




Face based person recognition mechanism using monogenic Binarized Statistical Image Features

Nour Elhouda Chalabi^{1,2}  · Abdelouahab Attia^{1,2} · Abderraouf Bouziane^{1,2} · Zahid Akhtar³

Received: 27 July 2021 / Revised: 7 February 2022 / Accepted: 10 March 2022 /
Published online: 23 March 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

These days, automated face recognition systems are hugely being applied in diverse applications ranging from personal use to border crossing. Feature extraction/representation is extremely vital module in any biometric systems, including face recognition. Thus, the main contribution of this paper is the proposition of a novel descriptor based on monogenic signal representation and Binarized Statistical Image Feature (BSIF) to extract quite distinctive relevant features from face image, named (M-BSIF). In fact, BSIF has not always efficient for face feature extraction, as it was not able to attain the best recognition rates. In order to enhance the capability of BSIF feature representation, our proposed feature description scheme, first applies band pass mechanism via log-Gabor filter on the image, then a monogenic filter is applied to decompose face image into three complementary parts, i.e., local amplitude, local phase, and local orientation. Next, BSIF is utilized to encode these complementary components in order to extract M-BSIF features. Experimental analyses on three publicly available databases (i.e., ORL database, AR database and JAFFE database) demonstrate the efficacy of the proposed M-BSIF descriptor. The proposed system outperforms a framework using only single BSIF.

Keywords Face recognition, monogenic signal representation · BSIF · Log Gabor filter · Feature extraction

✉ Nour Elhouda Chalabi
chalabi.houda94@gmail.com

¹ Computer Science Department, Mohamed El Bachir El Ibrahimi University, Bordj Bou Arreridj 34000, Algeria

² LMSE Laboratory, Mohamed El Bachir El Ibrahimi University, Bordj Bou Arreridj 34000, Algeria

³ Department of Network and Computer Security, State University of New York Polytechnic Institute, Utica, NY 13502, USA

1 Introduction

Feature extraction is a key component in any recognition system. The compact and quality representation is of a high importance, and has deep relation with the performance of the system. Several feature extraction methods and techniques have been introduced and utilized in the literature [21, 25]. For instance, Binarized Statistical Image Feature (BSIF) [18] has been used for feature extraction in variety of systems such as fingerprint recognition system, palmprint recognition system, iris recognition system, and many more. One of the main reasons is that BSIF texture representation has efficacy, e.g., Raja et al. [33] proposed a BSIF features based schema for visible spectrum iris and periocular verification systems. In another work, Doyle et al. [10] developed an iris recognition based on BSIF features. While, Younesi et al. [42] introduced a palmprint recognition system using Gabor filter and BSIF features; such that the region of interest (ROI) images are past to Gabor filters, then to BSIF for features extraction,

In a recent work, Mishra et al. [28] proposed a palmprint recognition framework based on BSIF features fused with BRISK features. Attia et al. [4] designed a finger knuckle print recognition frameworks utilizing a bank of multi scale BSIF features; different BSIF filters were used to create a separate BSIF filter bank. In another work, Attia et al. [5] have presented a multimodal biometric system using finger major and minor dorsal finger knuckle patterns for recognition. The BSIF was employed to extract the feature from the different modalities. In different framework, Attia et al. [6] adapted a finger knuckle print recognition system with deep rule-based classifier such that the BSIF and Gabor filter where employed for feature extraction. BSIF features extraction was also employed with face modality. For example, Ylioinas et al. [41] proposed a face recognition system, where BSIF features was used. However, BSIF was not employed on the whole image, it was rather block wise adaption. The BSIF was employed on different rejoin of the face image and BSIF histogram vectors were collected at the end. Ouamane et al. [31] developed a multimodal face verification algorithm. A local features fusion method was presented based on four different feature extraction techniques (i.e., LBP, SLF, BSIF, and LPQ features). Beside BSIF, new feature extraction method are proposed and used daily, e.g., a recently proposed work by Kumar et al. [20], where the a novel descriptor was employed for feature extraction named Dense Local Graph Structure (D-LGS). Cheng et al. [9] proposed another feature descriptor, where deep learning convolutional neural network was used to retrieve features via sparse representation. Yee et al. [40] proposed a face recognition system based on Laplacian Completed Local Ternary Pattern (LapCLTP) features. Leng et al. [22] in a different work has proposed an approach that enhance the efficacy of DCT feature extraction, and improve the accuracy for face and palmprint recognition.

In the literature, many works have been done in order to improve the recognition system. By enhancing the different methods used during the recognition process such as feature extraction, feature selection and many more [23, 24]. The monogenic signal was first introduced by Felsberg et al. [12], where a two-dimensional (2-D) generalization of the analytic signal was introduced. This whole monogenic signal was based on the Riesz transform as extension of the Hilbert transform. Several works have been presented in the literature using monogenic signal representation. For instance, Yang et al. [39] proposed a monogenic binary coding (MBC) for local feature extraction. In another work, a feature extraction and representation model for face recognition (MBP) was proposed by Yang et al. [38], where Local Binary Pattern (LBP) for monogenic magnitude encoding was used. Huang et al. [16] presented a facial expression recognition system with monogenic signal representation to obtain local phase orientation and magnitude from the image. Then, local binary pattern (LBP) was

employed to capture the texture and motion information. Whereas, Ohet al. [30] developed a micro-expression recognition system via multi scale monogenic signal representation. Motivated by the higher efficacy of the monogenic signal and feature extraction and representation capability of BSIF, we present, in this paper, a novel scheme named (M-BSIF) for face feature based on monogenic filter and BSIF method. Thus, M-BISF has been used to extract feature from face image. The Monogenic signal representation has proved its great potential in different application domains [15, 17].

Also, BSIF descriptor has been widely utilized in several recognition systems due to its capacity. However, BSIF was not always the best feature descriptor for certain trait [11, 34], i.e., it was not able to attain the best recognition rates. In order to enhance the capability of BSIF method, our proposed feature description scheme takes BSIF to the next level with enhancement yielded by monogenic filters. In this work, a new feature descriptor is proposed called M-BSIF. This new descriptor is realized based on the properties offered by the monogenic signal representation such as local features (local phase, local amplitude and local orientation). In addition, monogenic signal representation is cost effective; it takes less time and space. Also, it can maintain more image information compared to other techniques such as filters those on Gabor wavelets. Specifically, the proposed M-BSIF is realized by, first, applying a band pass mechanism via log-Gabor filter on the image, then a monogenic filter is applied to decompose face image into three complementary parts, i.e., local amplitude, local phase, and local orientation. Next, BSIF is utilized to encode these complementary components in order to extract more efficient features. The used method offers an effective template that represent person in the system, then in matching step it used to decide whether the user is genuine or impostor. The presented framework aims at further enhancing the effectiveness of BSIF features, by utilizing the properties of monogenic signal representation.

The rest of the paper is organized as follows. Section 2 presents outline of the presented system as well as detailed explanation on the proposed feature descriptor. Section 3 outlines the results and a discussion on the findings together with a comparative study. Finally, conclusion is presented Section 4.

2 Proposed M-BSIF based scheme for face recognition system

In this section, we describe the overall M-BSIF based face recognition system, as also shown in Fig. 1. Figure 2 shows the general process of the designed feature descriptor.

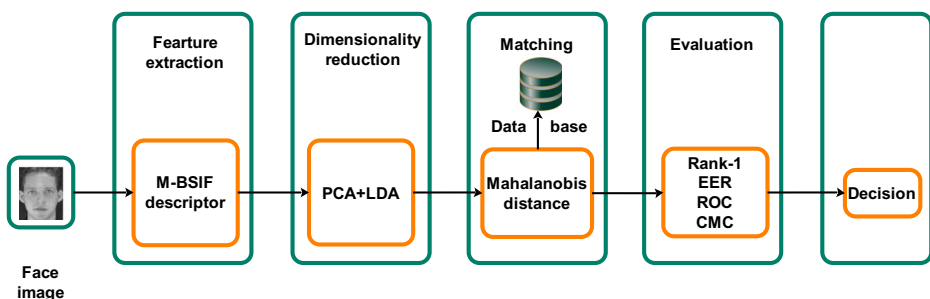


Fig. 1 The general diagram of the proposed system

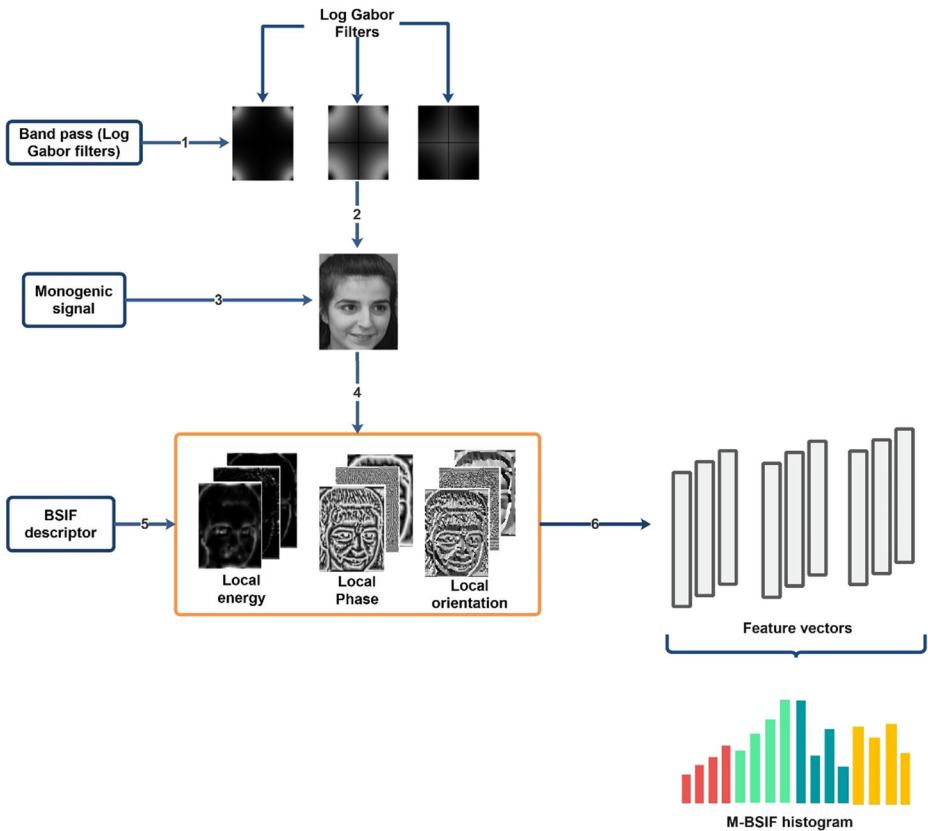


Fig. 2 The procedure of obtaining M-BSIF descriptor

As depicted in Fig. 1, the proposed system is composed of five steps. The first step is feature extraction. In this module, proposed M-BSIF descriptor is employed to extract the features. In the second module, dimensionality reduction is performed using PCA + LDA. In the next step/module, Mahalanobis distance is utilized for matching the face representations. Then, the final decision is represented. Lastly, the evaluation is conducted via four evaluation measures, i.e., Rank-1, EER, ROC and CMC. In the following sub section, each step of the presented system is detailed.

2.1 M-BSIF feature extraction

The proposed M-BSIF feature descriptor's steps are shown in Fig. 2. M-BSIF descriptor has three main steps: (i) band pass using Log-Gabor Filter, (ii) Monogenic signal representation, (iii) BSIF feature extraction. As shown in Fig. 2, before Monogenic signal representation, band-pass (Fig. 2(1)) is applied via Log Gabor filter (Fig. 2(2)) in order to obtain the frequency responses. Then, monogenic signal representation is employed (Fig. 2(3)), that results in (Fig. 2(4)) with three different components, i.e., local amplitude (energy), local phase and local orientation. These three components increase the recognition rate by providing more information. Lastly, BSIF descriptor is used (Fig. 2(5)) to achieve the feature vectors from the

previously resulting components. As shown in (Fig. 2(6)), features are obtained that will help in the enhancement of the systems performance with much relevant information. Finally, each image feature vectors are concatenated and the BSIF histogram is obtained.

2.1.1 Monogenic signal representation

The monogenic, introduced by Felsberg et al. [12], is a two-dimensional (2-D) generalization of the analytic signal. This whole monogenic signal is based on the Riesz transform as an extension of the Hilbert transform. One of the vital characteristics is preserving properties of a 1-D analytic signal while feature extraction. It adds more information such as local amplitude (energy), local phase, and local orientation.

Let $f(x)$ be the real-valued signal and $h(x)$ be the Hilbert transform kernel in the spatial domain, then the analytic signal $f_{anal}(x)$ is defined by:

$$f_{anal}(x) = f(x) + j \cdot h(x)f(x), \tag{1}$$

where $h(x) = 1/\pi x$. Overall, the Hilbert transform is performed in frequency domain and the response of $h(x)$ in frequency domain is $H(\omega) = -j \cdot \text{sign}(\omega) = -j\omega/|\omega|$.

An analytic signal can be decomposed in two elements, i.e., the local amplitude (energy) $A(x)$ and local phase $\phi(x)$ as follows:

$$A(x) = \|f_{anal}(x)\| = \sqrt{f^2(x) + \mathcal{H}^2}, \tag{2}$$

$$\phi(x) = \text{atan2}(\mathcal{H}, f), \phi(x) \in [0, 2\pi[, \tag{3}$$

where $\mathcal{H} = \langle f \rangle \{ f \}$.

To realize the generalization, the Riesz transform is used [35]. It is the scalar-to-vector signal transformation whose frequency response equals to $-j\omega / \|\omega\|$ in 2-D. Therefore, the Riesz transform can be expressed as follow:

$$f_R(s) = \left(\text{0pt}f_x(s)f_y(s) \right) = \left(\text{0pth}_x * f(s)h_y * f(s) \right), \tag{4}$$

where $f(s)$ with $s = (x, y)$ is the input signal. The filters h_x and h_y are characterized by the 2-D frequency responses $H_x = -j\omega_x / \|\omega\|$ and $H_y = -j\omega_y / \|\omega\|$, where $\omega = (\omega_x, \omega_y)$. Now, the spatial representation of Riesz kernel can be obtained as follows:

$$\begin{aligned} h_x &= \frac{x}{2\pi\|s\|^3} \\ h_y &= \frac{y}{2\pi\|s\|^3} \end{aligned} \tag{5}$$

The monogenic signal for an image $f(s)$ is defined as the combination of f and its Riesz transform as follows:

$$f_{\text{monogenic}}(s) = \left(f(s), f_x(s), f_y(s) \right), \tag{6}$$

With $f(s)$ as the real part of the monogenic signal. While, $f_x(s)$ and $f_y(s)$ represent the imaginary part. Thus, the image signal $f(s)$ can be decomposed into three different components, i.e., the local amplitude, local phase and local orientation [12]. The corresponding equation that defines those components is:

$$\begin{cases} A = \sqrt{f^2 + f_x^2 + f_y^2} \\ \phi = -\text{sign}(f_x) \text{atan2}\left(\sqrt{f_x^2 + f_y^2}/f\right) \\ \theta = \text{atan}\left(f_y/f_x\right) \end{cases} \tag{7}$$

where A is the local amplitude or energy of the monogenic signal. While, ϕ represents the local phase or structural information and θ the local orientation or geometric information.

2.1.2 Log Gabor filter

Before applying the Riesz transform, the image must go through band pass filtering in order to maintain the invariance/equivariance characteristics of the signal decomposition [14]. Thus, in this work, we employed log Gabor to perform the band pass filtering as in [39, 13]. The outline of the frequency response of log-Gabor filters is given as:

$$G(\omega) = \exp\left\{-[\log(\omega/\omega_0)]^2/2[\log(\sigma/\omega_0)]^2\right\} \tag{8}$$

In Eq. (8), ω_0 represent the center frequency and σ is the scaling factor of the bandwidth. In [19], it was suggested that it's better to set the ratio σ/ω_0 as constant for filters with constant shape ratio. But, in this work, we use different shape ratios between $[0, 1]$, in order to see the effect of a different ratios on the final results.

The band-pass monogenic signal representation is given as:

$$f_{\text{lg-monogenic}} = \left(f_{\text{lg}}(s), f_{\text{lg-x}}(s), f_{\text{lg-y}}(s) \right) = \left(f_{\text{lg}}(s), h_x * f_{\text{lg}}(s), h_y * f_{\text{lg}}(s) \right), \tag{9}$$

where $f_{\text{lg}}(s) = f(s) * F^{-1}(G(\omega))$, and F^{-1} is the 2-D inverse Fourier transform. The amplitude, phase and orientation of a signal f can be similarly computed [12] as shown in the following equation:

$$\begin{cases} A_{\text{lg}} = \sqrt{f_{\text{lg}}^2 + f_{\text{lg-x}}^2 + f_{\text{lg-y}}^2} \\ \phi_{\text{lg}} = -\text{sign}(f_{\text{lg-x}}) \text{atan2}\left(\sqrt{f_{\text{lg-x}}^2 + f_{\text{lg-y}}^2}/f_{\text{lg}}\right) \\ \theta_{\text{lg}} = \text{atan}\left(f_{\text{lg-y}}/f_{\text{lg-x}}\right) \end{cases} \tag{10}$$

In this work, we adapt a multiscale monogenic representation in order to fully describe the signal. Hence, in a multiscale log-Gabor the parameters ω_0 and σ could be given as:

$$\sigma = \sigma_{\text{ratio}}\omega_0, \omega_0 = (\lambda_{\text{min}}\mu^{s-1})^{-1}, \tag{11}$$

where λ_{min} as minimal wavelength, μ as the multiplication factor of wavelength, and s as the scale index. Figure 2 shows a multiscale monogenic signal representation on 3 scales.

2.1.3 BSIF feature extraction

Binarized Statistical Image Feature (BSIF) descriptor [18] is a textural local descriptor. BSIF uses a set of filters with a fixed size $\phi_i^{k \times k}$ are applied on an image X of size $m \times n$ to obtain binarized responses. The filter results are obtained as in Eq. (12):

$$r_i = \sum_{m,n} \phi_i^{k \times k} X(m, n), \tag{12}$$

Where $\phi_i^{k \times k}$ is a linear filter of size k and $I = \{1, 2, \dots, n\}$ denotes the number of statistically independent filters whose response can be computed together and binarized to obtain the binary string according to Eq. (13):

$$b_i = \begin{cases} 1 & \text{if } r_i > 0 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

Once the binarization is completed, the BSIF features then are obtained as the histogram of pixel’s binary codes that can effectively characterize the texture components in the image. Note that there are two main important factors in BSIF descriptor, i.e., the filter size(k)and the filter length (n).

2.2 Dimensionality reduction

Following the feature extraction phase using M-BSIF, the feature vector obtained is of a high dimension. With the purpose of reducing the vector’s dimension and tackling the large data, we employ a dimensionality reduction technique known as PCA + LDA. This feature reduction technique, due to its speed and simplicity, is considered one of the best ways to deal with large data. To apply this method, first the PCA is used to project the images into a lower data space [36]. Whereas, the LDA is used to maximize intra-class distances and minimize the intra-class distances. The matrix of transformation W can be found as follows for maximizes criterion [7]:

$$T(W) = W_{opt} = \text{Wargmax} \frac{|W^T S_B W|}{|W^T S_W W|} = [W_1 W_2 \dots W_d] \tag{14}$$

$T(W)$ is the fisher discriminant criterionthat is to be maximized. W is built by concatenation of the number d leading Eigenvectors. Note that W is obtained by resolving the following system:

$$S_W^{-1} S_B W_j = W_j \lambda_j, \tag{15}$$

where $j = 1, 2, \dots, d$

2.3 Matching module

For matching two face M-BSIF representations, nearest neighbour was employed as a classifier. In addition, the distance used by this classifier is Mahalanobis distance. Let two vectors V_i and V_j represent the feature vectors of query and images stored in database (i.e., template), respectively. The distance between V_i and V_j is calculated by Eq. (16):

$$d_{Ma}(V_i, V_j) = (V_i - V_j)^T C^{-1} (V_i - V_j) \quad (16)$$

where C represents the covariance matrix

3 Experiment results and discussion

In order to evaluate the proposed system, a series of experiments were conducted on three different public databases, i.e., ORL database, AR database and JAFFE database. Each database was used in three different sub experiments where different shape ratio as well as different scales were used.

3.1 Experiment I: ORL database

In this section, we report the results of the experiment conducted on the ORL database. Olivetti Research Laboratory (ORL) [8] database is one of the most used databases in face recognition literature. ORL is a publicly available database, which contains 40 subjects with 10 images per subject. In total, it has 400 images of a size 112×92 per image. ORL database images are grayscale with pose and slight light variations with small detailed facial changes. Moreover, the images were taken against uniform background.

This experiment's results are presented in Tables 2, 3 and 4 by setting scale index s to 1, 2 and 3, respectively, and shape ratio between $[0, 1]$. The tables report the results in terms of recognition rate (Rank-1) and Error Equal Rate (EER). In addition, for a more insightful results representation, Fig. 3 shows the Receiver Operating Characteristic (ROC), Cumulative Match Characteristic (CMC) and Error Equal Rate (EER) curves, where:

- CMC curve represents the performance of a face recognition system in the identification mode. ROR (Rate Of Recognition) is information about the percentage of traits, which represents the closest match in the stored database corresponding to the right identity, i.e., identity being among the top r ranked matches with $r = 1, 2, \dots, n$; n stands for the number of persons in the database. The CMC plot of the ranking ability in an identification system is regarded as to the rate of recognition.
- ROC Curve represents a graphical illustration of the evolution of the False Rejection Rate (FRR) against False Acceptance Rate (FAR) for all possible operating thresholds.
- EER is an important metric used to evaluate a biometric system, it is where the FAR and FRR are equal. Lower the EER is better the system is.

The parameters wavelength λ_{min} , multiplication factor μ and the number of scales used in all experiment initially were chosen randomly, because a fixed set of parameters that can produce

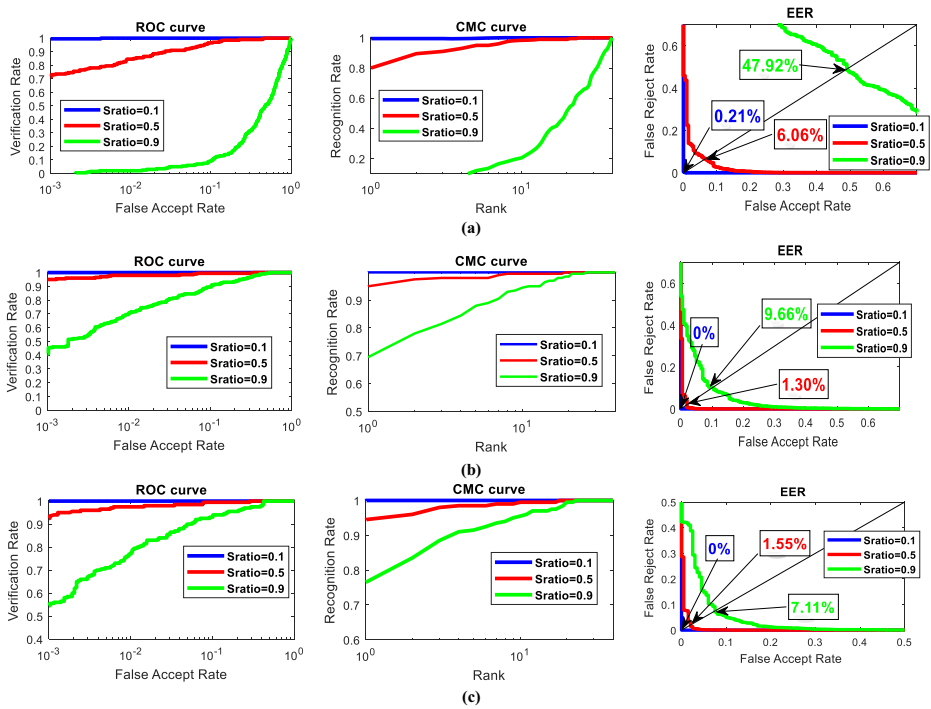


Fig. 3 ORL database- (a) ROC, CMC and EER for Nb scale = 1, (b) ROC, CMC and EER for Nb scale = 2, (c) ROC, CMC and EER for Nb scale = 3. (ROC helps with the visualization of the performance at different threshold, CMC curve represents the performance of a face recognition system in the identification mode)

the best performance is to be determined. Table 1 shows the parameters used in each table for each experiment.

Table 2 reports the performance on recognition rate Rank-1 and EER using 1 scale. In Table 2, we can that the best results belong to the one with shape ratio equals to 0.1, where the Rank-1 equals to 99.5% and EER equals to 0.21%. Overall, while gradually decreasing the shape ratio, we can see the increase from 4.5% for Rank-1 and 47.92% for EER with shape ratio 0.9 to 99.5% for Rank-1 and 0.21% EER for shape ratio 0.1. The big increase can be seen between 0.9 and 0.8 shape ratio, and between 0.5 and 0.4 shape ratio.

The results of Rank-1 and EER using 2 scales reported in Table 3. In from Table 3, the best results belong to the one with shape ratio equals to 0.1, where the Rank-1 equals to 100% and EER equals to 0%, as well as shape ratio 0.2, where the Rank-1 equals to 99% and EER equals to 0.12%. The reported results in Table 3 are better than in Table 2. While gradually decreasing the shape ratio, we can observe an increase in Rank-1 and a decrease in EER. Compared to

Table 1 Parameters set

Table number	λ_{min}	μ	s	Center frequency
2, 5, 8	1	1	1	[1]
3, 6, 9	1.4	2	2	[1.4, 2.8]
4, 7, 10	1	1.4	3	[1.4, 1.960, 2.744]

Table 2 Results for number of scales = 1 and frequency center = [1]

Shape ratio	Rank-1	EER
0.9	4.50%	47.92%
0.8	12.50%	34.59%
0.7	42.50%	20.43%
0.6	66.00%	9.69%
0.5	80.00%	6.06%
0.4	92.00%	2.37%
0.3	97.50%	0.56%
0.2	99.00%	0.17%
0.1	99.50%	0.21%

Table 2, the reported results in this experiment (Table 3), the performance starts at high for shape ratio of 0.9, where the Rank-1 equals to 69.5% while EER equals to 9.66%. Also, performances start improving from the shape ration 0.8, where 81% was given in Rank-1 and 6.30% in EER, which is better results compared to the results reported in Table 2 for the same shape ratio.

Table 4 presents the results for the experiment with 3 scales. We can notice that the best results belong to the last three shape ratios 0.3, 0.2 and 0.1, which were able to give 99.5% and 100% Rank-1, and 0.09%, 0.08%, and 0% EER, respectively. In this experiment, like the previous two tables, the results keep improving while decreasing the shape ratios. In Table 4, results start picking up from 0.8 shape ratio. Now comparing those three sub-experiments, the reported results improve highly using 3 scales.

In order to support those findings, the results are presented in terms of ROC, CMC and EER curves in Fig. 3. Moreover, we did select three shape ratios from each table to plot, i.e., 0.1, 0.5 and 0.9 in order to clearly observe and compare the results. In Fig. 3, we can see that the proposed descriptor is highly effective, and we can also see the impact of different shape ratios as well as the scales. We can also confirm that the best results are reported to the shape ratio to 0.1. It is worth noting that using higher number of scales improve the results overall.

3.2 Experiment II: AR face database

In this section, we report the results of experiments on AR database. First, AR face database [27] is a publicly available database. It contains 4000 color images of 126 individual. The images are of frontal views with different expressions, lighting conditions and occlusions

Table 3 Results for number of scales = 2 and frequency center = [1.4, 2.8]

Shape ratio	Rank-1	EER
0.9	69.5%	9.66%
0.8	81.00%	6.30%
0.7	90.00%	3.50%
0.6	93.00%	1.80%
0.5	95.00%	1.30%
0.4	97.50%	0.69%
0.3	99.00%	0.15%
0.2	99.00%	0.12%
0.1	100%	0.00%

Table 4 Results for number of scales = 3 and frequency center = [1.4, 1.960, 2.744]

Shape ratio	Rank-1	EER
0.9	76.50%	7.11%
0.8	85.00%	5.01%
0.7	89.00%	3.07%
0.6	94.00%	2.03%
0.5	94.50%	1.55%
0.4	97.00%	0.67%
0.3	99.50%	0.09%
0.2	99.50%	0.08%
0.1	100%	0.00%

(scarf and sun glasses). The pictures were taken in two sessions in two weeks with 26 images per subject. In this study, we considered the first 100 individual, i.e., 2600 images in total.

This experiment results are presented in Tables 5, 6 and 7 by setting scale index s to 1, 2 and 3, respectively, while varying shape ratio parameter between $[0, 1]$. The tables report the results in terms of recognition rate (Rank-1) and Error Equal Rate (EER). For a more insightful results representation, Fig. 4 shows the ROC, CMC and EER curves. The parameters wavelength λ_{min} and multiplication factor μ are same as in the previous experiment Table 1.

Table 5 reports the findings as recognition rate Rank-1 and EER using 1 scale. In Table 5, the best reported results belong to the one with shape ratio equals to 0.1, where the Rank-1 equals to 97.1% and EER equals to 0.93%. Overall, in Table 5, performance start pretty low, however gradually, while decreasing the shape ratio, we can see the improvement in terms of the recognition rate from 0.80% for Rank-1 and 47.47% for EER to 0.95% for Rank-1 to

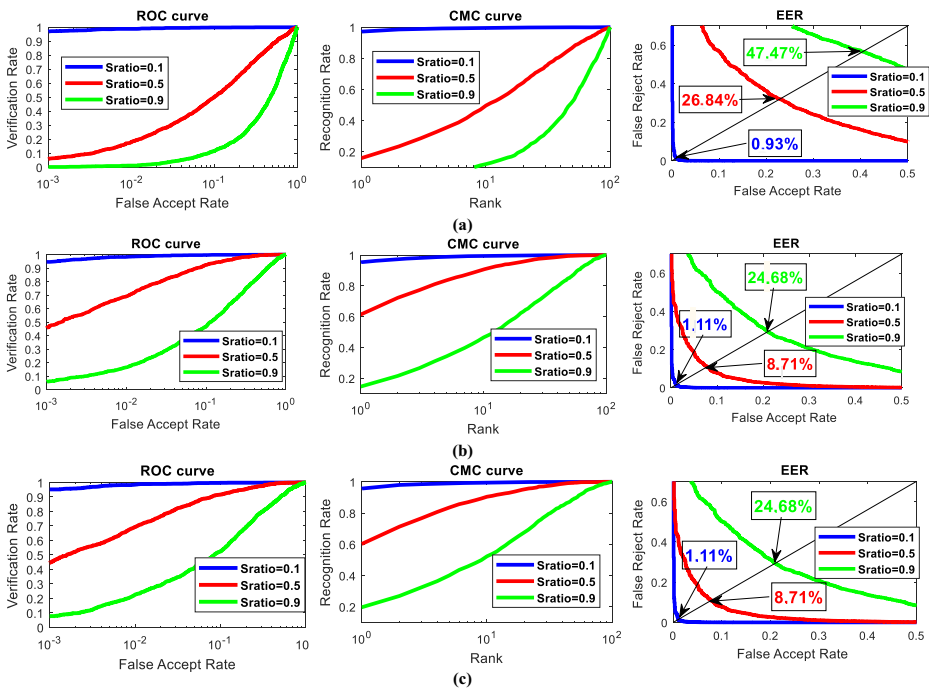


Fig. 4 AR database- (a) ROC, CMC and EER for Nb scale = 1, (b) ROC, CMC and EER for Nb scale = 2, (c) ROC, CMC and EER for Nb scale = 3

Table 5 Results for number of scales = 1 and frequency center = 1

Shape ratio	Rank-1	EER
0.9	0.80%	47.47%
0.8	3.30%	42.61%
0.7	6.60%	36.06%
0.6	10.50%	31.19%
0.5	15.80%	26.84%
0.4	32.30%	17.69%
0.3	78.90%	5.20%
0.2	93.90%	1.84%
0.1	97.10%	0.93%

97.1% and 0.93% for EER with shape ratio 0.1. The dramatic increase can be seen between 0.4 and 0.3 shape ratios, where the recognition rate Rank-1 jumps from 32.3–78.9% and EER from 17.96 to 5.20%.

Table 6 represents the results of recognition rate Rank-1 and EER using 2 scales. It is apparent from Table 6 that the best results belong to the one with shape ratio equals to 0.2, where the Rank-1 equals to 97.60% and EER = 0.81%. Overall, the reported results (like in Table 5), gradually, while decreasing the shape ratio, increase in Rank-1 and a decrease in EER. Compared to Table 5, the reported results in this experiment start of higher level, where the shape ratio of 0.9 gives 14.6% for Rank-1 and EER equals to 28.05%, which is a remarkable improvement. Moreover, the results start picking up and improvements start at shape ratio 0.4. Overall, the results in this table compared to the results reported in Table 5 are better, while the best reported results are similar.

The Table 7 provides results belonging to the experiment using 3 scales. We can observe that the best results are attained by the shape ratios = 0.2. It was able to yield 97.6% Rank-1 and EER = 0.82%. As in previous two tables, in this experiment as well the results keep improving while decreasing the shape ratios. We can see in Table 7 that the results start picking up from 0.4 shape ratio. Now comparing those three sub-experiments, we can observe the best performance in Table 7. Regardless of best reported results, the improvement in the performance overall was better while using 3 scales.

To further support our findings, the results are presented in terms of ROC, CMC and EER curves in Fig. 4. Also, we selected three shape ratios from each table, i.e., 0.1, 0.5 and 0.9 in order to clearly observe and compare the results. Therefore, from Fig. 4, we can see that the proposed descriptor is highly effective, and we can also see the impact of different shape ratios as well as the scales. We can also confirm that the best results are reported to the shape ratio to 0.1. It is worth noting that using higher number of scales improve the results overall.

Table 6 Results for number of scales = 2 and frequency center = [1.4, 2.8]

Shape ratio	Rank-1	EER
0.9	14.60%	28.05%
0.8	20.30%	23.02%
0.7	25.40%	20.41%
0.6	40.60%	15.42%
0.5	61.40%	8.71%
0.4	84.10%	3.95%
0.3	95.20%	1.40%
0.2	97.60%	0.81%
0.1	95.30%	1.23%

Table 7 Results for number of scales = 3 and frequency center = [1.4, 1.960, 2.744]

Shape ratio	Rank-1	EER
0.9	19.60%	24.68%
0.8	21.40%	22.23%
0.7	27.70%	20.05%
0.6	41.60%	15.20%
0.5	60.20%	8.71%
0.4	85.10%	4.08%
0.3	95.40%	1.26%
0.2	97.50%	0.82%
0.1	95.60%	1.11%

3.3 Experiment III: JAFFE face database

Japanese Female Facial Expressions (JAFFE) database [26] is another publicly available benchmark dataset that has been used for evaluation of different face recognition systems. This database has total of 213 images posed by ten subjects (Japanese female) while maintaining six different facial expressions plus a neutral expression. The images are in the gray level and the size of each image is 256×256 pixels.

This experiment results using JAFFE database are presented in Tables 8, 9 and 10 by setting scale index s to 1, 2 and 3, respectively, while varying shape ratio parameter between $[0, 1]$. The tables report the results in terms of recognition rate (Rank-1) and Error Equal Rate (EER). For a more insightful results representation, Fig. 5 shows the ROC, CMC and EER curves. The parameters wavelength λ_{min} and multiplication factor μ are same as in the previous two experiments (please see Table 1).

Table 8 represents the results of recognition rate Rank-1 and EER using 1 scale. It is apparent from Table 8 that the best results belong to the one with shape ratios from 0.5 to 0.1, where the Rank-1 equals to 100% and EER = 0%. Overall, the reported results start pretty low, but gradually improve within only 3 first shape ratios. We can also notice that the increase can be seen between 0.7 and 0.6 shape ratios, where the recognition rate Rank-1 jumps from 77 to 94% in recognition rate Rank-1 and from 7.22 to 2.72% for EER. All in all, the results reported in this experiment are excellent even only one scale is used.

Table 9 provides the results of recognition rate Rank-1 and EER using 2 scales. It is apparent from Table 9 that the all attained results are good, where the Rank-1 = 100% and EER = 0% on the shape ratios from 0.7 to 0.1. Also, even shape ratios 0.9 and 0.8 give good

Table 8 Results for number of scales = 1 and frequency center = [1]

Shape ratio	Rank-1	EER
0.9	16.00%	41.94%
0.8	43.00%	25.77%
0.7	77.00%	7.22%
0.6	94.00%	2.72%
0.5	100%	0.05%
0.4	100%	0.00%
0.3	100%	0.00%
0.2	100%	0.00%
0.1	100%	0.00%

Table 9 Results for number of scales = 2 and frequency center = [1.4, 2.8]

Shape ratio	Rank-1	EER
0.9	96%	0.05%
0.8	99%	1.27%
0.7	100%	0.00%
0.6	100%	0.00%
0.5	100%	0.00%
0.4	100%	0.00%
0.3	100%	0.00%
0.2	100%	0.00%
0.1	100%	0.00%

performances of Rank-1 equals to 96% and 99% and EER equals to 0.05 and 1.27%, respectively, which is a remarkable improvement compared to Table 8.

Table 10 the experiment using 3 scales. We can notice that all achieved results are excellent from 0.9 to 0.1 shape ratio, where the results are 99% for Rank-1 and 1% for EER. These performances are interestingly good, even better than previous two tables, therefore the results overall are better using 3 scales.

Lastly, for further more support the findings, the results are presented in terms of ROC, CMC and EER curves in Fig. 5. Like the previous experiments, we selected three shape ratios (i.e., 0.1, 0.5 and 0.9) from each table to plot in order to clearly observe and compare the results. In Fig. 5, we can see that, first the proposed descriptor is highly effective as well as similarly to the previous experiment the impact of different shape ratios as well as the scales. In this experiment, all results are remarkable regardless of the shape ratio; especially in the last experiment.

All in all, the results from different experiments indicate that the proposed feature extraction scheme is highly effective as shown in above reported tables and figures. Various effects of the shape ratio, number of scales and the center of frequency have been studied to show how performances of a face recognition system changes by varying these three parameters. Therefore, the best set of parameters according to our study is shape ratio = 0.1, and number of scales = 3.

3.4 Comparative study

In order to further demonstrate the effectiveness of the proposed system, a comparative study was conducted with same system using only BSIF feature extraction technique on three datasets.

Table 10 Results for number of scales = 3 and frequency center = [1.4, 1.960, 2.744]

Shape ratio	Rank-1	EER
0.9	99.00%	1.00%
0.8	100%	0.00%
0.7	100%	0.00%
0.6	100%	0.00%
0.5	100%	0.00%
0.4	100%	0.00%
0.3	100%	0.00%
0.2	100%	0.00%
0.1	100%	0.00%

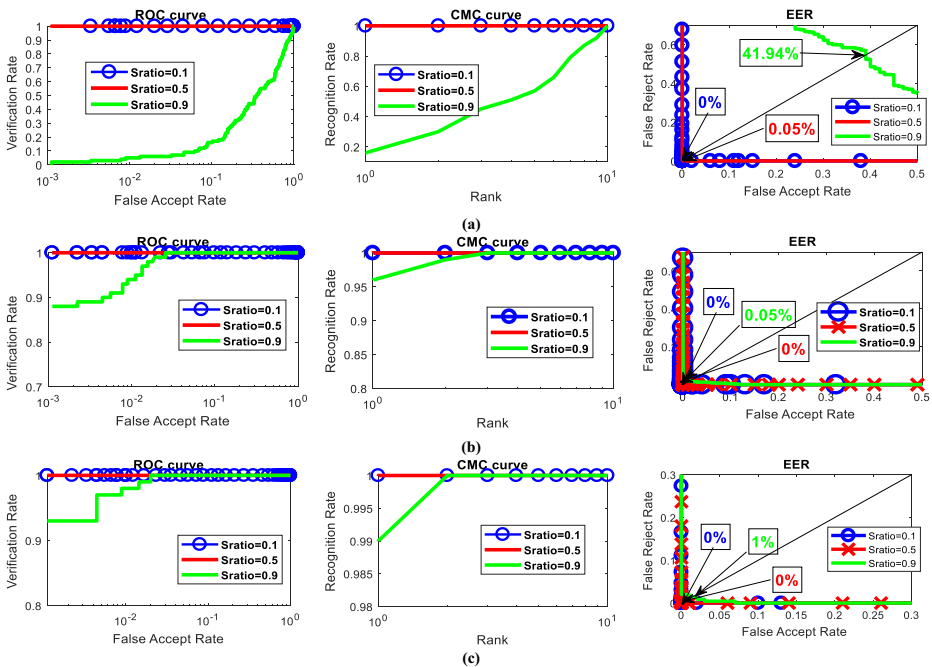


Fig. 5 JAFFE database- (a) ROC, CMC and EER for Nb scale = 1, (b) ROC, CMC and EER for Nb scale = 2, (c) ROC, CMC and EER for Nb scale = 3

As shown in Table 11, compared to only BSIF feature descriptions [25], the proposed system is highly effective. For example, we can clearly see the huge difference between the BSIF only features based system and the proposed M-BSIF based system, which attained 49.84% and 0.81% EER, respectively, using AR dataset. While, for ORL database the proposed system was able to obtain 4.00% recognition rate and 100% recognition rate while using only BSIF and M-BSIF systems, respectively. Similarly, on JAFFE database, the only

Table 11 Comparative study of proposed method with existing works

Reference	Database	Recognition rate Rank-1%	EER %
The proposed work			
M-BISF	ORL	100%	0.00%
	AR	97.60%	0.81%
	JAFFE	100%	0.00%
BSIF feature extraction	ORL	4.00%	48.82%
	AR	1.00%	49.84%
	JAFFE	10.00%	49.56%
Kumar et al.[20]	ORL	100%	0.00%
Ouyang et al. [32]	ORL	91.31%	-----
Annamalai [3]	ORL	98.92	-----
Yan et al. [37]	ORL	99.82	0.1%
Cheng et al.[9]	AR	95.85%	-----
Yan et al. [37]	AR	99.78%	0.3%
Yee et al. [40]	JAFFE	99.24%	-----
Munikrishna et al. [29]	JAFFE	100%	0.00%

BSIF and M-BSIF based frameworks yielded 10.00% and 100% recognition rates, and 49.56% and 0.00% EERs, respectively. While compared to [20] it did similarly good where both were able to yield a 100% recognition rate for ORL database. Further our proposed system was able to outperform [9] in terms of recognition rate for the AR database, where our system was able to obtain 97.60% while [9] obtained 95.85% and close results with [37] were this later was able to obtain 99.78%. In addition, our system was also able to give good results like [40] and [29] and that using JAFFE database where our system yield 100%. Similarly MBSIF was also able to outperform the recent work [32], while using the ORL database were this work was only able to attain 91.31% on recognition rate while using M-BSIF was able to obtain 100%. MBSIF was also able to perform better then [3] using ORL database even this later was able to obtain a good accuracy of 98.92%.

4 Conclusions

In this paper, a new descriptor M-BSIF was proposed for face feature extraction. The method is based on monogenic signal representation and BSIF descriptor. More specifically, the proposed face feature descriptor enhances the ability of the traditional BSIF feature extraction in face recognition system. The presented framework first applies a band pass scheme using log-Gabor filter, then the monogenic filter is used to obtain a set of three different components (i.e., local energy, local phase and local orientation) from each face image. Then, BSIF feature extraction technique is employed on the previously obtained components. The histograms code extracted from the component images are concatenated into a larger feature vector. Next, PCA + LDA technique is used to reduce the dimensionality of the feature vector. Finally, the cosine Mahalanobis distance is utilized to verify whether the user is genuine or impostor. The experiments were conducted on three public different datasets, i.e., ORL, AR and JAFFE. The experimental analyses show the high efficacy of the proposed M-BSIF feature extraction in face recognition system. Also, the proposed M-BSIF was able to attain high accuracy compared to existing schemes in the state-of-the-art. All in all, we can conclude that the proposed M-BSIF descriptor can be an efficient tool for feature extraction in face recognition, therefore we aspire to use it with different modalities. As a future work, we aspire to find the best set of parameters that can guarantee the quality of the findings on larger datasets. Also, we will study the robustness of the frameworks against different adversaries [1, 2].

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Akhtar Z, Alfarid N (2011) Secure learning algorithm for multimodal biometric systems against spoof attacks. In: International conference on information and network technology (IPCSIT), pp 1–6
2. Akhtar Z, Rattani A (2017) A face in any form: new challenges and opportunities for face recognition technology. *Computer* 50(4):80–90
3. Annamalai P (2020) Automatic face recognition using enhanced firefly optimization algorithm and deep belief network. *Int J Intell Eng Syst* 13(5):19–28

4. Attia A, Chaa M, Akhtar Z, Chahir Y (2020) Finger knuckle patterns based person recognition via bank of multi-scale binarized statistical texture features. *Evol Syst* 11(4):625–635
5. Attia A, Akhtar Z, Chahir Y (2020) Feature-level fusion of major and minor dorsal finger knuckle patterns for person authentication. *Signal, Image and Video Processing*, pp 1–9
6. Attia A, Akhtar Z, Chalabi NE, Maza S, Chahir Y (2021) Deep rule-based classifier for finger knuckle pattern recognition system. *Evol Syst* 12(4):1015–1029
7. Belhumeur PN, Hespanha JP, Kriegman DJ (1997) Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans Pattern Anal Mach Intell* 19(7):711–720
8. Cambridge AL (2002) The database of faces. Available: <https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>
9. Cheng E-J et al (2019) Deep sparse representation classifier for facial recognition and detection system. *Pattern Recognit Lett* 125:71–77
10. Doyle JS, Bowyer KW (2015) Robust detection of textured contact lenses in iris recognition using BSIF. *IEEE Access* 3:1672–1683
11. Dronky MR, Khalifa W, Roushdy M (2021) Using residual images with BSIF for iris liveness detection. *Expert Syst Appl* 182:115266
12. Felsberg M, Sommer G (2001) The monogenic signal. *IEEE Trans Signal Process* 49(12):3136–3144
13. Field DJ (1987) Relations between the statistics of natural images and the response properties of cortical cells. *JOSA A* 4(12):2379–2394
14. Granlund GH, Knutsson H (2013) *Signal processing for computer vision*. Springer Science & Business Media, Berlin
15. Han Y, Yu N (2021) Synthetic aperture radar target recognition based on joint classification of selected monogenic components by nonlinear correlation information entropy. *J Appl Remote Sens* 15(2):26502
16. Huang X, Zhao G, Zheng W, Pietikainen M (2012) Spatiotemporal local monogenic binary patterns for facial expression recognition. *IEEE Signal Process Lett* 19(5):243–246
17. Ji X, Sun Y (2021) SAR image target recognition based on monogenic signal and sparse representation. *Wirel Commun Mob Comput* 2021
18. Kannala E, Juho, Rahtu (2012) Bsif: Binarized statistical image features. *International conference on pattern recognition*, IEEE, pp 1363–1366
19. Kovési P (1996) *Invariant measures of image features from phase information*. University of Western Australia
20. Kumar D, Garain J, Kisku DR, Sing JK, Gupta P (2020) Unconstrained and constrained face recognition using dense local descriptor with ensemble framework. *Neurocomputing* 408:273–284
21. Leng L, Zhang J (2013) Palmhash code vs. palmphasor code. *Neurocomputing* 108:1–12
22. Leng L, Zhang J, Khan MK, Chen X, Alghathbar K (2010) Dynamic weighted discrimination power analysis: a novel approach for face and palmprint recognition in DCT domain. *Int J Phys Sci* 5(17):2543–2554
23. Leng L, Zhang J, Chen G, Khan MK, Alghathbar K (2011) Two-directional two-dimensional random projection and its variations for face and palmprint recognition. In: *International conference on computational science and its applications*, pp 458–470
24. Leng L, Zhang S, Bi X, Khan MK (2012) Two-dimensional cancelable biometric scheme. In: *2012 International Conference on Wavelet Analysis and Pattern Recognition*, pp 164–169
25. Leng L, Li M, Kim C, Bi X (2017) Dual-source discrimination power analysis for multi-instance contactless palmprint recognition. *Multimed Tools Appl* 76(1):333–354
26. Lyons M, Akamatsu S, Kamachi M, Gyoba J (1998) Coding facial expressions with gabor wavelets. In: *Proceedings Third IEEE international conference on automatic face and gesture recognition*, pp 200–205
27. Martinez AM (1998) *The AR face database*. CVC Technical Report 24
28. Mishra A, Agrawal R, Khan MA, Jalal AS (2019) A robust approach for palmprint biometric recognition. *Int J Biometrics* 11(4):389–408
29. Munikrishna DC, Raja KB, Venugopal KR (2018) Spatial domain face recognition system using convolution of PDV and LBP. In: *International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, vol 3, pp 321–326
30. Oh Y-H, Le Ngo AC, See J, Liong S-T, Phan RC-W, Ling H-C (2015) Monogenic Riesz wavelet representation for micro-expression recognition. In: *IEEE International Conference on Digital Signal Processing (DSP)*, pp 1237–1241
31. Ouamane A, Benakcha A, Belahcene M, Taleb-Ahmed A (2015) Multimodal depth and intensity face verification approach using LBP, SLF, BSIF, and LPQ local features fusion. *Pattern Recognit Image Anal* 25(4):603–620
32. Ouyang A, Liu Y, Pei S, Peng X, He M, Wang Q (2020) A hybrid improved kernel LDA and PNN algorithm for efficient face recognition. *Neurocomputing* 393:214–222

33. Raja KB, Raghavendra R, Busch C (2014) Binarized statistical features for improved iris and periocular recognition in visible spectrum. In: 2nd International Workshop on Biometrics and Forensics, pp 1–6
34. Sharma D, Selwal A (2021) HyFiPAD: a hybrid approach for fingerprint presentation attack detection using local and adaptive image features. *Vis Comput* 1–27
35. Stein EM, Weiss G (2016) Introduction to fourier analysis on euclidean spaces (PMS-32), vol 32. Princeton University Press, Princeton
36. Turk M, Pentland A (1991) Eigenfaces for recognition. *J Cogn Neurosci* 3(1):71–86
37. Yan K, Huang S, Song Y, Liu W, Fan N (2017) Face recognition based on convolution neural network. In: 36th Chinese Control Conference (CCC), pp 4077–4081
38. Yang M, Zhang L, Zhang L, Zhang D (2010) Monogenic binary pattern (MBP): A novel feature extraction and representation model for face recognition. In: 20th International Conference on Pattern Recognition, pp 2680–2683
39. Yang M, Zhang L, Shiu SC-K, Zhang D (2012) Monogenic binary coding: An efficient local feature extraction approach to face recognition. *IEEE Trans Inf Forensics Secur* 7(6):1738–1751
40. Yee SY, Rassem TH, Mohammed MF, Awang S (2020) Face recognition using Laplacian completed local ternary pattern (LapCLTP). In: *Advances in Electronics Engineering*. Springer, Berlin, pp 315–327
41. Ylioinas J, Kannala J, Hadid A, Pietikäinen M (2015) Face recognition using smoothed high-dimensional representation. In: *Scandinavian conference on image analysis*, pp 516–529
42. Younesi A, Amirani MC (2017) Gabor filter and texture based features for palmprint recognition. *Procedia Comput Sci* 108:2488–2495

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