Improved Fuzzy Clustering (IFC) and Correlation Based user Threshold Selection with TRI-Branch for Finger Vein Recognition

K. Santhosh Kumar, D. Maheswari

Abstract—During the process of template matching with regard to finger vein identification, the probe used will get permitted only when the count of its vein points overlapping with the registered client is higher compared to the predetermined threshold. But, the admittance might be incorrect due to the neglect of the structure of the vein pattern. In the earlier works, the extraction of the vein structure (tri-branch) is done out of the vein pattern, and then combined with the entire vein pattern using a user-oriented threshold dependent filter setup. It renders a greater value of false acceptances, due to the User-oriented Threshold obtained from filter. In the step of branch tracking, the closest points are falsely detected, and therefore few tracking algorithms are needed. In order to resolve this, user-specific Threshold is generated on the basis of the correlation filter based selection combined with genetic algorithm. In the step of branch tracking, the closest points between the samples are decided by using the improved fuzzy clustering algorithm. It is observed that the local branches of the vein close to the segregation point of the vein pattern differ hugely from the fake pictures. The results of experiment carried out on two publicly available databases show the efficiency obtained of the novel design for boosting the performance achieved with respect to vein pattern based finger vein identification.

Keywords: Branch tracking, Finger vein, genetic algorithm, improved fuzzy clustering algorithm, Predefined threshold, TRI-branch vein structure, User-specific Threshold and Vein pattern

I. INTRODUCTION

Biometrics that includes the identification of the face, fingerprint, voice, iris, and finger-vein has found frequent application in different fields, like door access controls, online banking, personal recognition for computer access, automated teller machines (ATMs), and border control checks. Finger-vein identification helps in the distinct identification of persons on the basis of the vascular patterns present inside their fingers. This technique offers the benefits of sleek device size, superior accuracy, and ease of use. Also, one more unique benefit using the finger vein involves the alive body recognition, which implies that just the vein present in a finger that is alive could be detected [1].

In the case of finger vein identification, the available feature extraction techniques shall be categorized into two sets, which include non-vein pattern dependent techniques and vein pattern dependent techniques. The former type of techniques performs feature extraction from the entire finger image with no regard to the difference existing between the vein pattern and the non-vein area. But, the latter sort of techniques only makes use of the vein pattern extracted or else its feature for the identification task. Since just the vein pattern, with no vein area is utilized, vein pattern dependent techniques are found to give much better information in the intrinsic finger vein identification [2, 3].

This research work highlights on the performance improvement in the identification associated with the template matching process. During the template matching process of finger vein identification, the probe will get allowed only when the number of its vein points overlapping with the registered user is higher compared to the predetermined threshold. But, the allowance might be incorrect since the arrangement of the probe’s vein pattern and the vein pattern that is enrolled may exhibit difference even in the case of a huge count of overlapping vein points. Hence, the vein pattern may be capable of enhancing the template matching [4, 5, 6].

In the research works carried out recently, the vein structure close to the segmentation point of the vein pattern, called as the tri-branch vein structure, is investigated and it is used for improving the performance achieved of template matching using the novel user-based threshold dependent filter setup. The tri-branch vein pattern are obtained out of the entire vein structure using the vein network. It yields a bigger value of false acceptances, due to the User-based Threshold from the filter. In the step of branch tracking, the closest points are falsely detected, and therefore few tracking algorithms are needed [7].

In order to resolve this problem in this work, the user-specific Threshold is generated on the basis of the correlation filter based selection combined with genetic algorithm. In the step of branch tracking, the closest points between the samples are decided by using improved fuzzy clustering algorithm. It is observed that the local vein branches close to the segmentation point of the vein pattern differ hugely in the fake images[8].

II LITRATURE REVIEW

Qin et al [9] introduced a novel approach to boost the performance achieved with finger-vein recognition systems. At first, a vein pattern extraction technique for extracting the shape and orientation features of the finger-vein is introduced.
Next, in order to make place for the powerful local and global differences simultaneously, a region-dependent matching approaches is explored using the Scale Invariant Feature Transform (SIFT) matching technique. At last, the shape, orientation and SIFT features of the finger-vein are merged for performance improvement further.

Liu and Kim [10] introduced an effective finger-vein extraction algorithm that depends on random forest training and regression scheme using a strong local binary pattern feature. Through the integration performed using a vein pattern matching technique that offers robustness towards finger misalignment, benchmark finger-vein identification is achieved.

Muntak ET AL [11] studied the skeleton of a vein pattern identification system, which was designed with Verilog HDL (Hardware Description Language) considering FPGA (Field Programmable Gate Array) implementation. Developing countries can reap huge benefits utilizing the FPGA-based security systems since FPGAs can be reconfigured, are very quick though with less consumption of power and therefore are more cost-efficient compared to general purpose processors. The already available algorithms are slightly improved algorithms in order to program most of the portions of this system.

Fayyaz et al [12] suggested user Finger Vein Authentication (FVA) in the form of a biometric system. Making use of the distinct features for the classification of theses finger veins is one among the important tips, which differentiate the associated works, and therefore a set of representative features is proposed to be learnt, on the basis of auto encoders. The user finger vein is modelled with the help of Gaussian distribution.

RamaPrabha et al [13] investigated computational intelligence technique for step of personal authentication in a biometric system. The realization of a personal identification system can be done with the help of finger vein biometric trait. Finger vein identification offers several benefits in comparison with other biometric traits. This system is trained with Genetic based Principal Component Analysis for extracting the vein pattern and Back Propagation Network to achieve pattern matching of the finger vein.

III. PROPOSED METHODOLOGY

This section explains about the novel finger vein recognition work elaborately. This model comprises of six components and the first one involves thinning and denoising, second one includes bifurcation detection, third one involves branch tracking, fourth one is Morphological dilation and dot product, fifth one includes the selection of User-specific threshold and the sixth one is identification. Figure 1.illustrates the structure of the proposed technical work.

![Finger Vein images](image1.png)

**Figure 1. Architecture of the proposed system**

This step starts with the Extraction of the tri-branch Vein Structure. Local vein branches close to the bifurcation point differ hugely between the fake images. In this, the bifurcation is used and its branches are extracted to boost the performance of vein pattern based technique. Since the bifurcation point detection and vein branch tracking in multiple-pixel sized vein pattern is found difficult, the tri-branch vein structure extraction is done with the help of single-pixel sized vein network, which implies the thinned vein pattern.

![Input image](image2.png)

**Figure 2. Input image**

3.1. Thinning and denoising

Thinning may be achieved by getting the skeleton of the area employing skeletonization, which is also called as thinning [16]. It computed in this work with the help of mathematical morphology operators. In this work, it corresponds to the hit or miss transform, which can be used for formulating its expression. The expression for the thinning of an image I with a structuring element $J = (J_1, J_2)$ is provided as:
\[ Thin(i, j) = i - (i \otimes j) \]  

(1)

where, the subtraction indicates the logical subtraction expressed by:

\[ X-Y=X \cap NotY \]  

(2)

The operation is applied iteratively till it makes no more modifications to the image (i.e., until the point of convergence). The structuring element order J is user specified. The image gets thinned using the structuring element pairs \((J_1^1, J_2^1)\), \(i = 1\.8\) in an order. A connected skeleton of the image is achieved out of this. The single-pixel sized vein network is obtained from the vein pattern using the operation of morphological thinning which is shown in figure 3. [17].

\[ N_e = \sum_{i=1}^{n} |p_i + 1 - p_i| \]  

(3)

Where \(p_0 = p_1\)

Result of the bifurcation detection is shown in figure 4.

3.2. Bifurcation detection.

Usually, three vein branches are connected to one single bifurcation point. Otherwise said, six is the switching number of the dividing line. According to this, the technique of bifurcation detection gets developed [20]. Supposing the present point and its eight neighbour points are represented by \(p(x, y)\) and \(\{P_i = p_1, p_2, \ldots p_8\}\), correspondingly [21]. The point \(p(x, y)\) is considered to be a division point, if \(N_e\) is equivalent to 6, which is expressed as below:

\[ N_e = \sum_{i=1}^{n} |p_i + 1 - p_i| \]  

(4)

Result of the bifurcation detection is shown in figure 5.

3.3. Branch tracking using Improved Fuzzy c-means Clustering algorithm

Fuzzy c-mean clustering is one among the unsupervised clustering algorithms, which is extensively employed in image processing and computer vision due to its easy implementation and better clustering performance.

The primary drawbacks of FCM algorithm are:

1. FCM algorithm exhibits sensitivity to the point that is isolated.
2. FCM algorithm requires the number of clustering \(c\) and the fuzzy weighted index \(m\) in advance, however \(c\) and \(M\) has a direct impact on the clustering result.
3. The FCM algorithm settles into the local extreme point or saddle point quickly and the optimum result could not be achieved.

To overcome those issues in this work improving the basic fuzzy c means algorithm.

3.3.2. An improved fuzzy C-means clustering algorithm

In this section, an improved fuzzy C-means clustering algorithm is presented. The fundamental concept is to include a weighted value to the data element’s membership degree, and then including the fuzzy clustering validity function to the algorithm for optimizing the clustering number \(c\). The IFCM Algorithm is given below:

Initialization: Given the clustering number \(c, 2 \leq c \leq n\), \(n\) refers to the amount of image pixel, determining the iteration threshold \(\epsilon\), actual clustering pattern \(P^{(0)}\), and iteration counter \(b = 0\).

**Step 1** computation or update done on the partition matrix \(U^{(b)}\)

\[ \mu_{ik}^{(b)} = \left( \sum_{j=1}^{c} \left[ \frac{d_{ik}^{(b)}}{d_{jk}^{(b)}} \right]^{\frac{2}{m-1}} \right)^{-1} \]  

(5)

If \(\exists i, r\), cause \(d_{ir}^{(b)}\) is small, then \(\mu_{ik}^{(b)} = 1\) and if \(i \neq r\), \(\mu_{ik}^{(b)} = 0\).

Obtain the modified membership grade

\[ N_{ik} = \mu_{ik} - \frac{(\epsilon - \mu_{ik})}{\epsilon} \]  

(6)
Step 2: update clustering pattern matrix $P^{(b+1)}$ for $i = 1, \ldots, c$

$$p_i^{(b+1)} = \frac{\sum_{k=1}^{N_i} (n_k^{(b+1)})^m}{\sum_{k=1}^{N_i} (n_k^{(b+1)})^m}$$  \hspace{1cm} (7)

Step 3: If $p^{(b)} - p^{(b+1)} < \varepsilon$, then the algorithm is stopped, and the partitions matrix $U$ and clustering pattern $P$ get exported, else let $b = b + 1$, move to Step 1.

Step 4: compute the value of $FP^{(b+1)}(U; c)$, if $FP^{(b+1)}(U; c) < FP^{(b)}(U; c)$, then $c = c + 1$, move to step 1, else, the value of validity function $FP(U; c)$ becomes minimal, and the clustering number $c = c - 1$, and clustering process is terminated.

In accordance with the steps, the single-pixel sized tri-branch vein arrangement (i.e., the division and its local branches) is identified which is shown in figure 6.

3.4. Morphological dilation and dot product.

A morphological operation performed on a binary image generates a new binary image in which the pixel’s value is a non-zero only when the test succeeds at that position in the input image. The operation of morphological dilation is carried out on the single-pixel sized tri-branch vein arrangement in this research work, in which the structuring object is basically an $8 \times 8$ matrix having an elemental value of 1. The dot product between the dilated setup and the entire vein pattern is carried out to get the map of tri-branch vein arrangement for the purpose of matching. The extracted vein structure is shown in figure 7.

3.5. User-specific Threshold based Filter

This section introduces the learning process of the user-specific threshold and the use of the threshold to filter out the unnecessary. Taking into consideration that the significance of the resultant matching score obtained of every biometric trait is different among the users; the user-specific weights are utilized for combining the multi-trait for various users in a multi-biometric system. Motivated by the learning of the user-based weight, the user-oriented threshold is done and then utilized for filtering out the imposters present in this framework.

3.5.1. User-Specific Threshold Based On Correlation Filter Based Selection with Genetic Algorithm

The training images are used for learning the user-specific thresholds. Supposing that $N$ registered users exist in the database, and every registered user possesses $m$ training images. During the training stage, the similarities between any two of the images amongst the $m$ training images are compared for every client, using the correlation coefficient of the image pixels.

If the correlation between two images is adopted in the form of a goodness measure, in which an image is considered good when it has high correlation to the other image but does not high correlation with any of the other image. Otherwise said, in case the correlation between image pixels is sufficiently high to render it to be relevant, and the correlation range between image’s pixel and any other relevant one does not attain a level such that it can be assumed by any of the other images, then it will be considered to be a proper measure for the classification task.

In this condition, the threshold selection problem needs an appropriate measure of correlations range between image’s pixel and a strong procedure to choose the threshold depending on this measure. [24]. Under the conventional linear correlation scheme correlation coefficient expressed by the formula:

$$correlation(r) = \frac{\sum_{x=1}^{N} \sum_{y=1}^{N} (XY - \sum_{x=1}^{N} X \sum_{y=1}^{N} Y)(X - \sum_{x=1}^{N} X)(Y - \sum_{y=1}^{N} Y)}{\sqrt{\sum_{x=1}^{N} X^2 - \sum_{x=1}^{N} X}^2 \sum_{y=1}^{N} Y^2 - \sum_{y=1}^{N} Y}$$  \hspace{1cm} (8)

Where $X$ and $Y$ specifies the pixels of two finger vein Images.

3.5.2. Genetic Algorithm and threshold selection:

The genetic algorithm (GA) is a form of an optimization and search strategy that depends on the concepts of genetics and natural selection. A genetic algorithm primarily comprises of three kinds of operators, which are selection, crossover, and mutation. With the selection operator, a good string (based on fitness) is chosen for breeding another new generation;

Crossover merges superior quality strings to produce an offspring that is much better; mutation locally modifies a string in order to preserve the genetic diversity existing between different generations of a population consisting of chromosomes. In every generation, the population is assessed and then evaluated for stopping the algorithm execution. In case the termination condition is not met, then the three GA operators again act upon on the population and then re-evaluate them. The GA cycle goes on till the termination condition is attained.

In the case of user specific threshold selection, Genetic Algorithm (GA) is utilized in the form of a random selection algorithm, which has the capacity of efficiently investigating
massive search spaces, generally needed in scenarios of user specific threshold selection.

This research work introduces a technique for user specific threshold selection depending on the correlation employing GA. Correlation range between the image’s pixels will determine the fitness of the individual to participate in the process of mating. Fitness function for GA is actually a very ordinary function that fixes a rank to every image based on the correlation coefficients [25].

3.5.3. Genetic Representation of Image’s pixel (Encoding) and creating initial population

The GA images need to be ciphered in the form of chromosomes for the application of GA operations on them. Encoding is defined as a process of representation involving the genes individually (pixels in image). The procedure can be carried out with bits, numbers, trees, arrays, lists or any other kind of objects. The images are represented in bits here.

3.5.4. Finding the correlation between pixels of finger vein image and rating them as per their fitness

The fitness of every single chromosome is measured based on its correlation coefficients along with other properties present in the population.

3.5.5. Applying GA Operations and building the optimum subset of Images

Crossover operation with the probability value, Pc=0.8 and Mutation with the probability value, Pm=0.01 is used on the consecutive populations to generate the next generation’s population. The algorithm terminates when no change is observed in the population’s best fitness for a predefined number of generations. In case the maximum of generation has been attained much earlier than the specific number of generation and there are no variations, then the process will terminate. Once the GA process ends, three fittest ones are available. As explained before, the fitness function for GA is just an ordinary function that designates a rank to every pixel based on the correlation coefficients.

The fitness function considered here is:

\[ f(x) = 1 - \min(r_x) \]  

(9)

Where, \( min(r_x) \) refers to the minimum value of the correlation coefficient associated with any image’s pixel. Depending on the user selected, the particular threshold filters out the imposter with success to be the matching score obtained between user and probe is lesser compared to the minimum real scores of the concerned user.

IV RESULT AND DISCUSSION

The results of the experiments carried out on the newly presented schemes have been introduced in this section. The implementation of the proposed scheme is carried out in MATLAB2012 running on a PC with CPU 3.60GHz and 12.00G memory capacity and then its comparison is done with the available Anatomy Structure Analysis-based vein extraction (ASAVE), user specific threshold based filtering (USCBF) and the proposed novel correlation filter based selection with genetic algorithm (CFBGN) in terms of precision, recall and accuracy and false rejection rate for the Hong Kong Polytechnic University (HKPU) database finger vein images [26].

In this database, each one of the first 210 fingers makes 12 images, acquired in two sessions, and each one of the final 102 fingers makes 6 images, acquired in one session. Each one of the images are composed in a 8-bit gray level BMP file having a resolution of 513_256 pixels.

Figure: 8. Sample finger vein image

Evaluation metrics

Precision

Precision is defined to be the ratio of the real positives against both true positives and false positives results obtained for multimodal biometric images. It is expressed below

\[ Precision = \frac{T_p}{T_p + F_p} \]

Sensitivity

Sensitivity assesses the ratio of actual positive that classifies the real image to be actual. The sensitivity is expressed as below:

\[ Sensitivity = \frac{T_p}{T_p + F_n} \]

Where \( T_p \) refers to the face image rightly as the face image. \( F_p \) indicates the image (non-face) incorrectly to be the non face image. \( F_n \) refers to the image (non face) incorrectly to be the face. \( T_n \) specifies the non face correctly to be non face.

Accuracy

Accuracy is defined to be the overall correctness achieved of the model and is computed to be the sum of original classification parameters \( (T_p + T_n) \) divided by the overall sum of classification parameters \( (T_p + T_n + F_p + F_n) \)

\[ Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \]  

(12)

TABLE: 1 performance comparison result values

<table>
<thead>
<tr>
<th>Method</th>
<th>Precision</th>
<th>Sensitivity</th>
<th>Accuracy</th>
<th>Error Rate</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASAVE</td>
<td>81.1900</td>
<td>83.4000</td>
<td>83.2500</td>
<td>16.750</td>
<td>82.2950</td>
</tr>
<tr>
<td>USCBF</td>
<td>88.9181</td>
<td>88.9725</td>
<td>89.0244</td>
<td>10.975</td>
<td>88.9453</td>
</tr>
<tr>
<td>CFBGN</td>
<td>91.8363</td>
<td>92.0033</td>
<td>91.8699</td>
<td>8.1301</td>
<td>92.0033</td>
</tr>
</tbody>
</table>
Improved Fuzzy Clustering (IFC) and Correlation Based user Threshold Selection with TRI-Branch for Finger Vein Recognition

Figure 9. Precision results of different method

Figure 9 depicts the results of the performance comparison of the novel CFBGN and the available ASAVE technique and USCBF technique in terms of Precision. It is concluded from the results that the novel CFBGN model generates much better precision results of 91.83 % while the available ASAVE technique and USCBF achieves just 88.91 %, 81.19% correspondingly.

Figure 10. Recall results of different method

Figure 10 depicts the outcomes of the performance comparison carried out of the novel CFBGN technique and the available ASAVE scheme and USCBF technique in terms of Recall. It can be concluded from the results that the novel CFBGN model yields higher Recall results of 92.00% while the available ASAVE technique and USCBF yields just 83.40%, 88.97% correspondingly.

Figure 11. Accuracy results of different method

Figure 11 illustrates the performance comparison results of the novel CFBGN technique and the available ASAVE and USCBF techniques in terms of accuracy. It can be concluded from the results that the novel CFBGN model yields much better accuracy results of 91.86% while the available ASAVE scheme and USCBF yields just 83.25%, 89.02% correspondingly.

V CONCLUSION AND FUTURE WORK

This research paper studies the extraction of the finger vein pattern to boost the identification performance achieved of the template matching process. In the branch tracking step, the closest points between the samples are decided by using the improved fuzzy clustering algorithm. In this research work, user-specific Threshold is set on the basis of the correlation filter dependent selection with genetic algorithm. It is observed that a local vein branch close to the segmentation point of vein pattern differs hugely between the fake images. It can be concluded from the result that the novel system yields much better performance in terms of precision, recall and accuracy and false rejection rate. As a future approach, the implementation of this system can be done with more diverse features yielding much better
security and lesser verification time.

REFERENCES


