



# RIMFRA: Rotation-invariant multi-spectral facial recognition approach by using orthogonal polynomials

Taner Cevik<sup>1</sup> · Nazife Cevik<sup>2</sup>

Received: 27 September 2018 / Revised: 15 March 2019 / Accepted: 21 May 2019 /

Published online: 11 June 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

This paper proposes a novel rotation-invariant multi-spectral facial recognition approach (RIMFRA) by using orthogonal polynomials. In the first step, a rotation, illumination and noise invariant local descriptor (RinLd) is proposed to represent the texture patterns of a face image. Color channels of the images embodies non-trivial information about the characteristic of the image. Hence, the local descriptor matrices are extracted among the color channels. The corresponding new descriptor matrices for the red, green and blue channels of the image are extracted. Afterwards, co-occurrence matrices are obtained from the six combinations of the corresponding color channel descriptor matrices, that are red-red, blue-blue, green-green, red-blue, green-blue and red-green. Finally, these matrices are decomposed by using the orthogonal polynomials to achieve a more reliable and characteristic pattern extraction. The coefficients obtained as a result of the decomposition process are used as the ultimate features for the classification of the images. Extensive simulations are conducted over benchmark datasets. As presented by the simulation results, the ultimate features yield very high discriminating performance as well as providing resistance to rotation and illumination variations.

**Keywords** Facial recognition · rotation invariant · multi-spectral · orthogonal polynomial

## 1 Introduction

Biometry has attracted a great deal of attention in recent years and has been widely used for its high performance in many areas such as surveillance, identification, and human-computer interaction [5, 15, 18, 30, 31, 34, 40, 50, 58, 60, 71]. Individuals have biological characteristics, also called metrics, that distinguish them from others [27]. Extracting behavioral and/or physiological characteristics of individuals to make discrimination is called as biometrics

---

✉ Taner Cevik  
tanercevik@aydin.edu.tr

<sup>1</sup> Department of Software Engineering, Istanbul Aydin University, Istanbul, Turkey

<sup>2</sup> Department of Computer Engineering, Istanbul Arel University, Istanbul, Turkey

recognition. Face, iris, retina, ear, palm are the prominent common discriminative physiological characteristics. Besides, voice, typing rhythm and gait are the behavioral characteristics, which are called as biometrics [48].

Face is one of the leading biometrics preferred for individual discrimination because it can distinguish individuals with high accuracy and less human participation. Face data can be easily collected and processed in real time using remote devices such as cameras without the need for any human intervention [14, 26].

As with many other images, face data are also exposed to disruptive external factors such as noise, illumination, pose variations and rotation. Variations in the pose, illumination, direction and the presence of random noise inhibit a pixel-by-pixel comparison among the images. Therefore, facial recognition has attracted great interest from researchers to overcome the above mentioned challenges. At the point where pixel-to-pixel comparison does not perform well, texture helps in image classification. Texture plays a key role in computer pattern recognition, especially in image related applications [32, 33]. Besides Although not a globally accepted definition, texture can be defined as the result of recurring local patterns throughout the picture [47]. As with other types of images, there is also a texture in the face images. Features are extracted from the texture of the face images and then analyzed and classified to distinguish individuals. To qualify a feature set as high quality, it must provide two criteria. The first is that the need for computer processing complexity should be low, so it can be used in real-time applications. The other criterion is that it should be able to express the properties of the texture in the best possible way so that it can do the splitting well during classification between the textures [41].

Numerous studies have been conducted to suggest a high-performance identifier for low computational complexity and high representational power. These methods can basically be grouped under two headings: holistic and local appearance features [39]. Holistic techniques analyze the entire face image and extract global information to recognize a subject. This global information is obtained by analyzing pixel relationships along the entire image and corresponding features are extracted. These features represent the global characteristic of the image that uniquely discriminates the face from others [64]. The most well-known of the holistic approaches are Principal Component Analysis (PCA) [65], Linear Discriminant Analysis (LDA) [7] and Independent Component Analysis (ICA) [10]. The hallmark of PCA is that it reduces dimensionality by transition from a high dimensional image space to a low dimensional orthogonal space. The transition is performed by considering the lowest mean square reconstruction error and also by applying a linear transformation. LDA focuses on finding linear transformation that maximizes inter-class variance and minimizes intra-class variation. ICA, first adopted by Herault and Jutten [69], seeks for a linear transformation to minimize the statistical dependencies of the components of a vector. Many of the following studies [22, 23, 28, 29, 54, 56, 63, 72] are based on these fundamental methods and have struggled to improve their performance by introducing new ideas among them.

Local approaches, unlike the holistic ones, reveal local distinctive features that are more resistant to changes in expression and illumination. In this aim, several research studies (LBP [2, 59], LPQ [73], LDP [25], LDN [52, 53], HoG [11], LTP [62], Gabor [43, 74]) have been performed to satisfy the task of obtaining a high level of distinctive and proper texture representation [8]. Among these, LBP has been a promising pioneer for follow-up studies due to its high performance and calculation efficiency [12]. LBP identifies local textures by comparing each pixel with its  $3 \times 3$  local neighborhood. Each pixel is then replaced by the

eight-bit-stream-result of the comparison step. Each bit in the bit stream represents the magnitude-comparison-result of the corresponding neighbor to the reference pixel. If the intensity value of the reference pixel is less than the neighboring pixel, the corresponding bit is assigned to zero, otherwise assigned to one.

One of the most obvious concerns in facial recognition is undoubtedly the failure of the proposed feature descriptors when the image is rotated. A robust descriptor, regardless of local or holistic, should work independently of the direction of the image, i.e. reflect the same image characteristics in all conditions. As is known, basic LBP does not consider rotational variations, thus, a number of follower improvements [70, 78] have been proposed to gain resistance to rotational variations. Furthermore, color channels contain significant information yet a great deal of the studies to date have derived characteristics from the monochrome images. In this paper, we propose a compound method that blends three main distinguishing subjects. First, a rotationally invariant local identifier is proposed which is also resistant to variations of light and facial expression. Second, the power of the color channels is used to investigate the statistical properties of matrices created by taking into account the occurrences of the local identifier in the multi-spectral area. Finally, the information stored in the multi-spectral occurrence matrices is represented by the orthogonal polynomial coefficients to reinforce the discriminative power of the proposed method.

The rest of this paper is organized as follows. Section II briefly describes the proposed method and gives basic information about pioneering ideas. Section III shows the results of the simulations and also refers to the discussions. Finally, Section IV completes the paper.

## 2 RIMFRA

This section describes the details about the proposed method in detail. The main steps of the general process are given in the following. At the outset, the multi-spectral rotation invariant local descriptor matrices are calculated from the RGB bands of the raw image. After the formation of the descriptor matrices, multi-spectral co-occurrence matrices are calculated. In the last step, orthogonal polynomial coefficients are obtained from each co-occurrence matrix. Finally, the coefficients obtained from the co-occurrence matrices are concatenated to form the ultimate feature vector for each facial image. Figure 1 depicts the operation of the complete process.

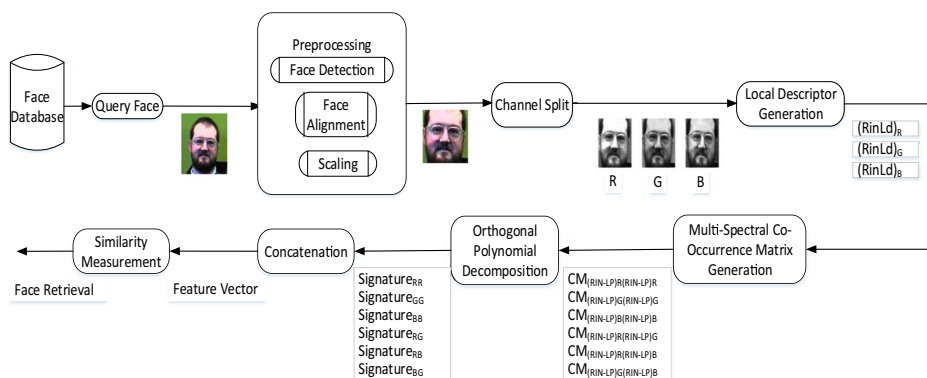


Fig. 1 The block diagram of the entire process

### 2.1 RinLd

The local texture descriptors, the prominent ones of which have been mentioned in the previous section, have provided promising discriminatory performances. However, the two of the most critical issues expected from these descriptors are that they should be rotationally invariant and resistant to changes that may occur in illumination. As mentioned earlier, LBP is one of the basic and leading local descriptors that figures out the local structure of the images. As is known, LBP defines the relationship between the central pixel and its neighboring pixels in an  $N \times N$  block ( $N$  indicates the width and length of the block). However, the initial LBP does not concern with the rotational variations. That is, the value of the descriptor calculated from a sub-portion of the image changes when the image is rotated. Another deficiency of LBP is that it does not consider the intensity value of the central or reference pixel. Therefore, it is possible for some pixels with different intensity values to be represented by identical values in the new domain. The state of this undesirable identical representation of the different pixels may be fixed by taking into account the intensity of the reference pixel.

In this paper, we propose a new local descriptor that is resistant to rotational and illuminative variations. Instead of working on monochrome images, the method we offer here operates on RGB images. The three color bands of the image are divided into  $N \times N$  blocks. Subsequently, the adjacent pixels of any reference pixel are sorted on a vector in descending order relative to their density values, as shown below:

$$SI_{N \times N} = \text{sort}_{dsc}(I_{N \times N}) \tag{1}$$

where  $SI_{N \times N}$  and  $I_{N \times N}$  express the sorted and unsorted neighboring pixels' intensity values respectively. The intensity value of the reference pixel is subtracted from each element of the sorted vector. If the absolute value of the result is greater than the threshold value ( $T$ ), a  $1$ , otherwise a  $0$ , is assigned to the corresponding position of a new binary vector.

$$B_{I_c} \begin{cases} 1 & \text{if } |SI_i - I_c| > T, i = 1, 2, \dots, N-1 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The threshold value is not held constant throughout the image, in the contrary, is dynamic and depends on the mean intensity value in the block. The intensity of the reference pixel is also taken into account while calculating the mean value in the block.  $T$  is calculated as follows:

$$T = |I_c - \frac{\sum_{i=1}^N SI_i}{N}| \tag{3}$$

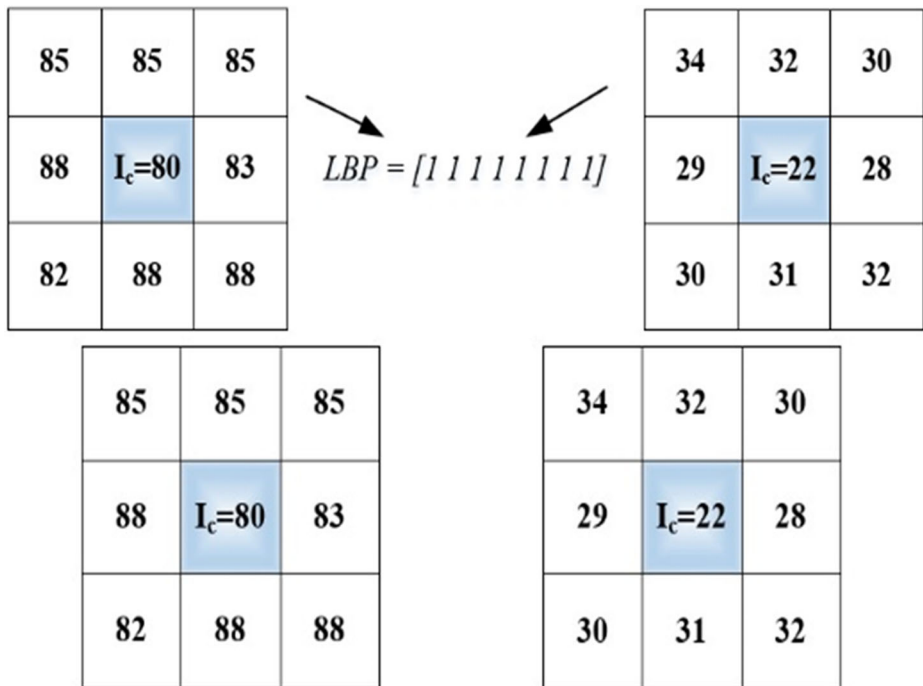
The resulting binary vector represents the comparison between the reference pixel and its neighbors in terms of intensity levels yet it does not solve the challenge of multiple pixels having different intensity values to be represented with identical values in the new domain.

Thus, the resulting value is recalculated by taking into account the intensity value of the reference pixel as follows:

$$SB_{I_c} = B_{I_c} \times (I_c/255) \tag{4}$$

The basic LBP and some of its derivatives do only consider the relationship between the reference pixel and its neighbors. However, the information concealed in the magnitude of the difference is being discarded in this way. Because of that, it is possible to encounter the challenge of two pixels with different intensities having identical values in the new domain. The most competent way to address this situation is to take into account the intensity value of the reference pixel. RIMFRA remedies the expressed challenge in two separate steps by calculating the value scaled according to the intensity value of the reference pixel while keeping the threshold value dynamic. Figure 2 illustrates the challenge of having the same binary patterns for pixels with different intensities and how RIMFRA handles this situation.

For each image, a total of three local descriptor matrices ((RinLd)<sub>R</sub>, (RinLd)<sub>G</sub>, and (RinLd)<sub>B</sub>) are created, one for each color band of the image. Following this stage, the constructed matrices are fed into the next step in the process described in the next section.



$$SI_{NxN} = [88\ 88\ 88\ 85\ 85\ 85\ 83\ 82] \quad SI_{NxN} = [34\ 32\ 32\ 31\ 30\ 30\ 29\ 28]$$

$$m = 85.5 \quad T = 5.5 \quad m = 30.5 \quad T = 8.75$$

$$B_{I_c} = (111100000) = 224, \quad SB_{I_c} = 71 \quad B_{I_c} = (1111100000) = 240, \quad SB_{I_c} = 21$$

Fig. 2 a Identical LBP codes assigned to different patterns b RIMFRA overcomes the miss-assignment situation

## 2.2 Multi-spectral co-occurrence matrices

The modern image acquisition and processing systems are capable of expressing and operating on colors in different spaces, namely RGB (Red, Green, Blue), HSV (Hue, Saturation, Value) and CIE Lab. Commonly, color images are represented as RGB. In fact, the information carried in the image is the brightness level of each band. Although many applications require RGB to be converted to other color domains, such as HSV or others, RGB based computer vision systems are simpler and more economical than others. Although the RGB color space is machine-dependent, which seriously disrupts uniformity [49], it still performs very successfully in areas such as calibration and classification [24]. For example, researchers have analyzed the images of apples and have successfully estimated the amount of fruit they contain [61]. In addition, some researchers have correctly predicted some of the geometric properties of different crop species by applying RGB-based image processing techniques [19, 35, 37, 45, 46, 68]. Moreover, it has been verified that methods based on color analysis are a reliable way of discrimination and retrieval in facial detection and monitoring. Furthermore, although humans vary according to skin color, the main distinguishing parameter has been shown to be density rather than chrominance [77].

Gray-level co-occurrence matrix (GLCM), which was introduced firstly by Haralick [21] at the beginning of 1970s, has been proven an efficient way of texture representation [13]. GLCMs are formed by considering the number of occurrences of intensity value patterns in the image. Haralick proposed a set of statistical features obtained from GLCMs that achieved a success rate of %84 at a higher operational speed [3, 20]. Although it is an ancient method, it has been the reference and inspiration in many fields such as iris recognition [75], image segmentation [1] and CBIR [9, 36] in videos. However, base GLCM runs on gray-level images and discards the information carried by color bands. Arvis et al. [6] have proposed a method, which incorporates the color bands of the pixels during the construction of the co-occurrence matrices. As is known, GLCM contains information on the spatial relationships of intensity values and their formation amounts. Let  $f$  is an image whose intensity values vary in the range  $[0, L-1]$ . The value on row  $i$ , column  $j$  in GLCM indicates the number of times that the pixel pair  $(z_i, z_j)$  occurs in  $f$  with orientation  $Q$ . The orientation represented with  $Q$  eventually refers to a displacement vector  $d = (dx, dy \mid d_x = d_y = d_g)$ , where  $d_g$  is the number of gaps between the pixels of interest. For the situation of adjacency,  $d_g = 0$ . The orientation can also be represented with two parameters as the distance  $d$  that the intensities  $z_i, z_j$  apart from each other with angle  $\alpha$ .  $d$  theoretically can take values from  $0$  to  $L-2$ . The orientation of the pixel pattern can be in four different directions as  $0^\circ, 45^\circ, 90^\circ$  and  $135^\circ$ . That is, each image can have four different GLCMs (for each angle  $0^\circ, 45^\circ, 90^\circ$  and  $135^\circ$ ) for a given  $d$ . The size of a GLCM matrix depends on the discrete intensity values in the image. That is, if the intensity values of the image vary in the range  $[0, L-1]$ , then the size of the GLCM is  $(L-1) \times (L-1)$ . With the basic GLCM method, four separate matrices are formed, one for each of the directions  $0^\circ, 45^\circ, 90^\circ, 135^\circ$ . Considering the color bands of the image, a total of twenty-four GLCMs are generated, six at a time for each direction. In RIMFRA, local descriptive matrices generated for each band of the first stage image are fed into the multi-spectral-co-occurrence matrix construction process, rather than directly entering the raw image as done in previous studies. The output

of the process is the multi-spectral co-occurrence matrices, i.e.  $CM_{(RinLd)R(RinLd)R}$ ,  $CM_{(RinLd)G(RinLd)G}$ ,  $CM_{(RinLd)B(RinLd)B}$ ,  $CM_{(RinLd)R(RinLd)G}$ ,  $CM_{(RinLd)R(RinLd)B}$ , and  $CM_{(RinLd)G(RinLd)B}$ .

### 2.3 Orthogonal polynomial decomposition

In the proposed framework, the final stage of the feature extraction is the orthogonal polynomial decomposition process. Orthogonal polynomials, such as Tchebichef, have been shown to be an efficient means of representation of 2D functions [57]. In addition, some orthogonal polynomials such as Hermite and Zernike have also been used in previous studies during texture extraction and classification [38, 67]. However, Tchebichef polynomials have been identified to pose better performance compared to others [4]. Previous studies fed the raw image as input directly to the orthogonal polynomial decomposition process. However, as described in the simulation results section, inputting multi-spectral co-occurrence matrices instead of the raw image provides higher performance in terms of facial discrimination. Since the majority of information about the image and structure is concealed in the first few moments and the details are thought to be expressed in higher order moments, the second order statistical data, which does not have a high degree of importance can be eliminated by means of the limited opening. Thus, unnecessary complexity is eliminated.

The decomposition of the input matrix into moment orders  $M_{pq}$  is given in the following:

$$M_{pq} = \frac{1}{\rho(p)\rho(q)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} m_p(x)w(x)m_q(y)w(y)f(x,y) \tag{5}$$

where  $0 \leq p, q, x, y \leq N-1$ ;  $m_n(x)$  represents a set of orthogonal polynomials,  $w(x)$  and  $\rho()$  denote the weight function and rho respectively.

The mathematical representation of the Tchebichef orthogonal polynomials is given in the following equation:

$$m_n(x) = n! \sum_{k=0}^n (-1)^{n-k} \binom{N-1-k}{n-k} \binom{n+k}{n} \binom{x}{k} \tag{6}$$

$$\rho(n) = (2n)! \binom{N+n}{2n+1} \tag{7}$$

$$w(x) = 1 \tag{8}$$

where  $m_n(x)$ ,  $\rho(n)$ ,  $w(x)$  denote the  $n^{th}$  Tchebichef polynomial, rho and weight functions respectively. In this study, the number of coefficients calculated for a single input matrix with the size  $N \times N$  is equal to  $2 N-2$ , hence, a signature comprising  $6(2 N-2)$  coefficients is generated ultimately. Figure 3 depicts the orthogonal polynomial decomposition of a sample matrix:

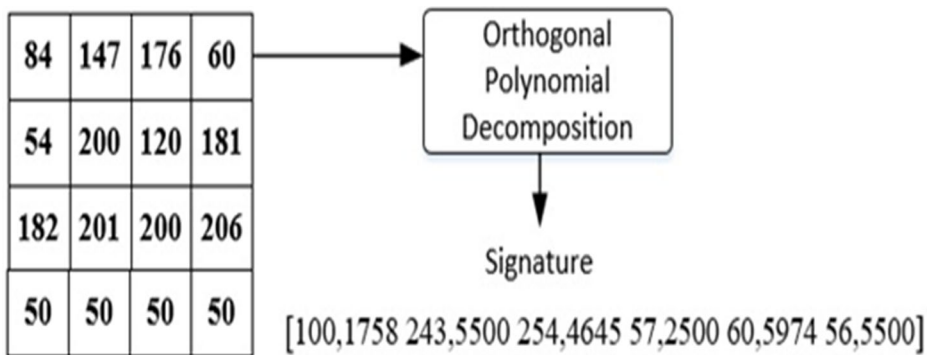


Fig. 3 Tchebichef signature of a sample matrix

### 3 Simulation results and discussions

Various experiments are conducted to measure and analyze the performance of the proposed framework under different circumstances. The evaluation of the proposed framework is performed on five benchmark databases, namely, Face94 [16], JAFFE [42], YALE (<http://vision.ucsd.edu/content/yale-face-database>), CAS-PEAL-R1 [17], ORL [55].

To ensure uniformity, some preprocessing is applied to each image. Each image is first scaled to a size of  $64 \times 64$ . Following the scaling stage, the face extraction is performed using the Viola Jones [66] algorithm to eliminate the effect of unnecessary background and foreground factors. To accurately measure and analyze the performance of the proposed framework and compare to the state-of-the-art methods in the area, hold out testing is utilized. That is, if an individual has  $N$  images in a data set, %80 of them and their average image are used for training. The rest is used for testing. When creating the average face image, each individual's images are aligned taking the individual's eyes into consideration. The sample average face calculation of an individual is given in Fig. 4:

The performance analysis of the proposed framework is performed in two folds. In the first step, the stability and resistance of the method to rotation, illumination changes and noise effects are clarified. Later, the recognition performance of the proposed method is analyzed and compared to the the-state-of-art methods such as LBP, LDP, LDNP (Local Directional Number Pattern) [51], Gabor Features, HoG (Histogram of Gradients), LTP, LTeTP (Local Tetra Pattern) [44] and LDrvP (Local Derivative Pattern) [76] by conducting extensive simulations.

Fig. 4 Sample face images of an individual and his average face image





### 3.1 The stability analysis

As mentioned earlier, rotational changes, variation of illumination and noise significantly affect recognition performance. Hence, the proposed local descriptor and the overall architecture should not fail under challenging circumstances and should remain intact to fulfill the recognition task satisfactorily. First, it is shown how the recommended local identifier remains intact against rotation changes. Following this, RIMFRA's performance analysis is carried out to verify its resistance with extensive simulations against lighting changes. The last section in this section shows the analysis results of the simulations performed to see the behavior of RIMFRA under changing noisy conditions.

#### 3.1.1 Rotation-variation resistance analysis

A solid texture descriptor should be stable and produce similar features even if the original image is rotated, because the content of the image does not change and belongs to the same person. It is therefore important to demonstrate and verify the behavior of the proposed method if the image is subject to rotational changes. Figures 5 and 6 show the stability performance of RIMFRA. In Fig. 5, a sample matrix, which symbolizes a block of an image is demonstrated. As depicted in the figure, the local descriptor content extracted from the block does not change even if the matrix is rotated 90° counter clockwise.

Figure 6 shows the face images of two people in the Face94 database and 90° rotated versions thereof.

Similarity performance analysis of the proposed method under rotational variation is made and compared with the state-of-the-art methods in the literature. Firstly, the resulting textural features of RIMFRA and other methods are obtained from the face images and their 90° variants. Next, the similarity analysis is done by calculating the Mean Square Error (MSE) between these sets of features. Figure 7 shows the histograms of the feature sets produced by RIMFRA for the face images of two sample individuals in the Face94 database and 90° rotated versions thereof.

Histograms presented in the first column belong first individual's face images and histograms in the right column belong to the second individual. Obviously, the histograms in the same column are very similar, which is the desired situation that confirms the robustness of RIMFRA against rotational changes. The similarity analysis is conducted on images from different databases to verify and compare the results fairly. Tables 1, 2, 3, 4 and 5 and Figs. 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 and 20 demonstrate the results of the similarity and correlation values between the feature sets of the sample images and their 90° rotated versions that are selected from different datasets. No image adjustment or enhancement technique is applied to images to see the performance of the methods used in raw images.

Figure 9 shows the face images and their 90° rotated versions of two individuals in the YALE database.

Figure 10 shows the histograms of the feature sets produced by RIMFRA for the face images of two sample individuals in the YALE database and 90° rotated versions thereof.

Figure 12 shows the face images and their 90° rotated versions of two individuals in the CAS-PEAL-R1 database.

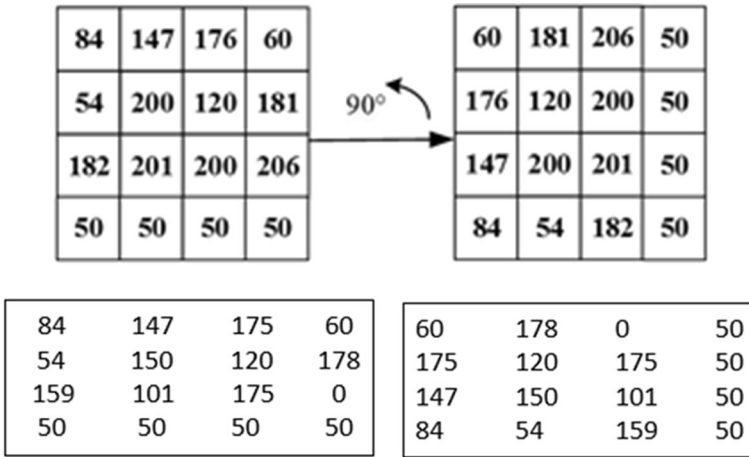


Fig. 5 Demonstration of the robustness of the proposed method to the rotational changes in an exemplary matrix

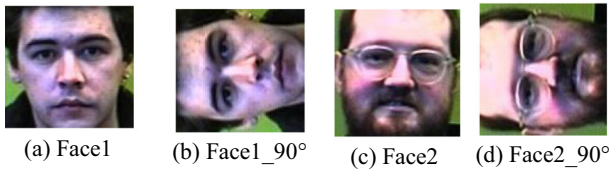


Fig. 6 Sample images and their rotated versions of two individuals in the Face94 database

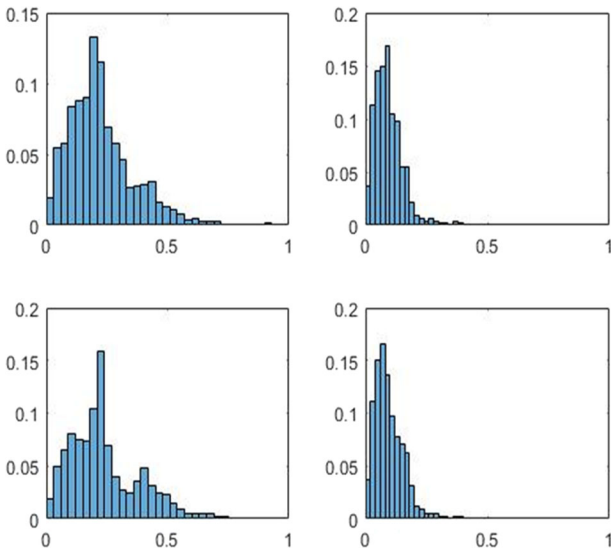


Fig. 7 RIMFRA feature set histograms of the images given in Fig. 6(a-d)

**Table 1** The MSE values calculated in the Face94 sample images

Method	MSE1	MSE2	MSE3	MSE4
RIMFRA	0.010	0.018	0.990	0.826
LTrnP	0.701	0.990	0.010	0.396
LDNP	0.990	0.881	0.010	0.336
LDrvP	0.010	0.990	0.737	0.514
LDP	0.990	0.268	0.685	0.010
LBP	0.010	0.990	0.597	0.597
LTetP	0.010	0.990	0.812	0.483
Gabor	0.010	0.568	0.990	0.990
HoG	0.701	0.990	0.010	0.031

**Table 2** MSE values calculated in the YALE sample images

Method	MSE1	MSE2	MSE3	MSE4
RIMFRA	0.011	0.010	0.990	0.940
LTrnP	0.990	0.990	0.010	0.138
LDNP	0.990	0.632	0.010	0.259
LDrvP	0.869	0.990	0.056	0.010
LDP	0.774	0.990	0.010	0.044
LBP	0.519	0.990	0.010	0.010
LTetP	0.519	0.990	0.029	0.010
Gabor	0.515	0.990	0.010	0.010
HoG	0.571	0.990	0.088	0.010

**Table 3** MSE values calculated in the CAS-PEAL-R1 sample images

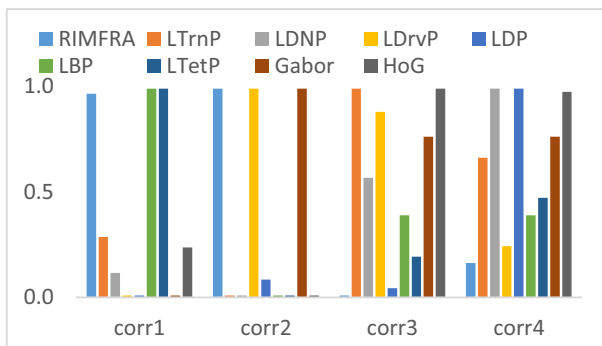
Method	MSE1	MSE2	MSE3	MSE4
RIMFRA	0.010	0.010	0.990	0.977
LTrnP	0.834	0.990	0.422	0.010
LDNP	0.377	0.990	0.067	0.010
LDrvP	0.010	0.990	0.151	0.266
LDP	0.682	0.990	0.010	0.347
LBP	0.990	0.010	0.877	0.882
LTetP	0.120	0.010	0.990	0.896
Gabor	0.063	0.010	0.990	0.990
HoG	0.129	0.990	0.174	0.010

**Table 4** MSE values calculated in the JAFFE sample images

Method	MSE1	MSE2	MSE3	MSE4
RIMFRA	0.010	0.027	0.927	0.990
LTrnP	0.990	0.658	0.010	0.044
LDNP	0.727	0.990	0.010	0.203
LDrvP	0.778	0.990	0.010	0.100
LDP	0.841	0.990	0.041	0.010
LBP	0.990	0.209	0.010	0.010
LTetP	0.990	0.256	0.089	0.010
Gabor	0.010	0.990	0.093	0.093
HoG	0.990	0.834	0.028	0.010

**Table 5** MSE values calculated on the ORL sample images

Method	MSE1	MSE2	MSE3	MSE4
RIMFRA	0.010	0.010	0.973	0.990
LTrnP	0.010	0.990	0.223	0.245
LDNP	0.690	0.990	0.589	0.010
LDrvP	0.990	0.982	0.010	0.307
LDP	0.530	0.990	0.010	0.732
LBP	0.347	0.010	0.990	0.990
LTetP	0.353	0.010	0.617	0.990
Gabor	0.990	0.086	0.010	0.010
HoG	0.297	0.990	0.231	0.010



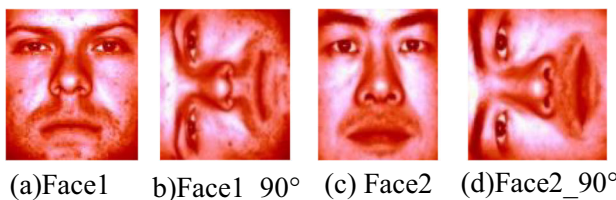
**Fig. 8** Graphical representation of the correlation between feature sets of sample images and their 90° rotated versions in Face94 dataset

Figure 13 demonstrates the histograms of the feature sets produced by RIMFRA for the face images of two sample individuals in the CAS-PEAL-R1 database and 90° rotated versions thereof.

Figure 15 shows the face images and their 90° rotated versions of two individuals in the JAFFE database.

Figure 16 demonstrates the histograms of the feature sets produced by RIMFRA for the face images of two sample individuals in the JAFFE database and 90° rotated versions thereof.

Figure 18 shows the face images and their 90° rotated versions of two individuals in the ORL database.



**Fig. 9** Sample images and their rotated versions of two individuals in the YALE database

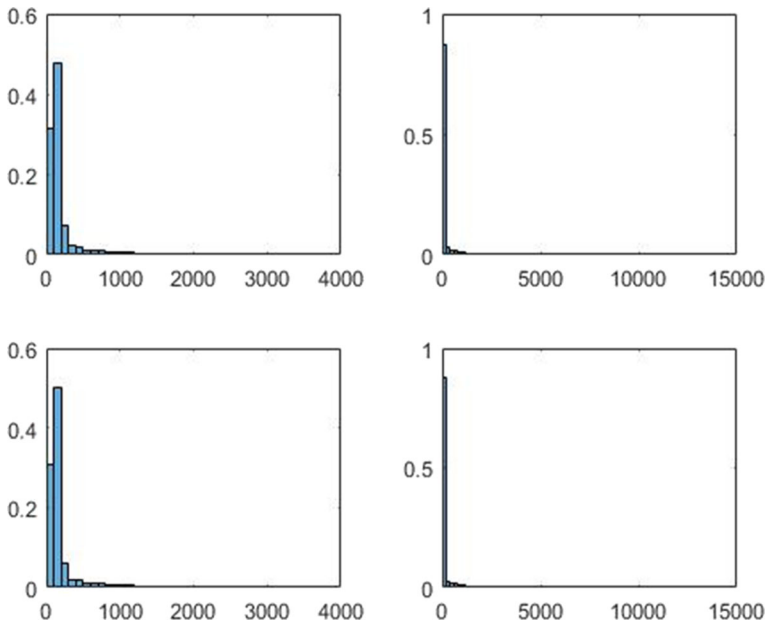


Fig. 10 RIMFRA feature set histograms of the images given in Fig. 9(a-d)

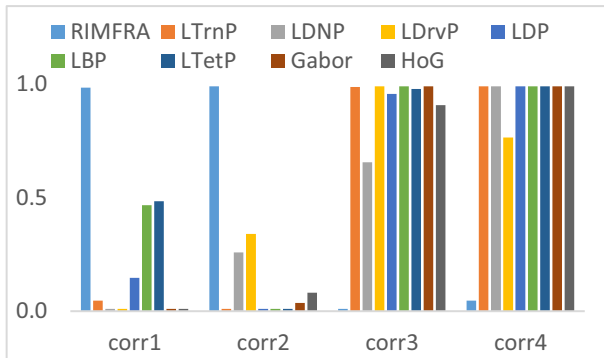


Fig. 11 Graphical representation of the correlation between feature sets of sample images and their 90° rotated versions in YALE dataset

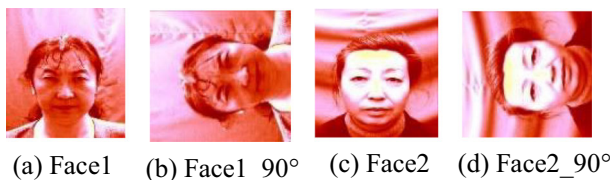


Fig. 12 Sample images and their rotated versions of two individuals in the CAS-PEAL-R1 database

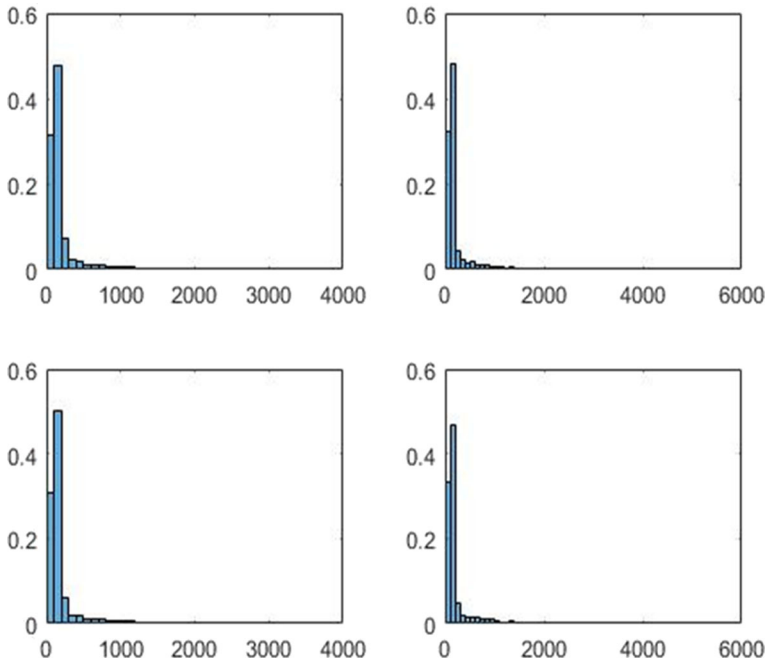


Fig. 13 RIMFRA feature set histograms of the images given in Fig. 12(a-d)

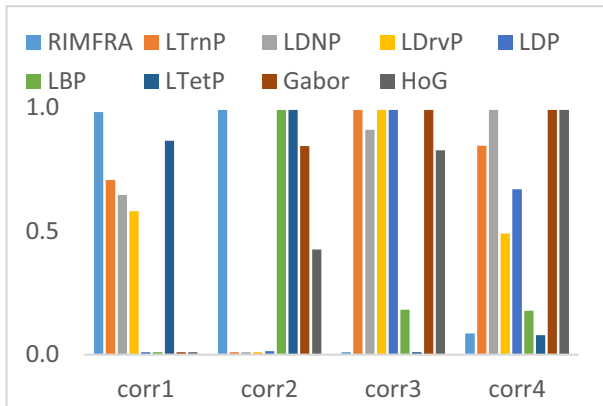


Fig. 14 Graphical representation of the correlation between feature sets of sample images and their 90° rotated versions in CAS-PEAL-R1 dataset



(a) Face1 (b) Face1\_90° (c) Face2 (d) Face2\_90°

Fig. 15 Sample images and their rotated versions of two individuals in the JAFFE database

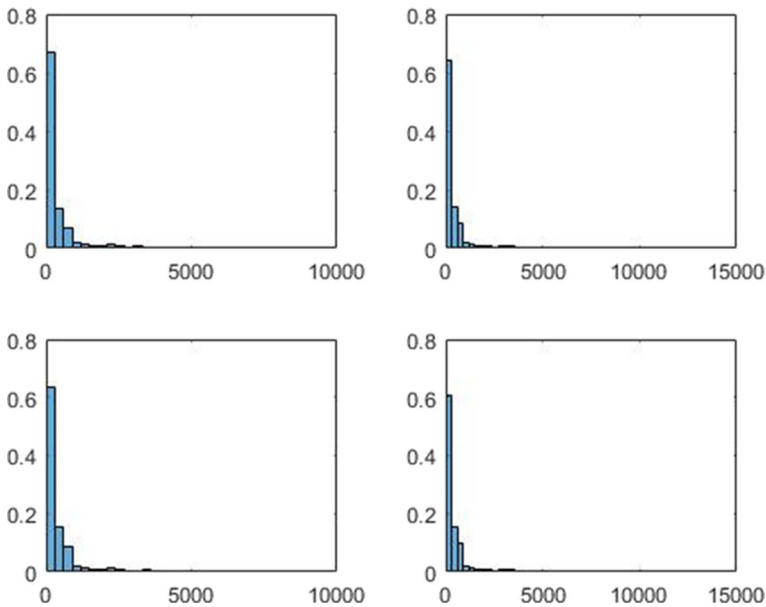


Fig. 16 RIMFRA feature set histograms of the images given in Fig. 15(a-d)

Figure 19 demonstrates the histograms of the feature sets produced by RIMFRA for the face images of two sample individuals in the ORL database and 90° rotated versions thereof.

In all tables and figures, columns 1 and corr1 values represent the similarity analysis and correlation results between the first image and its 90° rotated version respectively. Columns 2 and corr2 indicate the similarity analysis and correlation results between the second image and its 90° rotated version respectively. Columns 3 and corr3 values refer to the similarity analysis and correlation results between the first image and the 90° rotated version of the second image respectively. Columns 4 and corr4 values represent the similarity analysis and correlation results between the second image and the 90° rotated version of the first image respectively. Inherently, a

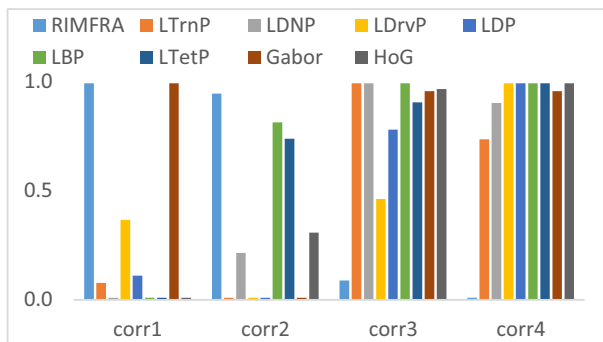
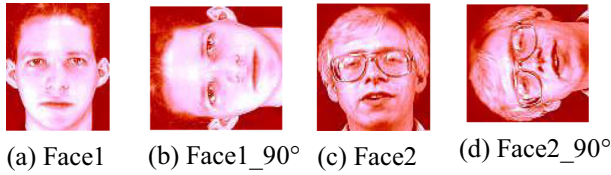


Fig. 17 Graphical representation of the correlation between feature sets of sample images and their 90° rotated versions in JAFFE dataset



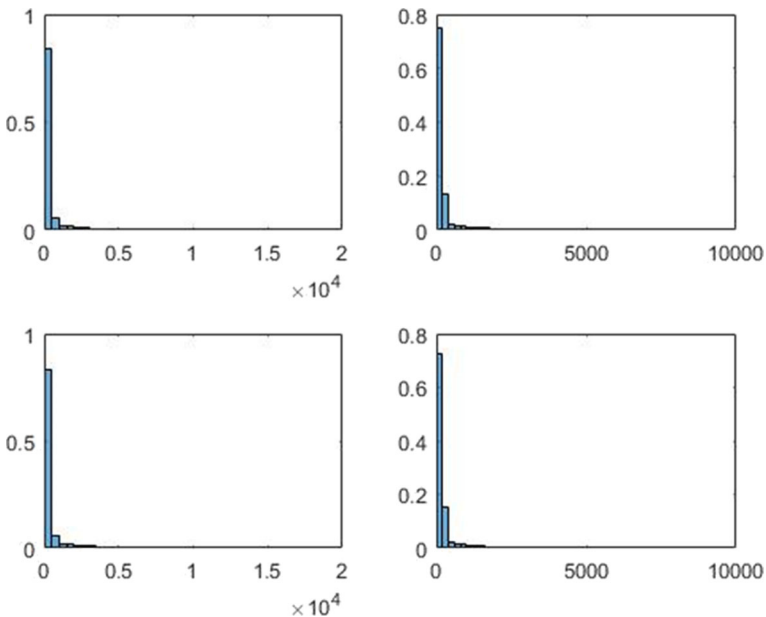
**Fig. 18** Sample images and their rotated versions of two individuals in the ORL database

high-representative texture descriptor of an individual's face image should remain similar even when the image of that individual is rotated. Furthermore, the dissimilarity between the images belonging to two different individuals should be high to differentiate the individuals. In addition, the correlation value between images of the same individual is high, but should be low among the images of different individuals. As clearly shown in the tables representing the results of the different datasets, RIMFRA achieves consistent and highly accurate performance, producing results that meet the above-mentioned considerations.

### 3.1.2 Illumination-variation resistance analysis

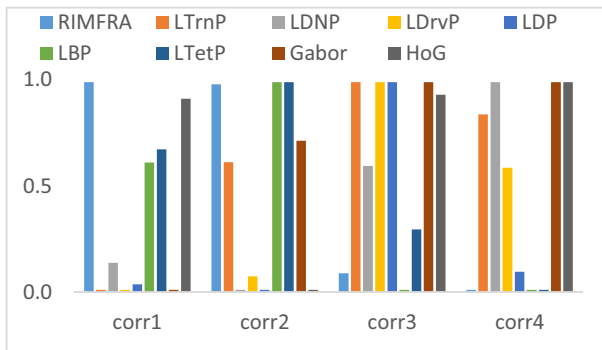
The second stage of the analysis process involves the performance test under illuminating variations and the comparison of the proposed method with other texture descriptors. Two types of analysis are performed to investigate the performance of our method. First, tests are performed on the face images in the CAS-PEAL-R1 data set (Fig. 21) exposed to natural lighting variations.

Table 6 and Fig. 22 show the MSEs and correlation values between the feature sets of the first image and others. Although CAS-PEAL-R1 is one of the most demanding data sets due to



**Fig. 19** RIMFRA feature set histograms of the images given in Fig. 18(a-d)





**Fig. 20** Graphical representation of the correlation between feature sets of sample images and their 90° rotated versions in ORL dataset

facial images containing compelling variations for texture descriptors, RIMFRA competes with the most modern descriptors proposed in the literature.

In the second stage of the illumination robustness analysis, an artificial, non-linear and non-uniform, third-order polynomial-based artificial illumination effect is created and included in the images in each dataset. Table 7, 8, 9, 10 and 11 and Fig. 23, 24, 25, 26 and 27 demonstrate the MSEs and correlation values calculated between the feature sets of the original image of an individual from each dataset and its artificially illuminated versions.

As presented in the tables above, RIMFRA offers promising performances in comparison to the latest technology in terms of robustness against illuminating variations.

### 3.1.3 Noise resistance analysis

Another compelling to consider during performance analysis of a texture descriptor is how it is resistant to noise effects without applying any noise filtration. Therefore, any pre-treatment method to mitigate the effects of noise is not applied in simulations in order to accurately analyze the resistance to noise. Two types of noise, i.e. salt-pepper and Gaussian, are applied to the images in each database. Firstly, the salt-



**Fig. 21** Sample images of an individual in the CAS-PEAL-R1 database under different lighting conditions

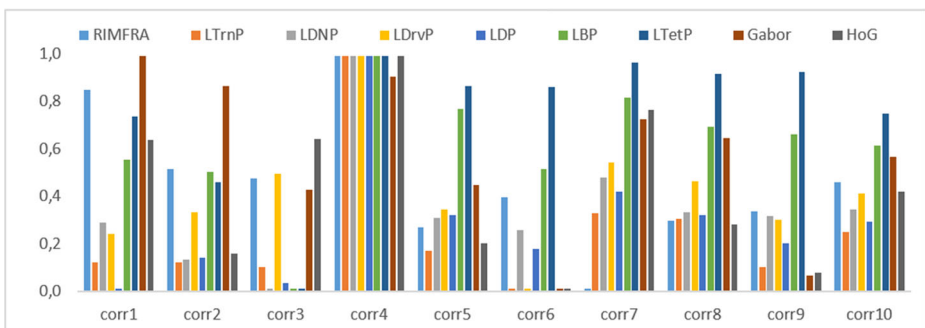
**Table 6** MSE values calculated on the CAS-PEAL-R1 sample images

Method	MSE1	MSE2	MSE3	MSE4	MSE5	MSE6	MSE7	MSE8	MSE9	AvgMSE
RIMFRA	0.010	0.678	0.900	0.294	0.116	0.265	0.643	0.197	0.442	0.394
LTrnP	0.400	0.574	0.707	0.058	0.392	0.900	0.010	0.151	0.718	0.435
LDNP	0.900	0.695	0.569	0.012	0.572	0.663	0.010	0.438	0.722	0.509
LDrvP	0.627	0.429	0.756	0.063	0.504	0.900	0.010	0.169	0.455	0.435
LDP	0.900	0.515	0.473	0.433	0.701	0.765	0.010	0.282	0.834	0.546
LBP	0.233	0.856	0.900	0.010	0.294	0.694	0.154	0.255	0.644	0.449
LTetP	0.010	0.900	0.268	0.390	0.123	0.158	0.453	0.190	0.269	0.307
Gabor	0.052	0.148	0.145	0.010	0.252	0.603	0.180	0.204	0.900	0.277
HoG	0.107	0.222	0.456	0.010	0.301	0.900	0.085	0.334	0.752	0.352

pepper noise is handled. Figure 28 shows an exemplary image, salt-pepper-noise exposed version, as well as RIMFRA feature set histograms of both. Clearly, the histogram of the multiple spectral-orthogonal signature of the original image and the histogram of the multiple spectral-orthogonal signature of the noisy version are very similar.

Table 12 and Fig. 29 show the similarity and dissimilarity values between a sample image in each dataset and their versions affected by the salt-pepper noise. For a method that is confirmed to be resistant to salt pepper noise, the MSE which expresses the difference between the feature sets of the image and the feature sets of the salt pepper noisy version should be low. In contrast, the correlation between the feature sets of the image and the feature sets of the salt pepper noisy version should be high. As a result, low MSE and high correlation values inherently points how the method is resistant to noise. As clearly seen, RIMFRA is one of the bests among the methods in terms of low MSE and high correlation.

The second noise resistance analysis is performed by incorporating Gaussian noise. Gaussian noise with different variance ( $\sigma^2$ ) values is applied to the images in each dataset. As in the salt-pepper noise effect analysis, the similarity and correlation values are measured between the feature sets of images and feature sets of their Gaussian noise exposed versions. Tables 13, 14, 15, 16 and 17, Figs. 30, 31, 32, 33 and 34 demonstrate the dissimilarity and correlations between the feature sets of the original images and noisy ones respectively.

**Fig. 22** Graphical demonstration of the correlation between the feature sets of a sample individual's self-illuminated images in CAS-PEAL-R1 dataset

**Table 7** MSE values calculated between the feature sets of the image of a sample individual in the Face94 data set and the feature sets of the artificially illuminated versions of this image

Method	MSE1 ( $\alpha = 0.1$ )	MSE2 ( $\alpha = 0.2$ )	MSE3 ( $\alpha = 0.3$ )	MSE4 ( $\alpha = 0.4$ )	MSE5 ( $\alpha = 0.5$ )
RIMFRA	0.010	0.010	0.010	0.010	0.990
LTmP	0.010	0.010	0.010	0.010	0.990
LDNP	0.010	0.010	0.010	0.010	0.990
LDrvP	0.010	0.010	0.010	0.010	0.990
LDP	0.010	0.010	0.010	0.010	0.990
LBP	0.010	0.010	0.010	0.010	0.990
LTetP	0.010	0.010	0.010	0.010	0.990
Gabor	0.010	0.010	0.010	0.010	0.990
HoG	0.010	0.010	0.010	0.010	0.990

**Table 8** MSE values calculated between the feature sets of the image of a sample individual in the YALE data set and the feature sets of the artificially illuminated versions of this image

Method	MSE1 ( $\alpha = 0.1$ )	MSE2 ( $\alpha = 0.2$ )	MSE3 ( $\alpha = 0.3$ )	MSE4 ( $\alpha = 0.4$ )	MSE5 ( $\alpha = 0.5$ )
RIMFRA	0.010	0.010	0.417	0.822	0.723
LTmP	0.010	0.426	0.682	0.813	0.990
LDNP	0.010	0.564	0.677	0.922	0.990
LDrvP	0.010	0.476	0.720	0.730	0.990
LDP	0.010	0.413	0.706	0.806	0.990
LBP	0.010	0.549	0.822	0.920	0.990
LTetP	0.010	0.558	0.750	0.896	0.990
Gabor	0.010	0.290	0.484	0.542	0.990
HoG	0.010	0.470	0.596	0.775	0.990

### 3.2 The recognition performance analysis

The recognition performance analysis of RIMFRA is done in two ways: 1- Training-based recognition performance analysis 2- Similarity-based recognition performance analysis. Since RIMFRA runs on color images, all images in each non-colored dataset are initially converted to RGB color space. To do this, a conversion map is generated by taking a reference colored image and its non-colored version. A best possible map is tried to be created for the conversion purpose. As is known, it is not possible to

**Table 9** MSE values calculated between the feature sets of the image of a sample individual in the CAS-PEAL-R1 data set and the feature sets of the artificially illuminated versions of this image

Method	MSE1 ( $\alpha = 0.1$ )	MSE2 ( $\alpha = 0.2$ )	MSE3 ( $\alpha = 0.3$ )	MSE4 ( $\alpha = 0.4$ )	MSE5 ( $\alpha = 0.5$ )
RIMFRA	0.010	0.010	0.010	0.010	0.990
LTmP	0.010	0.010	0.010	0.010	0.990
LDNP	0.010	0.010	0.010	0.010	0.990
LDrvP	0.010	0.010	0.010	0.010	0.990
LDP	0.010	0.010	0.010	0.010	0.990
LBP	0.010	0.010	0.010	0.010	0.990
LTetP	0.010	0.010	0.010	0.010	0.990
Gabor	0.010	0.010	0.010	0.010	0.990
HoG	0.010	0.010	0.010	0.010	0.990

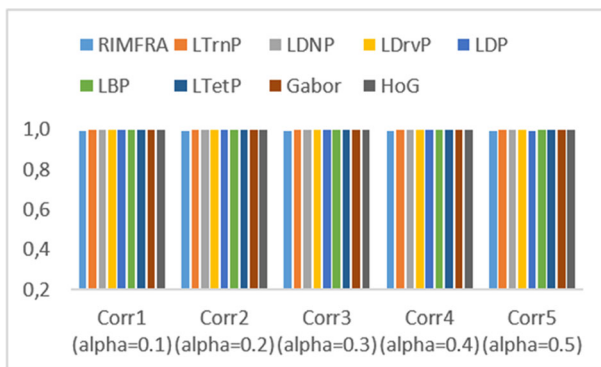
**Table 10** MSE values calculated between the feature sets of the image of a sample individual in the JAFFE data set and the feature sets of the artificially illuminated versions of this image

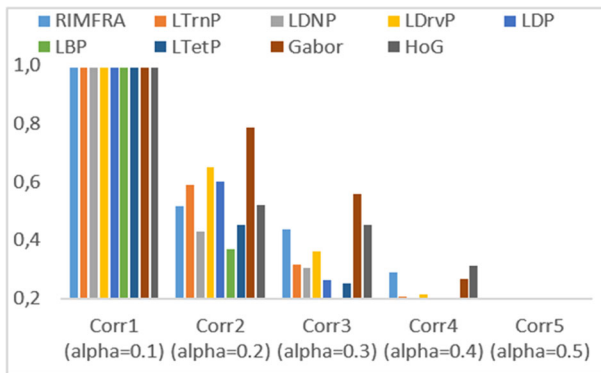
Method	MSE1 ( $\alpha = 0.1$ )	MSE2 ( $\alpha = 0.2$ )	MSE3 ( $\alpha = 0.3$ )	MSE4 ( $\alpha = 0.4$ )	MSE5 ( $\alpha = 0.5$ )
RIMFRA	0.010	0.010	0.010	0.010	0.990
LTrnP	0.010	0.010	0.010	0.010	0.990
LDNP	0.010	0.010	0.010	0.010	0.990
LDrvP	0.010	0.010	0.010	0.010	0.990
LDP	0.010	0.010	0.010	0.010	0.990
LBP	0.010	0.010	0.010	0.010	0.990
LTetP	0.010	0.010	0.010	0.010	0.990
Gabor	0.010	0.010	0.010	0.010	0.990
HoG	0.010	0.010	0.010	0.010	0.990

**Table 11** MSE values calculated between the feature sets of the image of a sample individual in the ORL data set and the feature sets of the artificially illuminated versions of this image

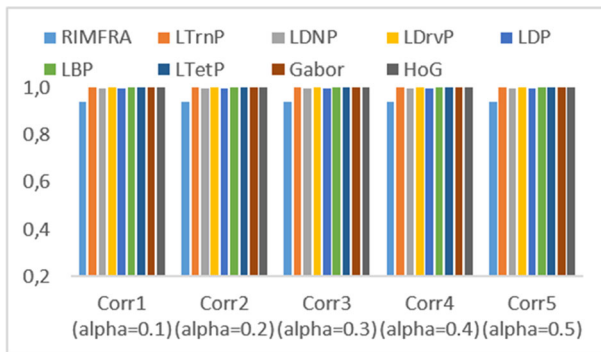
Method	MSE1 ( $\alpha = 0.1$ )	MSE2 ( $\alpha = 0.2$ )	MSE3 ( $\alpha = 0.3$ )	MSE4 ( $\alpha = 0.4$ )	MSE5 ( $\alpha = 0.5$ )
RIMFRA	0.010	0.010	0.010	0.010	0.990
LTrnP	0.010	0.010	0.010	0.010	0.990
LDNP	0.010	0.010	0.010	0.010	0.990
LDrvP	0.010	0.010	0.010	0.010	0.990
LDP	0.010	0.010	0.010	0.010	0.990
LBP	0.990	0.990	0.990	0.990	0.010
LTetP	0.010	0.010	0.010	0.010	0.990
Gabor	0.010	0.010	0.010	0.010	0.990
HoG	0.010	0.010	0.010	0.010	0.990

find a complete conversion map from gray to RGB conversion. Therefore, it is tried to find the best possible conversion map. Because the images of the Face94 data set are originally in a colored form, they give the best results during simulations. However, because other data sets consist of non-colored images, these images are first converted to RGB and then processed. The conversion process is unclear as it is approximate, which naturally affects the results.

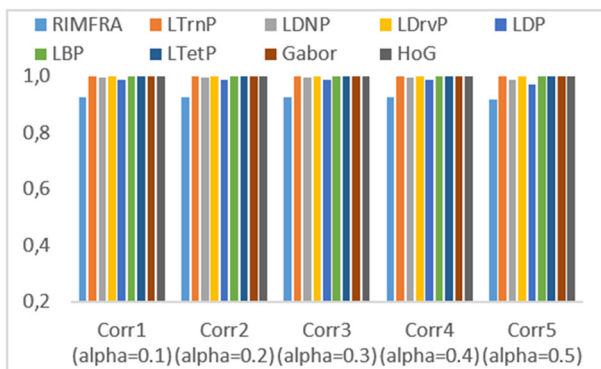
**Fig. 23** Graphical demonstration of the correlation between the feature sets of the artificially illuminated versions of an image in Face94 dataset



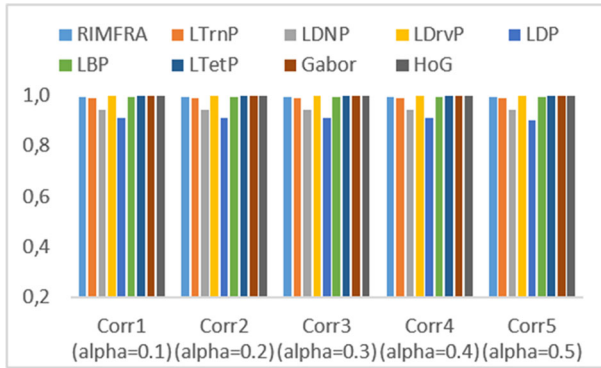
**Fig. 24** Graphical demonstration of the correlation between the feature sets of the artificially illuminated versions of an image in YALE dataset



**Fig. 25** Graphical demonstration of the correlation between the feature sets of the artificially illuminated versions of an image in CAS-PEAL-R1 dataset



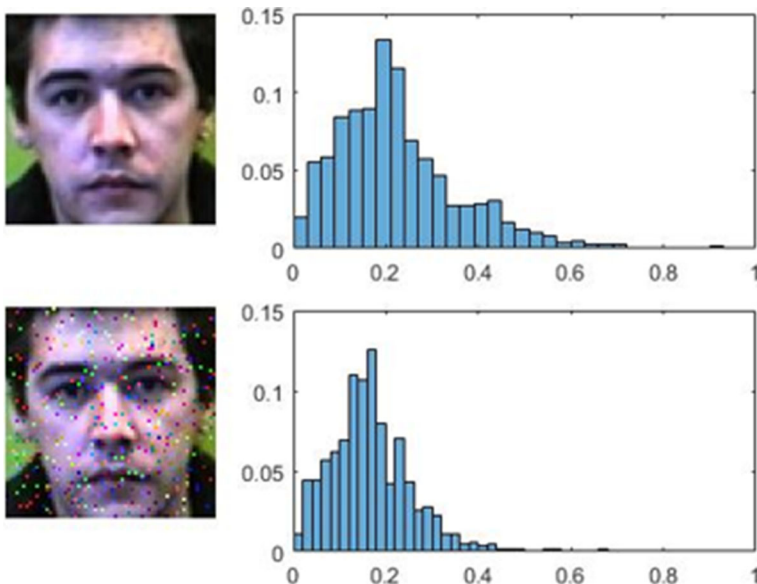
**Fig. 26** Graphical demonstration of the correlation between the feature sets of the artificially illuminated versions of an image in JAFFE dataset



**Fig. 27** Graphical demonstration of the correlation between the feature sets of the artificially illuminated versions of an image in ORL dataset

### 3.2.1 Training-based recognition performance analysis

At this stage, supervised learning is used during classification. %80 of each individual’s images in each dataset is used for training and the remaining images of the individuals are used for testing. Table 18 shows the performance results of RIMFRA and the-state-of-the-art methods in terms of recognition accuracy. As it can be seen in Table 18, RIMFRA performs promisingly well when compared to other methods in terms of classification accuracy analysis using supervised learning. RIMFRA performs remarkably even on the challenging datasets CAS-PEAL-R1, JAFFE and ORL.



**Fig. 28** Demonstration of the robustness of the proposed method against salt-pepper noise

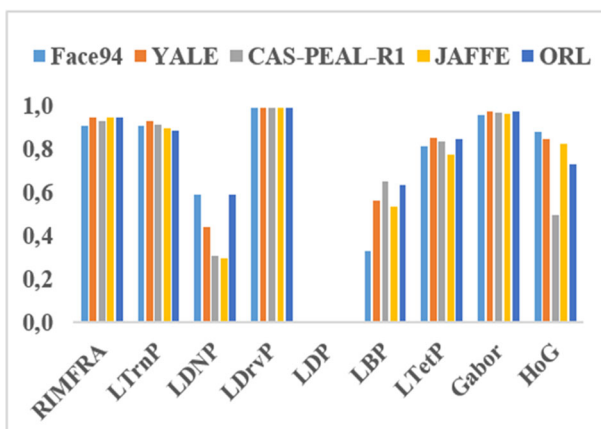
**Table 12** MSE values calculated between the feature sets of a sample individual's image in each database and its artificially salt-pepper noisy version

Method	MSEs				
	Face94	YALE	CAS-PEAL-R1	JAFFE	ORL
RIMFRA	0.141	0.038	0.039	0.010	0.031
LTmP	0.026	0.032	0.037	0.010	0.111
LDNP	0.062	0.120	0.167	0.010	0.111
LDrvP	0.010	0.010	0.010	0.010	0.010
LDP	0.886	0.990	0.990	0.990	0.990
LBP	0.990	0.967	0.657	0.010	0.876
LTetP	0.023	0.024	0.024	0.010	0.031
Gabor	0.030	0.037	0.026	0.010	0.046
HoG	0.113	0.301	0.678	0.010	0.528

### 3.2.2 Similarity-based recognition performance analysis

At this stage, recognition performance measurement of RIMFRA and the-state-of-the-art methods are done by implementing similarity analysis between the feature sets of the images. That is, the feature set of the image that is being searched is calculated and then compared to the feature sets of each image in the dataset. If the tag of the most similar image found matches up with the tag of the image that is being searched, that shows a hit (true-positive), otherwise a miss (false-positive). Table 19 figures out the recognition accuracy performances of each method in each dataset. As clarified in the table, RIMFRA competes with the other methods even on the challenging datasets without any training, that is without any knowledge.

The final step of the simulations is to measure recognition accuracy when images are subject to rotational changes. At this point, the desired image is rotated by 90° and then the feature set is extracted. The feature set is then compared with the feature set of the non-rotated images. Not surprisingly, RIMFRA shows remarkable performance under the circumstance of rotational change as presented in Table 20.

**Fig. 29** Graphical demonstration of the correlation between the feature sets of sample images in each database and their salt-pepper noisy versions

**Table 13** MSE values calculated between the feature sets of a sample individual's image in Face94 dataset and feature sets of its artificially Gaussian noise exposed versions

Method	$\sigma^2 = 0.01$	$\sigma^2 = 0.02$	$\sigma^2 = 0.03$	$\sigma^2 = 0.04$	$\sigma^2 = 0.05$
RIMFRA	0.010	0.103	0.250	0.636	0.990
LTrnP	0.010	0.406	0.780	0.898	0.990
LDNP	0.010	0.476	0.645	0.990	0.979
LDrvP	0.010	0.474	0.740	0.886	0.990
LDP	0.010	0.247	0.613	0.762	0.990
LBP	0.010	0.436	0.688	0.828	0.990
LTetP	0.010	0.463	0.699	0.990	0.929
Gabor	0.010	0.238	0.614	0.632	0.990
HoG	0.010	0.512	0.468	0.777	0.990

**Table 14** MSE values calculated between the feature sets of a sample individual's image in YALE dataset and feature sets of its artificially Gaussian noise exposed versions

Method	$\sigma^2 = 0.01$	$\sigma^2 = 0.02$	$\sigma^2 = 0.03$	$\sigma^2 = 0.04$	$\sigma^2 = 0.05$
RIMFRA	0.010	0.064	0.182	0.587	1.000
LTrnP	0.010	0.480	0.707	0.964	1.000
LDNP	0.010	0.538	0.710	1.000	0.993
LDrvP	0.010	0.484	0.665	0.825	1.000
LDP	0.010	0.427	0.602	0.817	1.000
LBP	0.010	0.552	0.798	0.987	1.000
LTetP	0.010	0.457	0.774	0.865	1.000
Gabor	0.010	0.374	0.782	0.557	1.000
HoG	0.010	0.349	0.671	0.927	1.000

**Table 15** MSE values calculated between the feature sets of a sample individual's image in JAFFE dataset and feature sets of its artificially Gaussian noise exposed versions

Method	$\sigma^2 = 0.01$	$\sigma^2 = 0.02$	$\sigma^2 = 0.03$	$\sigma^2 = 0.04$	$\sigma^2 = 0.05$
RIMFRA	0.010	0.413	0.674	0.850	1.000
LTrnP	0.010	0.525	0.681	0.820	1.000
LDNP	0.010	0.444	0.738	0.872	1.000
LDrvP	0.010	0.640	0.781	0.972	1.000
LDP	0.010	0.506	0.561	0.707	1.000
LBP	0.010	0.497	0.774	1.000	0.865
LTetP	0.010	0.486	0.732	0.743	1.000
Gabor	0.010	0.265	0.658	0.934	1.000
HoG	0.010	0.331	0.836	0.813	1.000

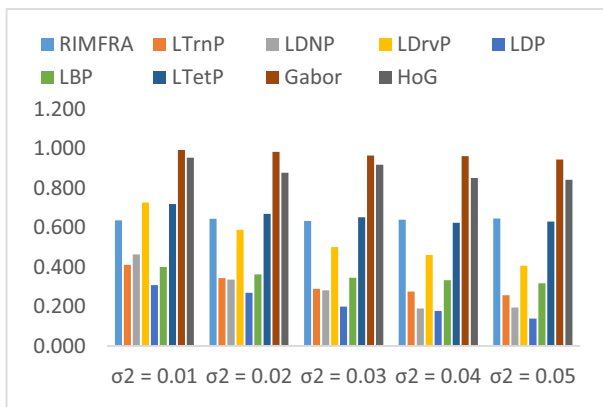
**Table 16** MSE values calculated between the feature sets of a sample individual's image in CAS-PEAL-R1 dataset and feature sets of its artificially Gaussian noise exposed versions

Method	$\sigma^2 = 0.01$	$\sigma^2 = 0.02$	$\sigma^2 = 0.03$	$\sigma^2 = 0.04$	$\sigma^2 = 0.05$
RIMFRA	0.010	0.271	0.483	0.323	1.000
LTrnP	0.010	0.442	0.748	0.836	1.000
LDNP	0.010	0.315	0.676	0.925	1.000
LDrvP	0.010	0.555	0.680	0.776	1.000
LDP	0.010	0.514	0.781	0.924	1.000
LBP	0.010	0.608	0.834	1.000	0.996
LTetP	0.010	0.682	0.790	0.841	1.000
Gabor	0.010	0.131	0.701	0.938	1.000
HoG	0.010	0.412	0.521	0.972	1.000

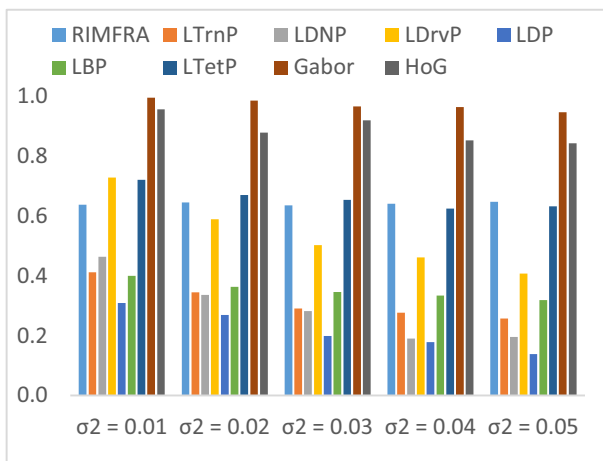


**Table 17** MSE values calculated between the feature sets of a sample individual’s image in ORL dataset and feature sets of its artificially Gaussian noise exposed versions

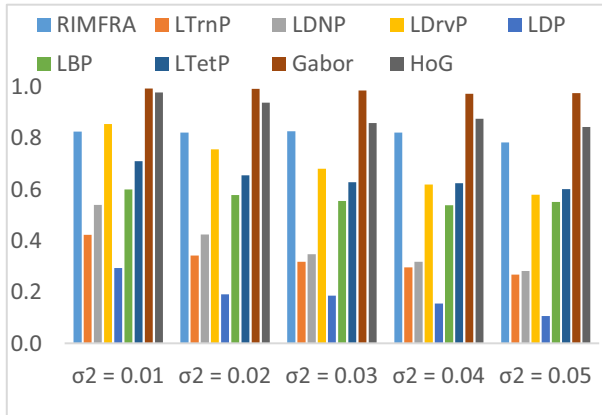
Method	$\sigma^2 = 0.01$	$\sigma^2 = 0.02$	$\sigma^2 = 0.03$	$\sigma^2 = 0.04$	$\sigma^2 = 0.05$
RIMFRA	0.010	0.298	0.558	0.689	1.000
LTrnP	0.010	0.517	0.817	0.996	1.000
LDNP	0.010	0.394	0.802	0.805	1.000
LDrvP	0.010	0.535	0.717	0.816	1.000
LDP	0.010	0.235	0.595	0.665	1.000
LBP	0.010	0.535	0.929	0.929	1.000
LTetP	0.010	0.433	0.717	0.866	1.000
Gabor	0.010	0.378	0.345	0.478	1.000
HoG	0.010	0.525	0.690	0.907	1.000



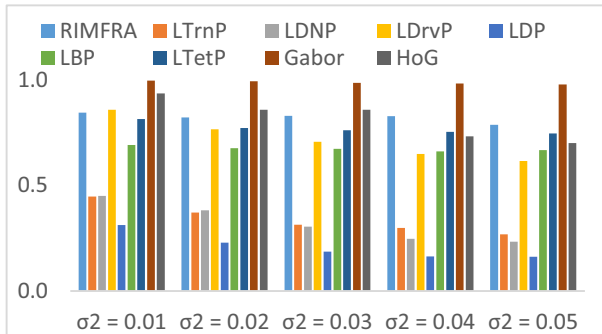
**Fig. 30** Graphical demonstration of the correlation between the feature sets of a sample image in Face94 dataset and feature sets of its Gaussian noisy versions



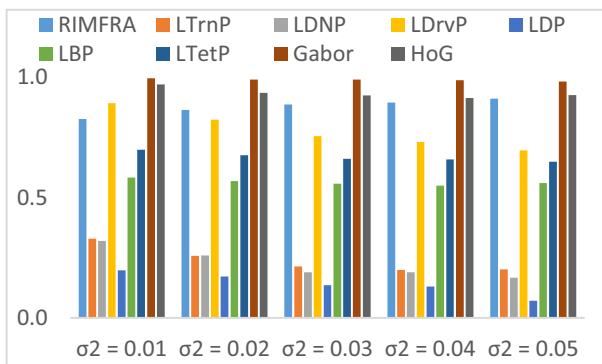
**Fig. 31** Graphical demonstration of the correlation between the feature sets of a sample image in YALE dataset and feature sets of its Gaussian noisy versions



**Fig. 32** Graphical demonstration of the correlation between the feature sets of a sample image in JAFFE dataset and feature sets of its Gaussian noisy versions



**Fig. 33** Graphical demonstration of the correlation between the feature sets of a sample image in CAS-PEAL-R1 dataset and feature sets of its Gaussian noisy versions



**Fig. 34** Graphical demonstration of the correlation between the feature sets of a sample image in ORL dataset and feature sets of its Gaussian noisy versions

**Table 18** The recognition performance results regarding supervised training

Method	Face94	CAS-PEAL-R1	JAFFE	ORL
RIMFRA	0.997	0.891	0.920	0.775
LTrnP	0.991	0.947	1.000	0.775
LDNP	0.996	0.950	1.000	0.750
LDrvP	0.991	0.889	0.900	0.638
LDP	0.991	0.881	0.980	0.725
LBP	1.000	0.912	0.960	0.825
LTetP	0.998	0.836	0.780	0.725
Gabor	1.000	0.897	0.820	0.863
HoG	1.000	0.987	1.000	0.888

**Table 19** The recognition performance results regarding similarity analysis

Method	Accuracies							
	Face94		CAS-PEAL-R1		JAFFE		ORL	
	TP	FP	TP	FP	TP	FP	TP	FP
RIMFRA	93	0	350	27	10	0	36	4
LTrnP	93	0	364	13	10	0	33	7
LDNP	93	0	362	15	10	0	35	5
LDrvP	93	0	362	15	10	0	36	4
LDP	93	0	341	36	10	0	33	7
LBP	93	0	353	24	10	0	35	5
LTetP	93	0	336	41	10	0	40	0
Gabor	93	0	352	25	10	0	40	0
HoG	93	0	375	2	10	0	40	0

**Table 20** The recognition performance results regarding similarity analysis

Method	Accuracies							
	Face94		CAS-PEAL-R1		JAFFE		ORL	
	TP	FP	TP	FP	TP	FP	TP	FP
RIMFRA	93	0	349	38	9	1	18	22
LTrnP	1	92	0	377	0	10	4	36
LDNP	1	92	2	375	0	10	2	38
LDrvP	0	93	0	377	0	10	5	35
LDP	0	93	0	377	0	10	2	38
LBP	4	89	0	377	0	10	1	39
LTetP	4	89	0	377	0	10	5	35
Gabor	13	80	0	377	3	7	6	34
HoG	0	93	0	377	0	10	7	33

## 4 Conclusion

This paper proposes a rotation-invariant multi-spectral facial recognition approach, which is highly resistant especially to rotational variances, as well as illumination changes and noise effects. Nearly all methods proposed so far have based on gray-level domain that ignores the information embodied in the color bands. The traditional view during texture extraction is considering the relationships of the pixels only in the colorless domain. However, during texture extraction, significant discriminative features can be obtained by considering the relationships between different color bands of the neighboring pixels. With this in mind, RIMFRA explores the multi-spectral relationships of pixels with their neighbors. Orthogonal polynomials are an effective way of representing 2D matrices. Thus, the resulting matrices produced in the previous step are fed to the orthogonal polynomial decomposition stage. The first few coefficients of the polynomial are the ones that the most information about the 2D matrix and also help reduce the size of the feature set. Simulation results encourage us to take the idea a step further by considering not only the RGB space but also other color spaces and combining the features of different color spaces as a compound feature set in future studies.

## References

1. Abutaleb AS (1989) Automatic Thresholding of Gray-level Pictures Using Two-dimensional Entropies. *Computer Vision Graphics Image Processing* 47:22–32
2. Ahonen T, Hadid A, Pietikäinen M (2006) Face description with local binary patterns: Application to face recognition. *IEEE Trans Pattern Anal Mach Intell* 28(12):2037–2041
3. Allam S, Adel M, Refregier P (1997) Fast Algorithm for Texture Discrimination by Use of a Separable Orthonormal Decomposition of the Co-occurrence Matrix. *Appl Opt* 36:8313–8321
4. An Approach to Textile Recognition, *Pattern Recognition*, Peng-Yeng Yin (Ed.), ISBN: 978-953-307-014-8, InTech. Available from: <http://www.intechopen.com/books/pattern-recognition/anapproach-to-textile-recognition>.
5. Andreu Y, García-Sevilla P, Mollineda RA (2014) Face gender classification: a statistical study when neutral and distorted faces are combined for training and testing purposes. *Image Vis Comput* 32(1):27–36
6. Arvis V, Debain C, Berducat M, Benassi A (2004) Generalization of the Co-occurrence Matrix for Color Images: Application to Color Texture Classification. *Image Analysis and Stereology* 23:63–72
7. Belhumeur P, Hespanha J, Kriegman D (1997) Eigenfaces vs. fisher-faces: Recognition using class specific linear projection. *IEEE Trans Pattern Anal Mach Intell* 19(7):711–720
8. Byungyong R, Rivera AR, Kim J, Chae O (2017) Local Directional Ternary Pattern for Facial Expression Recognition. *IEEE Trans Image Process* 26(12):6006–6018
9. Cheong M, Loke KS (2008) Textile Recognition Using Tchebichef Moments of Co-occurrence Matrices. In: Huang DS., Wunsch D.C., Levine D.S., Jo KH. (eds) *Advanced Intelligent Computing Theories and Applications. With Aspects of Theoretical and Methodological Issues*. ICIC. Lecture Notes in Computer Science, vol. 5226. Springer, Berlin, Heidelberg
10. Comon P (1994) Independent component analysis - a new concept? *Signal Process* 36:287–314
11. Dahmane M, Meunier J (2011) Emotion recognition using dynamic gridbased HoG features. *IEEE Int Conf Autom Face Gesture Recognit Workshops (FG)*:884–888
12. Dan Z, Chen Y, Yang Z, Wu G (2014) An improved local binary pattern for texture classification. *Optik* 125:6320–6324
13. Davis LS (1981) Image Texture Analysis Techniques - A Survey. In: Simon JC, Haralick RM (eds) *Digital Image Processing*. D. Reidel, Dordrecht
14. Dubey SR (2017) Local Directional Relation Pattern for Unconstrained and Robust Face Retrieval. arXiv: 1709.09518 [cs.CV]
15. Eskandari M, Toygar O, Demirel H (2014) Feature extractor selection for face-iris multimodal recognition. *Signal Image Video Process* 8(6):1189–1198
16. Face Recognition Data, University of Essex, UK, Face 94, [http://cswww.essex.ac.uk/mv/all\\_faces/faces94.html](http://cswww.essex.ac.uk/mv/all_faces/faces94.html).
17. Gao W, Cao B, Shan S, Chen X, Zhou D, Zhang X, Zhao D (2008) The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations. *IEEE Trans on System Man, and Cybernetics (Part A)* 38(1):149–161

18. Hadid A, Dugelay JL, Pietikäinen M (2011) On the use of dynamic features in face biometrics: recent advances and challenges. *Signal Image Video Processing* 5(4):495–506
19. Hahn F, Sanchez S (2000) Carrot volume evaluation using imaging algorithms. *J Agric Eng Res* 75:243–249
20. Haralick RM (1979) Statistical and structural approach to texture. *Proc IEEE* 67(5):786–804
21. Haralick RM, Shanmugan K, Dinstein I (1973) Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics* 3:610–621
22. He X, Cai D, Yan S, Zhang H (2005) Neighborhood preserving embedding. *IEEE Int Conf Comput Vis: 1208–1213*
23. He X, Yan S, Hu Y, Niyogi P, Zhang H (2005) Face recognition using laplacian faces. *IEEE Trans Pattern Anal Mach Intell* 27(3):328–340
24. Ishikawa Y, Hirata T (2001) Color change model forbroccoli packaged in polymeric films. *Transactions of the ASAE* 44:923–927
25. Jabid T, Kabir MH, Chae O (2010) Robust facial expression recognition based on local directional pattern. *ETRI J* 32(5):784–794
26. Jafri R, Arabnia HR (2009) A Survey of Face Recognition Techniques. *Journal of Information Processing Systems* 5(2):41–68
27. Jain A, Hong L, Pankanti S (2000) Biometric Identification. *Commun ACM* 43(2):91–98
28. Jain AK, Ross A (2008) Introduction to Biometrics. In: Jain, AK; Flynn; Ross, A. *Handbook of Biometrics*. Springer. pp. 1–22, ISBN 978–0–387-71040-2
29. Jian M, Lam KM (2013) Simultaneous Hallucination and Recognition of Low-Resolution Faces Based on Singular Value Decomposition. *Pattern Recogn* 46(11):3091–3102
30. Jian M, Lam KM (2014) Face-Image Retrieval Based on Singular Values and Potential-Field Representation. *Signal Process* 100:9–15
31. Jian M, Lam KM, Dong J (2014) Facial-Feature Detection and Localization Based on a Hierarchical Scheme. *Inf Sci* 262:1–14
32. Jian M, Lam KM, Dong J (2014) Illumination-insensitive Texture Discrimination Based on Illumination Compensation and Enhancement. *Inf Sci* 269:60–72
33. Jian M, Lam KM, Dong J, Zang W (2018) Comprehensive Assessment of Non-Uniform Illumination for 3D Heightmap Reconstruction in Outdoor Environments. *Comput Ind* 99:110–118
34. Kaya Y, Ertugrul OF (2017) Gender classification from facial images using gray relational analysis with novel local binary pattern descriptors. *Signal Image and Video Processing* 11:769–776
35. Khojastehnazhand M, Omid M, Tabatabaeefer A (2009) Determination of orange volume and surface area using image processing technique. *International Agrophysics* 23:237–242
36. Kim K, Jeong S, Chun BT, Lee JY, Bae Y (1999) Efficient Video Images Retrieval by Using Local Co-occurrence Matrix Texture Features and Normalised Correlation. *Proceedings of The IEEE Region 10 Conf* 2:934–937
37. Koc AB (2007) Determination of watermelon volume using ellipsoid approximation and image processing. *Postharvest Biol Technol* 45(3):366–371
38. Krylov AS, Kutovoi AV (2002) Texture Parameterization with Hermite Functions. *International Conference Graphicon, Nizhny Novgorod*
39. Lei Z, Liao S, Pietikäinen M, Li SZ (2011) Face recognition by exploring information jointly in space, scale and orientation. *IEEE Trans Image Process* 20(1):247–256
40. Li B, Lian XC, Lu BL (2012) Gender classification by combining clothing, hair and facial component classifiers. *Neurocomputing* 76(1):18–27
41. Liu L, Fieguth P, Guo Y, Wang X, Pietikainen M (2017) Local Binary Features for Texture Classification: Taxonomy and experimental study. *Pattern Recogn* 62:135–160
42. Lyons MJ, Akamatsu S, Kamachi M, Gyoba J (1998) Coding Facial Expressions with Gabor Wavelets. 3rd IEEE International Conference on Automatic Face and Gesture Recognition, Nara
43. Melendez J, Garcia MA, Puig D (2008) Efficient distance-based per-pixel texture classification with Gabor wavelet filters. *Pattern Anal Applic* 11(3):365–372
44. Murala S, Maheshwari RP, Balasubramanian R (2012) Local Tetra Patterns: A New Feature Descriptor for Content-Based Image Retrieval. *IEEE Trans Image Process* 21(5):2874–2886
45. Nambi VE, Thangavel K, Rajeswari KA, Manickavasagan A, Geetha V (2016) Texture and rheological changes of Indian mango cultivars during ripening. *Postharvest Biol Technol* 117:152–160
46. Nambi VE, Thangavel K, Shahir S, Thirupathi V (2016) Comparison of various RGB image features for nondestructive prediction of ripening quality of alphonso mangoes for easy adoptability in machine vision applications: a multivariate approach. *J Food Qual* 39:816–825
47. Nanni L, Brahnam S, Ghidoni S, Menegatti E, Barrier T (2013) Different Approaches for Extracting Information from the Co-Occurrence Matrix. *PLoS One* 8(12):1–9
48. Nisenson M, Yariv I, El-Yaniv R, Meir R (2003) Towards Behaviometric Security Systems: Learning to Identify a Typist. *Lect Notes Comput Sci*:363–374

49. Quevedo R, Aguilera J, Pedreschi F (2010) Color of salmon fillets by computer vision and sensory panel. *Food Bioprocess Technol*:637–643
50. Rai P, Khanna P (2014) A gender classification system robust to occlusion using Gabor features based (2D) PCA. *J Vis Commun Image Represent* 25(5):1118–1129
51. Rivera AR, Castillo JR, Chae O (2012) Local Directional Number Pattern for Face Analysis: Face and Expression Recognition. *IEEE Trans Image Process* 22(5):1740–1752
52. Rivera AR, Castillo R, Chae O (2013) Local directional number pattern for face analysis: Face and expression recognition. *IEEE Trans Image Process* 22(5):1740–1752
53. Rivera AR, Chae O (2015) Spatiotemporal directional number transitional graph for dynamic texture recognition. *IEEE Trans Pattern Anal Mach Intell* 37(10):2146–2152
54. Roweis S, Saul L (2000) Nonlinear dimensionality reduction by locally linear embedding. *Science* 290(22):2323–2326
55. Samaria F, Harter A (1994) Parameterization of a Stochastic Model for Human Face Identification. 2nd IEEE Workshop on Applications of Computer Vision, Sarasota
56. Schölkopf B, Smola A, Müller KR (1999) Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput* 10:1299–1319
57. See KW, Loke KS, Lee PA, Loe KF (2007) Image reconstruction using various discrete orthogonal polynomials in comparison with DCT. *Appl Math Comput* 193(2):346–359
58. Shan C (2012) Learning local binary patterns for gender classification on real-world face images. *Pattern Recogn Lett* 33(4):431–437
59. Shan C, Gong S, McOwan PW (2009) Facial expression recognition based on local binary patterns: A comprehensive study. *Image Vis Comput* 27(6), pp. 803–816. Available: <http://www.sciencedirect.com/science/article/pii/S0262885608001844>
60. Shih HC (2013) Robust gender classification using a precise patch histogram. *Pattern Recogn* 46(2):519–528
61. Stajanko D, Rakun J, Blanke M (2009) Modelling apple fruit yield using image analysis for fruit color, shape and texture. *Eur J Hortic Sci* 74(6):260–267
62. Tan X, Triggs B (2010) Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans Image Process* 19(6):1635–1650
63. Tenenbaum J, Silva V, Langford J (2000) A global geometric framework for nonlinear dimensionality reduction. *Science* 290(22):2319–2323
64. Tseng S (2003) Comparison of holistic and feature based approaches to face recognition. MSc Thesis, Royal Melbourne Institute of Technology University, Melbourne, Victoria
65. Turk MA, Pentland AP (1991) Eigenfaces for Recognition. *J Cogn Neurosci* 3(1):71–86
66. Viola P, Jones MJ (2004) Robust real-time face detection. *Int J Comput Vis* 57:137–154
67. Wang L, Healey G (1998) Using Zernike Moments for the Illumination and Geometry Invariant Classification of Multispectral Texture. *IEEE Trans Image Process* 7(2):196–203
68. Wang W, Li C (2014) Size estimation of sweet onions using consumer-grade RGB-depth sensor. *J Food Eng* 142:153–162
69. Wang X, Tang X (2004) A unified framework for subspace face recognition. *IEEE Trans Pattern Anal Mach Intell* 26(9):1222–1228
70. Wolf L, Hassner T, Taigman Y (2011) Effective unconstrained face recognition by combining multiple descriptors and learned background statistics. *IEEE Trans Pattern Anal Mach Intell* 33(10):1978–1990
71. Xia B, Amor BB, Drira H, Daoudi M, Ballihi L (2015) Combining face averageness and symmetry for 3D-based gender classification. *Pattern Recogn* 48(3):746–758
72. Yan S, Xu D, Zhang B, Zhang H, Yang Q, Lin S (2007) Graph embedding and extensions: A general framework for dimensionality reduction. *IEEE Trans Pattern Anal Mach Intell* 29(1):40–51
73. Yang S, Bhanu B (2011) Facial expression recognition using emotion avatar image. *IEEE Int Conf Autom Face Gesture Recognit Workshops (FG)*:866–871
74. Yin QB, Kim JN (2008) Rotation-invariant texture classification using circular Gabor wavelets based local and global features. *Chin J Electron* 17(4):646–648
75. Zaim A, Sawalha A, Quweider M, Iglesias J, Tang R (2006) A New Method for Iris Recognition Using Gray-level Co-occurrence Matrix. In: *IEEE International Conf. on Electro/Information Technology*, pp. 350–353
76. Zhang B, Gao Y, Zhao S, Liu J (2010) Local Derivative Pattern Versus Local Binary Pattern: Face Recognition with High-Order Local Pattern Descriptor. *IEEE Trans Image Process* 19(2):533–543
77. Zhang Q, Zhang J (2009) RGB Color Analysis for Face Detection. In: *Book: Advances in Computer Science and IT*, pp. 109–125, InTech
78. Zhao G, Ahonen T, Matas J, Pietikäinen M (2012) Rotation invariant image and video description with local binary pattern features. *IEEE Trans Image Processing* 21(4):1465–1477

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Taner Cevik** received the Ph.D. degree from Istanbul University in 2012. He joined the Software Engineering Department at Aydin University in 2017 and continues to work as an Associate Professor. His research interests are Image Processing and Wireless Multimedia Sensor Networks.



**Nazife Cevik** obtained the Ph.D. degree from Istanbul University in 2015. She joined the Computer Engineering Department at Istanbul Arel University in 2015 and continues to work as an Assistant Professor. Her research interests are Image Processing, Machine Learning and Bioinformatics.