Appl Intell (2016) 44:913–930
DOI 10.1007/s10489-015-0735-1

Expression invariant face recognition using semidecimated DWT, Patch-LDSMT, feature and score level fusion

Hemprasad Patil[1] · **Ashwin Kothari**[1] · **Kishor Bhurchandi**[1]

Published online: 19 December 2015

Abstract This paper addresses the issue of human face recognition in presence of expression variations, which pose a great challenge to face recognition systems. Typically, the discriminant features lie in both spatial as well as transform domain. In this paper, we propose combination of Discrete Wavelet Transform (DWT) and proposed Semi-decimated Discrete Wavelet Transform (SDWT) to develop an expression invariant face recognition algorithm followed by a novel wavelet coefficients enhancement function. The wavelet coefficients are boosted using the proposed coefficients enhancement function and extracted using the Weber Local Descriptors (WLD). This enhances weak skin edges based features, resulting in increased probability of recognition. The proposed algorithm also exploits spatial domain features using our customized version of Complete Local binary patterns (CLBP) named Patch Local Difference Sign Magnitude Transform (Patch-LDSMT) applied on complete images and physiologically meaningful overlapping regions of human facial images for the first time. Feature level fusion of the wavelet based features and Patch-LDSMT yields a robust feature vector whose dimensionality is reduced using Linear Discriminant Analysis (LDA). Comprehensive experimentation is carried out on the JAFFE, CMU-AMP, ORL, Yale, Cohn-Kanade (CK) and database collected by us. Benchmarking analysis illustrates that the proposed face recognition algorithm offers much better rank one recognition performance when compared with the current state-of-the-art expression invariant face recognition approaches.

Keywords Face recognition · Expression variations · Patch-LDSMT · Semidecimated DWT · DWT coefficient enhancement and WLD

1 Introduction

Face recognition is a biometric technology for recognition of an individual from an image or a video using statistical, frequency domain or spatial geometric features. The human brain can recognize human faces with little or no effort. However, it is a challenging problem for computer based face recognition systems especially when variations in expressions, pose, age and illumination are involved in the facial region. Many approaches based on statistics, neural networks and transform based feature computations and subsequent matching have been researched so far. Still there is a lot of scope as far as the practical performance matching with human capabilities is concerned. The scope resides in multiple aspects like accuracy, robustness of matching under real life conditions and the response time in case of big databases. Such systems are far from embedded implementations. Eigen faces based approach abstracts the dominant patterns in a given image data matrix using multivariate data analysis [1]. It obtains a set of basis functions and represents facial images with linear combination of those basis functions. Eigenfaces are widely used for dimensionality reduction and similarity model computation in face recognition applications. Fisherfaces utilize Fisher's linear discrimination technique to obtain widely separated labels for classes in low-dimensional subspaces [2]. Independent

✉ Hemprasad Patil
hemprasadpatil@gmail.com

[1] Department of Electronics and Communication Engineering, Visvesvaraya National Institute of Technology, Nagpur, Pin-440010 India

Component Analysis (ICA) [1] theory assumes that most of the perceptually separable information about person identification is located in higher order statistical data of the facial images. Basis functions used in ICA are higher order statistically independent. This results in sparse representation of facial images localized in low dimensional spaces. Neural network based approach merged a convolutional neural network and a Self Organizing Map (SOM) neural network with local image sampling [4]. The SOM is employed as a dimensionality reduction tool responsible for minor variations in an input face. The convolutional neural network is constrained to provide partial robustness to scale, translation and rotation variations. Bayesian techniques involve density estimation using PCA decomposition [5]. The training facial images are modelled with a mixture of multivariate Gaussian models. Subsequently, those models are transformed to a maximum-likelihood estimation system for face recognition. A few detailed surveys of 2-D face recognition techniques are presented by Zhao et al. [2], Zhang and Gao [3] and Kong et al. [4].

Verbal communications and emotions are the major reasons behind variations in the facial expression. The expressive facial query images act as a major challenge and affect the accuracy of state-of-the-art face recognition systems. Robust feature extraction methods play a vital role in the enhancement of accuracy of expression invariant face recognition systems. Thus, in this paper, we concentrate more on enhancing the spatial and frequency domain features using the proposed approaches while maintaining the simplest distance based approach for a required classifier. Existing approaches to expression invariant face recognition can be classified into two major groups:

1. spatial domain techniques and
2. transform domain techniques. Spatial domain techniques can further be categorized as
   (i) subspace based techniques and
   (ii) appearance model based techniques.

Chen and Lovell utilized wrapped PCA subspace for enhanced intra-class separation [5]. They performed filtering and whitening operations on the scattering matrix to obtain an adaptive PCA subspace. Li et al. proposed a new method with detached geometric and texture feature extraction [6]. Further, they projected it onto distinct PCA subspaces for feature extraction. Riaz et al. [7] proposed a method that involved Active Appearance Model (AAM) based framework. Their hybrid feature vector constituted features from the AAM parameters, temporal and optical flow. Lee and Kim [8] utilized the AAM along with expression state analyser for recognition of expressions and translation to a neutral facial feature vector. A combination of subspace and transform based techniques has also been utilized for expression invariant face recognition. The proposed work lies in this category. Abbas et al. [9] applied the DWT on facial images to obtain approximate and detailed coefficients. Further, all detailed wavelet coefficients' values were equated to zero and the images were reconstructed to address expression invariant face recognition problem. Vankayalapati and Kyamakya [10] presented a method using fusion of the DWT and radon transform based features. Radon transform yielded effective local features and the wavelet transform resulted in multiresolution representation. These complementary features built an effective feature vector upon feature level fusion. Wang et al. [11] applied thresholding function on contourlet coefficients and utilized SVM based feature classifier for face recognition. However, use of the sophisticated SVM classifier may slow down a face recognition system due to the required computations, if the database is big. Local Binary Patterns (LBP) are widely utilized in face recognition methods [12, 13] because of their ability to perform micro-pattern analysis in spite of their limitation in handling the local difference magnitudes. Local Difference Sign Magnitude Transform (LDSMT) is a part of Complete Local Binary Patterns specially designed for texture analysis and matching [14]. CLBP is an extension of LBP that can incorporate local mean, sign and magnitude of the local differences at every pixel position. In face recognition, the mean component is dropped as the face recognition is expected to be illumination intensity invariant. The modified CLBP named Patch-LDSMT is applied on small patches of physiologically meaningful facial regions containing left eye, right eye, nose and mouth to yield more meaningful features in the form of patch histograms unlike uniform sized patches used by [12] only on the complete image. Thus, a robust feature vector can be obtained upon fusion of the above discussed enhanced composite wavelet transform based WLD features and spatial domain Patch-LDSMT features.

Recently, non-orthogonal transforms like radon transform have been combined with DWT to derive ridgelet [15], curvelet [16] and contourlet transforms [17]. These transforms use optimum number of coefficients to represent the curves in the images effectively. They are also known to possess better directionality preserving capability compared to DWT due to the use of directional filter banks [17]. Facial curves can be expressed as sets of small contiguous but weak line segments formed by skin-on-skin edges aligned at different orientations. Thus, the curves are represented better using ridgelet, curvelet and contourlet transforms enhancing the feature representation and subsequent enhanced feature detection leading to better recognition rates. Human faces essentially have many soft edges [18], rather human faces are significantly discriminated on the basis of the difference between these soft and multidirectional edges [19–21]. Hence, the recent trend

of using curvelet or contourlet transform for facial feature detection is well justified. However, ridgelet, curvelet and contourlet transforms are computationally very complex. This makes the training and testing processes using these transforms very slow, especially for big databases. In this paper, we enhance the directionality preserving capability of DWT by combining it with its own semidecimated version. At every level of the transform, conventional DWT represents horizontal, vertical and diagonal ($\pm 45°$) edges effectively [22]. By combining it with its semidecimated version, it is empowered to represent directional edge features at angles $\tan^{-1}(\pm 2)$and $\tan^{-1} \pm(0.5)$ at each level. As the semi-decimated DWT is an intermediate result of the conventional DWT computation, computational efficiency is the main advantage of this approach unlike curvelet and contourlet transform. As already discussed, the patch-LDSMT, effectively computes intensity immune spatial features of different patches. The features in Transform domain and spatial domain are concatenated to achieve feature level fusion. Further, we use LDA to reduce the feature space dimensionality and obtain a final feature vector. Every patch of a query is matched with the corresponding database image patch individually and the patch matching results are used to arrive at a final decision resulting in score level fusion to obtain final matching score. Cosine similarity measure that finds its applications in data mining for searching a required structure in huge databases is used for matching the image feature vectors. The database image with maximum matching score is a rank one match. The proposed system is tested on the databases that involve drastic expression variations like JAFFE, CMU-AMP, ORL, Yale, CK and our own acquired face database. The proposed method consistently achieves excellent rank-one recognition rates with all the databases at equal or less training. This section presented a brief introduction to the field, literature survey and current state of the art followed by contribution of this work. Rest of the paper is organized as follows: In Section 2, the overview of 2-D orthogonal discrete wavelet transform, un-decimated wavelet transform, WLD and the modified CLBP, i.e. Patch-LDSMT, are presented. Semi-decimated Wavelet Transform (SDWT) and our adaptive wavelet transform coefficient enhancement algorithm are presented in Section 3. A novel feature level fusion based expression invariant face recognition algorithm based on the above discussed methods is proposed in Section 4 followed by discussion on score level fusion. In Section 5, results are presented and performance of the proposed face recognition algorithm is evaluated on the above databases along with their benchmarking. Finally, Section 6 presents conclusion.

2 Methods and materials

This section briefly presents the so far published fundamental work on which the proposed work is based.

2.1 2-D orthogonal discrete wavelet transform

Let the 2-D image be $\{x[i, j], i = 1, 2, ..., M; j = 1, 2, ...N\}$. The values of M and $N$are integer and power of two. The 2-D discrete orthogonal wavelet transform is implemented with orthogonal filtering on rows and subsequently on columns of an image/subband. Let $X = \{x[i, j]\}$ be the matrix representation to be operated upon by the wavelet transform kernel. $Y = W \cdot X$ is a forward wavelet decomposition [23], where Y denotes forward wavelet coefficient matrices and W represents wavelet filters. Level one 2-D orthogonal wavelet transform is computed and organized as shown in Fig. 1. In this work, we keep our discussions on DWT restricted to 2-tap Haar Wavelet transform to minimize computations.

$G(Z)$ and $H(Z)$ denote low pass and high pass wavelet filters respectively. The original image X is filtered either along rows and then along columns or vice-versa. LL_K is a low pass subband which is processed iteratively, for $K = 1, 2, ..P$ decomposition levels. For $K = 1, LL_K = X$. The forward DWT decomposition of LL_K results in

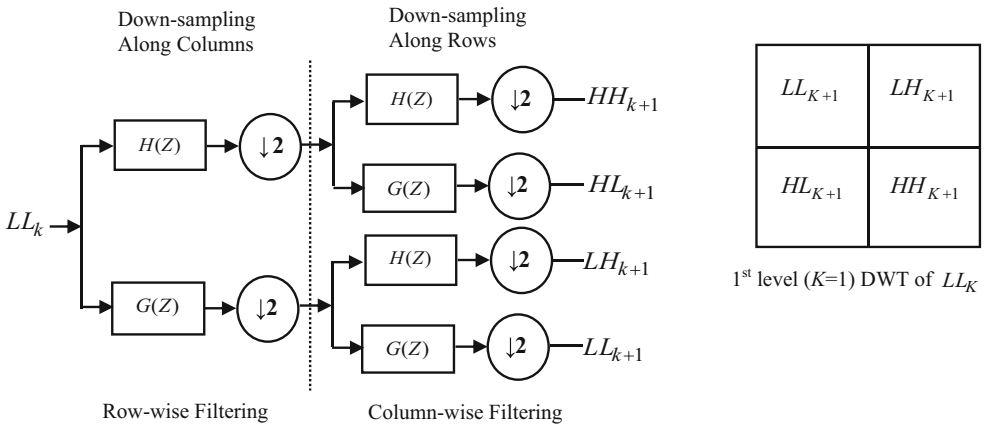

Fig. 1 Single stage forward DWT decomposition

LL_{K+1} (Low pass-Low pass), LH_{K+1} (Low pass-High pass), HL_{K+1}(High pass-Low pass) and HH_{K+1} (High pass-High pass) subbands. Any wavelet subband at level K has a size of $(M/2^K)\times(N/2^K)$ [23]. The orthogonally computed, i.e. first along rows and then along columns, wavelet transform is inefficient in preserving the other directional features of practical images. Figure 2 shows a sample image containing binary multidirectional edges and its level one conventional DWT decomposition is depicted in Fig. 3.

It can be easily observed in Fig. 3 that, the horizontal and vertical edges in the *LH* and *HL* bands perfectly represent the horizontal and vertical edges by white lines. However, the edges at other angles, though they are of the same magnitude, appear discontinuous and in gray shades. The other direction edges in *LH* band appear bright close to the horizontal axis and diminish as the angle with the horizontal axis increases. Similarly, the other direction edges in *HL* band appear bright close to the vertical axis and diminish as the angle with the vertical axis increases. This proves that the conventional DWT is not very good in preserving the multidirectional edges.

2.2 Undecimated wavelet transform

The standard Undecimated Wavelet Transform (UWT) [24] is computed to extract translation invariant wavelet coefficients. The UWT coefficients are effective for computer vision based techniques such as deconvolution, denoising, data analysis and filtering. Since the decimation operation is dropped, the size of each wavelet subband is same as that of the original input image ($M \times N$) at any particular scale. Hence, the undecimated wavelet transform has four redundancy compared to the 2-D orthogonal discrete wavelet transform. The advantage of the UWT is that it carries all the signal features in their original form as it does not decimate or down sample the filtered bands at all. The

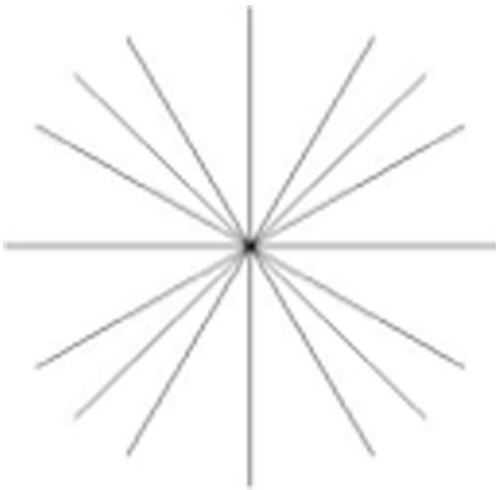

Fig. 2 An image containing multidirectional edges [size: 128 x 128]

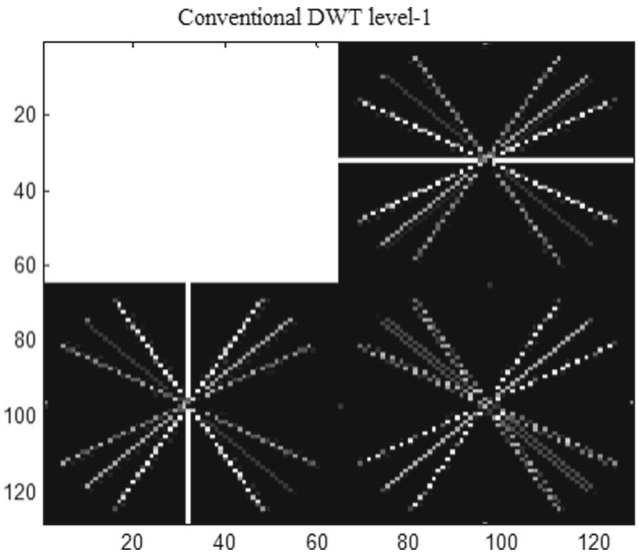

Fig. 3 Level one forward DWT subbands of an image containing multidirectional edges as indicated in Fig. 2

forward undecimated wavelet transform coefficients in time domain [24] are given by (1).

$$\begin{aligned} LL_{K+1} &= (g^{(K)}g^{(K)} * LL_K) \\ LH_{K+1} &= (g^{(K)}h^{(K)} * LL_K) \\ HL_{K+1} &= (h^{(K)}g^{(K)} * LL_K) \\ HH_{K+1} &= (h^{(K)}h^{(K)} * LL_K) \end{aligned} \tag{1}$$

Where, LL_K is the input image/ low pass (only) subband at level K and $h[n]$ and $g[n]$ are octave band analysis high pass and low pass filters for undecimated wavelet transform. For directional feature extraction, the proposed semidecimated wavelet transform (SDWT) is computed by alternate subsampling of either rows or columns resulting in two distinct mechanisms for forward decomposition as presented in Section 3.

2.3 Weber local descriptor

Weber Local Descriptor (WLD) is an effective feature extraction technique based on the Weber's law, which signifies that the ratio of just noticeable difference (JND) to the background intensity is always a constant [25]. Thus, the ratio matters more compared to individual intensity magnitudes for representing features. The WLD used here has two distinct components: differential excitation (ϑ) and orientation (Ξ) representing magnitude and direction respectively at each pixel position. Steps for computing ϑ and Ξ are presented in the following subsections. These parameters are computed in transform domain i.e. DWT and SDWT for each patch and their histograms are further used as patch features.

2.3.1 Differential excitation (ϑ)

The perceptual variations in the image to human eyes can be indicated by the magnitude variations between current pixel (x_m) and its neighbors ($x_n, n = 0, 1, ..., P-1$). The differential excitation component for the current pixel (x_m) in the [3x3] neighborhood (x_n) is given by (2) as in [25]. Usually the neighborhood is defined in terms of a square tile x_n. For example, for a 3×3 tile x_n, n =[0,1,... 8] and of course P =9.

$$\vartheta(x_m) = \arctan\left[\sum_{n=0}^{P-1} \frac{(x_n - x_m)}{x_m}\right] \tag{2}$$

2.3.2 Orientation (Ξ)

The orientation parameter (Ξ) considers a ratio of vertical and horizontal directional gradients using the same notations as (ϑ) for a [3x3] neighborhood in a tile $x_n = x_0, x_1, ..., x_8$ as in [25].

$$\Xi(x_m) = \arctan\left[\frac{x_7 - x_3}{x_5 - x_1}\right] \tag{3}$$

$\Xi \in [-\pi/2, \pi/2]$ is mapped to $\Xi' \in [0, 2\pi]$ and quantized into T dominant angles given by (4).

$$\Phi_b = f_q(\Xi') = \frac{2b}{T}\pi \quad , where,\ b = \mathrm{mod}\left(\left\lfloor \frac{\Xi'}{2\pi/T} + \frac{1}{2} \right\rfloor, T\right) \tag{4}$$

2.3.3 WLD histogram

Initially, for every pixel (x_m) in each patch of the original image, the differential excitation (ϑ_d) and quantized orientation (Φ_b) are computed using 3x3 neighborhood to form a two dimensional histogram, where, $\{(d = 0, 1, ..., N-1), (t = 0, 1, ..., T-1)\}$, N is the total number of pixels in the original image patch and T stands for number of dominant angle orientations. Further, each 2-D histogram is cascaded to form a single histogram H for representing the WLD feature vector of each patch [25].

2.4 Local difference sign magnitude transform (LDSMT)

The Local Binary Patterns (LBP) is widely used for face recognition and texture classification. A more generalized local texture feature extraction method named complete LBP is presented in [14] for texture recognition. The Local Difference Sign Magnitude Transform (LDSMT) computes sign differences in *CLBP* ($CLBP\_S$) and magnitude differences in *CLBP*($CLBP\_M$) [14]. Let the central pixel g_m be surrounded with P neighbours at city block distance R, where $p = 0, 1, ...P-1$. The local difference between g_m and g_p is given by (5).

$$d_p = g_p - g_m \tag{5}$$

The local difference d_p comprises of sign and magnitude components and is represented by (6).

$$\begin{aligned} &d_p = sign_p \times mag_p \\ &where, \\ &sign_p = \begin{cases} +1, d_p \geq 0 \\ -1, d_p < 0 \end{cases} \\ &mag_p = |d_p| \end{aligned} \tag{6}$$

Inspired with local sign and magnitude differences, $CLBP\_S$ and $CLBP\_M$ components are calculated as given in (7) and (8) respectively.

$$\begin{aligned} &CLBP\_S = \sum_{p=0}^{P-1} s(g_p - g_m)2^P \\ &where, \\ &s(diff) = \begin{cases} 1, diff \geq 0 \\ 0, diff < 0 \end{cases} \end{aligned} \tag{7}$$

$$\begin{aligned} &CLBP\_M = \sum_{p=0}^{P-1} t_h(mag_p, c_{th})2^P \\ &where, \\ &t_h(x, c_{th}) = \begin{cases} 1, x \geq c_{th} \\ 0, x < c_{th} \end{cases} \end{aligned} \tag{8}$$

The comparative threshold c_{th} is the mean magnitude value of the entire difference image [14]. We incorporate two modifications in CLBP [28] at this stage. The first is that the CLBP is applied on carefully separated overlapping square patches and called Patch-LDSMT and the c_{th} is now replaced by patch c_{th}, i.e. mean of the magnitude values of the patch difference image. This is obvious as different facial regions have different physiological and textural properties. For example, forehead skin and cheek skin have different textures, especially in case of male faces. The second modification is that the image intensity mean component in CLBP is eliminated as the facial features are expected to be robust to illumination variations over the facial regions. Thus in our customized form adapted for human facial feature detection, the customized CLBP for facial image recognition is named here as Patch-LDSMT.

3 Proposed adaptive wavelet coefficient enhancement function

In this section, we propose the semidecimated DWT and a novel wavelet coefficients enhancement function to enhance the conventional DWT and SDWT coefficients adaptively. It should be noted here that the enhancement functions proposed so far need parameter inputs for feature enhancement

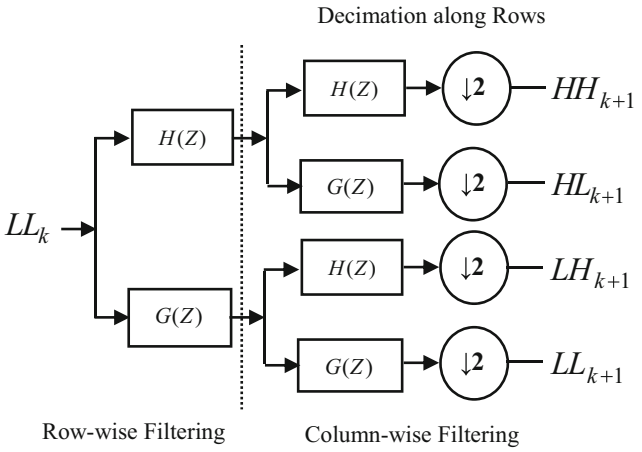

Fig. 4 Level one *R*-decimated (row subsampled) wavelet transform

[26]. The semidecimated wavelet transform are available in two parts, *R*-decimated and *C*-decimated.

3.1 R-decimated wavelet transform

Semidecimated wavelet transform decimates the subbands after filtering but only in one dimension. The *R*-decimated wavelet transform leads to a set of wavelet coefficients by filtering along rows as in DWT but alternate columns are not dropped. Rather directly, filters are applied along columns and row subsampling or decimation by two is done. The resultant representation is shown in Fig. 4. The size of each subband is $(M/2) \times N$.

Figure 5 shows the *R*-decimated level 1 DWT of the test image in Fig. 2. The *HL* component clearly represents the angular edges closer to the horizontal direction, i.e. at angle $\tan^{-1}(\pm 0.5)$, while the components become gray as the angle approaches $\pi/2$ or $-\pi/2$. However the continuity of the lines in *LH* component closer to the horizontal direction i.e. angle 0, π and 2π is the poorest. In *HL* component, the lines closer to the vertical axis, i.e. at angle $\tan^{-1}(\pm 2)$, are represented with bright intensities and their continuity is also good.

Thus, in *R*-decimated DWT the features are more accurately represented by the *HL* or vertical component. The *HL* and *HH* components also represent some features but it is clearly evident from Fig. 5 that the features may not be very accurate or dominant. However, we consider all the subband features for achieving better accuracy or final recognition rate.

3.2 C-decimated wavelet transform

The *C*-decimated wavelet transform produces a set of wavelet coefficients by filtering along rows followed by decimation along columns by two. Further, filters are applied along columns without any subsequent decimation. Thus, alternative rows are not decimated but columns are decimated. Figure 6 depicts the single stage decomposition for *C*-decimated wavelet transform. Each subband has a size of $M \times (N/2)$.

Figure 6 shows the C-decimated level 1 DWT of the test image in Fig. 2. The *HL* component clearly represents the angular edges closer to the horizontal direction, i.e. at angle $\tan^{-1}(\pm 0.5)$ while the components become gray as the angle approaches to $\pi/2$ or $-\pi/2$.

The continuity of the lines in *LH* component closer to the horizontal direction. The continuity decreases if the angle approaches $\pi/2$ or $-\pi/2$. In *HL* component, the lines closer to the vertical axis, i.e. at angle $\tan^{-1}(\pm 2)$, are represented with bright intensities but their continuity is not good. As the angle approaches 0, π and 2π, the continuity of lines improves. Thus, in *C*-decimated DWT, the features are more accurately represented by the *LH* or vertical component. The *HL* and *HH* component also represent some features but it is evident from Fig. 6 that the features may not be very accurate or

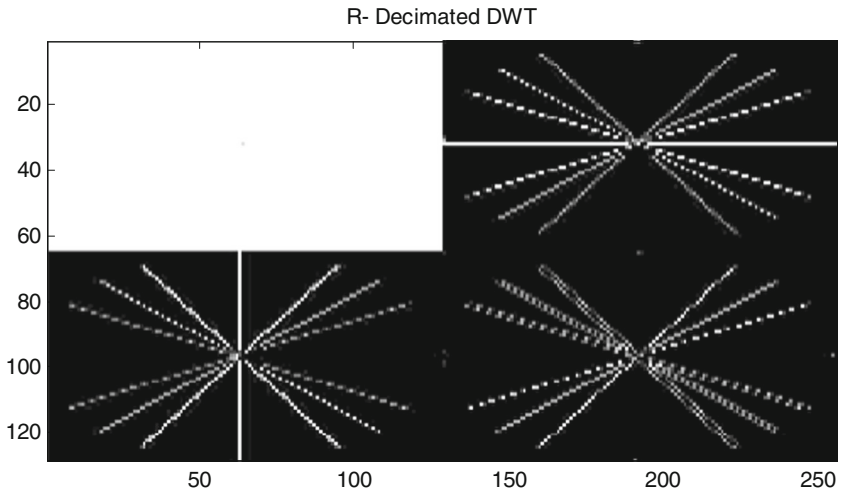

Fig. 5 Level one *R*- decimated subbands of a sample image in Fig. 2

dominant. However, we consider all the subband features for achieving better accuracy or final recognition rate.

The semidecimated row subsampled and column subsampled wavelet transforms both are two redundant. The advantage is that at any level K they can preserve the directionality at angles $\tan^{-1}(\pm 0.5)$, i.e. closer to horizontal axis, and $\tan^{-1}(\pm 2)$, i.e. closer to vertical axis. Most of the human facial features are situated across soft skin edges [27, 28] which are either close to vertical axis like nose features or horizontal axis like eyes, lips and mouth features. The conventional DWT and SDWT transform coefficients together can effectively capture most of such features. However, many discriminant features have less strength in terms of magnitude of the wavelet coefficients.

Hence, a robust enhancement function is required to strengthen such features. One such enhancement function for wavelet coefficients has been proposed by Velde [29]. But, this function requires parameter inputs from the user for the coefficient enhancements. We propose a fully automatic and adaptive wavelet subband coefficients' enhancement function based on the analysis of discriminant facial features of a large number of human facial images. It works perfectly well irrespective of race, ethnicity, illumination, pose and expression variations.

Steps for computing the proposed enhanced wavelet subband coefficients are illustrated below.

Step 1: For the wavelet subband W, compute maximum $W_{\max}$ and minimum $W_{\min}$ values.

Step 2: Perform normalization of wavelet subband W to the range 0 to 255 using (9) to obtain X.

$$X = 255 \times \frac{W - W_{\min}}{W_{\max} - W_{\min}} \tag{9}$$

Step 3: Computation of spans $\left[\lim_{low}, \lim_{high}\right]$ for uniform PDF transformation using (10–13).

$$\mu_X = mean(X) \tag{10}$$

$$\sigma_X = \text{standard deviation}(X) \tag{11}$$

$$\lim_{low} = \mu_X - \sigma_X \tag{12}$$

$$\lim_{high} = \mu_X + \sigma_X \tag{13}$$

Step 4: The proposed adaptive histogram equalization of an image in spatial domain is based on alteration of probability density function (PDF) of random variables [30]. If x is a random variable for input subband and y is a random variable for output i.e. enhanced wavelet subband, the PDF is denoted with $p_x(x)$ and $p_y(y)$ respectively. The transformation function is given by (14), where w is a dummy variable.

$$y = T(x) = \int_0^x p_x(w)dw \tag{14}$$

$$\frac{dy}{dx} = \frac{dT(x)}{dx} = \frac{d}{dx}\left[\int_0^x p_x(w)dw\right] = p_x(x) \tag{15}$$

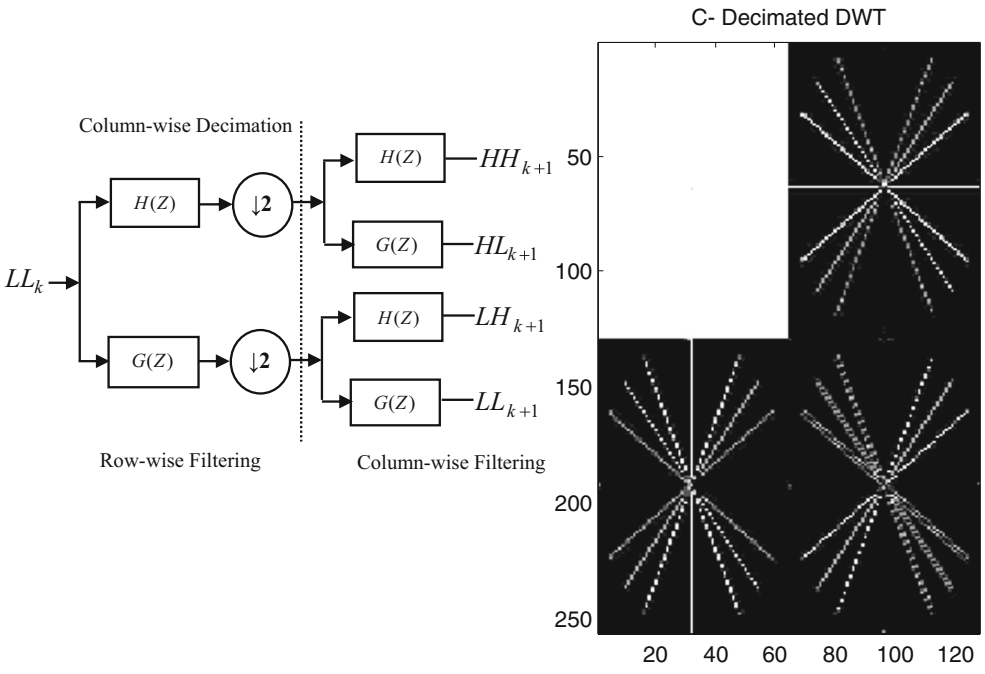

Fig. 6 Level one C-decimated subsampled wavelet transform and C-decimated subbands of image displayed in Fig. 2

$$p_y(y) = p_x(x)\left|\frac{dx}{dy}\right| = p_x(x)\frac{1}{p_x(x)} = 1 \qquad (16)$$

Thus, we obtain the uniform PDF $p_y(y)$ on transformation of subband x with $T(x)$. The limits of $T(x)$ can be changed and made adaptive as per the requirement of available query facial image for facial feature enhancement. The limits of uniform PDF in the range of 0 to x are changed to $\lim_{low}$ and $\lim_{high}$. In a discrete version, PDF of an input variable x is given by (17).

$$p_x(x) = \frac{n_k}{n}, \quad k = \lim_{low}, \lim_{low}+1, ..., \lim_{high} \qquad (17)$$

Where, n denotes total number of pixels in the span $[\lim_{low}, \lim_{high}]$ and n_k stands for frequency of pixels with intensity x_k.

$$y_k = T(x_k) = \sum_{j=0}^{k} p_x(x_j) = \sum_{j=0}^{k} \frac{n_j}{n}, \qquad (18)$$
$$k = [\lim_{low}, \lim_{high}]$$

Thus, the wavelet subband coefficients within the span $k = [\lim_{low}, \lim_{high}]$ are mapped uniformly again to $[\lim_{low}, \lim_{high}]$ keeping the other insignificant coefficients intact. The resultant enhanced wavelet subband is denoted with Y. Thus, we achieve selective and adaptive equalization of W.

Step 5: The enhanced wavelet subband Y is remapped to the original range using (19).

$$W_E = W_{\min} + \left[\frac{Y}{255} \times (W_{\max} - W_{\min})\right] \qquad (19)$$

The resultant feature enhancement is shown in Fig. 7.

Note the difference between the 3-D elevations corresponding to eyes, nose, lips etc. in Fig. 7a and b. The top view as in Fig. 7c and d is only marginally different though (d) is enhanced version of (c). However, the enhanced cheek profile curves alongwith facial texture in Fig. 7d can be easily observed.

4 Proposed feature and score level fusion

This section illustrates the proposed expression invariant face recognition technique based on feature and score level fusion. Each original image is processed using the proposed technique to detect the feature vector as discussed further.

Some of our heuristic concepts to divide the preprocessed facial image into overlapping regions to achieve score level matching are discussed further. The original facial image

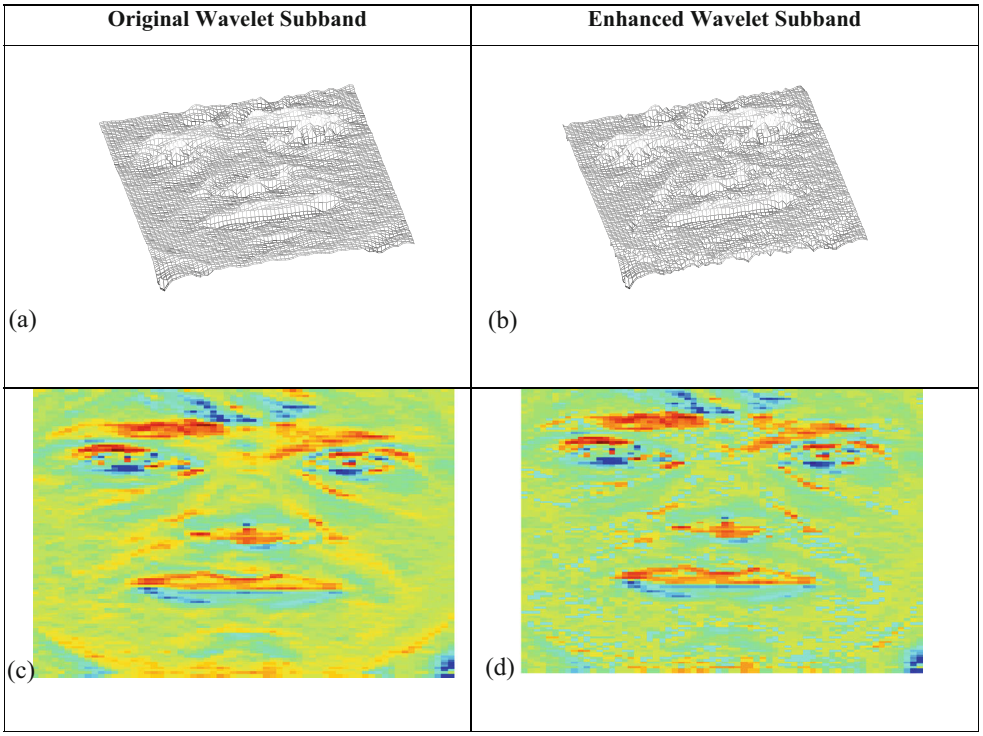

Fig. 7 Wavelet subband and boosting using proposed function (**a**) Original *LPHP* subband at level-1 in 3-D domain, (**b**) enhanced *LPHP* subband at level-1 in 3-D domain, (**c**) original *LPHP* subband in 2-D domain (**d**) Enhanced *LPHP* subband in 2-D domain. Note the clearly visible enhanced skin texture and edges in (**b**) and (**d**)

is split into six main regions namely, region A, region B, region C, region D, region E and region F as shown in Fig. 8. The 128x128 size facial images are divided by two along each dimension to achieve the regions A, C, D and F.

The region B is 50 % overlap of A and C while region E is 50 % overlap of D and F. Initially these regions are matched with the corresponding database entries and matching scores were computed using individual single regions. However, during these intermediate experiments it was found that, matching using region F yields fewer matching scores. This is obvious because, this region of the face is most vulnerable to expression variations. If we avoid this region during the feature detection, we are most likely to obtain expression invariant features. Also, the matching scores using regions D and E separately were found to be drastically less. In fact, the regions D and E mainly contain cheek and lips. Lips are the most sensitive parts of human face during expression changes.

The cheek and jaws are not much sensitive to expression changes as they do not carry prominent discriminatory features for expression invariant face recognition. The regions A, B and C were found to yield high individual matching scores due to the presence of forehead, eye sockets, nose bridge and major portion of nose, in spite of the deformation in eyes due to expression variations. Hence, we decided to use only regions A, B and C for face matching using score level fusion.

Figure 9 depicts the overall flow of the proposed method. The forehead, eye sockets and upper nose (nose bridge) area of the human face are less vulnerable to expression changes. The spatial domain Patch-LDSMT histogram is denoted by H_{LDSMT}. The adaptive wavelet transform coefficients WLD histogram is expressed with H_{EWT_WLD}. The concatenated $\{H_{LDSMT}, H_{EWT\_WLD}\}$ matrix represents a resultant feature vector for a complete original facial image. Dimensionality reduction is performed to reduce the computational complexity and to eliminate insignificant features. LDA [31] is employed for the dimensionality reduction of the resultant feature vector.

The patch-wise LDSMT feature histograms are utilized to extract the dense pixel level variations in the spatial domain for every patch and subsequently for the complete image. *2-D* orthogonal wavelet transform is applied on the original facial image regions followed by the proposed R and C-decimated wavelet transform in order to obtain multi-directional frequency domain features. The same is also applied on complete images. We boost all the high pass wavelet coefficients of individual patches and of complete images using the proposed adaptive wavelet coefficients enhancement function.

The multi-block WLD histograms of each enhanced wavelet subband coefficients represent a transform domain feature vector for every patch after concatenation. Similarly, the WLD histogram of the complete image works as a complete feature vector in transform domain for the image. Further, we have fused the spatial patch-LDSMT histograms and wavelet transform domain feature vectors using concatenation of the separate feature vectors [32]. The fixed size patch histogram feature of complete image is also computed using the LDSMT.

The similarity score is computed for region A, region B and region C respectively. The similarity scores are also computed for complete images. The contingency tables (confusion matrix) containing similarity scores are summed up to obtain a score level fusion of the three region scores. The ideal span of resultant contingency table is [0,3]. Similarity score level fusion results in slightly better recognition performance.

The test (query) feature vectors are compared against the feature vectors in the database using the cosine similarity distance based nearest neighbor classifier. The orientation based cosine similarity [33] between train and test feature vectors f_{train} and f_{test} is given by (20). The cosine

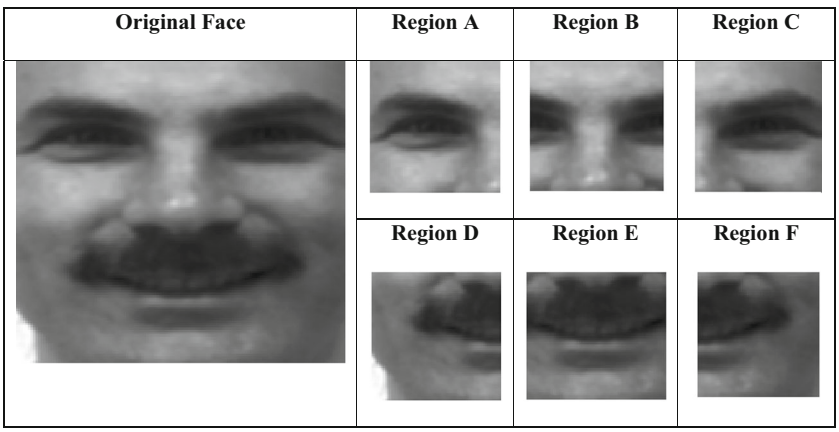

Fig. 8 Original facial image and regions *A, B, C, D, E* and *F*

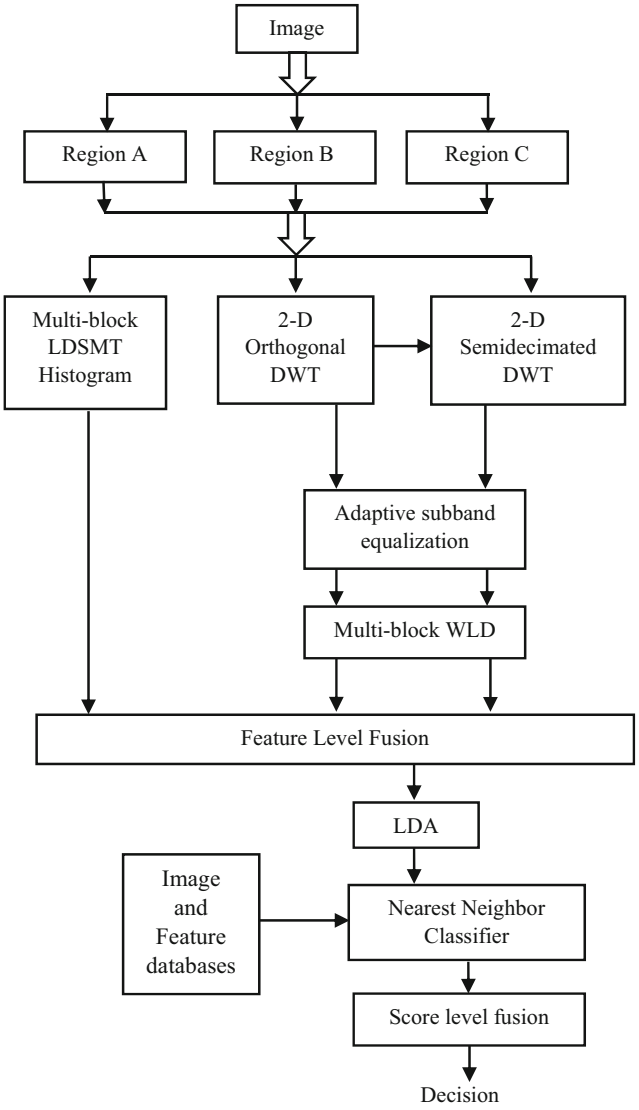

Fig. 9 Algorithm of the proposed robust expression invariant face recognition system

similarity measure is suitable to classify closely spaced samples in very large data spaces [34].

$$Sim_{\cos}(f_{train}, f_{test}) = \frac{\sum_{i=1}^{n} \vec{f}^{i}_{train} \cdot \vec{f}^{i}_{test}}{\sqrt{\sum_{i=1}^{n} \vec{f}^{i2}_{train}} \cdot \sqrt{\sum_{i=1}^{n} \vec{f}^{i2}_{test}}} \quad (20)$$

Figure 9 presents the algorithmic flow of the proposed algorithm for processing of the three regions depicted in Fig. 8. The same process flow is applicable to a complete image. However, score level fusion will not be applicable in case of processing of complete images. Rather, processing a complete image and matching it with the corresponding database entry yields the final matching score.

The algorithm for computation of expression invariant facial feature vector is proposed in terms of the following modules: (i) Pre-processing, (ii) Patch- LDSMT computation, (iii) Enhanced wavelet subband computation, (iv) WLD based feature extraction, (v) Feature level fusion, (vi) Dimensionality reduction, and (vii) Construction of feature vector database (viii) Matching using LDA. Typical preprocessing steps include initial skin detection, ears and neck region removal resulting in a cropped facial image. The cropped facial image in converted into gray image, resized to 128x128 size and used for further processing. As can be seen in Fig. 9, the facial image is subdivided into equal sized least expression sensitive regions A, B, and C while performing region based processing. In case of 'complete image' based processing, the region separation step is skipped. The input to pre-processing module is either regions of facial image or entire facial image. As the cropped facial image is resized to 128x128 square size image, the outputs of this module are square sized region A, B and C images. In the patch-LDSMT computation, the square size images are subdivided into small square patches (blocks). For each patch, the customized patch-LDSMT histogram is computed which represents a histogram of local texture in the spatial domain. The concatenated versions of the histograms computed for the patches are denoted by a feature vector H_{LDSMT} and is the output of this module. For obtaining frequency domain representation of the image, the conventional DWT and the proposed semi-decimated DWT are used followed by the coefficient equalization and multi-block Weber Local Descriptors for feature detection. For the enhanced wavelet subband computation module used for the equalization, wavelet subbands from the forward orthogonal DWT and semidecimated DWT act as inputs. Each subband is processed with the proposed adaptive wavelet coefficient equalization algorithm which yields the enhanced wavelet subband. A set of the enhanced wavelet subbands is the output of this module. Each wavelet subband acts as an input to the WLD based feature extraction module for each region A, B and C. The subband is subdivided into multiple square sized patched and WLD feature vector is computed for each patch of each band. The WLD feature vector is normalized so that it can be further fused with spatial domain LDSMT feature vector. The concatenated set of the feature vectors form a resultant feature vector H_{EWT_WLD} that is the output of this module. Outputs of module (ii) and (iv) act as inputs to feature level fusion module. In this module, an ultimate feature vector is derived from combination of all the spatial and subband the feature vectors H_{LDSMT} and H_{EWT_WLD}. This class of fusion belongs to 'fusion before matching' scheme. The Resulting feature vector f_{res} acts as an output of this module. In order to remove feature redundancy, the dimensionality reduction is performed on feature vector f_{res}. The dimensionality reduction module

maps the feature vector f_{res} to a lower dimensional subspace using Linear Discriminant Analysis which results in a feature vector $\vec{f}_{train}$. Subsequently, the feature vector $\vec{f}_{train}$ is stored along with its corresponding class in the feature vector database.

A query also goes through all these processing steps and the feature vector of the query is matched one by one with all the database entries using the cosine distance as the final step of LDA. The nearest neighbor is ranked as the first match to the query and so on. The procedure for computation of a training feature vector $\vec{f}_{train}$ is described in Algorithm 1.

Algorithm 1 Proposed algorithm for extraction of expression invariant features of a training image and a region belonging to the training image.

Input: A region I_{train} from a training image, Patch-LDSMT patch size L_p, WLD patch size D_p, wavelet transform forward decomposition levels W_L

Output: Feature vector $\vec{f}_{train}$

Step 1 (Preprocessing):

1.1 If a region I_{train} is in the RGB format; convert it to gray-scale.

1.2 Resize I_{train} to 128×128 pixels and denote it with J_{train}

Step 2 (Patch-LDSMT Computation):

2.1 Split the image J_{train} into multiple patches P_i each with a size of $L_p \times L_p$ pixels.

2.2 ***for*** each patch do

for each pixel within the patch do

2.3 Calculate Patch-LDSMT representation

end for

2.4 Calculate a 1-D histogram $H_i, i = [1,2,...P_i]$

2.5 Concatenate all 1-D histograms to configure a feature vector H_{LDSMT}

end for

Step 3 (Enhanced wavelet subband computation):

for each level $[1,..W_L]$, ***do***

compute 2-D orthogonal DWT coefficients and semidecimated wavelet coefficients using Haar wavelets

for each wavelet subband X, ***do***

boost it using the adaptive wavelet coefficient enhancement algorithm in order to obtain the enhanced wavelet coefficient Y

end for

end for

Step 4 (Feature extraction from the enhanced wavelet coefficients using WLD):

for each enhanced wavelet subband Y ***do***

4.1 Split Y into various patches P_q, with each patch of the size $D_p \times D_p$ coefficients.

end for

4.2 ***for*** each patch ***do***

for each wavelet coefficient within a patch ***do***

4.3 Compute the WLD 1-D histogram H_j, for each patch, where $j = [1,2,....,P_q]$

end for

4.4 Concatenate all histograms to construct a feature vector H_{EWT_WLD}

end for

Step 5 (Feature level fusion):

5.1 ***for*** each region and complete image ***do***

Compute a resultant feature vector f_{res} upon feature level fusion by concatenation of H_{LDSMT} and H_{EWT_WLD}

end for

5.2 ***for*** each region and complete image ***do***

Store patch features

Store image features

end for

Step 6 (Dimensionality Reduction):

6.1 ***for*** individual patch features ***do***

Perform the dimensionality reduction of f_{res} using the LDA to obtain $\vec{f}_{train}$.

end for

6.2 ***for*** complete image ***do***

Perform the dimensionality reduction of f_{res} using the LDA to obtain $\vec{f}_{train}$.

end for

Step 7 (Feature vector database):

Save the patches and image feature vector $\vec{f}_{train}$ along with its label into a feature vector database f_{db}

For a given test image, a test feature vector $\vec{f}_{test}$ is computed using steps 1-6 of Algorithm 1. The $\vec{f}_{test}$ is compared against the feature vectors $\vec{f}_{train} \in f_{db}$ to obtain a label using the cosine similarity based nearest neighbour classifier. Note that, the parameters W_L, D_p and L_p must be same for $\vec{f}_{train}$ as well as $\vec{f}_{test}$.

5 Experiments and results

In this section, we describe the experiments and evaluate the performance of the proposed approach on different expressive facial databases. Figure 10 displays the representative faces from the JAFFE, CMU-AMP, ORL, Yale, CK and our own database.

For all the databases, preprocessed original facial images are resized to 128×128 pixels and are stored in our system database. The *Algorithm 1* is applied on all the described regions as in Fig. 8 followed by the complete image. All the individual region and complete image features in transform

Fig. 10 Representative expressive faces from **a** the JAFFE facial expression database, **b** the CMU-AMP database, **c** the ORL database, **d** the Yale face database, **e** the Cohn-Kanade (CK) database and **f** Our own database

domain and spatial domain fused as described in Section 4 are stored in the system database. The image matching scores computed using score level fusion of region matching and feature level fusion for complete image matching are computed as discussed in Section 4.

A very robust benchmarking protocol has been used for performance evaluation of the proposed algorithm. For example, a facial image database has eight expression faces of each person. Initially, we select the first three faces as training images and remaining five faces as testing images for every person. For all the training images of every person, the features were computed and were stored in the system database. Thus with this facial image database, one can have minimum eight combinations of three training faces to select all the faces at least three times in their original sequence in the facial database. With every combination the remaining five faces were used as testing images or queries for every person. Every query image of each face was applied once. Thus, all such combinations of the three training facial images of each person were used for training and all the remaining five faces of the person were used for testing.

The internal patch size of the Patch-LDSMT and EWT_WLD are chosen to be $L_p = 16$ and $D_p = 8$ respectively after heuristic experimentation. The number of decomposition levels for the wavelet transform $W_L = 3$.

On every combination of training images, the recognition rate was computed and average of all such recognition rates for all the combinations has been reported as the rank 1 recognition rate performance of the proposed algorithm on the facial image database in the result section. This protocol was used for performance evaluation of the proposed algorithm on each of the listed databases.

Table 1 Comparison with other approaches (The JAFFE database)

Algorithm	Training images/person	Rank-one recognition rate
PCA-FLD-ANN [36]	10	84.90 %
B-JSM [37]	4	95.12 %
Higher Order-SVD [38]	6	92.96 %
LLE-Eigen [39]	18	93.93 %
Proposed approach without regions (complete image)	4	98.03 %
Proposed approach with regions	4	98.33 %

Table 2 Benchmarking of rank-one recognition rates on the CMU-AMP database

Algorithm	Training images/person	Rank-one recognition rate
B-JSM [37]	4	98.95 %
Proposed approach without regions (complete image)	4	99.02 %
Proposed approach with regions	4	99.27 %

5.1 Experimentation on the JAFFE face database:

The Japanese Female database [35] consists of a set of 213 facial images captured from 10 persons. The expression variations considered are neutral, sad, disgust, anger, happy, surprise and fear. This database is widely utilized for benchmarking of expression invariant face recognition approaches. We have sequentially separated four facial images/person as a training set and the remaining faces are used as a testing dataset. Table 1 shows the benchmarking of the proposed approach with regions and the complete image along with the other state-of-the-art methods.

5.2 Experimentation on the CMU-AMP face database:

The CMU-AMP database [40] contains a collection of facial images from 13 persons, with 75 images/person. This database is acquired under fixed lighting conditions and explicitly involves expression variations. The training dataset involves sequentially chosen 4 images/person and remaining images are treated as a testing dataset. Table 2 depicts the benchmarking of the proposed algorithm with state-of-the-art methods used on this database in the published literature.

5.3 Experimentation on the ORL face database

The ORL face database [41] contains 400 images from 40 distinct subjects. Our experimentation is performed with 3 images/subject as a training dataset, with rest of the faces treated as test images. Benchmarking with existing approaches is indicated in Table 3.

5.4 Experimentation on the Yale face database

The Yale face database [31, 46] contains a set of 165 images from 15 subjects. The expression variations like happy, sleepy, wink, normal, surprised, centre-light-neutral, no-glasses-neutral and sad are captured. We experimented on the Yale face database by successively choosing three images/person as a training dataset and the remaining images as a test dataset. Table 4 depicts the comparison of our approach with the state-of –the-art methods.

5.5 Experimentation on the Cohn-Kanade (CK) dataset

The Cohn-Kanade (CK) face expression dataset [50] consists of a collection of 8795 images from 97 distinct persons. The captured expressions include anger, disgust, happy, sadness, surprise, fear and neutral. All the expressions are varied from mild to severe strength. The experiments on the CK database are performed as per two distinct protocols as described in [7] and [37]. Nagesh and Li [37] utilized B-JSM compressive sensing based technique and performed the experimentation by varying the number of training images (J). The training images are chosen successively and recognition rates indicate average of nine distinct trials. As indicated in Table 5, we report higher average rank-one recognition rate in all the settings.

Riaz et al. [7] experimented with 4000 facial images from 62 persons. The training dataset consists of 66 % of the total facial images and remaining images are specified as test images. Benchmarking with two distinct methods as described in [7] is presented in Table 6.

Table 3 Comparison of rank-one scores on the ORL database

Algorithm	Training images/person	Rank-one recognition rate
GM-DE [42]	3	95.00 %
DWT-DDFFE [43]	4	96.82 %
Wavelet-Energy [44]	5	90.50 %
Gabor-MPCA [45]	3	88.90 %
Proposed approach without regions (complete image)	3	98.02 %
Proposed approach with regions	3	98.14 %

Table 4 Comparison with other approaches (the Yale database)

Algorithm	Training images/person	Rank-one recognition rate
FT-DCT-DWT [47]	3	82.50 %
Wavelet-DCT [9]	3	91.82 %
2D-GFD-FLD [48]	3	70.80 %
Gabor-NNDA [49]	5	90.00 %
Proposed approach without regions (complete image)	3	98.04 %
Proposed approach with regions	3	98.44 %

Table 5 Benchmarking of rank-one scores with B-JSM on the Cohn-Kanade data set

Training images per person (J)	Rank-1 recognition rate		
	B-JSM [37]	Proposed algorithm without regions (complete image)	Proposed algorithm with regions
5	95.47 %	96.2 %	96.32 %
6	95.93 %	97.00 %	97.11 %
7	96.15 %	97.81 %	97.95 %
10	97.14 %	98.4 %	98.49 %

Table 6 Benchmarking of rank-one recognition rates with active appearance model based technique [7] on the Cohn-Kanade data set

Algorithm	Rank-1 recognition rate	Number of training images	Number of testing images	Total images
AAM-BDT [7]	91.17 %	2679	1381	4060
AAM-BN [7]	98.69 %	2679	1381	4060
Proposed algorithm without regions (complete image)	98.93 %	2679	1381	4060
Proposed algorithm with regions	98.95 %	2679	1381	4060

Table 7 Validation of Rank-one recognition rates using the proposed approach on our own database

Training images/person	Testing images/person	Rank-one recognition rate without regions (complete image)	Rank-one recognition rate with regions
5	2	100 %	100 %
4	3	97.3 %	98.73 %
3	4	96.5 %	97.38 %
2	5	95.00 %	95.87 %

Table 8 Summary of Closed and Open Set experimentation

	Accept	Reject
True	As presented in Tables 1 to 7	0 %
False	0 %	100 %

5.6 Validation on our own dataset

Our own database consists of 560 facial images from 80 distinct persons. It includes expression variations like neutral, sad, anger, fear, smile, surprise and disgust. While validating our approach with this database, we have varied the training to testing faces ratio. The training images are chosen sequentially from available facial images. Table 7 depicts the rank-one performance using the proposed approach.

We have used a computer system with 32GB RAM, Intel i7-4770 CPU@3.4GHz and Windows 8 operating system for the experimentation using MATLAB. For the databases containing less than or equal to 400 total images the response time is less than 1 second and for databases containing images from 400 to 9000 the response time is roughly 4 seconds.

The results presented so far indicate the rank-1 recognition rates (i.e. true accept). We have further carried out the extension to find out false accept and false reject rate. The system was trained using three faces from the Yale database and the faces in JAFFE database were used as query images. The matching threshold for declaring the image as 'Not in the Database' was selected 75 %. For example, an image that is not available in the database will also yield some matching score with every image available in the database. If maximum of this matching score is less than 75 % of the max then, the query image is declared 'Not in the Database'. This experimentation is also called open set experimentation. The open set experimentation on these two databases yield 0 % false accept and true reject rates. The summary of closed and open set experimentation is presented in Table 8.

6 Discussion

The expression invariant face recognition can be effectively performed with the proposed technique as indicated by the experimental outcomes. We would like to point out that the proposed feature level fusion scheme outperforms the individual Patch-LDSMT and wavelet transform with WLD based approaches in terms of the accuracy of face recognition. The proposed technique is accurate even when the number of training images/person are sufficiently low.

This paper has presented a scheme for spatial domain and modified transform domain feature detection, fusion and matching for face recognition. Patch-LDSMT detects spatial domain features. The proposed combination of full decimated and semidecimated DWT offers directional facial features that are enhanced using the proposed adaptive histogram equalization and subsequent matching. The dimensionality reduction using LDA and consequent feature level fusion yields consistently high recognition rates.

Highlights of the contributions made by the proposed approach are as follows:

1. It uses orthogonal DWT and SDWT that is the computational by-product of DWT. Hence it is computationally simple against contourlet, curvelet and ridgelet transforms. SDWT computations require less memory compared to UDWT.
2. Use of semidecimated DWT preserves edges at around 60° and 30° against 0°, 90° and 45° with the use of DWT.
3. The conventional CLBP used for texture analysis has been modified in the form of patch-LDSMT for face recognition by dropping the mean value of the complete difference image. The mean value of complete difference image has hardly any significance in case of facial image. In texture, it is significant.
4. The feature concatenation using regular but arbitrary fixed size regions has been modified using human face anatomical overlapping regions. The overlap helps in obtaining robust matching score in case of possible less accurately registered query. Some of the overlapping boundary features of region A or C may be captured in region B and contribute to final matching score using score level fusion.
5. The local uniform size anatomically separated overlapping regions assist in having a robust matching score or decision in presence of non-uniform illumination variations and expression variations due to region-level feature detection and matching. This will also help in occluded facial image matching as one of the regions may completely match while the others are occluded.
6. A novel adaptive DWT and SDWT coefficient enhancement function has been proposed for feature boosting before subsequent extraction.
7. Use of cosine similarity that is generally employed for exploring and matching data structures in huge databases in case of data mining applications.
8. Use of rigorous recognition and verification protocols for performance benchmarking against the techniques proposed in the published literature.

7 Conclusion

The conventional DWT offers directional features at 0^0, 90^0 and $\pm 45^0$. Its semidecimated version protects directional features at around 30^0 and 60^0 in horizontal and vertical subbands at each level. Thus, at three levels of the conventional and semidecimated transforms, directional features at 16 different angles are offered for matching. Patch-LDSMT and WLD offer illumination independent features in spatial and frequency domain respectively. Finally, we conclude

that DWT and SDWT preserve both, spatial as well as frequency domain information, for minute feature detection of objects like human faces without any computational overhead unlike contourlet or curvelet transforms.

In this work, we proposed a novel approach using spatial Patch-LDSMT as well as wavelet domain WLD features for expression invariant face recognition. One of the major contributions of this paper lies in proposing the adaptive wavelet transform coefficient enhancement algorithm. The 3-level 2-D orthogonal DWT and SDWT coefficients yield multiresolution and localized features. The local level dense pixel descriptors Patch-LDSMT and WLD are utilized for feature extraction in spatial and transform domain respectively for the first time for face recognition. WLD can effectively process negative valued coefficients, which is advantageous for feature extraction from the wavelet transform subbands. After extensive experimentation on the JAFFE, CMU-AMP, ORL, Yale and CK database, it is observed that the system provides much better rank-one recognition rates than the state-of-the-art methods. Validation with our own grabbed facial images proves its generality. Finally, the proposed SDWT, the adaptive wavelet coefficient enhancement function, overlapping region feature detection using Patch-LDSMT and WLD, feature and score level fusion make the proposed algorithm very robust and useful for researchers working in the area of occluded and expression invariant face recognition, micro-expression analysis and human-computer interaction.

References

1. Bartlett MS, Movellan JR, Sejnowski TJ (2002) Face recognition by independent component analysis. IEEE Trans Neural Netw 13(6):1450–1464. doi:10.1109/TNN.2002.804287
2. Zhao W, Chellappa R, Phillips PJ, Rosenfeld A (2003) Face recognition: a literature survey. ACM Comput Surv 35(4):399–458. doi:10.1145/954339.954342
3. Zhang X, Gao Y (2009) Face recognition across pose: a review. Pattern Recogn 42(11):2876–2896. doi:10.1016/j.patcog.2009.04.017
4. Kong SG, Heo J, Abidi BR, Paik J, Abidi MA (2005) Recent advances in visual and infrared face recognition—a review. Comput Vis Image Underst 97(1):103–135. doi:10.1016/j.cviu.2004.04.001
5. Chen S, Lovell BC (2004) Illumination and expression invariant face recognition with one sample image. In: 17th International Conference on Pattern Recogn, 23–26 Aug., Cambridge, England, UK. doi:10.1109/ICPR.2004.1334112
6. Xiaoxing L, Mori G, Hao Z (2006) Expression-invariant face recognition with expression classification. In: 3rd Canadian conference on computer and robot vision, 07–09 June., Quebec, Canada. doi:10.1109/CRV.2006.34
7. Riaz Z, Mayer C, Wimmer M, Beetz M, Radig B (2009) A model based approach for expressions invariant face recognition. In: Tistarelli M, Nixon M (eds) Advances in biometrics, vol 5558. Lecture notes in computer science. Springer Berlin Heidelberg, pp 289–298. doi:10.1007/978-3-642-01793-3_30
8. Lee H-S, Kim D (2008) Expression-invariant face recognition by facial expression transformations. Pattern Recogn Lett 29(13):1797–1805. doi:10.1016/j.patrec.2008.05.012
9. Abbas A, Khalil MI, Abdel-Hay S, Fahmy HM (2008) Expression and illumination invariant preprocessing technique for face recognition. In: International conference on computer engineering & systems. 25–27 Nov., Cairo, Egypt. doi:10.1109/ICCES.2008.4772966
10. Vankayalapati HD, Kyamakya K (2009) Nonlinear feature extraction approaches with application to face recognition over large databases. In: 2nd international workshop on nonlinear dynamics and synchronization. 20–21 July. Klagenfurt, Austria. doi:10.1109/INDS.2009.5227967
11. Yi W, Jian-Ping L, Jie L, Lin L (2008) The contourlet transfrom and SVM classification for face recognition. In: The international conference on apperceiving computing and intelligence analysis, 13–15 Dec.Chengdu, China. doi:10.1109/ICACIA.2008.4770006
12. Ahonen T, Hadid A, Pietikainen M (2006) Face description with local binary patterns: application to face recognition. IEEE Trans Pattern Anal Mach Intell 28(12):2037–2041. doi:10.1109/TPAMI.2006.244
13. Yang B, Chen S (2013) A comparative study on local binary pattern (LBP) based face recognition LBP histogram versus LBP image. Neurocomputing 120(1):365–379. doi:10.1016/j.neucom.2012.10.032
14. Zhenhua G, Zhang L, Zhang D (2010) A completed modeling of local binary pattern operator for texture classification. IEEE Trans Image Process 19(6):1657–1663. doi:10.1109/TIP.2010.2044957
15. Do MN, Vetterli M (2003) The finite ridgelet transform for image representation. IEEE Trans Image Process 12(1):16–28. doi:10.1109/TIP.2002.806252
16. Starck JL, Candes EJ, Donoho DL (2002) The curvelet transform for image denoising. IEEE Trans Image Process 11(6):670–684. doi:10.1109/TIP.2002.1014998
17. Do MN, Vetterli M (2005) The contourlet transform: an efficient directional multiresolution image representation. IEEE Trans Image Process 14(12):2091–2106. doi:10.1109/TIP.2005.859376
18. Wang J, Ying Y, Guo Y, Peng Q (2006) Automatic foreground extraction of head shoulder images. In: Nishita T, Peng Q, Seidel H-P (eds) Advances in computer graphics. Lecture notes in computer science, vol 4035. Springer, Heidelberg, pp 385–396. doi:10.1007/11784203_33
19. Bai-ling Z, Haihong Z, Shuzhi SamG (2004) Face recognition by applying wavelet subband representation and kernel associative memory. IEEE Trans Neural Netw 15(1):166–177. doi:10.1109/TNN.2003.820673
20. Cheng Y, Wang CL, Li ZY, Hou YK, Zhao CX (2010) Multiscale principal contour direction for varying lighting face recognition. Electronics Letters 46(10):680–682. doi:10.1049/el.2010.0696
21. Haifeng H (2015) Illumination invariant face recognition based on dual-tree complex wavelet transform. IET Comput Vis 9(2):163–173. doi:10.1049/iet-cvi.2013.0342
22. Weeks M (2010) Digital signal processing using MATLAB and wavelets, 2nd edn. Jones and Bartlett Publishers Inc
23. Chang SG, Bin Y, Vetterli M (2000) Spatially adaptive wavelet thresholding with context modeling for image denoising. IEEE Trans Image Process 9(9):1522–1531. doi:10.1109/83.862630
24. Starck JL, Fadili J, Murtagh F (2007) The undecimated wavelet decomposition and its reconstruction. IEEE Trans Image Process 16(2):297–309. doi:10.1109/TIP.2006.887733
25. Jie C, Shiguang S, Chu H, Guoying Z, Pietikainen M, Xilin C, Wen G (2010) WLD: a robust local image descrip-

tor. IEEE Trans Pattern Anal Mach Intell 32(9):1705–1720. doi:10.1109/TPAMI.2009.155
26. Sengar PS, Rawat TK, Parthasarathy H (2013) Color image enhancement by scaling the discrete wavelet transform coefficients. In: The annual international conference on emerging research areas and international conference on microelectronics, communications and renewable energy (AICERA/ICMiCR), 4–6 June. Kanjirappally, India. doi:10.1109/AICERA-ICMiCR.2013.6575994
27. Yongsheng G, Leung MKH (2002) Face recognition using line edge map. IEEE Trans Pattern Anal Mach Intell 24(6):764–779. doi:10.1109/TPAMI.2002.1008383
28. Hansung L, Yunsu C, Jang-Hee Y, Chulho W (2012) Face recognition based on sparse representation classifier with gabor-edge components histogram. In: Eighth international conference on signal image technology and internet based systems, 25–29 Nov. Naples, Italy. doi:10.1109/SITIS.2012.26
29. Velde KV (1999) Multi-scale color image enhancement. In: International conference on image processing Kobe, Japan. doi:10.1109/ICIP.1999.817182
30. Gonzalez RC, Woods RE (1992) Digital image processing, 1st edn. Addison-Wesley Longman Publishing Co., New York
31. Belhumeur PN, Hespanha JP, Kriegman D (1997) Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans Pattern Anal Mach Intell 19(7):711–720. doi:10.1109/34.598228
32. Rattani A, Kisku DR, Bicego M, Tistarelli M (2007) Feature level fusion of face and fingerprint biometrics. In: First IEEE international conference on biometrics: theory, applications, and systems, 27–29 Sept. Crystal City, VA. doi: 10.1109/BTAS.2007.4401919
33. Pirlo G, Impedovo D (2013) Cosine similarity for analysis and verification of static signatures. IET Biom 2(4):151–158. doi:10.1049/iet-bmt.2013.0012
34. Chen C, Tang J (2012) Simulation study on the performance of several classifiers in face recognition. In: The 9th international conference on fuzzy systems and knowledge discovery, 29–31 May 2012. Sichuan, China. doi:10.1109/FSKD.2012.6233940
35. Lyons M, Akamatsu S, Kamachi M, Gyoba J (1998) Coding facial expressions with Gabor wavelets. In: The 3rd IEEE international conference on automatic face and gesture recognition, 14–16 Apr 1998. Nara, Japan. doi:10.1109/AFGR.1998.670949
36. Tsai P, Jan T (2005) Expression-invariant face recognition system using subspace model analysis. In: The IEEE international conference on systems, man and cybernetics, 10–12 Oct. 2005. Waikoloa, HI. doi:10.1109/ICSMC.2005.1571395
37. Nagesh P, Li B (2009) A compressive sensing approach for expression-invariant face recognition. In: The IEEE conference on computer vision and pattern recogn, 20–25 June 2009. Miami, FL. doi:10.1109/CVPR.2009.5206657
38. Hua-Chun T, Yu-Jin Z (2008) Expression-independent face recognition based on higher-order singular value decomposition. In: The international conference on machine learning and cybernetics, 12–15 July 2008. Kunming, China. doi:10.1109/ICMLC.2008.4620893
39. Abusham EE, Ngo D, Teoh A (2005) Fusion of locally linear embedding and principal component analysis for face recognition. In: The pattern recogn and image analysis. 1 Jan. 2005. Portugal. doi:10.1007/11552499_37
40. Liu X, Chen T, Kumar BVKV (2003) Face authentication for multiple subjects using eigenflow. Pattern Recogn 36(2):313–328. doi:10.1016/S0031-3203(02)00033-X
41. Samaria FS, Harter AC (1994) Parameterisation of a stochastic model for human face identification. In: The 2nd IEEE workshop on applications of computer vision, 5–7 Dec 1994. Sarasota, FL. doi: 10.1109/ACV.1994.341300
42. Maheshkar V, Agarwal S, Srivastava VK, Maheshkar S (2012) Face recognition using geometric measurements, directional edges and directional multiresolution information. Procedia Technol 6(0):939–946. doi:10.1016/j.protcy.2012.10.114
43. Ajit Krisshna NL, Deepak VK, Manikantan K, Ramachandran S (2014) Face recognition using transform domain feature extraction and PSO-based feature selection. Appl Soft Comput 22(1):141–161. doi:10.1016/j.asoc.2014.05.007
44. Cunjian C, Jiashu Z (2007) Wavelet energy entropy as a new feature extractor for face recognition. In: The 4th international conference on image and graphics, 22–24 Aug. 2007. Sichuan, China. doi: 10.1109/ICIG.2007.60
45. Gudur N, Asari V (2006) Gabor wavelet based modular PCA approach for expression and illumination invariant face recognition. In: The 35th IEEE applied imagery and pattern recogn workshop. 11–13 Oct. 2006. Washington, DC. doi:10.1109/AIPR.2006.24
46. Georghiades AS, Belhumeur PN, Kriegman D (2001) From few to many: illumination cone models for face recognition under variable lighting and pose. IEEE Trans Pattern Anal Mach Intell 23(6):643–660. doi:10.1109/34.927464
47. Lihong Z, Cheng Z, Xili Z, Ying S, Yushi Z (2009) Face recognition based on image transformation. In: The WRI global congress on intelligent systems. 19–21 May 2009. Xiamen, China. doi: 10.1109/GCIS.2009.308
48. Mutelo R, Woo W, Dlay S (2008) Discriminant analysis of the two-dimensional Gabor features for face recognition. IET Comput Vis 2(2):37–49
49. Kirtac K, Dolu O, Gokmen M (2008) Face recognition by combining Gabor wavelets and nearest neighbor discriminant analysis. In: The 23rd international symposium on computer and information sciences. October 27–29. Istanbul, Turkey. doi:10.1109/ISCIS.2008.4717916
50. Kanade T, Cohn JF, YingLi T (2000) Comprehensive database for facial expression analysis. In: The 4th IEEE international conference on automatic face and gesture recognition. March 28–30. Grenoble, France. doi:10.1109/AFGR.2000.840611

Hemprasad Patil received B.E. and M. Tech. degrees in Electronics and Telecommunication Engineering in 2007 and 2009 respectively. He is currently pursuing Ph.D. degree from Visvesvaraya National Institute of Technology, Nagpur, India. He is a member of IAENG. He has more than 15 publications to his credit. His research areas include intelligent face recognition, biometrics, signal processing, algorithm design, statistical pattern recognition, big data analytics and computer vision. He is a recipient of the JTF award from CICS, India.

Ashwin Kothari received B.E. and M. Tech. degrees in Electronics Engineering in 1994 and 2005. He further obtained Ph.D. degree from Visvesvaraya National Institute of Technology, Nagpur, India in 2010, where he is currently working as an associate Professor. He is principal investigator of funded major research project in the area of Image Processing, Communication and embedded Systems titled COMMBEDDED SYSTEMS. He is member of IEEE, IRSS, IAENG and ISTE. He has more than 40 publications to his credit. His research areas include Pattern classifiers, Rough-Neuro hybrid approach, intelligent face recognition, Rough set based processor design.

K. M. Bhurchandi received B.E. and M.E. degrees in Electronics Engineering in 1990 and 1992. He further obtained Ph.D. degree from Visvesvaraya Regional College of Engineering (VNIT), Nagpur University, Nagpur, India in 2002, where he is currently working as a Professor. He has more than 45 publications to his credit. He is principal investigator of a funded research project titled Scale Invariant Face Recognition. He is also a Principal Investigator of a funded project in the area of Image Processing, Communication and embedded Systems titled COMMBEDDED SYSTEMS. He is the co-author of a popular book titled Advanced Microprocessors and peripherals published by McGraw Hill, India. His research interests include color image processing and analysis, computer vision, digital signal processing, embedded systems and computer architectures.