Applied Computing and Informatics 15 (2019) 163–171

Contents lists available at ScienceDirect

Applied Computing and Informatics

journal homepage: www.sciencedirect.com

Original Article

Local binary patterns based on directional wavelet transform for expression and pose-invariant face recognition

Mohd. Abdul Muqeet [a,*], Raghunath S. Holambe [b]

[a] Electrical Engineering Department, Muffakham Jah College of Engineering and Technology, Hyderabad, TS 500034, India
[b] Department of Instrumentation Engineering, S.G.G.S Institute of Engineering and Technology, Nanded, Maharashtra 431606, India

ARTICLE INFO

Article history:
Received 1 July 2017
Revised 21 October 2017
Accepted 15 November 2017
Available online 13 December 2017

Keywords:
Face recognition
Directional wavelet transform (DIWT)
Local binary patterns (LBP)

ABSTRACT

Face image variations such as expression and pose radically increase the intra-class variations which affect the performance of feature extraction methods. It is desirable to extract more robust local discriminative features to effectively represent such face variations. This paper proposes a novel facial feature extraction method which utilizes interpolation-based directional wavelet transform (DIWT) and local binary patterns (LBP). An efficient direction assessment method based on quadtree partitioning is implemented to facilitate adaptive direction selection in the local regions from the face images to obtain DIWT sub-bands. LBP histogram features are extracted from selected top-level DIWT sub-bands to obtain local descriptive feature set. Experimental results are simulated on ORL database, GT database, and FEI database. A comparison with various contemporary methods which include holistic, local descriptors and LBP-based non-adaptive multiresolution analysis methods illustrate the efficacy of the proposed method.

1. Introduction

In the security-heightened world, face recognition has emerged as a popular biometric recognition method with non-intrusive and low precision equipment for image acquisition. Prominent holistic face recognition methods are PCA [1], LDA [2], and locality preserving projections (LPP) [3]. Local descriptors [4–7] are attaining much interest due to their ease of implementation and robustness to noise. Local binary patterns (LBP) [4] are effectively applied for facial feature extraction owing to its texture discrimination capability and computational efficiency [5]. Jabid et al. [6] proposed a robust local facial feature descriptor based on local directional patterns (LDP) and confirmed good results compared to LBP. But LDP exhibit sensitivity to noise and rotation. Weber local descriptors (WLD) [7] are more descriptive and computationally efficient than LBP. Recently, Zhang et al. [8] used the WLD to extract local facial features at predefined facial landmarks and effectively captured pose-invariant features. Ullah et al. [9] applied WLD for gender classification from face images. Although such local descriptors achieve higher recognition results than holistic methods, their performance restricts due to variations in expression, pose, illumination and occlusion.

DWT-based methods are increasingly admired for face recognition and assure good recognition results with certain modification in the way the DWT sub-bands are utilized [10–12]. But due to isotropic property and less directionality, these methods do not assure superior recognition results for moderate to extreme variations in expression and pose. Numerous multiresolution analysis (MRA) methods with anisotropic scaling and better directionality are combined with LBP to extract MRA-based local descriptive features. Local Gabor binary patterns (LGBP) [13] combine the Gabor filters with LBP and provide high accuracy at the cost of high dimensionality histogram features. Arousi et al. [14] presented a face recognition technique under illumination-variant conditions using the SPT decomposition and LBP. Recently, Zou et al. [15] evolved a new feature descriptor based on the curvelet transform and LBP. Here, authors used the LBP coded image of low-frequency approximation sub-band and mid-frequency subbands to form the feature set. These MRA [10–12] and LBP-based MRA methods [13–15] despite capturing the directional information lack the adaptation in selecting directions based on image characteristics. Moreover, suffer from various issues such as sub-band selection to form an efficient feature set, a high computational rate,

* Corresponding author at: Electrical Engineering Department, Muffakham Jah College of Engineering and Technology, Banjara Hills, Hyderabad 500034, TS, India.
E-mail address: ab.muqeet2013@gmail.com (M.A. Muqeet).
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.aci.2017.11.002
2210-8327/

and complex filter design. Face variations mainly affect the edge manifolds of an image and if directional information and orientation of such edge manifolds can be approximated adaptively it will significantly improve the performance.

Adaptive directional lifting wavelet transforms [16–18] due to the benefit of directional lifting and adaptation in the direction selection as per characteristics of the image have been effectively used for image compression. Chang et al. [16] proposed a direction adaptive discrete wavelet transform (DA-DWT) for image compression. Only one pair of lifting step and non-interpolated distant integer samples are used to realize directional lifting. Here Neville filters [21] are used as the prediction and update filters. Ding et al. [17] suggested adaptive directional lifting (ADL) for image coding and used interpolated fractional samples. ADL method could be implemented with one or two pairs of lifting steps with 5/3 or 9/7 tap filters [20]. Maleki et al. [18] proposed directional wavelets (DIW) with megaquad partitioning algorithm in the adaptive directional lifting based framework to efficiently capture edge features.

Furthermore, due to lifting based factorization, perfect reconstruction is also assured and the resultant multiresolution image is completely compatible with that of the conventional 2-D DWT multiresolution image. Recently, Muqeet and Holambe proposed a face recognition method [19] using DA-DWT [16] for face verification and established its effectiveness compared to various famous subspace and non-adaptive MRA methods.

To effectively deal with the expression and pose variation in face images this paper consider the underlying theory of DIW and proposes interpolation-based directional wavelet transform (DIWT). In the implementation of DIWT, we used quadtree partitioning instead of megaquad partitioning which is used in DIW [18]. LBP is applied to the selected top-level DIWT sub-bands which eventually extract the descriptive histogram features. Finally, LDA is applied to obtain discriminant features in reduced space. Nearest Neighbor (NN) [2] classifier is used in an identification process. The outline of the paper is depicted as follows. A review of the related methodology is explained in Section 2 including the DIWT and LBP. Further, in Section 3, proposed facial feature extraction method is explained. In Section 4 experimental results are performed on three eminent face databases and results of the proposed method are compared with various contemporary methods. Conclusions based on the method are presented in Section 5.

2. Material and methods

The primary motivation of the proposed method is to obtain efficient MRA-based local descriptive features from the face images. The subsections will illustrate the DIWT with the improved quadtree partitioning method, and review LBP and its importance in proposed facial feature extraction method.

2.1. Directional wavelet transform (DIWT)

The essence of directional wavelet transform (DIWT) is to perform the lifting wavelet transform steps at a viable variety of directions while preserving the characteristics of multiresolution, localization, and isotropy intact. The DIWT still performs the conventional lifting wavelet transform but consist of modification in the prediction and update step. The 2-D DIWT can be realized with only one pair of lifting step, which means only one prediction step followed by one update step.

Let $X = \{(i,j)|i,j \in \mathbb{Z}^2\}$ signify an image defined on a 2-D orthogonal sampling grid. Let $x^e = x(2i,j)$ and $x^o = x(2i+1,j)$ represent even and odd samples respectively. Here we consider row sub-sampling. Column sub-sampling follows the same relations with exchange of row and column variables.

In the prediction step, odd samples $x^o(i,j) \in X_o$ are predicted from even samples $x^e(i,j) \in X_e$. The prediction of each odd sample is a linear combination of neighboring even samples with a strong correlation. We express that the pixels have a strong correlation in the direction d. Here samples from six even rows are selected to take part in prediction step given in (1) and the generated high-pass sub-band $H(i,j)$ is depicted in (2),

$$P(x^o(i,j)) = \sum_{n=-N_p}^{N_p-1} K_n^p \cdot x^e(i+n, j+ntan^{-1}d) \tag{1}$$

$$H(i,j) = g_H \cdot [x^o(i,j) - P(x^o(i,j))] \tag{2}$$

where $2N_p$, K^p, and g_H are the length of the prediction filter, the coefficient of the prediction filter, and scaling factor respectively. Now in update step, even samples $x^e(i,j)$ are updated from odd row samples of high pass signal along the same optimal direction d. Similarly samples from six nearest even rows take part in update step. The update step and the low-pass sub-band $L(i,j)$ are given in (3) and (4) respectively,

$$U(x^e(i,j)) = \sum_{n=-N_u}^{N_u-1} K_n^u \cdot \begin{bmatrix} x^o(i+n, j+ntan^{-1}d) - \\ P(x^o(i+n, j+ntan^{-1}d)) \end{bmatrix} \tag{3}$$

$$L(i,j) = g_L \cdot [x^e(i,j) + g_H^{-1} \cdot [U(H(i,j))]] \tag{4}$$

where $2N_u$, K^u, and g_L are the length of the update filter, the coefficient of the update filter, and scaling factor respectively. Neville filters [21] with six vanishing moments are used for prediction and update filter [16], i.e. $N_p = N_u = 3$. These filters possess linear phase characteristics with high vanishing moments which increase their texture discrimination capability. From [21], the coefficients of the prediction filter can be obtained as $K_n^p = [3, -25, 150, 150, -25, 3]/2^8$ and the coefficients of update filter can be obtained as $K_n^u = [3, -25, 150, 150, -25, 3]/2^9$. For our feature extraction method, we consider five directions aligned at $\{45°, 72.5°, 90°, 112.5°, 135°\}$ [18]. This direction set is used to demonstrate the efficacy of DIWT to capture directional multiresolution features. The term $j + ntan^{-1}$ in (1) and (3) may not always be an integer sample and does not exist on the original sampling grid. Consequently, an interpolation scheme is carried out to estimate the intensity at this non-integer sample [18],

$$x(i+n, j+ntan^{-1}d) = \sum_{l=-N_c}^{N_c-1} \alpha_l \cdot x(i+n, [j+ntan^{-1}d - l]) \tag{5}$$

where $[\cdot]$ denotes the integer part, $N_c = 1$ and $\alpha_{-1} = \alpha_0 = 0.5$ which selects the two closest even samples for interpolation [18]. The integer samples used to interpolate the non-integer samples at the optimal direction d have to be even sampled. If the optimal direction crosses over the integer sample the value is estimated by the nearest even sample otherwise, the fractional sample is calculated from interpolation of two nearest even samples as given in (5). The sub-pixel interpolation improves the directional orientation property of the image and maximises the compaction of image energy into low frequency sub-band.

Due to face variations, each face image contains a varying amount of information located at different local pixel regions. To consider these local regions the quadtree partitioning scheme is applied which facilitate effective direction assignment. Here each face image is quadtree partitioned into non-overlapping blocks and a direction from the direction set is selected for each block based upon the minimum value of the prediction error energy. All the pixels in a quadtree block will have the same direction.

The proposed partitioning scheme used in our proposed method is explained here. Let each face image is quadtree partitioned into non-overlapping blocks with initial block size as S_{ini} and minimum block size as S_{min} where quadtree partitioning need to stop. Each block is denoted by $S_b \in X$ and sample from each block is denoted as $x_{S_b} \in S_b$. The energy summation of the prediction error for each block is calculated as,

$$PE(S^d_{b,i}) = \sum_i \|x_{S_{b,i}} - P_{S_{b,i}}(x_{S_{b,i}})\|_2^2 + \lambda D \tag{6}$$

The complexity of this partitioning is controlled by a value of Lagrangian multiplier λ, specifically if we set it to zero; we may obtain a full tree and if we set it to ∞ does not allow any partitioning [18]. D is the number of bits spent on signaling the selection of direction. The direction which gives the minimum prediction error is obtained as,

$$d_b = \text{argmin}_i\{PE(S^d_{b,i})\} \tag{7}$$

To check each of quadtree partition is an optimal partition or not, an improvement in the quadtree partitioning is presented. Now further quadtree partitioned each block S_b into four sub-blocks $S_{b,r}, r = 1,\ldots,4$ and for each $S_{b,r}$ calculate the prediction error $PE(S^d_{b,r})$ and consecutively the direction $d_{b,r}$. If $PE(S^d_{b,r}) \geqslant PE(S^d_b)$ and size of sub-block $S_{b,r}$ reaches to S_{min} then stop partitioning and store the optimal direction and optimal block otherwise, repeat the previous step of quadtree partitioning. Quadtree partitioned blocks improves the adaptive direction selection of pre-assigned five directions and thus the edge manifolds present at the diagonal directions including horizontal and vertical direction can be effectively approximated in DIWT implementation.

The above-mentioned 1-D process can be easily extended to the 2-D case where second dimension lifting is again performed on $H(i,j)$ high pass and $L(i,j)$ low pass sub-bands generating four sub-bands $LH(i,j)$, $LL(i,j)$, $HH(i,j)$, and $HL(i,j)$. The schematic of 2-D DIWT is presented in Fig. 1. Partitioning structure with direction estimation is presented in Fig. 2 for a face image from GT database [23].

2.2. Local binary patterns

The local binary patterns (LBP) [4] has emerged as an efficient local feature descriptor for face recognition [5,13–15]. To assign a label for each pixel, the LBP operator uses its intensity value as a threshold and compares it against pixel values in a 3×3 neighborhood and considers the result as a binary number. Generally, the LBP is computed with P sampling points $(x_p \in (0,\ldots,P-1))$ in the neighborhood of center pixel $x_m(i_c, j_c)$ at a radial distance given by R [4],

$$LBP_{P,R} = \sum_{p=0}^{P-1} t_s \cdot (x_p - x_m) \cdot 2^p \tag{8}$$

$$t_s(diff) = \begin{cases} 1, & (diff) \geqslant 1 \\ 0, & (diff) < 0 \end{cases} \tag{9}$$

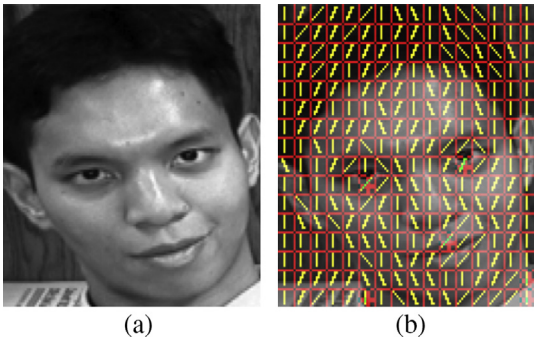

Fig. 2. (a) Face image (GT face database). (b) Quadtree partitioning structure with direction estimation.

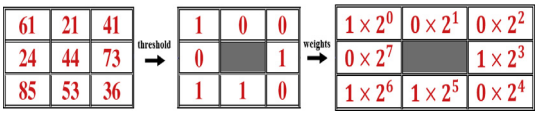

Fig. 3. Example of LBP computation.

where $t_s(diff)$ is a threshold function. Fig. 3 illustrates the LBP operator and the consequent label for center pixel x_m. If the sampling points p's are not mapped in the neighborhood of the center pixel, they are bi-linearly interpolated [4]. Ojala et al. [4] introduced uniform patterns, where a binary pattern is uniform if it contains at most two bitwise transitions from 0 to 1, or vice versa when the binary pattern is considered circularly,

$$U(LBP_{P,R}) = |s(x_{p-1} - x_m) - s(x_0 - x_m)| + \sum_{p=1}^{P-1} |s(x_p - x_m) - s(x_{p-1} - x_m)| \tag{10}$$

It was verified that "uniform" patterns are fundamental patterns of local image textures [4]. In the mapping of $LBP_{P,R}$ to $LBP^{u2}_{P,R}$, subscript $u2$ means that the uniform patterns $U(LBP)$ have a value of at most 2. There are $P \times (P-1) + 2$ uniform patterns and remaining non-uniform patterns are accumulated into one single bin resulting in $P \times (P-1) + 3$ feature dimension. After LBP labeling of the image, codes of all pixels of an input image $x_L(i,j)$ are collected and formed into a histogram [4] given as,

$$H_l = \sum_{i,j} F\{x_L(i,j) = l\}, \quad l = 0,1,2,\ldots,n-1 \tag{11}$$

$$F\{A\} = \begin{cases} 1, & \text{if } A \text{ is true} \\ 0, & \text{if } A \text{ is false} \end{cases} \tag{12}$$

where n is the number of different labels produced by the LBP operator. With $LBP^{u2}_{8,1}$, the feature dimension is 59. The LBP histogram is a composition of micro-patterns and provides information about the local distribution of spots, edges over the entire image. LBP

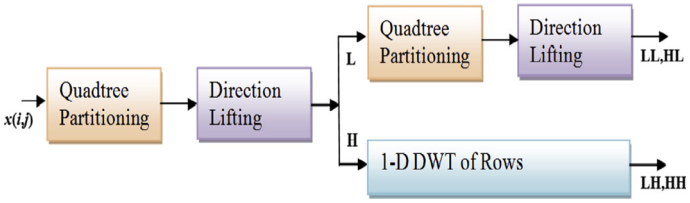

Fig. 1. Schematic of 2-D DIWT.

has tolerance to illumination and offer simplicity in implementation and can be effectively used for local feature descriptor in our proposed method.

3. Proposed face recognition method

In the proposed method, LBP histograms are extracted from DIWT sub-bands to form a novel MRA-based local feature descriptor. To achieve this, we perform DIWT decomposition and consider the top-level low-frequency approximation sub-band *LL* and high-frequency sub-bands *HL*, and *LH*. Due to the adaptive direction selection within the quadtree partitioned blocks and directional lifting, dependencies found over image discontinuities can be effectively de-correlated which concentrate most of the energy of high-frequency sub-bands into low-frequency *LL* sub-band [17]. But the sub-bands *HL* and *LH* also contain edge and contour details of face images significant in extracting pose and expression relevant features with the aid of LBP. We ignored the high-frequency *HH* sub-band as it mostly contains the noise with negligible feature details. Histogram encoded for the whole sub-band describes only the occurrences of the micro-patterns without describing their location information. To preserve the spatial characteristics and to a form a robust local feature descriptor, multi-region $LBP^{u2}_{8,1}$ uniform pattern-based histograms features [4] are obtained from non-overlapping regions of DIWT sub-bands {*LL*, *HL*, *LH*}. $LBP^{u2}_{8,1}$ patterns are statistically significant and offer reduced dimensionality with increased robustness to noise.

Each of the sub-band {*LL*, *HL*, *LH*} is equally divided into m non-overlapping rectangle regions $R_0, R_1, \ldots, R_m$, each of size $x \times y$ pixels. From each of these m regions we extract local histogram features LBP_R_k each with 59 labels separately. Local histograms features from successive regions are concatenated to form a multi-region single enhanced histogram feature with dimension $59 \times m$. Let $LL\_H_{l,k}$, $HL\_H_{l,k}$, and $LH\_H_{l,k}$ denote the concatenated histogram features for *LL*, *HL*, and *LH* sub-bands respectively,

$$LL\_H_{l,k} = \sum_{i,j} F\{LL_L(i,j) = l\} LL\{(i,j) \in R_k\} \tag{13}$$

$$HL\_H_{l,k} = \sum_{i,j} F\{HL_L(i,j) = l\} HL\{(i,j) \in R_k\} \tag{14}$$

$$LH\_H_{l,k} = \sum_{i,j} F\{LH_L(i,j) = l\} LH\{(i,j) \in R_k\} \tag{15}$$

where $l = 0, 1, 2, \ldots, n-1, k = 0, \ldots, m-1$.

Furthermore $LL\_H_{l,k}$, $HL\_H_{l,k}$, and $LH\_H_{l,k}$ histogram features are concatenated to form *LBP_FV* feature set representing DIWT-based local histogram feature,

$$LBP\_FV = [LL\_H_{l,k}, HL\_H_{l,k}, LH\_H_{l,k}] \tag{16}$$

We have maintained 128×128 pixels uniform size for face images of all the face databases. Each of the DIWT sub-bands {*LL*, *HL*, *LH*} is of size 32×32 pixels. In our proposed method each of these sub-bands is divided into $m = 16$ regions with $x \times y = 8 \times 8$ pixels regions with collective dimension of *LBP_FV* is $59 \times m \times 3 = 2832$. The consideration of size of each region $x \times y = 8 \times 8$ and obtaining 16 regions from the DIWT sub-band coefficients provides a trade-off between recognition performance and feature vector length. The multi-region histogram feature *LBP_FV* provides a novel DIWT-based local descriptive feature from the sub-band coefficients. LDA is applied to *LBP_FV* to obtain reduced dimension discriminant feature space. Similarly, DIWT and consecutively LBP are applied on test face images to obtain test histogram features. Next, LDA is applied to obtain the reduced dimension test features. Finally, we apply the NN classifier to classify test features in the reduced space for face identification. The overall flow of the proposed method is presented in Fig. 4.

4. Experimental results

All experiments are conducted using Matlab 2014a on a standard i3-330 2.13 GHz machine with 2.0 GB RAM. To evaluate the efficacy of the proposed expression and pose-invariant feature extraction method, we consider three face databases such as ORL database [22], Georgia Tech (GT) database [23] and FEI face database [24,25]. Face feature extraction methods such as LDA [2], LPP [3], LBP [5], LDP [6], WLD [7], LGBP [13], and LSPBPS [14] and CTLBP [15] have been compared against our proposed method.

4.1. Parameter settings

In this sub-section, we discuss parameter settings for implementation of our proposed and different comparative methods. With face image size 128×128, the DIWT decomposition level J for all the databases is selected two as optimal [19].

A set of experiment is also conducted on databases to select the optimal value of S_{ini} and λ. The parameter S_{ini} is the initial block size required to start the quadtree partitioning. The parameter λ controls the quadtree partitioning. With reference to [19], for quadtree partitioning we select three different values of Lagrangian multiplier, specifically $\lambda = 5$, $\lambda = 7$ and $\lambda = 9$ to obtain the optimal value. Again considering [19], we performed the proposed method for two different block sizes $S_{ini} = 8 \times 8$ and $S_{ini} = 16 \times 16$ individually. We also consider $S_{min} = 4 \times 4$ pixels. We randomly selected five images as training and rest images as test images for all the three databases. Moreover, we repeat the individual experiment on individual database ten times and noted the average rank-one

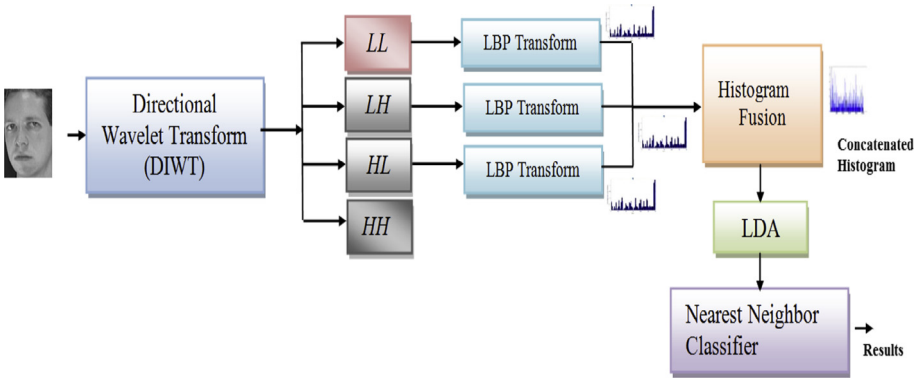

Fig. 4. Diagram of the proposed method.

Table 1
Comparison of different sub-block size and value of Lagrangian multiplier (%).

Block Size/Database	$\lambda = 5$		$\lambda = 7$		$\lambda = 9$	
	8×8	16×16	8×8	16×16	8×8	16×16
ORL	95.19	94.34	95.81	95.00	97.00	96.22
GT	67.12	66.44	67.89	67.11	68.40	67.23
FEI	81.78	80.25	82.60	81.00	84.67	83.50

recognition rate as depicted in Table 1. Based upon the results, we consider $S_{ini} = 8 \times 8$ pixels and $\lambda = 9$ as optimal, which demonstrate their effectiveness to capture directional features as regards to local face variations.

For LDA, and LPP based methods the NN classifier is used with Euclidean distance measure. For LBP and LGBP, images are partitioned into the 8×8 pixel region. For LGBP, we used filters at five different scales and eight orientations [13]. $LBP^{u2}_{8,1}$ with NN classifier and Chi-square distance measure is used for both the methods [5]. For LDP, 56 bin histogram features are extracted from 8×8 non-overlapped regions of face images. Chi-square dissimilarity measure is used to compare two spatially encoded LDP histograms features. For WLD, a patch size of 3×3 pixel is considered for coding and generated histograms of differential excitation and orientaion component are concatenated into a 2-D histogram. Later this 2-D histogram is encoded into a 1-D histogram to be used as the feature vector and Chi-square dissimilarity measure is used to compare two histograms. For LSPBPS, multi-region $LBP^{u2}_{8,1}$ is obtained from all the generated sub-bands to form the feature set and Chi-square dissimilarity measure is used to compare two LSPBPS histogram features. For CTLBP, the LBP coded image of low-frequency sub-band with mid-frequency sub-bands are used to form the feature set and LPP is used for dimensionality reduction [15]. To evaluate the performance we use cumulative match characteristic (CMC) curves [26,27]. CMC curve is used as a measure of 1: m identification system performance. It judges the ranking capabilities of our proposed method. Due to random selection approach, we run the experiment ten times on each database and only report the average rank-one recognition rates in all following experiments.

4.2. Experiment on the ORL face database

The ORL database entails 400 images composed from 40 dissimilar subjects with 10 diverse images of each subject. The face images exhibit changes in the capture time, lighting, head position, facial expressions such as eyes open or closed, smiling or not smiling. All the face images are resized to 128×128. Some samples images of one subject are shown in Fig. 5.

We randomly choose N ($N = 2, 3, 4, 5$) images of each subject for training and the rest images for testing. Table 2 shows the average rank-one recognition rates of different comparative methods. Selecting $N = 5$ randomly, the average CMC for different methods are also depicted in Fig. 6. It is visible both from Table 2 and Fig. 6 that the proposed method yields superior results. A substantial improvement of 11.77%, 10.72% can be observed over LDA and LPP methods for $N = 2$. We can also observe an improvement of 12.59%, 11.55%, 9.25% over LBP, LDP, and WLD respectively for $N = 2$. Compared with LBP-based non-adaptive MRA methods such as LGBP, LSPBPS, and CTLBP our method exhibit higher rank-one recognition rates even for fewer numbers of training images. The proposed method makes an adaptive selection of best lifting direction from the direction set and uses interpolated samples at the spatial resolution of five directions and preserves the local details of expression and pose-variant features. Such details, when extracted in terms of LBP histogram features, improves the overall performance.

4.3. Experiments on the GT face database

The Georgia Tech (GT) face database comprises of 750 JPEG color images of 50 distinct persons. All images are captured against a cluttered background with dissimilar facial expression, lighting conditions, and scale. All images are converted to grayscale images and resized to 128×128 pixels. Some sample images of a subject are shown in Fig. 7. We randomly consider N ($N = 3, 4, 5, 6, 7$) images of each subject for training, respectively, and in every case, the remaining images for testing. Comparative results are depicted in Table 3. The average cumulative match curves for different methods are described in Fig. 8 for $N = 7$. It is observed that the proposed method not only outperforms LDA and LPP but also excels with LBP, LDP, and WLD. As compared to LGBP, LSPBPS, and CTLBP for $N = 3$, our method offers improvement of 16.80%, 8.84%, and 7.30% respectively. Specifically, the improvement is significant for less number of training images. LGBP, CTLBP, and LSPBPS do not provide adaptation in selecting a direction within a block of samples. Whereas, DIWT performs directional lifting with the adaptive directional selection and considers the edge manifolds relating to expression and pose variations effectively.

Table 2
Benchmarking of the rank-one recognition rates on the ORL face database (%).

Number of training samples per subject	2	3	4	5
LDA	73.63	76.29	83.33	87.00
LPP	74.50	78.00	87.75	90.50
LBP	72.94	80.61	86.39	90.49
LDP	73.81	83.22	85.58	91.00
WLD	75.73	84.60	89.37	92.50
LGBP	79.20	83.90	90.43	92.52
LSPBPS	81.27	84.50	92.10	94.50
CTLBP	82.50	84.86	93.33	95.00
Proposed method (DIWTLBP)	83.45	88.26	94.17	97.00

Fig. 5. Samples face images of a subject from the ORL face database.

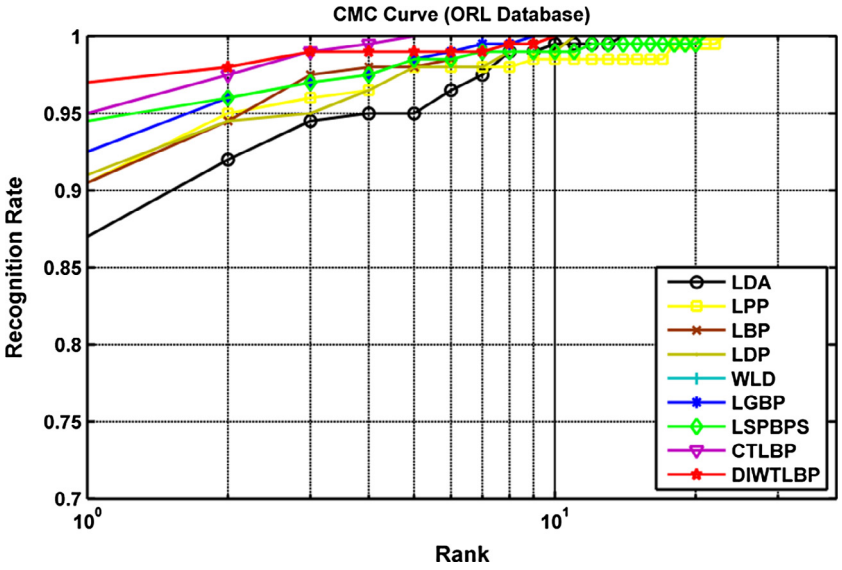

Fig. 6. CMC curves of comparative methods for ORL database.

Fig. 7. Samples face images of a subject from the GT face database.

Table 3
Benchmarking of the rank-one recognition rates on the GT face database (%).

Number of training samples per subject	3	4	5	6	7
LDA	31.50	45.45	48.40	52.22	58.50
LPP	45.17	54.55	57.60	64.89	70.75
LBP	43.20	51.64	56.20	62.78	68.74
LDP	47.33	56.71	61.45	66.40	72.75
WLD	52.00	60.82	64.40	71.22	75.49
LGBP	48.67	58.91	63.80	70.00	74.76
LSPBPS	53.33	61.36	66.20	73.56	77.50
CTLBP	54.23	62.35	67.60	74.89	78.73
Proposed method (DIWTLBP)	58.50	65.66	68.40	77.33	82.25

Moreover, LBP histogram features from DIWT sub-bands contain most significant details which improve the recognition rate.

4.4. Experiments on the FEI face database

The FEI database contains 14 color images of 200 subjects taken against a white homogenous background. The images are in upright front position but a profile rotation of about 180° is considered while capturing the images which increase the complexity of this database [25]. Sample images of a subject are shown in Fig. 9.

All the images are manually cropped and resized to 128 × 128 pixels and converted to grayscale images. We randomly choose N (N = 3, 4, 5, 6, 7) images of each subject for training and the rest images for testing. Table 4 illustrates the average rank-one recognition rates for different methods. The average CMC curves for different methods are depicted in Fig. 10 for N = 7. LDA and LPP being holistic methods fail to consider the extreme expression and pose variations. We observe an improvement of 21.10%, 14.24%, and 10.89% over LBP, LDP, and WLD respectively for N = 3. As compared to LSPBPS and CTLBP which considers more sub-bands to achieve good results we used only three top-level sub-bands of DIWT to effectively generate descriptive features using the LBP histogram. When N = 6, there is an improvement of 6.81%, 5.96%, and 4.82% over LGBP, LSPBPS, and CTLBP respectively. Thus the proposed

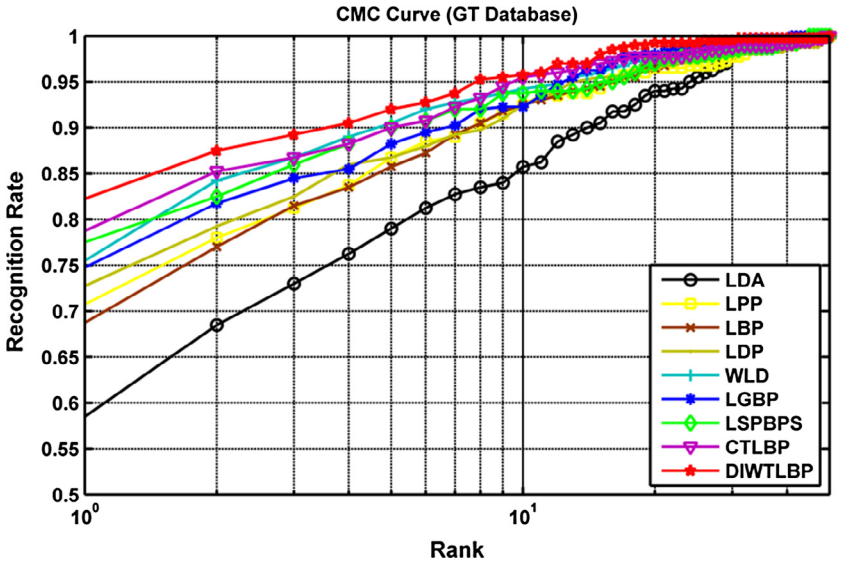

Fig. 8. CMC curves of comparative methods for GT face database.

Fig. 9. Samples face images of a subject from the FEI face database.

Table 4
Benchmarking of the rank-one recognition rates on the FEI face database (%).

Number of training samples per person	3	4	5	6	7
LDA	46.73	48.20	56.89	68.50	74.29
LPP	48.55	51.20	59.78	72.00	78.00
LBP	51.36	56.40	65.33	74.25	84.57
LDP	55.82	65.32	67.61	77.25	86.29
WLD	58.00	67.60	70.19	78.57	86.86
LGBP	54.73	63.00	75.11	82.00	87.71
LSPBPS	59.45	66.20	76.67	82.75	88.29
CTLBP	60.91	64.40	78.89	83.75	90.29
Proposed method (DIWTLBP)	65.09	73.80	84.67	88.00	91.14

method also works well for the database with a large number of face images under pose and expression variations.

4.5. Comparison with different adaptive directional lifting methods

In order to demonstrate the efficacy of DIWT with five directions, we compare it with two adaptive directional lifting methods such as ADWT with nine directions [19] and ADL [17] with nine directions with 9/7 tap filters. To implement ADWT and ADL, similar settings from Section 4.1 are considered such as $J = 2$, $S_{ini} = 8 \times 8$, $S_{min} = 4 \times 4$ and $\lambda = 9$. Furthermore, for a fair comparison similar LBP setting are also applied for histogram feature extraction. For ORL database $N = 5$, and for GT and FEI face database $N = 7$ images are randomly considered as training images and rest as testing images. Table 5 illustrates the comparative results. ADWT incorporates block-based partitioning and do not considers the interpolation of fractional samples. In our proposed method quadtree partitioning scheme and incorporation of interpolation mechanism for fractional samples considers the face image characteristics more efficiently. In ADL, adaptation of nine directions increases computational complexity.

4.6. Computational complexity

The proposed method also has a comparable computational time for feature extraction compared with different methods. Fig. 11 shows the computation time of some of the comparative method to process an ORL database face image of size 128×128

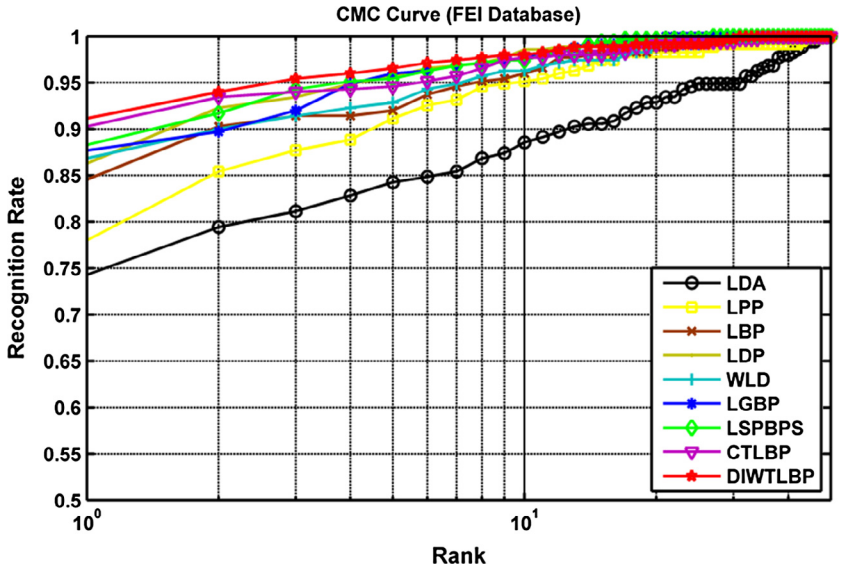

Fig. 10. CMC curves of comparative methods for FEI face database.

Table 5
Rank-one recognition rates for different LBP-based adaptive directional transform methods (%).

Database	ADWTLBP	ADLLBP	DIWTLBP
ORL	95.00	95.50	97.00
GT	81.72	81.00	82.25
FEI	89.67	90.80	91.14

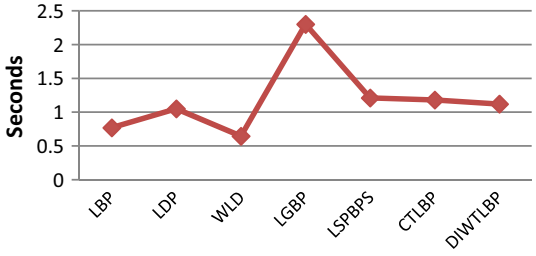

Fig. 11. Computation time for ORL face image.

pixel resolution and to produce the features and no further step of dimensionality reduction and classification is applied. We mainly considered those methods which require the feature extraction at the pre-processing stage. The LGBP method has the highest computation time.

5. Conclusion

This paper offered a novel expression and pose-invariant facial feature extraction method primarily based on DIWT-based LBP histogram features. To implement the DIWT an effective quadtree partitioning scheme is implemented. DIWT provides adaptation in directional selection based on image characteristics and efficiently represents image edge manifolds. Multi-region LBP histogram features from the top level sub-bands {*LL*, *HL*, *LH*} form an efficient feature set. According to Tables 2–4, it is proven that the proposed method demonstrates superior discrimination ability and yields the best rank-one recognition results for the selected face databases. From the results, we can signify that the proposed method not only excels with the holistic method such as LDA and LPP but also demonstrate superiority against various local descriptors such as LBP, LDP, and WLD methods for face images with lesser to extreme expressions and pose variations. Our experimental results verify that the proposed method also outperforms some non-adaptive LBP-based MRA methods such as LGBP, LSPBPS, and CTLBP. We also demonstrated the effectiveness of our method over other adaptive directional lifting methods.

References

[1] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cognit. Neurosci. 3 (1) (1991) 71–86.
[2] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. Fisherfaces: recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 711–720.
[3] X. He, S. Yan, Y. Hu, P. Niyogi, H. Zhang, Face recognition using laplacianfaces, IEEE Trans. Pattern Anal. Mach. Intell. 27 (3) (2005) 328–340.
[4] T. Ojala, M. Pietikäinen, T. Maenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Anal. Mach. Intell. 24 (7) (2002) 971–987.
[5] T. Ahonen, A. Hadid, M. Pietikäinen, Face description with local binary patterns: application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 28 (12) (2006) 2037–2041.
[6] T. Jabid, M.H. Kabir, O. Chae, Local directional pattern (LDP)-a robust image descriptor for Object Recognition, in: Proceedings of 7th IEEE International Conference on Advanced Video and Signal Based Surveillance, 2010, pp. 482–487.
[7] J. Chen, S. Shan, C. He, et al., WLD: a robust local image descriptor, IEEE Trans. Pattern Anal. Mach. Intell. 32 (9) (2010) 1705–1720.
[8] Z. Zhang, L. Wang, Q. Zhu, S.K. Chen, Y. Chen, Pose-invariant face recognition using facial landmarks and weber local descriptor, Knowl.-Based Syst. 84 (2015) 78–88.
[9] I. Ullah, M. Hussain, G. Muhammad, H. Aboalsamh, G. Bebis, A.M. Mirza, Gender recognition from face images with local WLD descriptor, in: 19th International Conference Systems, Signals and Image Processing (IWSSIP), pp. 417, 420, 11–13 April 2012.
[10] J.T. Chien, C.C. Wu, Discriminant waveletfaces and nearest feature classifiers for face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 24 (12) (Dec. 2002) 1644–1649.
[11] Z.H. Huang, W.J. Li, J. Wang, T. Zhang, Face recognition based on pixel-level and feature-level fusion of the top-level's wavelet sub-bands, Info. Fusion 22 (2015) 95–104.
[12] Z.H. Huang, W.J. Li, J. Shang, J. Wang, T. Zhang, Non-uniform patch based face recognition via 2D-DWT, Image Vision Comput. 37 (2015) 12–19.
[13] W. Zhang, S. Shan, W. Gao, H. Zhang, Local Gabor binary pattern histogram sequence (LGBP): a novel non-statistical model for face representation and recognition, in: Proceedings of IEEE International Conference and Computer Vision, 2005, pp. 786–791.
[14] M. El Aroussi, M. El Hassouni, S. Ghouzali, M. Rziza, D. Aboutajdine, Local steerable pyramid binary pattern sequence LSPBPS for face recognition method, Int. J. Signal Process. 5 (4) (2009) 281–284.

[15] L. Zhou, W. Liu, Z.M. Lu, T. Nie, Face recognition based on curvelets and local binary pattern features via using local property preservation, J. Syst. Software 95, 209–216. https://doi.org/10.1016/j.jss.2014.04.037.
[16] C.L. Chang, B. Girod, Direction-adaptive discrete wavelet transform for image compression, IEEE Trans. Image Process. 16 (5) (May 2007) 1289–1302.
[17] W. Ding, F. Wu, X. Wu, S. Li, H. Li, Adaptive directional lifting-based wavelet transform for image coding, IEEE Trans. Image Process. 16 (2) (Feb. 2007) 416–427.
[18] A. Maleki, B. Rajaei, H.R. Pourreza, Rate-distortion analysis of directional wavelets, image processing, IEEE Trans. Image Process. 21 (2) (2012) 588–600.
[19] M.A. Muqeet, R.S. Holambe, Local appearance-based face recognition using adaptive directional wavelet transform, J. King Saud Univ. – Comput. Inform. Sci. (2017), https://doi.org/10.1016/j.jksuci.2016.12.008.
[20] W. Sweldens, The lifting scheme: construction of second generation wavelets, SIAM J. Math. Anal. 29 (2) (1998) 511–546.
[21] J. Kovacevic, W. Sweldens, Wavelet families of increasing order in arbitrary dimensions, IEEE Trans. Image Process. 9 (3) (2000) 480–496.
[22] [Online] ORL database, http://www.uk.research.att.com/pub/data/att_faces.zip.
[23] [Online] GT Database, http://www.anefian.com/research/facereco.htm.
[24] [Online] FEI Database, http://fei.edu.br/cet/facedatabase.html.
[25] G. Thomaz, C. Eduardo, G. Antonio, A new ranking method for principal components analysis and its application to face image analysis, Image Vis. Comput. 28 (2010) 902–913.
[26] P. Phillips, P. Grother, R. Michaels, D. Blackburn, T. Elham, J. Bone, FRVT 2002: Facial Recognition Vendor Test, Technical report, DoD, April 2003.
[27] V. Štruc, N. Pavešic, The Complete Gabor-Fisher Classifier for Robust Face Recognition, EURASIP Advances in Signal Processing, vol. 2010, 2010, https://doi.org/10.1155/2010/847680.