Journal of Artificial Intelligence and Data Mining (JAIDM), Vol. 9, No. 4, 2021, 571-582.

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research paper

Robust Vein Recognition against Rotation using Kernel Sparse Representation

Ali Nozaripour and Hadi Soltanizadeh*

Department of Electrical/Computer Engineering Semnan University, Semnan, Iran.

Article Info

Article History:
Received 15 November 2020
Revised 07 May 2021
Accepted 11 June 2021

DOI:10.22044/JADM.2021.10253.2164

Keywords:
Sparse Representation, Kernel Trick, Dorsal Hand Vein Pattern, Region of Interest, Floating Region of Interest, Classification.

**Corresponding author: h_soltanizadeh@semnan.ac.ir (H. Soltanizadeh).*

Abstract

Sparse representation due to the advantages such as noise-resistance and having a strong mathematical theory has been noticed as a powerful tool in the recent decades. In this work, using the sparse representation, kernel trick, and the different technique of region of interest (ROI) extraction presented in our previous work, a new and robust method against rotation is introduced for the dorsal hand vein recognition. In this method, in order to select ROI, by changing the length and angle of the sides, the undesirable effects of hand rotation during taking images are largely neutralized. Thus depending on the amount of hand rotation, ROI in each image will be different in size and shape. On the other hand, due to the same direction distribution on the dorsal hand vein patterns, we use the kernel trick on sparse representation for classification. As a result, most samples with different classes but the same direction distribution will be classified properly. Using these two techniques leads to introduce an effective method against hand rotation for dorsal hand vein recognition. An increase of 2.26% in the recognition rate is observed for the proposed method when compared to the three conventional SRC-based algorithms and three classification methods based on sparse coding that use dictionary learning.

1. Introduction

Among all the biometrics used for identification, the pattern of the subcutaneous veins due to its great advantages such as the diagnosis of being alive (due to the flow of blood in the veins), easy and contactless acquisition of the image, and also high security (due to being inside the body) is of interest to many researchers [1-3].

In order to design the identification systems (ID), one of the most important concerns is to select a similar and uniform area in all images called the region of interest (ROI), which will have a huge impact on the final result. Due to the number of dorsal hand veins being less than the palm and finger veins, in the ID systems, through the dorsal hand vein pattern, this area should be selected with most of the veins, and does not change the coordinates during all the rotation modes [4, 5].

The easiest way to prevent the unwanted effects of the hand spin and relocation is to use a docking device when taking an image [6]. Some articles have also used points between fingers in order to select a square or rectangular ROI to eliminate the docking [7, 8]. However, in all of these works, some restrictions on the hand position must be applied while taking an image [9].

In this work, in order to eliminate all the restrictions, a robust method against hand spin proposed in our previous work [10] was used. In this technique, unlike other methods, the shape of ROI is not constant, and the length and angle of the sides will be changed depending on the angle of the hand spin. Another problem with the ID systems is the use of a reliable algorithm in order to classify and compare the images. In the last few years, we have witnessed the rapid growth of the

sparse representation theory and algorithms and the success in applications such as image compression and restoration [11-14], compressed sensing [15, 16], image classification [17-25], and denoising [26, 27].

According to the sparse representation, the classification consists of two steps: coding and classification. First, the query signal/image is coded according to a dictionary with a sparsity constraint. After that, the classification operation is performed based on this coding, which is somewhat sparse [17]. In some of the sparse representation works, the dictionary can be pre-designed. For example, in [28], Haar Wavelet and Gabor Wavelet have used as a dictionary. However, these dictionaries are not usually useful for specific types of images such as facial images and dorsal hand veins.

Another type of dictionary is created by learning algorithms. A large number of dictionary learning (DL) methods have been proposed for image processing [11-16, 29, 30] and classification [17-20]. For example, one of the most famous dictionary learning methods is the K-SVD algorithm [12], in which an over-complete dictionary of patches related to the training dataset is learned. Since K-SVD is just able to reconstruct the training sample, this algorithm is not suitable for classification. Thus Zhang *et al*. [31], by adding a discriminative reconstruction constraint to the dictionary learning model (D-KSVD), made a separable dictionary, which was used for face recognition (FR). One of the other dictionary learning methods for classification is making a dictionary for each class separately. In this method, the query sample is classified base on the reconstruction error associated with each class [32].

Some methods have used the training samples as the dictionary atoms directly. For instance, Wright *et al*. in [23] have proposed the SRC method for the coding and classification of a query face image. The advantage of these methods is that there is no need for dictionary learning but it cannot well-classify the data that has the same direction distribution. In other words, the methods like SRC cannot classify the sample vectors belonging to different class distributions in the same vector direction [24].

Thus due to the same direction distribution on the dorsal hand vein patterns, we used the kernel trick on sparse representation for classification. Furthermore, as mentioned earlier, we used our previous ROI extraction method [10] named Floating ROI (FROI) for ROI extraction. Therefore, in this work, using the sparse representation, kernel trick, and FROI, a new and robust method against rotation is introduced for dorsal hand vein recognition.

The rest of this paper is organized as what follows. In Section 2, we briefly review the SRC method and introduce the kernel trick. Section 3 presents the proposed method with its algorithm. Our method is compared with other classifiers on the Bogazici University database in Section 4. Finally, conclusions are presented in Section 5.

2. Sparse Representation

2.1. Overview

Sparse representation, as a powerful and efficient replacement tool for the classical transformations, has received much attention in the recent decades. The goal of this method is to represent a query signal as a linear combination of several specified basis signals, in which most of its coefficients are zero or very small. In other words, suppose a signal as $y \in \mathbb{R}^m$ and the basis signals $\{d_i\}_{i=1}^n$. Now we want to obtain the vector $\mathbf{x} \in \mathbb{R}^n$ from the system of the linear equation as:

$$y = \sum_{i=1}^{n} x_i\, d_i = Dx, \tag{1}$$

where $D = \{d_1, d_2, \dots, d_n\} \in \mathbb{R}^{m \times n}$, the transform matrix whose columns are the basis signals. If $m = n$ and D is orthogonal, the solution of the above equation can be easily obtained by the following equation:

$$x = D^{-1}y = D^{T}y. \tag{2}$$

However, in many cases, the solution will not be sparse. In the sparse representation, the number of basis signals is more considered than its dimension $n > m$. Thus we are faced with an underdetermined system since the number of equations is less than the unknowns. If y is not in the span of the D columns, the system does not have any solution; otherwise, there are numerous solutions. In order to avoid undesirable states (no solution), the matrix D is assigned as a full rank matrix. Now the sparest solution must be selected, i.e. a solution with the fewest non-zero entries. Thus we have the following optimization problem:

$$\begin{aligned} &min_x \|x\|_0 \\ S.t. \quad &y = Dx \end{aligned} \tag{3}$$

in which $\|x\|_0$ denotes the $\ell_0 -$ pseudo norm of the vector $x \in \mathbb{R}^n$, which counts the number of non-zero entries. This is for the ideal conditions. Considering the noise vector $\xi \in \mathbb{R}^m$ with the energy $\|\xi\|_2 < \varepsilon$, the following relation will

result:

It should be noted that in the sparse representation, the D matrix is called a dictionary and each column of it is called an atom.
Since the problems (3) and (4) are non-convex and $\|x\|_0$ is not smooth, solving them requires a combinatoric search, and therefore, it turns to an NP-hard problem for the higher dimension. Thus in many research works, $\ell_0 -$ norm has been used as a convex problem and a good approximation of the $\ell_0 -$ norm. In the following section, we briefly discuss one of the most popular of these methods.

2.2. Sparse Representation-based Classifier

Assuming that there are sufficient training samples n_i for the ith class:

$$D_i = \{d_{i,1}, d_{i,2}, \dots, d_{i,n_i}\} \in \mathbb{R}^{m \times n_i} \quad (5)$$

Therefore, each test sample of $\boldsymbol{y} \in \mathbb{R}^m$ belonging to class ith must approximately be laid in the linear span of the training samples belonging to the ith class:

$$y = x_{i,1} d_{i,1} + x_{i,2} d_{i,2} + \dots + x_{i,n_i} d_{i,n_i} \quad (6)$$

where $x_{i,j}$ is the scalar. The dictionary D is formed by placing all the training samples together:

$$D = \{D_1, D_2, \dots, D_c\} = \left[d_{1,1}, \dots, d_{1,n_1}, \dots, d_{c,1}, \dots, d_{c,n_c}\right] \in \mathbb{R}^{m \times n}, \quad (7)$$

where c represents the number of classes, and therefore, $n = \sum_{i=1}^{c} n_i$. According to the above, y can be written as a linear combination of all the training samples as follows:

$$y = x_{1,1} \times d_{1,1} + x_{1,2} \times d_{1,2} + \dots + x_{c,n_c} \times d_{c,n_c} = D \times x\,, \quad (8)$$

As mentioned in the previous section, due to the NP-hard problem (4), $\ell_1 -$ norm can be used as an approximation of the $\ell_0 -$ norm. Therefore, relation (4) can be re-written as follows:

$$\begin{aligned} & min_x \|x\|_1 \\ S.t. \quad & \|y - Dx\|_2 \leq \varepsilon. \end{aligned} \quad (9)$$

Thus if the test sample y belongs to the ith class, by solving problem (9), the entries of x are expected to be zero, except some of those associated with the ith class, namely:
$x = [0, \dots, 0, x_{i,1}, x_{i,2}, \dots, x_{i,n_i}, 0, \dots, 0]^T$.
It has been shown in [22] that the class of the test sample can be found by computing the reconstruction error for each class and finding a minimum of them.

$$\begin{aligned} \hat{C} = & \; argmin_{i=1}^{c} \|y - D\hat{x}_i\|_2, \\ S.t. \quad & \|y - D\hat{x}\|_2 \leq \varepsilon \end{aligned} \quad (10)$$

where y is the test sample, $\hat{x}_i$ is a new vector with the same dimension of the x whose only non-zero entries are the entries in x that is associated with the ith class. It should be noted that vector x is obtained by solving problem (9).
As observed, SRC is a non-parametric learning method that requires no training process and can determine the test sample class directly. In SRC, the training samples are considered as the sparse representation matrix columns, and then the sample size is reduced by applying a transfer matrix. Random projection can be a good choice for dimensional reduction because in SRC, a precise choice of the feature space is not important and the random features contain enough information to achieve the desired result [24].
However, this method is weak in dealing with the data for the various classes that have the same direction distribution. The main reason for this weakness is that after the normalization operation, the data with the same direction will overlap with each other. In order to solve this problem, we transfer the data to a higher space with the kernel.

2.3. Kernel Trick

Typically, the methods of feature extraction and classification can be divided into linear and non-linear methods. When a database is not linearly separated due to complexity, the non-linear methods such as kernel-based methods are used. Using kernels, the linear algorithms can be generalized to the non-linear applications.
For example, the kernel principal component analysis (KPCA) and generalized discriminate analysis (GDA) are generated by applying the kernel method to the principal component analysis (PCA) and linear discriminant analysis (LDA), respectively [33].
In the kernel-based methods, at first, the database is mapped to a high dimensional space to be linearly segregated in that space.
The main idea of these methods is that the inner product of the mapped data by a non-linear function can be estimated with a kernel function called the kernel trick.

$$k(x, x') = \Phi(x)^T . \Phi(x'), \quad (11)$$

where T shows the transpose of a vector or matrix, x and x' are the feature vectors in the initial space, and Φ represents an implicit non-linear mapping associated with the kernel function $k(.,.)$. It should be noted that in the kernel

methods, it does not require to know what Φ is and just adapt the kernel function (11). Some of the most important kernels that apply to the Mercer condition are the linear kernel, polynomial kernel, radial base kernel, and circular kernel.

$$Linear\ kernel: \quad k(x,x') = x^T.x', \tag{12}$$

$$Polynomial\ kerne: \quad k(x,x') = (x.x'+1)^P \tag{13}$$

$$RBF\ kernel \quad k(x.x') = exp(-\gamma\|x-x'\|_2^2), \tag{14}$$

$$Circular\ kernel \quad k(x.x') = tanh(kx.x'-\delta), \tag{15}$$

3. Proposed Method

3.1. ROI Selection and Vein Extraction of Infrared Images

In this work, we used the dorsal hand vein images of the Bogazici University database that contained 1200 infrared images belong to the left hand of 100 people [33]. Since the infrared light penetration is not the same on the various skins, and the noise and shadings in the images are different, the contrast of images taken from different individuals is not the same. Therefore, it is necessary to increase the image contrast firstly.

Thus using a median filter, the noise of the images are eliminated, and then in order to normalize the intensity distribution in the different images, the normalization operation is performed using Equations (16), (17), and (18) to get the same mean and variance for all images [10].

$$M = \frac{1}{p\times q}\sum_{i=0}^{p-1}\sum_{j=0}^{q-1} I(i,j) \tag{16}$$

$$VAR = \frac{1}{p\times q}\sum_{i=0}^{p-1}\sum_{j=0}^{q-1}(I(i,j)-M(I))^2 \tag{17}$$

$$G(i,j) = \begin{cases} M_0 + \sqrt{(\frac{VAR_0}{VAR}\times (I(i,j)-M)^2} & I(i,j) \\ M_0 - \sqrt{(\frac{VAR_0}{VAR}\times (I(i,j)-M)^2} & I(i,j) \end{cases} \tag{18}$$

where, *M and VAR* are the mean and variance of the input image, respectively, and M_0 = 150 and VAR_0 = 255 are the mean and variance of the interest, respectively.

After that, for the ROI extraction from the smooth infrared images related to a hand, we have done as follows.

When we have a wrist spin of the human's hand, one side will be compressed and the other side will be extended. As shown in Figure 1(b), when we have a wrist spin to the left, a compression occurs on the left side. Therefore, the distance between points A and C will be shorter compared to the normal state. On the contrary, when we have a wrist spin to the right, the distance between points B and D will be shorter because a compression occurs on the right side (Figure 1(c). Therefore, to select a similar area in the smooth images, we are used FROI [10].

According to the FROI technique, the lateral side length in the compressed area of the hand with the angles of two sides, toward the upper side (θ in Figure 2(c)), should be changed relative to the amount of the hand spin. Thus by changing the upper two angles of FROI, as well as changing the length of the lateral side in the compressed part, a similar area with the maximum vein pattern will be obtained [10].

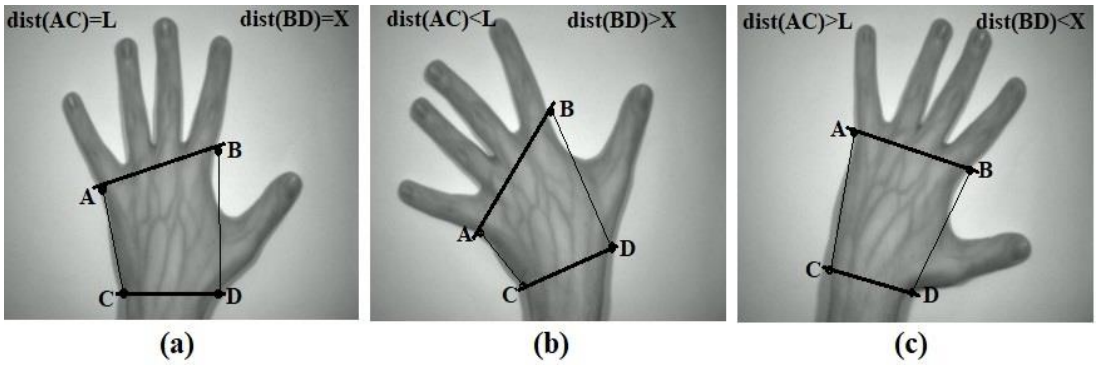

Figure 1. Images of a hand in three situations of the wrist spin. (a) Normal situation with no wrist spin. The lengths of sides are defined as $\overline{AC} = L$ and $\overline{BD} = X$. (b) Wrist spin to the left, so $\overline{AC} < L$ and $\overline{BD} > X$. (c) Wrist spin to the right, therefore $\overline{AC} > L$ and $\overline{BD} < X$.

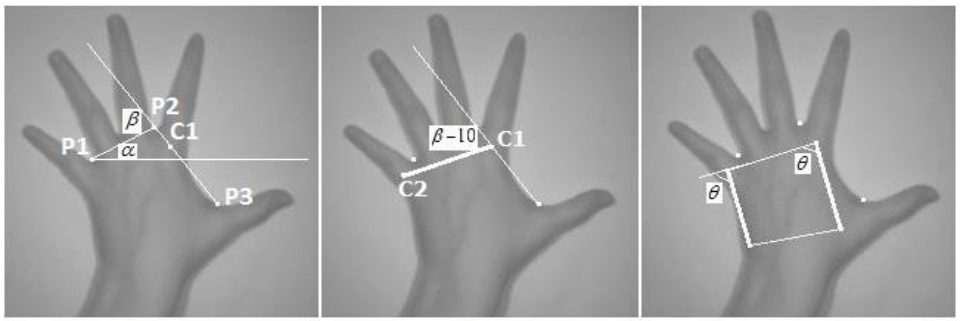

Figure 2. Process of FROI extraction. (a) Finding the angle β and the point C_1. (b) Finding the point C_2 and forming the upper side of FROI. (c) Determining FROI via the relationships (19) and (20).

For this, we first find point C_1 (as the first corner of quadrilateral ROI) based on the points $P_1, P_2,$ and P_3. This point (C_1) designates on $\overline{P_2P_3}$ with $\frac{1}{4}$ of the length of $\overline{P_2P_3}$ from P_2 (Figure 2(a)). Then a line (d_1) is drawn from C_1 with the length of 1.25 of $\overline{P_1P_2}$ at an angle β-10 degree with the line $\overline{P_2P_3}$ (Figure 2(b)). After that, we used Equations (19) and (20) for the lateral side length of FROI in the compressed area and the angles of two sides toward the upper side of FROI, respectively (Figure 2(b)).

$$d_2 = \left(\frac{100-(10\,log|\alpha-10|}{100}\right) \times d_1. \qquad (19)$$

In above relation, d_2 is the lateral side length in the compressed area of the hand, d_1 is the length of the upper side and the other side of FROI, and α is the angle that the line $\overline{P_1P_2}$ makes with the horizon (Figure 2(a)).

The relation (20) is also defined in order to find the angle of the two sides toward the upper side of FROI:

$$\theta = \begin{cases} 90 & 0 < \alpha < 20 \\ 180 - (\beta + 10\,log|\alpha - 10|) & oth, \end{cases} \qquad (20)$$

where θ is the left angle that the lateral sides make with the upper side of FROI, and β is the angle that the line $\overline{P_1P_2}$ makes with the line $\overline{P_2P_3}$ (Figure 2(a, c)). As it can be seen, in this approach, unlike [7, 8], the final form of ROI is not always a square or rectangular. Figure 3 shows a better performance of this technique (FROI) than the technique [7] in which for all hand spins ROI is a fixed square.

For a better comparison, on the veins pattern of the first-hand spin images (Figure 3 (Sa1, Fa1)), we signed 4 points, all of which were located inside ROIs, and in the corresponding image with a different spin (Figure 3 (Sa2, Fa2)), the location of such corresponding points is marked. By comparing the corners of ROIs in Figure 3 (Sa1) and Figure 3 (Sa2)(the second column of Figure 3), it was observed that the hand spin caused the corners to be displaced (take a look at L, L', and L" in Figure 3 (Sa1, Sa2)).

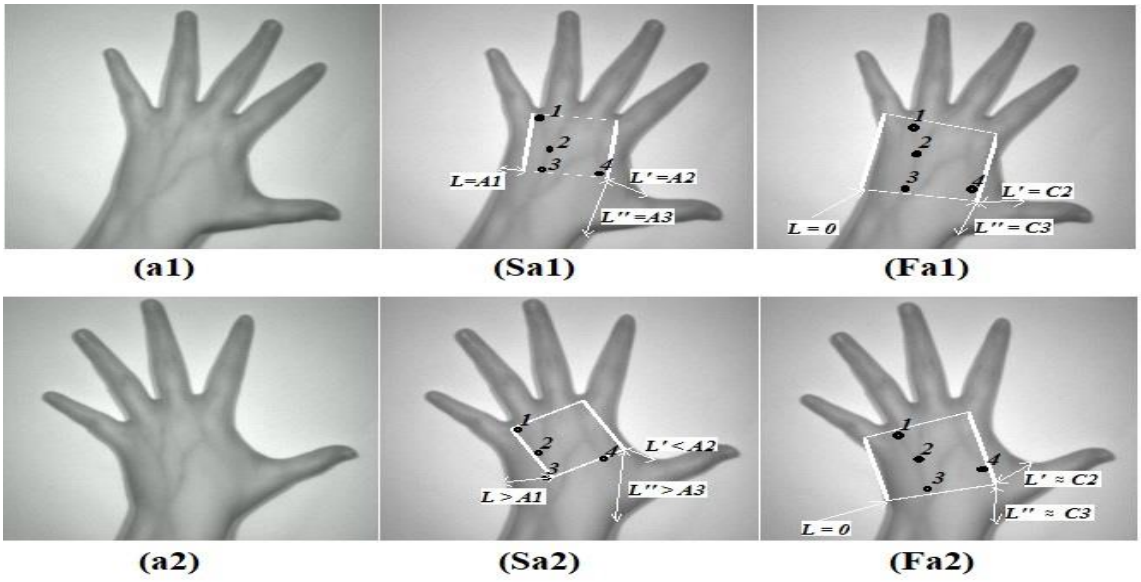

Figure 3. Comparing FROI) [10 and Stable ROI (SROI) [7]. The first column (a1, a2) shows the images of one hand in two situations of the hand spin. The second column illustrates them in which ROIs are extracted by the SROI technique, and the third column shows ROI extraction with FROI.

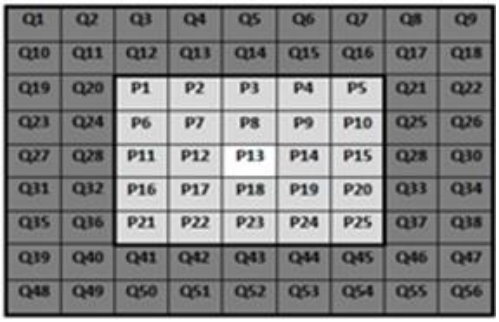

Figure 4. Designed mask for extracting the vein pattern from ROI.

Also, after the hand spin (Figure 3 (Sa2)), from the four specified points that we signed in ROI (Figure 3 (Sa1)), two points of them are located out of the region (see points 3 and 4 in Figure 3 (Sa2)). Therefore, the vein patterns in both ROIs (for Figure 3 (Sa1, Sa2)) will not be similar. However, in the third column (Figure 3 (Fa1, Fa2)), in both situations of the hand spin, the positions of the corners of the quadrilaterals are approximately constant, so all the four points in both images (Figure 3 (Fa1, Fa2)) are inside ROI, i.e. the veins located in both situations of the hand spin are similar. The FROI technique with its details is given by [10].

After selecting and extracting ROI, a mask designed in Figure 4 with Equation (21) is used in order to extract the vein pattern.

$$\begin{cases} P = \frac{1}{25}\sum_{i=1}^{25} P_i \\ Q = \frac{1}{56}\sum_{j=1}^{56} Q_j \end{cases} \quad (21)$$

This mask moves over each pixel belonging to the extracted ROI, and in any case, if $P > Q$, the pixel of ROI located under the central cell of the mask (P13) is recognized as a vein; otherwise, the pixel is considered as the background.

Since the width of the dorsal hand veins in the images is usually between 3 and 5 pixels, in order to find the best mask for extracting the skeleton of the veins from ROI, different modes of the two masks P and Q (in terms of size and proportion), to 50 Random ROIs are applied. Then the extraction rate of the vein pixels for each one of them is calculated manually.

In other words, the values 1, 3, 5, and 7 for the P mask and the values 3, 5, 7, 9, 11, and 13 for the Q mask are considered, and by combining them, 18 different masks are designed. Then each mask is applied to the 50 ROIs selected randomly from the database, and the extraction rate of the vein pixels for each ROI is calculated manually.

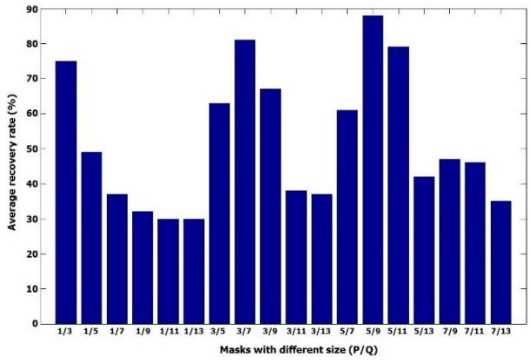

Figure 5. Performance of different masks to extract the vein pattern from ROIs. On the horizontal axis, the numerator represents the size of P mask and the denominator indicates the size of Q mask.

Figure 5 shows the average accuracy of 50 ROIs for each mask.

As it can be seen in this figure, the highest value is obtained in the proportion of 5 to 9 (i.e. *P*-mask with the size of 5×5 and *Q*-mask with the size of 9×9) with an approximate average recovery rate of 88%. Therefore, in this work, this mask (Figure 4) is used to extract the skeleton of the veins from ROIs.

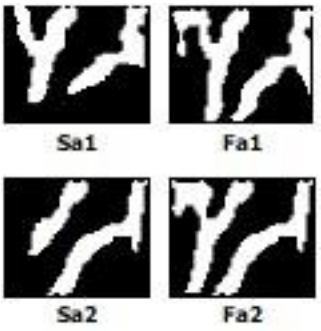

Figure 6. Veins extracted from ROIs of Figure 3. As it can be seen, the veins extracted by the proposed method (right column) are much more similar, due to the floating region based on the hand spin rate. However, the veins extracted by the method [7] (left column) are different due to the fixed ROI and not considering the hand spin.

Figure 6 shows a sample of veins extracted from the images of Figure 3. As expected by hand spin, the samples of veins extracted from the third column of Figure 3 (Fa1, Fa2) are very close to each other, whereas the veins extracted from the second column (Sa1 and Sa2) are not very similar. Thus using fixed ROI for all of the hand spins is not suitable for identification in the next step. Therefore, it can be concluded that if there is no restriction on the hand spin at the time of image acquisition, to extract a similar vein pattern of a person, the shape of ROI must be changed relative to the spin.

$$K = [k(.,a_{1,1}), k(.,a_{1,2}), \dots, k(.,a_{c,n_c})] \in \mathbb{R}^{n\times n}. \quad (26)$$

3.2. Comparison

After extracting the vein pattern for all images and converting to the same size images (in this work, 64*80 pixels), as mentioned in the previous

$$\min_w \|w\|_1 \quad S.t. \quad k(.,y) = Kw. \quad (27)$$

$$\Phi(a) = [\Phi_1(a), \Phi_1(a), \dots, \Phi_M(a)]^T \in \mathbb{R}^M, \quad (22)$$

section, we used the sparse representation with the kernel trick for vein pattern classification. To do this, all of the extracted vein images are converted to the vectors. Then the vectors are transmitted by a Φ mapping to a new higher-dimensional space.

$$\Phi(y) = w_{1,1}\times\Phi(a_{1,1}) + \dots + w_{c,n_c}\times\Phi(a_{c,n_c}) = \Phi(D)w \quad (23)$$

$$k(a_i, a_j) = \exp(-\gamma\|a_i - a_j\|_2^2) \quad (28)$$

so that $M \gg m$. Therefore, the training images represented in SRC with a_i will be here, shown as

$$\Phi^T(D)\Phi(y) = \Phi^T(D)\Phi(D)w \Rightarrow k(.,y) = Kw \quad (24)$$

$\Phi(a_i),\ i = 1,2,\dots,n$. Now, like the SRC method,

$$k(.,y) = [k(a_{1,1},y), k(a_{1,2},y), \dots, k(a_{c,n_c},y)]^T \in \mathbb{R}^n, \quad (25)$$

the test image can be represented as a linear combination of training images as follows:

By multiplying Φ^T(D) from the left and considering Equation (11), we can write:

where:

Now, if the equation constraint (9) is replaced in (24):

K is the kernel Gram matrix, which is symmetric and positive semi-definite, and $K_{ij} = k(a_i, a_j)$. Thus we can obtain K and $k(.,y)$ when given a kernel function $k(.,.)$. After various tests on different kernels, we found that the RBF kernel was more suitable. Therefore, the kernel function $k(a_i, a_j)$ could be expressed as:

where $\gamma = median(\frac{1}{\|a_p - \bar{a}\|_2^2})$ and $\bar{a} = \frac{1}{n}\sum_{p=1}^{n} a_p$.

Since Equation (27) is high dimensionality, we have to do a dimensional reduction. Thus a random matrix $B \in \mathbb{R}^{m'\times n}$ (generally, $m' \le n$) is multiplied on both sides of Equation (24) to convert Equation (27) into the following:

$$\min_w \|w\|_1 \quad S.t. \quad k'(.,y) = K'w, \quad (29)$$

where:

$$k'(.,y) = Bk(.,y) \in \mathbb{R}^{m'} \quad K' = BK \in \mathbb{R}^{m'\times n}. \quad (30)$$

For the similar equation (4) by considering the noise vector $\xi \in \mathbb{R}^{m'}$ with the energy $\|\xi\|_2 < \varepsilon$, the following relation will result:

$$\min_w \|w\|_1 \quad S.t. \quad \|k'(.,y) - K'w\|_2 \le \varepsilon. \quad (31)$$

Finally, by solving the dimensional-reduced Equation (29) or (31), similar to the SRC method and according to Equation (32), the class of each test sample can be determined.

$$\hat{C} = \min_{i=1,2,\dots,c} \|k'(.,y) - K'\hat{w}_i\|_2, \quad (32)$$

where $k'(.,y)$ is the sample of the query test in the kernel space after the dimensional-reduction, whose class should be determined, $\hat{w}_i$ is a vector by the same size w, and only the entries of the ith class are equal to the entries of w, and the rest of them are zero. It should be noted that the vector w is obtained from problem-solving (29).

A summary of the proposed method for dorsal hand vein recognition is given in Algorithm 1.

Algorithm 1: Proposed algorithm

1- Prepare the images according to the sub-section 3-1 and relations (16)- (18).
2- ROI selection and vein pattern extraction using the relations (19)-(21).
3- Convert all the extracted vein pattern images to vector.
4- Select two-third of the vectors as the training samples and one-third of them as the test samples, randomly.
5- Select a Mercer kernel $k(.,.)$ and its parameters (in our work, we used the RBF kernel). Then compute the matrix K and vector $k(.,y)$ in (24), (25) and (26).
6- Select a projection method and get the corresponding matrix$B$. Then compute K' and $k'(.,y)$ by (30).
7- Normalize the columns of K' and $k'(.,y)$ to have unit ℓ_2 −norm.
8- Solve the ℓ_1 −minimization problem (29) or (31) to get the coefficient vector w.
9- Find the class of the test sample by (32).

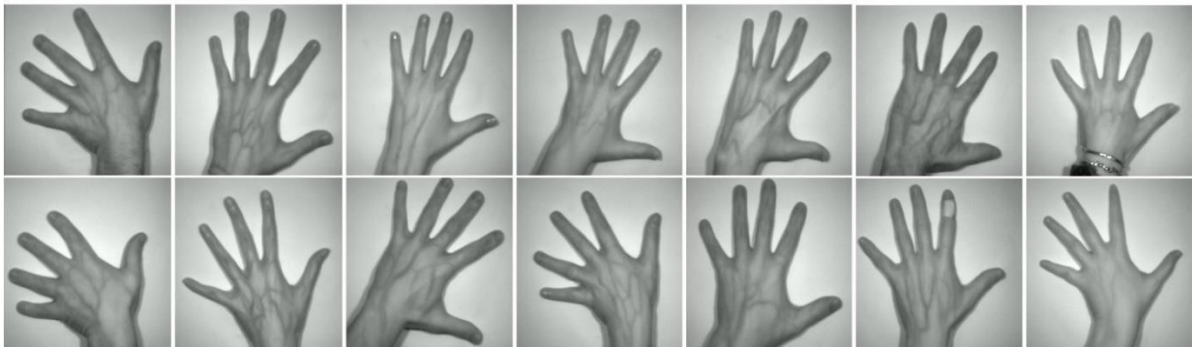

Figure 7. Some image samples of Bogazici database.

4. Experimental Results

As mentioned earlier, for testing the proposed method, we used 1200 infrared images belonging to the left hand of 100 people of the Bogazici database because they did not have any wrist spin restriction [34]. These images captured on the four situations (normal situation, after the squeezing an elastic ball, repetitively, after carrying a 3 kg bag and after holding an ice pack for one minute).

Figure 7 shows some sample images of this database. In order to make the dictionary, two-third of the total database images are randomly selected as training (800 images) and the rest of them are used as the test images (400 images). As noted in the previous sections, the proposed method consists of two parts: ROI extraction and classification by sparse coding and kernel trick. Therefore, in order to show the efficiency of the proposed method, the comparisons will also be included in two parts.

In the ROI extraction part, FROI [10] is compared with the method of [7] as a famous technique in fixed ROI extraction techniques, which will be from now and is called stable ROI (SROI). Then in the classification part (finding the class of a query image), given that the proposed method is an optimized model of SRC, first, the results of the proposed method are compared with the results of similar methods such as conventional SRC [23], mutual SRC (MSRC) [3], and genetic algorithm SRC (GASRC) [35]. After that, since the proposed method is a classifier-based sparse representation, we compared our results by three classification method-based sparse coding that used dictionary learning such as discriminative Fisher embedding dictionary learning (DFEDL) [20] and Fisher discriminative KSVD (FD-KSVD) [19], and Dictionary Pair Learning (DPL) [17].

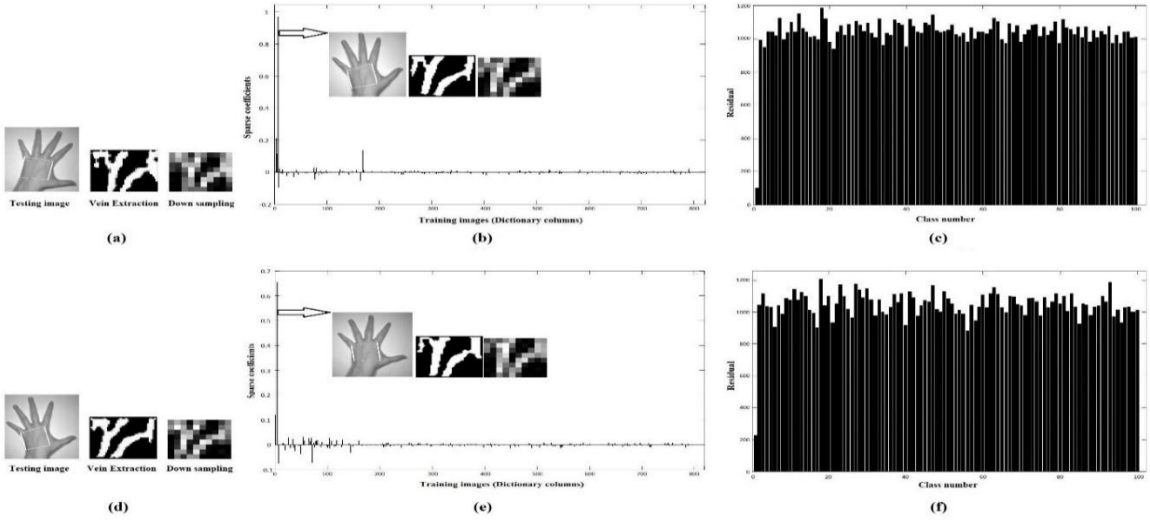

Figure 8. Two valid test images with algorithm 1.
(a) A test image belongs to subject 1 with its vein pattern and 8×10 downsampled image as features. (b) Sparse coefficients. The training sample corresponding to the largest coefficients is illustrated. (c) The residuals of the test image. As it can be seen, using relation (31), the lowest residual value is achieved for the first class. Therefore, this sample is recognized in the first class. (d) Another testing image belongs to subject 1 with different wrist spins. (e) Sparse coefficients along with a similar training sample that corresponds to the largest coefficients. (f) The residuals of the testing image. As shown, this sample is also recognized in the first class. Thus by changing the shape of ROI, we could get rid of the destructive effects of the wrist spin.

On the other hand, both the ROI extraction and classification parts are complementary, i.e. the classification operation must be performed on the ROI extracted from the database images. Thus in order to make a fair comparison, at first, we extracted ROI of the database images by both the FROI and SROI methods.

After that, we applied these methods (SRC, MSRC, GASRC, DFEDL, FD-KSVD, DPL, and the proposed method) once on FROIs and again on SROIs, and recorded the results. We evaluated the experimental results based on the recognition rate. A platform with an Intel Core i7-7500U CPU and 8.0 GB RAM was used to run the experimental analysis by the Matlab 2013a software.

4.1. An Example of Comparison of SROI with FROI

In this section, in order to show the performance of the proposed method in the ROI extraction part as well as the effects of hand rotation in the final recognition result, two test images (with different wrist spins) belonging to the first person (first-class) of the database are used. (These images were previously used in Figure 3 and Figure 6).

As mentioned earlier, after extracting ROIs (with both the FROI and SROI techniques) and extracting the vein patterns from them, all the images of the vein patterns will convert to 64×80 pixels (the nearest size to all ROIs extracted). Therefore, after vectorization, we get 5120 features. Then by sub-sampling, they convert to size 8×10 pixels, and at the end, after vectorization, we have 80-D vectors. Therefore, according to the number of training images (800), by using their vectorized form, the matrix D (dictionary) will be formed with 800 columns and 80 rows, every 8 columns belonging to one class. Hence, the system $\boldsymbol{y} = \mathbf{Dx}$ is underdetermined, and according to the relation (1), m = 80, and n = 800. Figure 8 shows the sparse coefficients recovered by algorithm 1 for the test images in the second column of Figure 6 (Fa1, Fa2) along with the residuals respect to the 100 classes. It has previously been mentioned that these are vein patterns extracted from the third column images of Figure 3 (Fa1, Fa2), where ROIs are extracted with the FROI technique). As illustrated in Figure 8, the largest coefficients are associated with the training samples from subject 1 (Figure 8 (b, e)). Thus despite the different wrist spins for the two test samples (Figure 8 (a, d)), given the less residual (Figure 8 (c, f)), both of the test samples are correctly recognized in class 1.

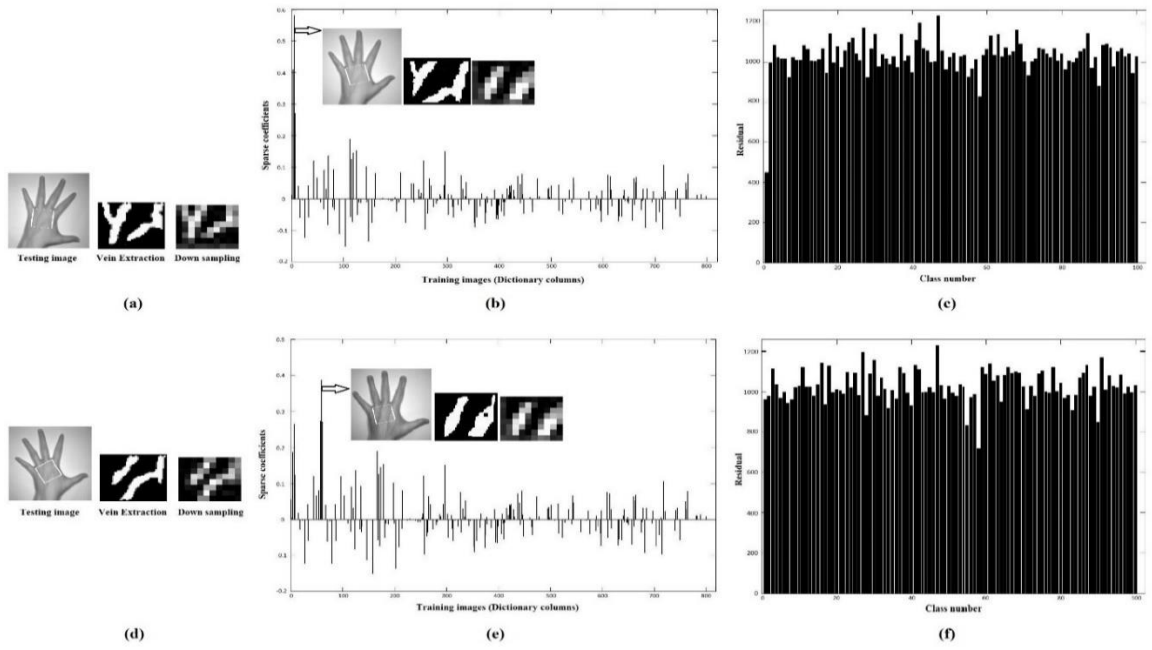

Figure 9. Valid and invalid test images based on SROI extraction. (a) A test image belongs to subject 1 with its vein pattern and 8×10 down-sampled image as the features. (b) Sparse coefficients. The training sample of the largest coefficient is illustrated. (c) Residuals of the test image. As it can be seen, due to the constant as well as small ROI, the major vein patterns are lost. However, due to the little wrist spin, the class of the test image has been correctly recognized, although in comparison to Figure 8, the residual of the first class is increased. (d) Another testing image belongs to subject 1 with different wrist spins. (e) Sparse coefficients along with a wrong recognized training sample that corresponds to the largest coefficients. (f) Residuals of the test images. As shown, due to the large wrist spin, this sample is wrongfully recognized in the 58th class.

On the other hand, in Figure 9, the sparse coefficients and the residuals for the same test images of Figure 6 are shown. However, the difference is that the ROI extraction technique is SROI [7]. As illustrated in Figure 9, for the first test image (Figure 9 (a)), due to the little wrist spin, the class has been correctly recognized. However, due to the small and constant ROI, the major vein patterns are lost. Thus compared to Figure 8, the residual elevation of the first class is increased. Likewise, for the second test image (Figure 9 (d)), due to the large wrist spin, this sample is recognized in a wrong class. We cannot use a constant ROI when we have various large wrist spin.

4.2. Comparison of Proposed Method with SRC-based Classifiers

In this section, a comparative analysis of the proposed method with some of the reported SRC-based algorithms such as conventional SRC [23], Genetic Algorithm SRC (GASRC) [35], and Mutual SRC (MSRC) [3] is provided. If there is no other statement, we set the parameters of these methods according to the following description. For the conventional SRC and MSRC, the $\ell_1 -$ norm minimization problem is solved by exploiting the $\ell_1 -$ MAGIC software package and let $\varepsilon = 0.001$. In GASRC, the generation is set to 50, and the number of individuals is set to 20. Also for GASRC, the ratio of sample number in the representation set to the total training samples is set to 0.1.

As mentioned earlier, after vein pattern extraction (using the FROI or SROI techniques), all the images of the vein patterns will be converted to 64×80 pixels (the nearest size to all ROIs extracted). For each method, we compute the recognition rates with the feature space dimensions 30, 80, and 320. These numbers correspond to the down-sampling ratios of 1/12, 1/8, and 1/4, respectively. Table 1 shows the numerical values of the recognition rate for all the SRC-based methods in different dimensions when we use the FROI technique for the ROI extraction part. The bold italics highlight the best recognition rate on the training subset. We can see that the classification performance of the proposed method is better than the other four methods.

Also, as mentioned in the previous section, in order to illustrate the importance and effect of the ROI extraction part in the final result and the aftermath of wrist spin when used fixed ROI, we repeated the earlier trials for ROI extracted by the SROI technique.

Table 1. Recognition rate (%) of four methods in different dimensions while we use the FROI technique in the ROI extraction part.

Name of method	Feature sub-space dimension		
	30	**80**	**320**
SRC	71.01	81.23	85.01
MSRC	81.28	89.52	93.99
GASRC	72.25	85.12	91.02
Our method	**85.38**	**95.73**	**96.25**

Table 2 shows their result in different dimensions. Comparing to Table 1, we notice a decrease in the recognition rate for all methods. As described in the previous section, this is for losing a huge part of the vein patterns by cause of the wrist spin. For the same reason, we should not use the fixed ROI when the images have various wrist spins.

Table 2. Recognition rate (%) of four methods in different dimensions while we use the SROI technique in the ROI extraction part.

Name of method	Feature sub-space dimension		
	30	**80**	**320**
SRC	52.38	64.8	66.14
MSRC	59.12	68.13	69.04
GASRC	53.02	66.55	68.36
Our method	**63.2**	**71.32**	**73.11**

4.3. Comparison between Proposed Method and Classifiers based on Dictionary Learning

In this section, the results of the proposed method are compared with three classification methods that use dictionary learning for making the dictionary, e.g. Fisher Discriminative KSVD (FD-KSVD) [19], Dictionary Pair Learning (DPL) [17], and discriminative Fisher embedding dictionary learning (DFEDL) [20]. Similar to the previous section, these methods are applied to the ROIs extracted by the FROI and SROI techniques. In order to form the dictionary of these methods, dictionary learning was performed with the feature subspace dimensions of 30, 80, and 320. It is worth noting that for a fair comparison, similar to our dictionary, during the dictionary learning, the number of atoms for each class was considered 8 atoms.

Table 3 shows the numerical values of the recognition rate for four methods in different dimensions when we used the FROI technique, and Table 4 illustrates them for using the SROI technique in the ROI extraction part. Like before and as expected, due to the use of floating ROI instead of fixed ROI, the results of Table 3 got better than Table 4. Also as we can see in Table 3, the recognition rate of the proposed method is

better than the other three methods at 80D and 320D feature spaces. The bold italics highlight the best recognition rate on the training subset.

Table 3. Recognition rate (%) of four methods in different dimensions while we use the FROI technique in the ROI extraction part.

Name of method	Feature sub-space dimension		
	30	**80**	**320**
FD-KSVD	85.2	89.95	94.32
DPL	83.12	89.25	92.41
DFEDL	**88.21**	93.4	95.78
Our method	86.38	**95.73**	**96.25**

Table 4. Recognition rate (%) of four methods in different dimensions while we use the SROI technique in the ROI extraction part.

Name of method	Feature sub-space dimension		
	30	**80**	**320**
FD-KSVD	64.43	69.83	70.03
DPL	61.85	68.18	69.09
DFEDL	**69.13**	**72.46**	72.95
Our method	63.2	71.32	**73.11**

5. Conclusions

In this work, a robust and efficient algorithm against rotation to identify through the dorsal hand vein pattern was proposed. One of the important issues of a dorsal hand vein identification system is hand rotation and displacement because it causes difficulty to select identical and fixed ROI in all different modes of hand rotation. For this, the method proposed in the previous work [10], called FROI, was used in order to neutralize the adverse effects of hand rotation by varying the length and angle between the sides. Therefore, unlike other techniques, the shape of ROI is not fixed and the length and angle of the sides will be changed depending on the angle of the wrist spin. Note that since the dorsal hand vein pattern is not symmetrical, symmetry is not required for the ROI shape. For classification, since the dorsal hand veins pattern belonging to different classes (different people) have the same orientation, the kernel trick and sparse representation are used because the transmission of feature vectors to higher dimensions makes them more distinct. Thus the proposed algorithm consists of two parts: first, we extract a similar region from all images belonging to a class. Then by combination of the sparse representation and kernel trick, we find the distinctive sparse signal representation, which will be used for classification. We used the Bogazici University database in order to test the proposed method because there are no restrictions of hand rotation on the images. In order to better compare with the previous work, after extracting the vein patterns, the comparison is performed on three different dimensions of the feature space. The results obtained showed that the proposed method performed better in two dimensions than the other methods, and achieved a recognition rate of 96.25% at the 320D feature space.

References

[1] D. Huang, X. Zhu, Y. Wang, and D. Zhang, "Dorsal hand vein recognition via hierarchical combination of texture and shape clues," *Neurocomputing*, vol. 214, pp. 815-828, 2016.

[2] S. Joardar, A. Chatterjee, and A. Rakshit, "A real-time palm dorsa subcutaneous vein pattern recognition system using collaborative representation-based classification," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, pp. 959-966, 2014.

[3] S. Shazeeda and B. A. Rosdi, "Finger vein recognition using mutual sparse representation classification," *IET Biometrics*, vol. 8, pp. 49-58, 2019.

[4] H. Wan, L. Chen, H. Song, and J. Yang, "Dorsal hand vein recognition based on convolutional neural networks," in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017, pp. 1215-1221.

[5] D. Huang, Y. Tang, Y. Wang, L. Chen, and Y. Wang, "Hand-dorsa vein recognition by matching local features of multisource keypoints," *IEEE transactions on cybernetics*, vol. 45, pp. 1823-1837, 2014.

[6] A. M. Badawi, "Hand Vein Biometric Verification Prototype: A Testing Performance and Patterns Similarity," *IPCV*, vol. 14, pp. 3-9, 2006.

[7] C.-L. Lin, S.-H. Wang, H.-Y. Cheng, K.-C. Fan, W.-L. Hsu, and C.-R. Lai, "Bimodal biometric verification using the fusion of palmprint and infrared palm-dorsum vein images," *Sensors*, vol. 15, pp. 31339-31361, 2015.

[8] W.-Y. Han and J.-C. Lee, "Palm vein recognition using adaptive Gabor filter," *Expert Systems with Applications*, vol. 39, pp. 13225-13234, 2012.

[9] I. I. Morales-Montiel, J. A. Olvera-López, and I. O. Pineda, "An Image Rotation Approach for Hand Dorsal Vein Recognition," *Res. Comput. Sci.*, vol. 99, pp. 105-113, 2015.

[10] A. N. Pour, E. Eslami, and J. Haddadnia, "A new method for automatic extraction of region of interest from infrared images of dorsal hand vein pattern based on floating selection model," *International Journal of Applied Pattern Recognition*, vol. 2, pp. 111-127, 2015.

[11] K. Engan, S. O. Aase, and J. H. Husoy, "Method of optimal directions for frame design," in *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258)*, 1999, pp. 2443-2446.

[12] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries

for sparse representation," *IEEE Transactions on signal processing*, vol. 54, pp. 4311-4322, 2006.

[13] A. Sahoo and P. Das, "Dictionary Based Intra Prediction for Image Compression," in *RICE*, 2017, pp. 73-76.

[14] N. Zhou, H. Jiang, L. Gong, and X. Xie, "Double-image compression and encryption algorithm based on co-sparse representation and random pixel exchanging," *Optics and Lasers in Engineering*, vol. 110, pp. 72-79, 2018.

[15] J.-C. Ni, Q. Zhang, Y. Luo, and L. Sun, "Compressed sensing SAR imaging based on centralized sparse representation," *IEEE Sensors Journal*, vol. 18, pp. 4920-4932, 2018.

[16] E. Miandji, S. Hajisharif, and J. Unger, "A unified framework for compression and compressed sensing of light fields and light field videos," *ACM Transactions on Graphics (TOG)*, vol. 38, pp. 1-18, 2019.

[17] X. Wang, H. Li, J. Qiu, and C. Yu, "Palm vein recognition based on competitive code and DPL," in *Proceedings of the 3rd International Conference on Cryptography, Security and Privacy*, 2019, pp. 179-183.

[18] F. Mokhayeri and E. Granger, "A paired sparse representation model for robust face recognition from a single sample," *Pattern Recognition*, vol. 100, p. 107129, 2020.

[19] H. Zheng and D. Tao, "Discriminative dictionary learning via Fisher discrimination K-SVD algorithm," *Neurocomputing*, vol. 162, pp. 9-15, 2015.

[20] Z. Li, Z. Zhang, J. Qin, Z. Zhang, and L. Shao, "Discriminative fisher embedding dictionary learning algorithm for object recognition," *IEEE transactions on neural networks and learning systems*, vol. 31, pp. 786-800, 2019.

[21] S. Hu, C. Xu, J. Peng, Y. Xu, and L. Tian, "Weighted kernel joint sparse representation for hyperspectral image classification," *IET Image Processing*, vol. 13, pp. 254-260, 2019.

[22] Y. Peng, L. Li, S. Liu, J. Li, and H. Cao, "Virtual samples and sparse representation-based classification algorithm for face recognition," *IET Computer Vision*, vol. 13, pp. 172-177, 2019.

[23] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, pp. 210-227, 2008.

[24] L. Zhang, W.-D. Zhou, P.-C. Chang, J. Liu, Z. Yan, T. Wang, *et al.*, "Kernel sparse representation-based classifier," *IEEE Transactions on Signal Processing*, vol. 60, pp. 1684-1695, 2011.

[25] S. Mavaddati, "A novel face detection method based on over-complete incoherent dictionary learning," *Journal of AI and Data Mining*, vol. 7, pp. 263-278, 2019.

[26] G. Ma, T.-Z. Huang, J. Huang, and C.-C. Zheng, "Local low-rank and sparse representation for hyperspectral image denoising," *IEEE Access*, vol. 7, pp. 79850-79865, 2019

[27] X. Pu, Z. Li, B. Li, H. Lei, W. Gao, and J. Liu, "Sparse representation based medical ultrasound images denoising with reshaped-RED," in *Eleventh International Conference on Digital Image Processing (ICDIP 2019)*, 2019, p. 111790K.

[28] K. Huang and S. Aviyente, "Sparse representation for signal classification," *Advances in neural information processing systems*, vol. 19, pp. 609-616, 2006.

[29] C. F. Dantas, J. E. Cohen, and R. Gribonval, "Hyperspectral Image Denoising using Dictionary Learning," in *2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2019, pp. 1-5.

[30] H. Wang, P. Wang, L. Song, B. Ren, and L. Cui, "A novel feature enhancement method based on improved constraint model of online dictionary learning," *IEEE Access*, vol. 7, pp. 17599-17607, 2019.

[31] Q. Zhang and B. Li, "Discriminative K-SVD for dictionary learning in face recognition," in *2010 IEEE computer society conference on computer vision and pattern recognition*, 2010, pp. 2691-2698.

[32] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Fisher discrimination dictionary learning for sparse representation," in *2011 International Conference on Computer Vision*, 2011, pp. 543-550.

[33] A. Güven, K. Polat, S. Kara, and S. Güneş, "The effect of generalized discriminate analysis (GDA) to the classification of optic nerve disease from VEP signals," *Computers in Biology and medicine*, vol. 38, pp. 62-68, 2008.

[34] A. Yüksel, L. Akarun, and B. Sankur, "Biometric identification through hand vein patterns," in *2010 International Workshop on Emerging Techniques and Challenges for Hand-Based Biometrics*, 2010, pp. 1-6.

[35] Z. Fan, M. Ni, Q. Zhu, C. Sun, and L. Kang, "L0-norm sparse representation based on modified genetic algorithm for face recognition," *Journal of Visual Communication and Image Representation*, vol. 28, pp. 15-20, 2015.

روشی مقاوم در برابر چرخش دست جهت تشخیص هویت بر اساس نمایش تنک کرنلی

علی نوذری پور و هادی سلطانی زاده *

دانشکده مهندسی برق و کامپیوتر دانشگاه سمنان، سمنان، ایران.

ارسال ۲۰۲۰/۱۱/۱۵؛ بازنگری ۲۰۲۱/۰۵/۰۷؛ پذیرش ۲۰۲۱/۰۶/۱۱

چکیده:

نمایش تنک به دلیل مزایایی مانند مقاوم بودن در برابر نویز و داشتن پایه ریاضی قوی، به عنوان ابزاری قدرتمند در دهه های اخیر مورد توجه قرار گرفته است. در این کار، با استفاده از نمایش تنک، ترفند کرنل (هسته) و تکنیکی متفاوت جهت استخراج ROI که در کار قبلی ما ارائه شده است، روشی جدید جهت تشخیص هویت بر اساس الگوی رگ‌های پشت دست افراد و مقاوم در برابر چرخش دست معرفی شده است. در این روش، برای انتخاب ROI، با تغییر طول و زاویه بین اضلاع، اثرات نامطلوب چرخش دست که در هنگام تصویر برداری ایجاد می‌شود، تا حد زیادی خنثی شده است. بنابراین بسته به مقدار چرخش دست در هر تصویر، شکل و اندازه ROI متغیر خواهد بود. از سویی دیگر، به دلیل امکان وجود توزیع جهت یکسان الگوهای رگ در افراد مختلف، جهت طبقه بندی از ترفند کرنل (هسته) در نمایش تنک استفاده کرده‌ایم. با اینکار، بیشتر نمونه‌های موجود مربوط به کلاس‌های مختلف (الگوی رگهای پشت دست مربوط به افراد مختلف) که دارای توزیع جهت یکسان هستند به درستی و در کلاس مربوط به خود تشخیص داده می‌شوند. استفاده از این دو روش منجر به معرفی روشی مقاوم در برابر چرخش دست جهت تشخیص هویت افراد بر اساس الگوی رگهای پشت دست شده است. نتایج آزمایشات نشان می‌دهد که روش پیشنهادی در مقایسه با سه الگوریتم مبتنی بر SRC و سه روش طبقه بندی مبتنی بر کدگذاری تنک با استفاده از یادگیری دیکشنری، موفق به افزایش نرخ میزان تشخیص تا حدود ۲٫۲۶٪ شده است.

کلمات کلیدی: نمایش تنک، ترفند کرنل، الگوی رگهای پشت دست، ناحیه مطلوب متغیر، طبقه بندی.