ISSN (Print): 0974-6846 ISSN (Online): 0974-5645

# Finger Vein Identification based on the Fusion of Nearest Neighbor and Sparse Representation based Classifiers

#### Shazeeda\* and Bakhtiar Affendi Rosdi

Universiti Sains Malaysia, School of Electrical and Electronic Engineering, Engineering Campus, Seberang Perai Selatan, Nibong Tebal, Penang 14300, Malaysia; s12\_eee052@student.usm.my, eebakhtiar@usm.my

#### **Abstract**

**Objective:** In this study, a new approach for personal identification using finger vein pattern is presented, to improve nearest neighbour algorithm by combining k-nearest neighbour and sparse representation based classifiers (KNN-SRC). **Methods/Analysis:** In the proposed KNN-SRC method, K numbers of best nearest neighbor samples were selected based on *k* NN classifier. Subsequently, the selected Ksamples were considered as the train samples for SRC classifier. **Findings:** Finger vein is a cutting-edge technology in biometrics that attracts attention of researchers from worldwide. As compared to the conventional biometric traits such as face, fingerprint and iris, finger vein is more secured and difficult to forge, as the veins are embedded in human tissue. Despite the intensive progress in feature extraction techniques from the captured vein images, there is a critical urge to develop an effective method for classification of the extracted features. Results of the present study using our own database revealed that, the proposed KNN-SRC method accomplished the analysis significantly faster than the SRC method which could be attributed to the reduced number of training samples. In addition, the KNN-SRC method gives higher accuracy than the common nearest neighbour (*k* NN) and SRC methods, individually, which could be attributed to sparsely representation of the test sample. **Novelty /Improvement:** In this method, recognition rate of finger vein images is improved by combining two classification techniques. For the first time, the KNN-SRC classification method was used for finger vein images.

Keywords: Biometric, Finger Vein Recognition, Nearest Neighbour Classification, Sparse Representation Classifier

#### 1. Introduction

Security is one of the most important factors to transform the developing countries into developed ones. Personal identification is gaining high demand in corporate, government and public security systems. Therefore, a rigorous research has been focused to develop technologies such as Personal Identification Number (PIN), smart card and passwords. Nevertheless, such conventional authentication techniques have potentially a high risk of being duplicated and can be used by an imposter easily. Therefore, a novel technology called biometric, was emerged in late 90s. Biometrics exploits the physical or behavioral characteristics for personal identification and authentication<sup>1</sup>. Among the physical characteris-

tics, finger vein has recently become a central focus in biometrics authentication research. Biologically finger veins are the epidermal patterns of blood vessels. Along with the common requirement of biometric identification such as stability and durability<sup>2</sup>, the finger vein has unique network pattern. Finger vein based biometric is an emerging technology in the security systems due to its various advantages such as the vein pattern is embedded inside the skin and not visible to the human eye. Hence, it is extremely difficult to be forged<sup>3</sup>. In addition, as the vein image is provided by a living person only, thus the finger vein pattern is very difficult to be cheated by unreliable veins<sup>4</sup>.

So far, numerous studies have been reported to improve the accuracy of finger vein recognition. For

<sup>\*</sup>Author for correspondence

example, a finger vein extraction method using repeated line tracking<sup>5</sup>, phase only correlation method<sup>6</sup>, Local Binary Pattern (LBP)<sup>Z</sup>, Principal Component Analysis (PCA)8, Gabor Wavelets and Circular Gabor Filter9 were proposed. However, most of the reported approaches for finger vein recognition are mainly focused on the feature extraction process. Although, classification is a vital and crucial element in pattern recognition, yet very few studies have been focused on the classification process. For example, Wu et al. 10 proposed Support Vector Machine (SVM) to classify the finger vein images. Similarly, another study utilized a Sparse Representation based method in which the sparse solution was obtained by 11 minimization for the classification of the finger vein images<sup>11</sup>. Recently, our group has reported a classifier called Local Mean based K-nearest centroid neighbor (LMkNCN) on finger vein images which was based on nearest neighbors<sup>12</sup>.

However, many drawbacks are associated with the reported classifiers. For example, a biggest difficulty in the SVM is to generate the model from training data, as the classification outcome will be less than optimal if the parameters are not set properly. Whereas, in the Sparse Representation based method, the training samples are represented by sparse solution which has to be obtained by optimization of l1 norms which is an iterative process and thereby it increases complexity and processing time. In the Nearest Neighbor based classifier such as LMkNCN, the distances between the test and the train images will be computed by Euclidean distance method. The method is sensitive to the noisy and irrelevant attributes that leads to less meaningful distances which reduces the recognition accuracy<sup>13</sup>.

An algorithm was proposed by Nan et al. <sup>14</sup> on hand writing and face recognition using an algorithm based on the combination of the two standard classification techniques i.e. KNN and SRC (KNN-SRC). This algorithm resulted into high recognition rate compared to the individual classifiers, KNN and SRC in From this algorithm, the drawbacks of KNN method can be reduced by sparsely representing the nearest neighbors of the test sample and also the complexity of SRC can be reduced by reducing the number of training samples. The advantages of KNN-SRC algorithm was motivated us to apply it for the finger vein recognition in order to obtain high recognition rate with high processing speed.

The rest of this paper is organized as follows: in Section 2 an over view of the proposed classification techniques are described, in Section 3 the proposed algorithm

is explained, in Section 4 the experimental results and discussion are provided and in section 5 a conclusion is given for the overall study.

# 1.1 Overview of the Classification Techniques Used in the Study

#### 1.1.1 Nearest Neighbor (NN)

Nearest Neighbour algorithm is a non-parametric and a simple effective method of classification. The method is considered to be the instance based as it predicts the class label for the test sample by finding k closest points from the training samples. The k closest points are predicted by the distance weighting estimation from the test sample to all the train samples and predict the class label for the test sample<sup>13</sup>. Some of the common weighting schemes are Euclidean distance, Minkowski Distance, Manhattan distance, Chebychev distance and Hamming distance<sup>15</sup>.

Mathematically, the concept of NN is as follows. Let y be the test image for which the nearest neighbour in each class is  $T_{ij}$  may be estimated by distance measure. In the present paper the distance estimation is done by squared Euclidean distance method. The distance between the test sample y and the train sample  $T_{ij}$ , mathematically can be written as

$$\mathbf{d}_{\mathbf{i}}(\mathbf{y}) = \|\mathbf{y} - \mathbf{T}_{\mathbf{i}\mathbf{j}}\|^2 \tag{1}$$

where,  $T_{ij}$  is the nearest neighbour in each class.

Let us assume that the n th class is the class with minimum distance to the test sample y, then

$$d_n(y) = min \ d_i(y) \tag{2}$$

For 1-NN (k=1), the decision for the test sample  $\mathbf{y}$  is made by considering the minimum distance, where  $\mathbf{n}$  is the class to which the test sample  $\mathbf{y}$  belongs to. The 1-NN concept can be extended to k NN, suppose there are  $\mathbf{k}_i$  number of nearest neighbour samples for the test sample  $\mathbf{y}$ , then  $\mathbf{k}_n = max \, \mathbf{k}_i$ , i.e. the sample  $\mathbf{y}$  belongs to the class which has maximum number of nearest neighbors of nearest neighbors for the test sample by estimating the minimum distance class to the test sample. The class with minimum distance to the test sample is considered to be the class of the test sample.

# 1.1.2 Sparse Representation based Classifier (SRC)

Sparse representation is a process to search the most appropriate representation of the signal from the linear combination of atoms in a dictionary matrix. A set of dictionary matrix is created to represent the training samples and the test sample can be represented by the atoms of the dictionary matrix. The sparse solution can be obtained by optimization of  $l_i$  minimization problem  $^{11,16}$ . There are many techniques for solving the optimization problem such as Method Of Frames (MOF), Matching Pursuit (MP), the Best Orthogonal Basis (BOB), and Basis Pursuit (BP) $^{17}$ .

Mathematically, the sparse representation can be expressed as, let  $^{c}$  be the number of training classes with  $^{n}$  images in each class. We can arrange the  $^{i}$   $^{t}h$  train

ing sample as  $A_i = \{a_{i,1}, a_{i,2}, a_{i,3}, \dots, a_{i,n_i}\}$   $\in \mathbf{D}^{m \times n_i}$ , hence the column matrix  $A_i$  is the training images of the i th class. Hence the generalized train set can be defined as  $A = \{A_1, A_2, A_3, \dots, A_c\}$   $\in \mathbf{D}^{m \times n}$ 

where, 
$$n = \sum_{i=1}^{c} n_i$$
.

The sparse solution to y = Ax can be optimized by  $\hat{x}_0 = arg \min \|x\|_0$  subjected to y = Ax (3) where,  $\|x\|_0$  is  $l_o$  norm which gives the nonzero component of the vector x. By replacing  $l_o$  norm in equation (3) with the  $l_o$  norm defined as

 $\hat{x}_1 = arg \ min \ \|x\|_1 \ \text{subjected to } y = Ax$  (4) where,  $\|x\|_1$  is  $l_i$  norm and also if certain condition of sparsity is satisfied, the solution of equation (3) can be considered as the sparse solution for equation (4). The optimization problem can be solved by iterative methods through the standard linear programming algorithms. Consider  $Y_i(x)$  be the vector for the  $i \ th$  class whose nonzero entries are the entries of x, give as

$$Y_i(x) = \left\{0, 0, \dots \dots \alpha_{i,1}, \alpha_{i,2}, \alpha_{i,3} \dots \alpha_{i,n} \right\}$$

Now the test sample  $\hat{y}$  is reconstructed by considering the entries related to i th class as  $\hat{y} = AY_i(x)$ . The decision of the test sample as to which class it belongs

is obtained by computing  $r_i(y) = y - \hat{y}$  which is the reconstructed column matrix for the i th class. The minimum value in the reconstructed column vector is considered as the class of the test sample which is given as

$$r_c(y) = \min \|y - \hat{y}\| \tag{6}$$

where, c is decided as the class to which the sample y belongs to 16.

# 2. Materials and Methods

# 2.1 Proposed Algorithm: K-Nearest Neighbor based Sparse Representation Classifier (KNN-SRC)

In the present study, an attempt is made to combine the nearest neighbour and the sparse representation. The present study was motivated and based on the algorithm proposed by Nan et al.<sup>14</sup>. In the first stage, for the given test samples the nearest neighbors were computed by squared Euclidean distance as shown in equation 1. In the second stage the classification by sparse representation was applied by considering required number of nearest neighbors estimated as the training samples for the SR classifier. The main advantages of the proposed algorithm are, the number of train samples for the SRC was reduced to a large extent which in turn reduces the complexity and increases the processing speed and ultimately the accuracy of NN was improved. Figure 1 depicts the illustration of the present method.

For the given test sample y, the nearest neighbors for the test sample were estimated with respect to all the training samples. Let K be the nearest neighbors for the given test sample y, such that  $K_i$  be the neighbors

$$\left\{ \boldsymbol{\check{\alpha}}_{i,1}, \boldsymbol{\check{\alpha}}_{i,2}, \boldsymbol{\check{\alpha}}_{i,3}, \dots, \boldsymbol{\check{\alpha}}_{i,K_i} \right\}$$
  
for the  $\boldsymbol{i}$  th class  $K = \sum_{i=1}^{c} K_i$ . Let

$$\check{A}_i = \left\{ \check{\alpha}_{i,1}, \check{\alpha}_{i,2}, \check{\alpha}_{i,3}, \ldots, \check{\alpha}_{i,K_i} \right\} \in \mathbf{D}^{m \times K_i} \text{ and } \check{A}$$

is the matrix composed of K nearest neighbors for the given test sample. The sparse coefficients are computed as

$$\min \|x\|_1 \text{ subjected to } y = Ax \tag{7}$$

In the present study, the linear coefficients were computed by  $l_i$  norm based on the iterative method called

primary dual interior frame work by selecting K nearest neighbors as the training samples<sup>18</sup>.

Let  $\widetilde{Y}_i$  (x) be the nonzero coefficient vector for the i th class and the characteristic column vector is constructed as  $\widetilde{\beta} = \widetilde{A} \widetilde{Y}_i$  (x). The corresponding class residual reconstruction is obtained as

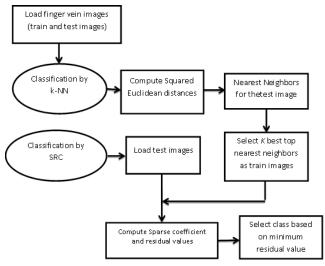
$$\widetilde{r_i}(y) = \|y - \widetilde{\beta}\| = \|y - \widetilde{A}\widetilde{Y_i}(x)\|$$
(9)

Decision of the class for the test sample was done based on the minimum value of the residual reconstruction which can be written as,

$$\widetilde{r_c}(y) = min \|y - \widetilde{\beta}\|$$
(8)

The classification results are computed in terms of accuracy based on the number of images classified correctly. The accuracy for the classification in terms of percentage can be given as

$$Accuracy = \frac{Number\ of\ test\ images\ that\ were\ classified\ correctly}{Total\ Number\ of\ test\ images} \times\ 100$$



**Figure 1.** Block diagram representation of KNN-SRC methodology.

# 3. Results and Discussion

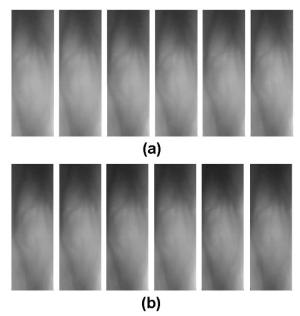
## 3.1 Experimental Results

The experiment was carried using MATLAB R2013a on WINDOWS 7 INTEL® CORE ™ i5 CPU, with operating frequency of 2.53G Hz and a usable memory of 4G Bytes. The database used in the present study was developed from our previous study¹9. The database consists of 492

classes with 6 images in each class, a total of 2952 number of images. The finger images are captured in two sessions at a difference of two-week time. Hence, the total number of images is 5904 (2952 from the 1st session and 2952 from the 2nd session). The region of interest (ROI) of the images for the finger vein recognition was obtained by the algorithm described in<sup>19</sup>. Figure 2 (a) shows the ROI of the finger vein images acquired in the 1st session and Figure 2 (b) shows the ROI of the finger vein images in the 2nd session.

In the present study, the 2952 finger vein images taken in the first session were considered as the training images and all 2952 images of second session as test images. The finger vein images of the train and test samples were resized to 0.1 and are normalized to the unit norm.

In the proposed algorithm (KNN-SRC), firstly the different number of K is considered where K is the number of nearest neighbors for the given test sample obtained from k NN classification. The selection of K is made by selecting best top minimum distance nearest neighbors.

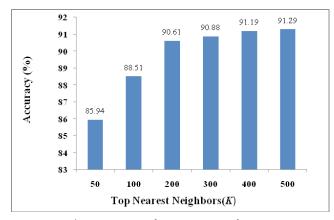


**Figure 2.** Finger vein images from the database FV-USM (a) 1<sup>st</sup> session (b) 2<sup>nd</sup> session.

Table 1 gives the accuracy and the processing time for the KNN-SRC classification. The results of accuracy and processing time are depicted in graphical representation as shown in Figures 3 and 4, respectively. From the analysis of the results it is observed that the accuracy increases when *K* is increased.

**Table 1.** Results for KNN-SRC classification for finger vein images

Best top nearest neighbors (K)	Accuracy (%)	Total Time (s)
50	85.94	76.93
100	88.51	111.65
200	90.61	171.10
300	90.88	234.39
400	91.19	300.46
500	91.29	352.62



**Figure 3.** The accuracies for KNN-SRC for K top nearest neighbors.

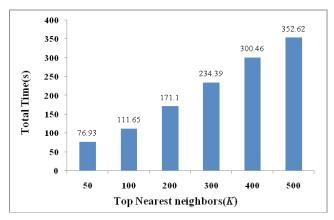


Figure 4. The processing time of KNN-SRC.

#### 3.2 Discussions

Overall comparative analysis of KNN-SRC with k NN (k = 1) and SRC classification techniques tested in the present study are given in Table 2. It is observed that the accuracy was improved in KNN-SRC when compared to 1-NN and SRC individually. In addition, the processing speed was found to be faster than SRC. Figures 5 and 6 graphically

depict the comparison of accuracy and processing time between all the tested classifiers.

From the analysis of the experimental results, it is inferred that KNN-SRC displayed significantly better accuracy than both 1-NN and SRC. Additionally, the KNN-SRC was found to be better than SRC in terms of processing speed. The drawback of the present study lies in the selection of the nearest neighbors for the classification by sparse representation. In the present study, among the total 2952 test samples, about 257 test samples were not identified correctly. The main reason for the misclassification is found to be due to the NN classifier which in turn due to the high sensitivity of this algorithm towards the noise and irrelevant attributes. However, it involves with simple calculation of distances thereby it takes very less processing time.

The analysis of the KNN-SRC algorithm for K = 500 considering the finger vein database is as follows.

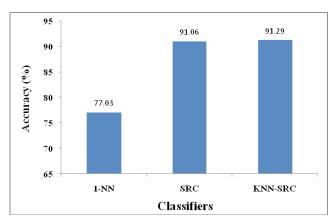
- Total number of train samples = 2952.
- Total number of test samples = 2952.
- Number of test samples identified correctly by 1-NN = 2274.
- Number of test samples not identified correctly by 1-NN = 678.
- Number of test samples identified correctly by KNN-SRC = 2695.
- Number of test samples not identified correctly by KNN-SRC = 257.
- Number of test samples not identified correctly only by SRC = 6 (when SRC is applied after k NN i.e. KNN-SRC).
- Number of test samples not identified correctly only by NN= 251.

Table 2. Comparison of accuracies among classifiers

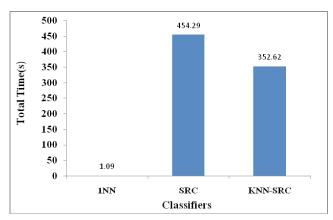
Classifier	Accuracy(%)	Total Time(s)
1-NN	77.03	1.09
SRC	91.06	454.29
KNN-SRC (K=500)	91.29	352.62

### 4. Conclusion

In this paper, a new algorithm by combining two standard classification techniques, k-NN and SRC were proposed and successfully implemented for finger vein database. The proposed classifier aims at enhancing the finger vein



**Figure 5.** Comparative analysis of accuracies of the classifiers.



**Figure 6.** Comparative analysis of processing time of the classifiers.

recognition by combining the k-nearest neighbor and sparse representation based classifiers (KNN-SRC). In the proposed KNN-SRC classification algorithm, the misclassification problem in k-NN was resolved by sparsely representing the test sample and the processing speed of SRC was improved by reducing the training samples. Results of the present study showed that the classification of the finger vein images was significantly improved by the combined classifier (KNN-SRC) than compared to the k-NN and SR classifiers, individually. The accuracies and processing speed obtained are 77.03% and 1.09s, 91.06% and 454.29s, and 91.29% and 352.62s for 1-NN, SRC and KNN-SRC, respectively. In the present method, the accuracy is improved compared to 1- NN and the SRC. In addition, the processing speed of the KNN-SRC is significantly faster than compared to the SRC. The present method can be further enhanced by selecting a better method for nearest neighbor classification and a suitable methodology for selecting K nearest neighbor. Thus, the further enhancement work is under investigation and will be reported in due course.

# 5. Acknowledgement

This work is supported by Universiti Sains Malaysia Research University Grant No. 1001/PELECT/814116 and Post Graduate Incentive Grant No. 1001/PELECT/8023013.

## 6. References

- Jain AK, Ross A, Prabhakar S. An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology. 2004 Jan;14(1):4–20.
- Yang J, Yihua S, Jinli Y, Lihui J. A novel finger-vein recognition method with feature combination. 16th IEEE International Conference on Image Processing (ICIP), Cairo; 2009. p. 2709–12.
- Yang J, Shi Y, Yang J, Jiang L. A novel finger-vein recognition method with feature combination. Proceedings of 16<sup>th</sup> IEEE International Conference on Image Processing (ICIP); 2009. p. 2709–12.
- Hashimoto J. Finger vein authentication technology and its future. Proceedings of 2006 IEEE Symposium on VLSI Circuits. Digest Technical Papers; 2006. p. 5–8.
- Miura N, Nagasaka A, Miyatake T. Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. Machine Vision and Applications. 2004;15(4):194–203.
- Mahri N, Suandi SA, Rosdi BA. Finger vein recognition algorithm using phase only correlation. 2010 International Workshop on Emerging Techniques and Challenges for Hand-Based Biometrics (ETCHB); 2010. p. 1–6.
- Wang KQ, Khisa AS, Wu XQ, Zhao QS. Finger vein recognition using LBP variance with global matching. 2012
   International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR); 2012. p. 196–201.
- 8. Beng TS, Rosdi BA. Finger-vein identification using pattern map and principal component analysis. 2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Malaysia; 2011. p. 530–4.
- Yang J, Yang J, Shi Y. Combination of gabor wavelets and circular gabor filter for finger-vein extraction. Emerging Intelligent Computing Technology and Applications. 2009; 346–54.
- 10. Wu JD, Liu CT. Finger-vein pattern identification using SVM and neural network technique. Expert Systems with Applications. 2011; 38(11):14284–9.
- 11. Xin Y, Liu Z, Zhang H, Zhang H. Finger vein verification system based on sparse representation. Applied Optics. 2012; 51(25):625228.
- 12. Amatriain X, Jaimes A, Oliver N, Pujol JM. Data mining methods for recommender systems. Recommender Systems Handbook: Springer US; 2011. p. 39–71.

- 13. Zhang N, Yang J. K nearest neighbor based local sparse representation classifier. Proceedings of 2010 Chinese IEEE Conference on Pattern Recognition (CCPR), China; 2010. p. 1–5.
- 14. Bonet I, Rodríguez A, Grau R, García MM, Saez Y, Nowé A. Comparing distance measures with visual methods. Proceedings of 2008 Intern. Conf. Artif. Intel. Mexico, Springer Berlin Heidelberg; 2008. p. 90–9.
- 15. Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y. Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2009; 31(2):210–27.

- 16. Chen SS, Donoho DL, Saunders MA. Atomic decomposition by basis pursuit. SIAM Review. 2001; 43(1):129–59.
- 17. Borsic A, Adler A. A primal–dual interior-point framework for using the L1 or L2 norm on the data and regularization terms of inverse problems. Inverse Problems. 2012; 28(9):095011.
- 18. Asaari MS, Suandi SA, Rosdi BA. Fusion of band limited phase only correlation and width centroid contour distance for finger based biometrics. Expert Systems with Applications. 2014; 41(7):3367–82.