5, 6, 7, 8] and industries [9]. These features of the human hand, such as principal lines and palm veins, are relatively stable and the hand images, either colour or infrared, can be acquired relatively easily.

In biometrics applications, subspace learning methods play a dominant role in image representation as they allow efficient characterization of a low-dimensional subspace within the overall space of raw image measurements. Once a low-dimensional representation of the target class (face, hand, etc.) has been obtained, then standard statistical methods can be used to learn the range of appearance that the target exhibits in the new, low-dimensional space. Because of the lower dimensionality, less examples are required to obtain a useful estimate of discriminant functions (or discriminant directions). Appearance based representations of palmprint, such as eigenpalm [4], Fisherpalm [20] and ICA [19], have been investigated. In this paper, a new appearance representation of a fused hand image is proposed. The feature, which we call “Laplacianpalm”, is extracted from the fused image by locality preserving projection (LPP) [3]. While the Eigenpalm method aims to preserve the global structure of the image space and the Fisherpalm method aims to preserve the discriminating information, our Laplacianpalm method aims to preserve the local structure of the image space. In many real world classification problems, the local manifold structure is more important than the global Euclidean structure, especially when nearest neighbour like classifiers are used for classification. LPP potentially has discriminative power although it is unsupervised.

The experimental results on a large database show that the Laplacianpalm outperforms Fisherpalm and Eigenpalm, and the multimodal outperform any of the individual modality.

The rest of the paper is organized as follows: Section 2 discusses the system and image fusion. Geometry normalization is discussed in Section 3. A feature extraction scheme is detailed in Section 4. Section 5 reports our experimental results. The conclusions and the future work are presented in section 6.

2. Fusion of Palmprint and Palm vein Images

We discuss our method of fusion. A colour camera and a low cost monochrome JAI CV-M50IR 1/2" CCD IR camera are used to capture the palmprint image and palm vein image respectively. Since the NIR camera is not sensitive enough to detect the IR radiation emitted by the human body (3000-12000 nm), an IR light source [33] is used to irradiate the palm. The two cameras are fixed mounted upwards on a fixture. The user puts his hand on a plane glass in front of the two cameras with the fingers spread out naturally at a fixed distance away from the cameras. When a person put his hand on our imaging system, a colour palmprint image and a vein image of his palm are captured simultaneously. The palmprint image and palm vein image provide complementary information and the integration of such complementary information is useful for a more robust system. The fusion of the two images has an advantage over the fusion of matching score as it is not easy to be attacked using spoofs.

One of the main issues in image fusion is image alignment or registration, which refer to pixel-by-pixel alignment of the images. Once the images are registered, they can be fused.

2.1. Image registration

Registration of these two different images is an issue by itself. Automatic registration in a less constrained environment will not be dealt with in this paper. We assume here that the palm is placed in a fixed plane, so the projection matrix (homography) from color image plane to the near-infrared image plane can be expressed as a 3×3 matrix. We obtained the matrix with a calibration grid as shown in Fig. 1. We set the colour image as the source image and the near infrared image as the target image. The actual images are then registered by transforming them using this matrix.

Some examples of image registration are shown in Fig. 2, where the palm vein images are registered to the palmprint image.
2.2. Line-preserving and contrast-enhancing image fusion

Image fusion is the process by which two or more images are combined into a single image while retaining important features from each of the original images. Wavelet-based image fusion has become the most widely used image fusion method [29]. The principle of image fusion using wavelets is to merge the wavelet decompositions of the two original images using fusion methods applied to the approximate coefficients and detail coefficients. A general wavelet transform-based fusion is described by:

\[ I(x, y) = \omega^{-1}(\omega(\omega(I_1(x, y)), \omega(I_2(x, y)))) \]  

where \( \omega \) is the wavelet transform, \( \phi \) is the fusion rule, \( I_1 \) and \( I_2 \) are images to be fused.

In this paper, we want to retain the palmprint and palm vein lines after the image fusion. We can achieve this goal by combining the multiscale edges in palmprint and palm vein images. Multiscale edge-based image fusion [1] and Multiscale edge-based image enhancement [31] have been investigated separately. They are more interested in highlighting the features of interest and for reconstructing the image for human visual interpretations. Also, with the wavelet transforms, they could delete a lot of unwanted codes to achieve data compression. Being different from the existing approaches [1, 31], we are more interested in distinguishing one person's palm from another using an automated means. Thus preserving and enhancing discriminant information is our major concern. In this paper, we proposed a novel fusion rule in which image fusion and enhancement based on multiscale edge are integrated.

Denote a smoothing function by any function \( \theta(x) \) whose integral is equal to 1 and that converges to 0 at infinity. Two wavelet functions are defined such that,

\[ \varphi^1(x, y) = \frac{\partial \theta(x, y)}{\partial x} \]  

\[ \varphi^2(x, y) = \frac{\partial \theta(x, y)}{\partial y} \]

Let

\[ \varphi^1_s(x, y) = \frac{1}{s} \varphi^1(\frac{x}{s}, \frac{y}{s}) \]  

\[ \varphi^2_s(x, y) = \frac{1}{s} \varphi^2(\frac{x}{s}, \frac{y}{s}) \]

The wavelet transform of an Image \( I(x,y) \) is

\[ \left( \varphi^1 I(x,y), \varphi^2 I(x,y) \right) = \left( \frac{\partial}{\partial x} (I * \theta)(x,y), \frac{\partial}{\partial y} (I * \theta)(x,y) \right) = \tilde{\nabla}(I * \theta)(x,y) \]

The wavelet transform modules of \( I(x,y) \) at scale \( s \) is defined as

\[ M_s I(x,y) = \sqrt{\left| \varphi^1_s I(x,y) \right|^2 + \left| \varphi^2_s I(x,y) \right|^2} \]

The angle of a gradient vector with the horizontal direction at scale \( s \) is defined as

\[ \theta^s(x,y) = \frac{\varphi^1_s(x,y)}{\sqrt{\varphi^1_s(x,y)^2 + \varphi^2_s(x,y)^2}} \]


\[ A_I(x, y) = \tan^{-1}\left(\frac{\omega_y^{-1}(x, y)}{\omega_x^{-1}(x, y)}\right) \]  \tag{8}

A multiscale edge detection is equivalent to finding the local maxima of the Mallat’s wavelet transform [30], i.e. the maxima of \( M \) along the gradient direction corresponds to the edge points of \( I(x, y) \) at scale \( s \).

In our approach, the modified multiscale edge representations of the palmprint and palm vein images are fused to enhance the image contrast and the intersection points of the palmprint and palm vein lines. The fused images are reconstructed from the fused multiscale edges.

In order to remove noise, a low-pass filter is applied to the fused image. We used Gaussian smoothing for this purpose. The low-pass approximation at the coarsest scale \( s \) can be fused pixel-wise by taking the average of the two subimages:

\[ L_s I(x, y) = \frac{L_{p,s} I(x, y) + L_{v,s} I(x, y)}{2} \]  \tag{9}

Once the image details are classified through the amplitude of their wavelet coefficients in (7), the edge points pertaining to both source images are merged using the following rule:

\[
M_s(i, j) = \begin{cases} 
\omega_x, & \text{if } \omega_x(i, j) \text{ is maxima} \\
\omega_y, & \text{if } \omega_y(i, j) \text{ is maxima} \\
\max(M_x(i, j), M_y(i, j)), & \text{if } M_x(i, j) \text{ and } M_y(i, j) \text{ are maxima} \\
0, & \text{else}
\end{cases} \]  \tag{10}

where \( M_{p,s} \) and \( M_{v,s} \) are the wavelet of palmprint and palm vein images at scale \( s \) respectively, \( M_s \) is the wavelet representation of the combined image. \( C_i \) is used to stretch the gradient maxima and is a function of the scale index \( s \), \( C_i \geq 1 \). The fusion consists of retaining only the modulus maxima of the wavelet coefficients from the two images and combining them. The contrast, the intersection points of the palmprint and palm vein lines correspond to the positions where both palmprint and palm vein contain maxima. Hence, the fusion rule in (10) enhances the image contrast and also the “point of importance”.

Mallat [30] developed an iterative method to reconstruct the image from its multiscale edges. The reconstruction errors are below human visual sensitivity. We adopted Mallat’s reconstruction algorithm to reconstruct a fused image from the combined multiscale edges. In our experiments, we found that only two iterations are required to reconstruct the wavelet coefficients. \( L_s I(x, y) \) in (9) is updated with iteration in order that the reconstruction not only reconstructs interior pixels between edges, but also make \( L_s I(x, y) \) consistent with modified multiscale edges in (10). In doing so, in addition to enhancing the edges, the contrast between the palmprint and palm vein lines and their surrounding areas is enhanced.

After the reconstruction, a synthetic image is obtained that contains the edge information from the two images simultaneously. An example of our fusion process can be seen in Fig. 3, where the scale parameter is set equal to \( 2^j \), \( j = 1, 2, 3 \). We determine the \( C_i \) in (10) based on the recognition results. In our experiments, we found that the best results can be obtained when \( C_1 = 3, C_2 = C_3 = 4 \). In the fused image in Fig 3, we can see that the contrast between the palm vein and their surrounding areas is enhanced. We can also see that the cross points of the prints and vein are enhanced using the fusion rule in (10), i.e. the images of both palmprint and palm vein contain maxima at these points. Some of the palmprints in the fused image are also enhanced because they also appear in the palm vein images. Some palmprints, palm vein images and fused images can be seen in Fig. 4. We can see that the lines are reconstructed well in the fused images.

3. Geometry normalization

After the two images have been fused, the resulting image has to undergo a normalization process to prepare it for feature representation/extraction. The normalization process can be executed on any one of the palmprint image, palm vein image and the fused image because they have been fully registered in section 2. In our approach, we use the palm vein image for this purpose because the background of palm vein image is cleaner than the other two images. The image can be binarized easily by a threshold determined by the method in [17]. The boundary near the webs between the fingers can be found using a boundary tracing algorithm. The two webs can be located by computing the tangent touching the vertices of the two contours. We use the distance between the two webs beside the middle finger to normalize the fused images. Once the two webs are located, a coordinate system is constructed based on the two webs. A region is selected according to the coordinate system. In Fig. 5, we show the normalization of a palm vein image. \((u_1, v_1)\) and \((u_2, v_2)\) are two inner points in the finger webs, \((u_0, v_0)\) is the middle point of \((u_1, v_1)\) and \((u_2, v_2)\). The Y-axis is defined to be the line passing through the two webs and the X-axis is defined as the line passing through \((u_0, v_0)\) and perpendicular to the Y-axis. The image is scaled based on the distance of the two webs. ABCD is the final region-of-interest.

4. “Laplacianpalm” Representation

For most pattern recognition problems, selecting an appropriate representation to extract the most significant features is crucial. In the context of the appearance-based paradigm for object recognition, Principal Component Analysis (PCA) has been widely adopted to extract features with the highest covariant for dimensionality
reduction. But the features extracted by PCA are "global" features for all pattern classes and are not necessarily suitable for discriminating one class from the others. Linear Discriminant Analysis (LDA) seeks to find a linear transformation by maximising the between-class variance and minimising the within-class variance. Both the PCA and LDA methods seek the holistic properties of the whole enrolment population and thus will miss crucial details.

In this paper, we use the LPP to extract features of the fused palmprint and palm vein images. We called it “Laplacianpalm” features.

Recently, He et al [3] proposed an appearance-based face recognition method called the Laplacianface approach. By using Locality Preserving Projections (LPP), the face images are mapped into a face subspace for analysis. The Laplacianfaces are obtained by finding the optimal linear approximation to the eigenfunctions of the Laplacian Beltrami operator on the manifold [10]. They are linear projective maps that arise by solving a variational problem that optimally preserves the neighbourhood structure of the data set. Unlike PCA and LDA, the Laplacianface method [3] finds an embedding that preserves local information, and obtains the face’s subspace that best detects the essential object manifold structure. Both PCA and LDA methods are shown to be two special cases of the LPP method with different similarity matrices [3]. However, since the LPP defines the local structure in the form of a global transform, it is not intuitive to select the appropriate similarity matrix. Experimental results suggest that the Laplacianface achieves lower error rates in face recognition [3].

In this paper, we use the LPP to extract features of the fused palmprint and palm vein images. We called it “Laplacianpalm” features.
Figure 6. Normalized palmprints, palm vein and fused images. First column: palmprints; second column: palm vein; third column: non-enhanced fused image; fourth column: enhanced fused image.

As can be seen, the fused images have additional features that can only be derived from the combined palmprint and palm vein features. Such additional features provide more information to distinguish a person which cannot be obtained using score level fusion.

In addition, we captured the palmprint and palm vein images almost simultaneously. As both are obtained from different sources, it will be more difficult for spoof attacks. Not considering channel attack, the only possible way is to provide both patterns in the same image. The palmprint and palm vein images can be checked for similarity and it has to be substantially different for a real hand. Since addressing the spoof hand attack is beyond the scope of this paper, it will be addressed in our future work.

5. Experimental results

To the best of our knowledge, we are the first to fuse palmprint and palm vein at the image level for person recognition. There is no open database that contains both palmprint and palm vein images. Therefore, we have to collect our own database to evaluate the algorithm. The database, collected over a period of six months, contains 106 subjects, with ten images for each subject. In our experiments, three of them are randomly chosen for training, while the remaining seven are used for testing. A total of 120 random trials were performed and the mean of these trials is used in the final recognition results.

The resolution of the palmprint and palm vein images are 768 x 576. The fused images are normalized to 128 x 128. Some samples in our database are shown in Fig. 6. The normalized palmprint and palm vein images are shown in the first and the second columns respectively. The fused images without enhancement, i.e. the fused results with $C_1=1$, $s_1=1$, 2, 3 in (10), are shown in the third column. The fourth column shows the fused image with enhancement, i.e. the fused results with $C_1=3$, $C_2 = C_3 =4$ in (10). Comparing the results in the third column and the fourth column, we can see that the contrast between the palmprint and palm vein lines and their surrounding areas in the fused images has been significantly enhanced using our fusion rule. The first ten “Laplacianpalm” features are shown in Fig. 7.

The results are compared with those of the Eigenpalm [23] and Fisherpalm [20] to illustrate the advantages of the “Laplacianpalm”. The average error rates of the 120 random trials versus the dimension of the feature subspace are given in Fig. 8 or Table 1 and Fig. 9 or Table 2. We call the feature extracted from the palmprint image by LPP “Laplacianpalmprint”, the feature extracted from palm vein image by LPP “Laplacianpalmvein” feature. The feature extracted from fused image by Fisher LDA (PCA + LDA) [2] is called Fisherpalm feature. The feature extracted from fused image by PCA is called Eigenpalm feature. It can be observed that the optimal recognition rate based on “Laplacianpalm” is better than the optimal result using palmprint alone, palm vein alone and Fisherpalm. Its optimum occurs at very low dimensions and is an added advantage over the others.

Figure 7. The first 10 “Laplacianpalm” features of our training data; the first to the tenth feature are shown from left to right, from top to down respectively.

The proposed method in this paper is implemented in Visual C++. It runs in real-time on a PC using Intel Pentium IV 3.4GHz processor with 1G RAM. The average computation time for testing the five templates of each user in our database is listed in Table 3.
Figure 8. Error rates comparison of the fusion of palmprint and palm vein images, palmprint image alone and palm vein image alone

<table>
<thead>
<tr>
<th>Approach</th>
<th>Dimension</th>
<th>Error rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palmprint alone</td>
<td>17</td>
<td>3.72%</td>
</tr>
<tr>
<td>Palm vein alone</td>
<td>35</td>
<td>2.7%</td>
</tr>
<tr>
<td>Fusion of palmprint and palm vein</td>
<td>27</td>
<td>1.016%</td>
</tr>
</tbody>
</table>

Table 1. Performance comparison of palmprint alone, palm vein alone and the fusion of them

Figure 9. Error rates comparison of the “Laplacianpalm”, Fisherpalm and Eigenpalm

<table>
<thead>
<tr>
<th>Approach</th>
<th>Dimension</th>
<th>Error rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenpalm</td>
<td>253</td>
<td>12.88%</td>
</tr>
<tr>
<td>Fisherpalm</td>
<td>170</td>
<td>6.78%</td>
</tr>
<tr>
<td>Laplacianpalm</td>
<td>27</td>
<td>1.016%</td>
</tr>
</tbody>
</table>

Table 2. Performance comparison of Eigenpalm, Fisherpalm and Laplacianpalm on the fused images

6. Conclusions and future work

We have presented a multimodal biometric system that recognizes images of the palmprints and palm vein patterns of the human hands captured using a colour and a NIR camera. Two contributions are given in this paper:

Firstly, we proposed a person recognition approach by combining palmprint and palm vein images at the image level. To the best of our knowledge, we are the first to fuse palmprint and palm vein at the image level for person recognition. We fused the two images by combining their modified multiscale edges. In doing so, the important feature in palmprint and palm vein images, line features and intersection point features are well preserved. In addition, the contrast between the feature lines and the surrounding areas is enhanced and new features from the intersection of palmprint and palm vein images are obtained. Another advantage of fusion is that it gives a better protection against spoof attacks because both palmprint and palm vein are required simultaneously by the system.

Secondly, we proposed a novel feature, “Laplacianpalm”, which is obtained by applying locality preserving projection (LPP) on the fused images. LPP features have been verified as a more accurate representation of palm images. We evaluated our algorithm using a large database. The results showed that the fused images are better than the results on the palmprint alone and palm vein alone. It is also better than the results using Fisherpalm or Eigenpalm features on the fused image.

For the future work, it will be important to develop a general registration method which can relax the assumption in this paper, i.e. not require the points in the hand are in the same plane. The comparison with the matching score level fusion could be investigated in the future. An important advantage of our image level fusion over the matching score level fusion is that some redundant information, e.g. the intersections of the palmprint lines and palm vein lines, could be found in the fused image. These new features are more reliable than the lines themselves. The detection of the new recognition pattern as well as the performance based on the new

<table>
<thead>
<tr>
<th>Operations</th>
<th>Execution time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preprocessing (Gaussian, convert color images to gray level images)</td>
<td>40 ms</td>
</tr>
<tr>
<td>Image registration</td>
<td>100 ms</td>
</tr>
<tr>
<td>Image fusion</td>
<td>310 ms</td>
</tr>
<tr>
<td>Normalization</td>
<td>45 ms</td>
</tr>
<tr>
<td>Feature extraction</td>
<td>18 ms</td>
</tr>
<tr>
<td>Matching</td>
<td>9 µs</td>
</tr>
</tbody>
</table>

Table 3. Execution time for testing the five samples of each user
pattern could be investigated in the future, including the robustness against spoof hand attacks.

References


