Feature Level Fusion of WLBP and HOG for Hand Dorsal Vein Recognition

N M Sri Sangeetha¹, T Anantha kumar², Dr A M Natarajan³
¹PG Scholar, ²Senior Research Fellow, ³Professor & Chief Executive
Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Sathyamangalam, Erode, District, Tamil Nadu, India
sangeemuthu@gmail.com1, anandmecse@gmail.com, amn@bannari.co.in

Abstract— In this paper, a new approach is proposed to extract features from the dorsal hand vein pattern. The modified Weber Local Binary Pattern (WLBP) is feature descriptor extracted, which effectively combines the advantages of WLD and LBP. WLBP feature vector consists of two components: Differential Excitation and LBP. The Differential excitation component derived based on Weber’s law, which extracts the local salient patterns. LBP is highly discriminative, computationally efficient, and extracts the local micro-patterns. By computing the two components, we obtain two images: differential excitation image and LBP image, from which a 2D histogram for WLBP is constructed. Histogram of Oriented Gradients (HOG) feature is also extracted from same image. The proposed method fuses WLBP feature and HOG feature and assessment are done for the hand dorsal vein recognition. Experimental results show that fusion of WLBP and HOG features performs better than WLBP. Experiments are piloted on NCUT database, which shows that proposed fusion of WLBP and HOG is more effective and powerful texture descriptor.

Keywords—Biometrics, dorsal vein, differential excitation, Region Of Interest, uniform patterns, Weber’s law, Histogram of Oriented Gradients.

1. INTRODUCTION

As the level of security breaches and transaction fraud increases, the need for highly secure identification and personal verification technologies is becoming apparent. Biometric is the most secure and convenient authentication tool. It can’t be borrowed, stolen, or forgotten, and forging one is practically impossible. Personal identification is fundamental in any identity based access control system, and there is an increasing use of biometric features to authenticate individuals by measuring some inherent physiological or behaviour characteristics. Behavioral biometrics includes behaviors such as signature, gait and typing rhythms. Physical biometrics encompasses fingerprint, iris, hand, voice, and face, for identification. Each of these traits has their own disadvantages. Ear and iris pose a problem during sample collection.

Not only is an expensive and highly attended system required for iris but it also has a high failure to enroll rate. In case of ear data, it is hard to capture a non-occluded image in real time environment. In case of the well-known face recognition systems there are some limitations like aging, background, etc. Fingerprints, though most reliable, still lack automation and viability as they are also susceptible to wear and aging. Signatures are liable to forgery.

Hand vein patterns are among the biometric traits being investigated today for identification purposes, attracting interest from both the research community and industry. Biometric identification with vein patterns is a more recent approach that uses the vast network of blood vessels underneath a person’s skin, which ensures the live identification with better security and reliability shown in Figure 1. Veins are found below the skin and cannot be seen with naked eyes.

Figure 1. Hand Dorsal Vein

The dorsal venous network of the hand is a network of veins formed by the dorsal metacarpal veins. It is found on the back of the hand and gives rise to veins such as the cephalic vein and the basilic vein. And under visible light, the hand-dorsa vein structure is hard to discern. It is difficult for someone to tamper with the vein pattern. This feature makes it a more reliable biometric for personal identification. Furthermore, the state of skin, temperature and humidity has little effect on the vein image, unlike fingerprint and facial feature acquirement.
The hand vein biometrics principle is secure in nature where dorsal hand vein pattern are used to prove the individual identity.

II. LITERATURE REVIEW

Pattern recognition system was invented which compares two IR image patterns of same individual and the system will be enabled by verifying the individual’s identity (Willmore 1994). Histogram of oriented gradient (HOG) descriptors was experimented for human detection; it significantly performs better than all existing features (Dalal & Triggs 2005). A new personal authentication system that uses thermal-imaged vein pattern for dorsal hand was developed (Wang & Leedham 2005). A novel image collection method using low-cost devices was proposed for vein image acquisition, followed by vein pattern segmenting and feature extraction process was put forward (Zhao et al 2008).

A new hand vein recognition method using Scale Invariant Feature Transform features are extracted from training and testing vein image samples and similarity measure is used to verify the individuality (Wang et al 2009). Partition Local Binary pattern (PLBP) was an extension on texture feature LBP was proposed. PLBP partitions the image into sub-images and LBP feature is extracted for each sub-image, each of them are combined to form feature vector. With experiments on NCUT database PLBP shows better result than other features (Wang et al 2010).

Recently, Chen et al. (2010) proposed a simple and powerful local descriptor, Weber Local Descriptor (WLD), which was inspired by a psychological law, Weber's law (Jain 1989). Experiments show that WLD impressively outperforms the other widely used descriptors (e.g., Gabor and SIFT). However, according to the definition of WLD, the orientation component only considers the four pixels on the horizontal and vertical directions, which cannot fully represent local information. In addition, the orientation and differential excitations are just averagely quantized into some intervals. This kind of simple quantization method not only leads to information loss further, but also limits the performance of WLD. A novel extension to original LBP was attempted, named extended local ternary pattern (ELTP). Investigations were performed and experimental results shows better efficacy of ELTP over the original LBP (Liao 2010).

Recently, a new method of feature coding based on back propagation neural network was proposed for better performance in classification of hand dorsal vein (Wang & Liao 2012).

A novel approach for hand dorsal vein identification was proposed by using both the texture and geometry features. It is done hierarchically in two steps, first the coarse step segments the vein region and calculates its skeleton and the Energy Cost extracted in the Thinning process (TEC) is also used to reduce a large number of false candidates, greatly improving the efficiency in the fine step, both texture and geometry informations are presented by Local Binary Patterns (LBP) and the graph composed by the crossing points and endpoints of vein pattern respectively, and the two parts of information extracted are at last combined for recognition process. The proposed method is experimented with the NCUT dataset containing 2040 hand dorsal vein images of 102 subjects, and the experimental results obviously highlight its effectiveness (Zhu & Huang 2012).

III. ROI EXTRACTION OF NCUT HAND DORSAL VEIN DATASETS

In this work NCUT dataset is used to evaluate the proposed algorithm. The hand vein images are roughly aligned and differed by slight translations and rotations. Back of the hand vein image captured with a resolution of 640 by 480. A dataset of 2040 hand vein images was acquired under the natural lighting condition. It was named as North China University of Technology hand dorsal vein dataset or NCUT dataset. In detail, 10 right and 10 left back of the hand vein images were captured from all 102 subjects, aged from 18 to 29, of which 50 were male while 52 were female. As the vein pattern is best defined when the skin on the back of the hand is stretched tight, subjects were asked to clench their fists as acquiring vein patterns.

The image coverage area is larger than the back of the hand as shown in Fig 1(a). In this work, the image centroid was identified to extract the ROI. Let \((x_0, y_0)\) be the centroid of vein image \(f(x, y)\) then

\[
x_0 = \frac{\sum_{i,j} i \times f(i, j)}{\sum_{i,j} f(i, j)}
\]

(1)

\[
y_0 = \frac{\sum_{i,j} j \times f(i, j)}{\sum_{i,j} f(i, j)}
\]

(2)

After finding the image centroid the image cropping is subsequently performed to yield a sub-image of 360x360 pixels. Figure 2(a) shows a back of the hand vein image captured with a resolution of 640 by 480. Fig 2(b) and 2(c) respectively show the centroid and ROI of an image.
IV. WEBER LOCAL BINARY PATTERN (WLBP)

The reason to select LBP as a component of WLBP is that LBP can extract more local structure information than Orientation. Practically, LBP has proven to be highly effective and discriminative. For the original differential excitation component, find that the operator calculating the intensity differences of a current pixel against its neighbors is actually Laplacian operator. Since Laplacian operator is sensitive to noise, replace it with Laplacian of Gaussian (LoG) in this method. As a well-known local descriptor, LBP has been proven to be highly discriminative, computationally efficient and invariant to monotonic gray level changes. It does better than the orientation component of WLD in extracting local features. With differential excitation component, extract local salient patterns on human perception. By LBP component, local micro-patterns corresponding to bright/dark spots, edges and flat areas etc. are computed.

A. Weber’s law

Most of us can easily catch a whispered voice in a quiet room, but in a noisy environment people may not hear someone shouting near our ear. This is the essence of Weber’s law, observed by German physiologist Weber (1795–1878). It states that the ratio of the smallest perceptual increment in a stimulus to the background stimulus intensity is a constant. This relationship, which is known as Weber’s law, is shown as the following equation:

\[
\frac{\Delta I}{I} = K
\]  
(3)

Where \(\Delta I\) represents the increment threshold (just noticeable difference for discrimination); \(I\) represents the initial stimulus intensity and \(K\) is a constant proportion. The fraction \(\Delta I/I\) is known as the Weber fraction.

B. Differential excitation

For WLD, the differential excitation \(\xi(x_c)\) of a current pixel \(x_c\) is calculated like this,

\[
\xi(x_c) = \text{arctan}[G_{\text{ratio}}(x_c)] = \text{arctan} \left[ \frac{\Delta I}{I} \right] = \text{arctan} \left[ \frac{\sqrt{g^2}}{I} \right]
\]  
(4)

Here \(x_i\) is the i-th neighbour pixel of current pixel \(x_c\). \(I\) equals current pixel \(x_c\), and \(\Delta I\) is the intensity differences of current pixel against its neighbours. In this approach, the differential excitation is different from that of WLD. The design of the differential excitation operator is optimized.

C. Local binary pattern (LBP)

Ojala et al (2002) proposed Local Binary Patterns (LBP) provides an efficient rotation-invariant texture descriptor, and an illustration of the basic LBP operator is shown in Fig 2. For each pixel in an image, its value is compared with all the neighbouring pixel values. The result of each comparison is coded as binary 1 if the center pixel value is smaller and binary 0 otherwise. The binary bits are then grouped in the clockwise direction starting from the top left pixel, and the arranged binary string is converted to a decimal number as the final LBP result for the center pixel.

Here is an example of basic LBP operator:

\[
\begin{bmatrix}
6 & 5 & 2 \\
7 & 6 & 1 \\
8 & 8 & 7 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & 0 \\
1 & 1 & 0 \\
1 & 1 & 1 \\
\end{bmatrix}
\]

Figure 3 Example of basic LBP operator

The LBP operator can be extended to neighbourhoods of different sizes. For circular neighbourhoods, the pixel values can be interpolated to allow any radius and number of pixels in the neighbourhoods. With LBPP,R denoting \(P\) sampling points on a circle with a radius of \(R\), some examples of the circular neighbourhoods are shown in Figure 4.
To remove the effect of image rotation resulting in different binary patterns to be generated, each LBP is rotated to a position that acts as the common reference for all rotated versions of the binary patterns, and this involves the use of the rotation invariant LBP operator, $\text{LBP}_{P,R}^{u}$ defined

$$
\text{LBP}_{P,R}^{u}(s, t) = \begin{cases} 
1 & \text{if } U(x) \geq 2, \text{ or } y \geq 0, \text{ or } z \geq 0 \\
0 & \text{otherwise} 
\end{cases}
$$

(5)

If $U(x)$ is smaller than 2, the current pixel will be labelled by an index function $I(x)$. Otherwise, it will be labelled as $(P-1)P + 2$. By the index function $I(x)$, each of the uniform patterns is assigned to a particular index. Uniform pattern is circular binary pattern if it contains at most two bitwise transitions from 0 to 1, or vice versa (Ojala et al 2002). For example, circular binary patterns of 00000000, 00011110 and 10000011 are uniform patterns. Ojala (2002) noticed that in the experiments with texture images, uniform patterns account for a bit less than 90% of all patterns when using the (8,1) neighbourhood and for 70% in (16,2) neighbourhood. Uniform pattern can be regarded as the basic properties of image, because most patterns of image are uniform pattern. Therefore, the “uniform” LBP was selected in this method.

D. WLBP descriptor

As one of the variables for many local descriptors, histogram is often utilized to express descriptors. In this work, histogram is utilized to express WLBP descriptor. WLBP consists of two components: differential excitation and LBP, for a given image, compute the pattern value for every pixel by using the uniform LBP operator $\text{LBP}_{P,R}^{u}$. And the differential excitation of every pixel is computed by formula (10). Then, obtaining two images: differential excitation image and LBP image. According to these images, we first construct the 2D histogram $\{\text{WLBP}(s,t)\}$, $(s = 1,\ldots,S, t = 1,\ldots,T)$ of the original image. The size of this 2D histogram is $T \times S$, where $S$ is the number of intervals of $\xi$, $T$ is the total number of LBP's patterns. In this 2D histogram, each column corresponds to a pattern $t$ of LBP, and each row corresponds to a differential excitation interval. Thus, the value of each cell $\{\text{WLBP}(s,t)\}$ corresponds to the frequency of the certain differential excitation interval $\xi$ and the LBP pattern $t$.

To enhance the discriminability, the 2D histogram $\text{WLBP}(s,t)$ is further encoded into 1D histogram $H$. Each row of 2D histogram is used to form a 1D histogram $H(s), (s = 1,\ldots,S)$. Each sub-histogram $H(s)$ corresponds to the differential excitation interval $\xi$. Concatenating the $S$ sub-histograms, obtaining the 1D histogram $H = \{H_s\}, s = 1,\ldots,S$. According to the equation of $T = (P-1)P + 3$, $T$ equals to 59 as the values of the two subscripts $P$ and $R$ in the uniform LBP operator $\text{LBP}_{P,R}^{u}$ are assigned with 8 and 2 respectively. While the parameters $S$ (intervals number of $\xi$) and $K$ (threshold) can be set freely according to the real condition.

V. HISTOGRAM OF ORIENTED GRADIENTS

Local shape information is often well described by the distribution of intensity gradients or edge directions even without precise information about the location of the edges themselves. This method is based on considering well-normalized local histograms of image gradient orientations in a grid. The distribution of local intensity gradients and edge directions are shown in Figure 5, even without precise knowledge of the corresponding gradient or edge positions.

![Figure 5 Edge directions of an image](image-url)

HOG for an image is calculated by following steps:

1. Divide image into small sub-images: “cells”.
2. Accumulate a histogram of edge orientations within that cell
3. The combined histogram entries are used as the feature vector describing the vein image.

Initially, before extracting HOG ensuring normalized color and gamma values is essential. Normalization achieves a better result. The most common method is to normalize is to apply the 1-D point discrete derivative mask in one or both of the vertical and horizontal directions.
Also this method needs filtering the intensity or color of the image with the following filter kernels:

\([-1,0,1]\) and \([-1,0,1]^T\)

The next step of process involves creating the histograms for each cell separately. Every pixel inside a cell weighted vote for an orientation-based histogram channel by calculating gradient computation of each pixel. The histogram channels are evenly spread over 0 to 180 degrees or 0 to 360 degrees, Depending on whether the gradient is “signed” or “unsigned”, histogram channels are spread across 0 to 180 degrees and 0 to 360 degrees. Pixel contribution on the vote weight can either be the gradient magnitude, or some kind of function based on the magnitude; in general the gradient magnitude alone generally produces the better results.

Local object appearance and shape can be recognized well by the distribution of local intensity gradients or edge detection. HOG features are calculated for local regions by taking orientation histograms of edge intensity. For example, Histograms of edge gradients with 16 orientations are calculated from each of \(4 \times 4\) local cells. The edge gradients and orientations are obtained by applying Sobel filters. Thus the total number of HOG features will be

\[256 = 16 \times (4 \times 4).\]

Although the images are normalized to position and scale, the positions of important features will not be registered with same grid positions. It is known that HOG features are robust to the local geometric and photometric transformations. If the translations or rotations of the object are much smaller than the local spatial bin size, their effect is small.

VI. EXPERIMENTAL RESULTS

The NCUT dataset contains dorsal hand vein images of 2040 images from which 1020 images from left and right hand. There are ten instances for each person sample. Of these, first eight instances of each person are considered as a training set and they should be trained priorly before classifying the testing samples. We conducted the experiment separately for left and right hand images.

WLBP feature vector and HOG feature vector are extracted separately and fused into a single vector. The training samples are arranged together in a single matrix. The similar procedure is repeated for the testing samples. The resulting weights are compared with the training vectors. The pair with the minimum distance is taken as the correct match. There are five different distance measures used in this work namely Euclidean, Cityblock, Minkowski and Chi square test distance. From the matched samples, the correctly matched and mismatched samples are segregated.

According to the Euclidean distance formula, the distance between two points in the plane \(a\) and \(b\) with \(k\) dimension is given by,

\[\text{Dist}_{\text{Euclidean}} = \sqrt{\sum_{j=1}^{k} (a_j - b_j)^2}\]  \(\text{(6)}\)

The City block distance is always greater than or equal to zero. The measurement would be zero for identical points and high for points that show little similarity. The City block distance between two points, \(a\) and \(b\), with \(k\) dimensions is calculated as,

\[\text{Dist}_{\text{Cityblock}} = \sum_{j=1}^{k} |a_j - b_j|\]  \(\text{(7)}\)

Notice that for the special case of \(p = 1\), the Minkowski metric gives the city block metric, for the special case of \(p = 2\), the Minkowski metric gives the Euclidean distance, and for the special case of \(p = \infty\), the Minkowski metric gives the Chebychev distance.

\[\text{Dist}_{\text{Minkowski}} = \left(\sum_{j=1}^{k} |a_j - b_j|^p\right)^{1/p}\]  \(\text{(8)}\)
The chi-squared distance is useful when comparing histograms. The chi-squared distance between two vectors is defined as,

$$\text{Dist}_{\text{Chi-Square}} = \sum_{j=1}^{k} \frac{(a_j - b_j)^2}{a_j + b_j}$$

(9)

The confusion matrix is created with the sample instances of testing and training in the row and column respectively. From this, the diagonal values constitute the true positives. The column total represents the false positives and the row total, the false negatives. The remaining values constitute the true negatives.

These steps are repeated for different threshold values. The True Positive Rate (TPR) also called as Genuine Acceptance Rate (GAR) is computed with the help of sensitivity formula and similarly the False Positive Rate (FPR) also called as the False Acceptance Rate (FAR) with the specificity formula. A graph is drawn for GAR with FAR and is called as the Receiver Operating Characteristic (ROC) curve. The area under the curve gives the accuracy of the system designed in this project. The sensitivity (TPR) is computed as,

$$\text{Sensitivity} = \frac{\text{No. of True Positives}}{\text{No. of True Positives} + \text{No. of False Negatives}}$$

(10)

$$\text{Specificity} = \frac{\text{No. of True Negatives}}{\text{No. of True Negatives} + \text{No. of False Positives}}$$

(11)

The Equal Error Rate (EER) curve can be drawn in two ways. Basically it is the meeting point of the two curves of FRR and FAR. The EER curve obtained for each distance measure is shown in Fig 10. Table I shows the EER value obtained for the hand vein dataset with each distance measure.

<table>
<thead>
<tr>
<th>Distance Measure</th>
<th>EER value obtained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square</td>
<td>0.11</td>
</tr>
<tr>
<td>Cityblock</td>
<td>0.13</td>
</tr>
<tr>
<td>Euclidean</td>
<td>0.18</td>
</tr>
<tr>
<td>Minkowski</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Figure 7 ROC curve for TPR vs. FPR

Figure 8 ROC curve showing FRR vs. FAR

Figure 9 CMC curve

ROC curve can also be drawn between False Rejection Rate (FRR) and FAR. The efficiency of the system depends on the area under the curves i.e the area under the ROC curve of TPR and FPR must be maximum as possible whereas the area under the ROC curve of FRR and FAR must be minimum. These two curves are shown in Fig 7 and Fig 8 respectively.
The tables II and table III show the recognition rate for fivefold cross validation for WLBP and proposed feature vector fusion of WLBP and HOG with various distance measures. Fusion vector WLBP and HOG feature outperforms the WLBP feature and gives improved recognition rate. It can be seen that the Chi-Square distance gives the best results when compared with other distance measures. The next best distance measure that gives satisfactory result is Cityblock distance also referred as Manhattan distance.

Table 2

<table>
<thead>
<tr>
<th>Distance Measure</th>
<th>Chi-Square</th>
<th>Cityblock</th>
<th>Euclidean</th>
<th>Minkowski</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left</td>
<td>Right</td>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td>K Fold Cross Validation</td>
<td>Testing Images</td>
<td>Recognition Rate (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K=1</td>
<td>204</td>
<td>88.23</td>
<td>88.23</td>
<td>87.25</td>
</tr>
<tr>
<td>K=2</td>
<td>204</td>
<td>95.09</td>
<td>98.03</td>
<td>93.62</td>
</tr>
<tr>
<td>K=3</td>
<td>204</td>
<td>96.56</td>
<td>98.05</td>
<td>95.09</td>
</tr>
<tr>
<td>K=4</td>
<td>204</td>
<td>97.05</td>
<td>95.05</td>
<td>97.54</td>
</tr>
<tr>
<td>K=5</td>
<td>204</td>
<td>97.05</td>
<td>99.01</td>
<td>93.13</td>
</tr>
<tr>
<td>Avg. Recognition Rate</td>
<td>94.79</td>
<td>96.07</td>
<td>93.32</td>
<td>92.05</td>
</tr>
</tbody>
</table>

Figure 10 EER curves for various distance measures
VII. CONCLUSION

In this research work a biometric classification system using dorsal hand vein patterns is proposed with fusion of WLBP feature and HOG feature descriptors. The advantages of taking hand vein patterns are that it proves the liveliness of the person, and since it is beneath the human skin, it is indeed difficult to forge with. In order to extract the dorsal hand vein features for personal identification, this paper proposes an efficient method that fuses two feature descriptors WLBP and HOG. This method uses various distance measures such as Chi-square, Cityblock, Euclidean, and Minkowski as similarity measure between training and testing images. Experimental results show that the proposed method gives better results compared with the existing methods. In future work, hand dorsal vein recognition can be done using multiple feature vectors which helps in improving the recognition rate.

REFERENCES


