

Robust face identification using DTCWT and PCA subspace based sparse representation

K. Selvakumar¹ · Jovitha Jerome¹ · Kumar Rajamani²

Received: 13 January 2015 / Revised: 17 August 2015 / Accepted: 24 August 2015 /

Published online: 15 September 2015

© Springer Science+Business Media New York 2015

Abstract This paper presents a robust method for recognizing human faces with varying illumination as well as partial occlusion. In the proposed approach, a dual-tree complex wavelet transform (DTCWT) is employed to normalize the illumination variation in the logarithm domain. In order to minimize the variations under different lighting conditions, appropriate low frequency DTCWT subbands are truncated and the rest of the directional subbands are used to reconstruct the stable invariant face. Using the fundamental concept that patterns from a single object class lie in a linear subspace, we develop class specific dictionaries using principal component analysis (PCA) based subspace learning on illumination invariant faces. By representing the pre-processed probe image against each dictionary using l_1 regularization into PCA reconstruction, target face and sparse noises are effectively factorized. Then, identification decision is made in favor of a class with minimum reconstruction error. Evaluations on challenging probe images demonstrate that the proposed method performs favorably against several state of the art methods.

Keywords Face identification · Illumination normalization · Partial occlusion · Dual-tree complex wavelet transform · Principal component analysis · Sparse representation

1 Introduction

Recently, face recognition research has seen a lot of success in less demanding commercial applications like Apple iphoto, Google Picasa, and Microsoft Photo Gallery. On the other hand, research in extremely demanding terrorist watchlist like applications in surveillance videos are still in premature stage. However, in between these two there are many currently

✉ K. Selvakumar
kskumaree@gmail.com

¹ Department of Instrumentation and Control Engineering, PSG College of Technology, Coimbatore, Tamilnadu 641004, India

² Robert Bosch Engineering and Business Solutions Limited, Bangalore, India

focused face identification applications like access control in secure buildings including offices, prisons, automatic teller machines, automobiles and computers, in which controlled gallery images can be collected in advance. In these kinds of identification applications, gallery subjects would act like allies not opponents [27]. For these applications, coupling among different variations like partial occlusion, blur, expression variation, illumination variation, pose change makes the face recognition problem extremely difficult, yet more meaningful [7]. Recently, many efficient face identification approaches have been reported to handle variations due to illumination [6, 8, 14, 26] and occlusion [21] separately. However, their performance degrades sharply due to the complexity in simultaneously handling multiple variations.

In face identification, illumination variation due to darkness or brightness is a very challenging issue to be addressed. It has been proven that intra-class difference due to illumination variation is very significant than inter-class difference [1]. Various methods that have been proposed in literature can be mainly classified into three categories; (i) illumination modelling (ii) illumination invariant feature extraction and (iii) pre-processing and normalization. In illumination modelling based approach, the main idea is to represent the illumination variations in a subspace and estimate the model parameters [2, 4]. But, the major drawbacks of this approach are requirement of larger amount of training data and huge computation time. In the invariant feature extraction approaches, the goal is to extract the facial features which are invariant to the illumination variations like gradient faces [26], local binary patterns (LBP) [24] and 2-D Gabor features [16]. However, experimental studies show that none of these methods are sufficient enough to represent the face images under extreme illumination variation [6]. The preprocessing and normalization approach remove the illumination variations in the images and extract the illumination invariant face image without any prior knowledge. In these approaches, histogram equalization [12] is the most commonly used one, in which the pixel intensities of the face image are redistributed by using a non-linear transformation function. Moreover, other non-linear image enhancement transforms like logarithmic and exponential functions are also used to correct the illumination variation to some extent [12]. In Chen et al. work [8], low frequency coefficients of discrete cosine transform (DCT) are discarded to eliminate the illumination variations because it mainly lies in the low-frequency bands. However, there are two major issues involved in this method; (i) how to select the appropriate number of DCT coefficients to be discarded; (ii) even after the normalization, illumination discontinuities caused by shadows will exist since it lies in the high-frequency bands. Goh et al. [11] proposed DWT decomposition based illumination normalization, by setting low frequency illumination component as zero. However, this method could not achieve better recognition performance due to aliasing and shift variance problems. To address these issues, recently, Haifeng [14] proposed multiscale illumination normalization approach using DTCWT for face recognition and achieved good recognition rate against illumination variations. However, experimental results of all these illumination invariant face recognition approaches show that good recognition rate was achieved only against the illumination variation not with other realistic issues like occlusion. On the other hand, for occlusion, the performance of the classical subspace learning based face recognition approaches like Eigenfaces and Fisherfaces along with the nearest neighbor are shown to be very poor recognition performance [3].

In recent developments, face recognition approaches based on linear representation have achieved state of the art performances against occlusion [23, 29]. Among them, linear regression classification (LRC) [23], and sparse representation classification (SRC) [29] are the most representative approaches. In LRC approach [23], face recognition has been cast as a

linear regression problem by representing a probe image as a linear combination of training images of each class. Reported results show that their modular approach algorithm has performed well against contiguous occlusion. However, this method lacks direct mechanism to handle partial occlusion. On the other side, to handle the partial occlusion, SRC approach has been extensively used for various applications including face recognition and object tracking [21, 29]. In the SRC based face recognition approach [29], it is considered to be a problem of finding the sparse representation of a test image in terms of all the training images and obtain the sparse error of that image which accounts for partial occlusion and random pixel corruption. However, SRC is not robust against severe illumination variation and contiguous occlusions such as shadow, sunglasses, and scarf. Moreover, it works fast only in low dimension feature vectors (e.g., 12×15 patches [21]) which may not capture the sufficient visual details to describe the face. Zhang et al., [31] argued that, in SRC approach, better discrimination is achieved mainly because of collaborative representation rather than l_1 norm based sparse error. With this argument, they proposed collaborative representation based classification using regularized least squares (CRC-RLS) which has better classification rate and significantly less computation complexity than the SRC approach. In CRC-RLS, fast computation is achieved due to l_2 regularization, as it has a close-form solution. Dong et al. [9] proposed a computationally efficient SRC approach by exploiting PCA subspace representation in l_1 regularization framework for tracking problems. Very recently, Xiao et al. [30] proposed a face identification method using robust principal component analysis, in which, given test face is decomposed into low rank face and error face for each subject, by solving the low rank matrix recovery optimization problem. However, in case of test images with large contiguous occlusion, discriminative information in the sparse error images used for classification will decrease. In spite of huge success in handling occlusion, all these sparse representation based methods cannot handle illumination variation effectively.

From the cited literature review on face recognition, handling multiple issues like illumination variation and occlusion simultaneously is still a major challenge. In order to address this problem, this paper proposes a novel face identification algorithm by combing DTCWT based illumination normalization technique with PCA subspace based sparse representation. First, the gallery images of all the subjects are preprocessed using DTCWT based illumination normalization technique. It is very effective to retain the geometrical structures in the facial image like the strength, orientation and contour. Then class-specific face model for all subjects is constructed by applying PCA on DTCWT faces in the gallery. Thereby, face recognition problem is defined as a problem of linear regression. This subspace representation helps to capture the rich and redundant properties compactly with relatively few basis vectors compared to a large number of training images. To handle the partial occlusion, sparse representation is used to represent the given probe image against all class models. This exploitation of PCA subspace within the sparse representation framework significantly reduces the computation time for higher resolution face images. Finally, the decision is ruled in favor of the class with minimum reconstruction error.

The remaining of this paper is organized as follows; Section II gives the DTCWT based illumination normalization process. Then, construction of class specific dictionaries, PCA subspace based sparse representation and proposed face identification algorithm are given in Section III. In section IV, experimental results are presented. Finally, conclusion of our work is drawn in section V.

2 Illumination normalization using DT-CWT

In this section, the DTCWT based illumination normalization technique and its effect on inter-class difference is given.

2.1 Illumination-reflectance model and logarithm transform

In general, we denote images by two-dimensional functions of the form $I(x,y)$. The function $I(x,y)$ can be characterized by two components (1) the amount of source illumination incident on the scene being viewed and (2) the amount of illumination reflected by the objects in the scene. These are called as illumination and reflectance components and are denoted by $L(x,y)$ and $R(x,y)$ respectively. The two functions combine as a product to form $I(x,y)$ [13, 15];

$$I(x,y) = L(x,y)R(x,y) \quad (1)$$

Based on the model given in (1), the following two assumptions are made on reflectance and illumination components. First, the reflectance component generally lies in the high frequency band in the given signal I , due to its abrupt changes. Second, the illumination component lies in the low frequency band of the given signal I , due its slow spatial variation.

Using the above assumptions, the goal is to extract the illumination invariant $R(x,y)$, which constitutes key geometrical structure in the facial image $I(x,y)$. In general, extracting R is quite difficult because of nonlinear variation of illumination. Equation (1) cannot be directly used to operate on the frequency components of reflectance and illumination because the transform of a product is not the product of transforms as shown below;

$$\mathcal{J}[I(x,y)] \neq \mathcal{J}[R(x,y)]\mathcal{J}[L(x,y)] \quad (2)$$

where $\mathcal{J}[]$ represents the frequency domain transform. Therefore, a logarithm of input image must be taken to approximate from the multiplicative reflectance-illumination model to the additive reflectance-illumination model [22] as given below;

$$\log I(x,y) = \log R(x,y) + \log L(x,y). \quad (3)$$

Using the $\log I(x,y)$, the reflectance component is extracted by DTCWT based multiscale illumination normalization technique.

2.2 Dual-tree complex wavelet transform

Conventional DWT mainly suffers from shift variance, and directional selectivity in denoising and pattern recognition applications [25]. In order to address these problems, Kingsbury introduced computationally efficient DTCWT [18]. Only the brief introduction of 2D DTCWT is given here and the comprehensive coverage of DTCWT and its relationship with other transforms can be found in [18]. For 2-D DTCWT, the separable implementation of 2-D wavelet $\psi(x,y)=\psi(x)\psi(y)$ is done by using filter-bank (FB) structure with associated complex wavelets as given below;

$$\left. \begin{aligned} \psi(x) &= \psi_h(x) + j\psi_g(x) \\ \psi(y) &= \psi_h(y) + j\psi_g(y) \end{aligned} \right\} \quad (4)$$

The expression for $\psi(x,y)$ is obtained as follows;

$$\begin{aligned} \psi(x,y) &= [\psi_h(x) + j\psi_g(x)] [\psi_h(y) + j\psi_g(y)] \\ &= \psi_h(x)\psi_h(y) - \psi_g(x)\psi_g(y) + j[\psi_g(x)\psi_h(y) + \psi_h(x)\psi_g(y)] \end{aligned} \tag{5}$$

The all other 2-D real orientation wavelets are obtained by repeating the above procedure with following 2-D complex wavelets: $\phi(x)\psi(y)$, $\psi(x)\phi(y)$, $\phi(x)\bar{\psi}(y)$, $\psi(x)\bar{\psi}(y)$, and $\psi(x)\bar{\phi}(y)$, where $\phi(x)=\phi_h(x)+j\phi_g(x)$ and $\phi(y)=\phi_h(y)+j\phi_g(y)$. The real parts of all these six oriented 2-D complex wavelets can be defined as

$$\left. \begin{aligned} \varphi_i(x,y) &= \frac{1}{\sqrt{2}}(\psi_{1,i}(x,y) - \psi_{2,i}(x,y)) \\ \varphi_{i+3}(x,y) &= \frac{1}{\sqrt{2}}(\psi_{1,i}(x,y) + \psi_{2,i}(x,y)) \end{aligned} \right\} \text{ for } i = 1, 2, 3. \tag{6}$$

where $\psi_{1,1}(x,y)=\phi_h(x)\psi_h(y)$, $\psi_{1,2}(x,y)=\psi_h(x)\phi_h(y)$, $\psi_{1,3}(x,y)=\psi_h(x)\psi_h(y)$, $\psi_{2,1}(x,y)=\phi_g(x)\psi_g(y)$, $\psi_{2,2}(x,y)=\psi_g(x)\phi_g(y)$, $\psi_{2,3}(x,y)=\psi_g(x)\psi_g(y)$. The normalization factor $\frac{1}{\sqrt{2}}$ is used to keep sum or difference as orthonormal operation. By the same way, the imaginary parts of six oriented 2-D complex wavelets are defined as follows;

$$\left. \begin{aligned} \xi_i(x,y) &= \frac{1}{\sqrt{2}}(\psi_{3,i}(x,y) - \psi_{4,i}(x,y)) \\ \xi_{i+3}(x,y) &= \frac{1}{\sqrt{2}}(\psi_{3,i}(x,y) + \psi_{4,i}(x,y)) \end{aligned} \right\} \text{ for } i = 1, 2, 3. \tag{7}$$

where $\psi_{3,1}(x,y)=\phi_g(x)\psi_h(y)$, $\psi_{3,2}(x,y)=\psi_g(x)\phi_h(y)$, $\psi_{3,3}(x,y)=\psi_g(x)\psi_h(y)$, $\psi_{4,1}(x,y)=\phi_h(x)\psi_g(y)$, $\psi_{4,2}(x,y)=\psi_h(x)\phi_g(y)$, $\psi_{4,3}(x,y)=\psi_h(x)\psi_g(y)$; $\phi_h(\cdot)$ and $\phi_g(\cdot)$ are the low pass functions of upper FB and lower FB respectively along first dimension. $\psi_h(\cdot)$ and $\psi_g(\cdot)$ are the high pass functions of upper FB and lower FB respectively along second dimension. Equations (6) and (7) produce the six directionally selective wavelets for each scale of the 2D DTCWT at approximately oriented angles of $\pm 15^\circ, \pm 45^\circ, \pm 75^\circ$. The impulse responses of real-part of each complex wavelet obtained using Eq. (6) are shown in Fig. 1. In this work, we have used the freely available DTCWT implementation by Cai and Li [5].

2.3 Illumination normalization algorithm

As the DTCWT holds better orientation selectivity, shift invariance, and less computation characteristics, it can be more efficient in eliminating illumination effect when compare to other representative methods. The block diagram of the DTCWT based illumination normalization technique is given in Fig. 2. Initially, the input image to this block diagram is enhanced by using the logarithm transform and then decomposed into three levels of frequency subbands by employing DTCWT. The number of levels of decomposition is decided based on the length of the filter coefficients and the dimension of the input image. The illumination component



Fig. 1 Impulse responses of six real valued directional wavelets associated with 2-D DTCWT obtained using Eq. (6)

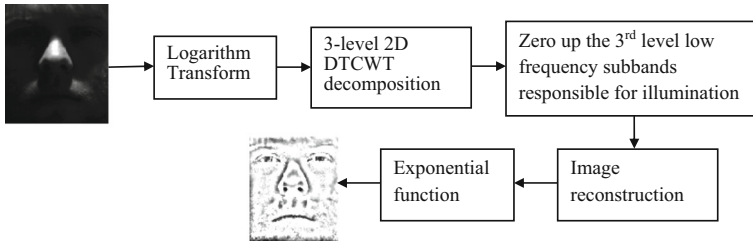


Fig. 2 DTCWT based illumination normalization block diagram

responsible for degrading recognition performance is eliminated by zeroing up the low frequency approximation subbands obtained at the highest level decomposition. Finally, the reflectance component is extracted by applying the inverse DTCWT over the high frequency directional subbands, followed by exponential transform on the reconstructed face image. Fig. 3a and b show the different lighting images and the corresponding illumination normalized images for Extended Yale-B dataset [10] and AR dataset [20] respectively.

Furthermore, in face identification problem, the illumination normalization step is mainly intended to reduce the intra-class difference and increase the inter-class difference. To demonstrate the ability of DTCWT illumination normalization step towards classification, we have used all training images of two example subjects i.e. class A and class B from the subset 5 of extended Yale-B dataset. Before and after illumination normalization, the training images of both class A and class B are provided to PCA, and then projected onto first two factors given by that dimensionality reduction method. Figure 4a and b show the training points projected onto PCA subspace before and after illumination normalization respectively. It is very clear that the separation between class A and class B after illumination normalization is superior.

3 Face identification using subspace based sparse representation

In this section, a new face identification algorithm is proposed using class specific face appearance models and subspace representation based l_1 regularization on illumination normalized feature space.

3.1 Class specific face appearance model

In linear face representation approaches, it is assumed that the faces belonging to one subject will lie in a linear subspace. Based on this principle, the proposed method uses PCA based subspace representation technique to learn face appearance model for each subject on illumination normalized feature space. The illumination normalized training faces of each subject are rescaled into same size feature vector $f \in \mathcal{R}^d$ to learn the appearance model. Let N be the total number of subjects and D_i be the dictionary that contains face images of particular subject as given below;

$$D_i = \{f_1, \dots, f_{n_i}\} \in \mathcal{R}^{d \times n_i}; i = 1, \dots, N; \tag{8}$$

where n_i is the number of training images for each subject. Using PCA, linear subspace model for each D_i is learned as explained below. To normalize the feature vectors, each dictionary is subtracted with their mean as



a



b

Fig. 3 Illustration of before and after illumination normalized sample input images of (a) AR dataset (b) Extended Yale B dataset

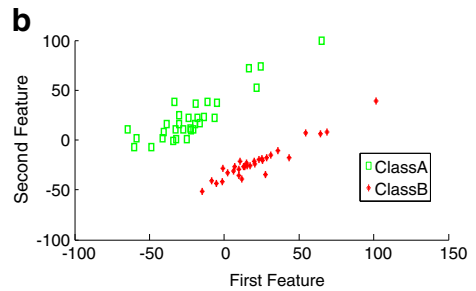
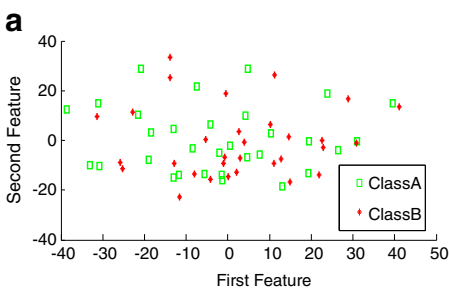


Fig. 4 PCA based 2D feature data distribution plot of class A and class B (a) before and (b) after illumination normalization

$$\bar{D}_i = D_i - m_i \times \mathbf{1}_{1 \times n_i}, \tag{9}$$

where $m_i \in \mathcal{R}^{d \times 1}$ is the mean of the n_i observations in D_i , and $\mathbf{1}_{1 \times n_i} = \overbrace{(\mathbf{1}, \dots, \mathbf{1})}^{n_i}$. The subspace model of the each subject is obtained by directly decomposing the \bar{D}_i using Singular Value Decomposition (SVD). The eigenvectors of the p_i largest singular values (i.e., diagonal elements) $\{\xi_i^1, \xi_i^2, \dots, \xi_i^{p_i}\}$ of \bar{D}_i define the projection matrix as given below;

$$U_i = \{u_1, \dots, u_{p_i}\} \in \mathcal{R}^{d \times p_i}; \tag{10}$$

where p_i is the number of principal components. The p_i can be chosen by

$$p_i = \arg \min_k \left\{ \begin{array}{l} \sum_{j=1}^k \xi_i^j \\ \frac{\sum_{j=1}^k \xi_i^j}{m_i} \geq T; k = 1, \dots, n_i; 0 \leq T \leq 1; \\ \sum_{j=1}^k \xi_i^j \end{array} \right\} \tag{11}$$

where T is the predefined threshold. A larger T will lead to larger p_i and corresponding subspace will preserve the most variation of the \bar{D}_i . All the p_i chosen principal components are the basis vectors of the corresponding subspace and all are orthogonal to each other. The learned subspace model of i_{th} subject can be explained by the projection matrix U_i and the mean m_i .

3.2 Face and noise factorization using subspace based sparse representation

Consider the probe image $y \in \mathcal{R}^{d \times 1}$, and it can be explained as given below;

$$y = U\sigma + e, \tag{12}$$

where σ indicates corresponding representation vector and e is error vector. In PCA, the error vector e is assumed to be Gaussian distributed with small variance and therefore σ is estimated as

$$\sigma = U^T y \tag{13}$$

By using σ , the reconstruction error can be approximated by $\|y - U\sigma\|_2^2$. However, this assumption is not valid when the given test face is significantly corrupted due to occlusion. Hence, to estimate the sparse noise term successfully, along with face templates, trivial templates are used in sparse representation based face recognition as given in (14). Trivial template set is a vector that has only one nonzero element, which explicitly codes the pixels being corrupted by noise or occlusion [29].

$$y = [A \quad I] \begin{bmatrix} \sigma \\ e \end{bmatrix} = Bc, \tag{14}$$

where A is the matrix which contains target templates and I is the identity matrix that represents trivial templates. The σ indicates corresponding target template coefficients and e is the noise term which can be viewed as the coefficients of trivial templates. By assuming that the probe image can be sparsely represented by target and trivial templates, the ill-conditioned Eq. (7) can be solved via l_1 regularization [29] as given below;

$$\min_c \frac{1}{2} \|y - Bc\|_2^2 + \lambda \|c\|_1 \tag{15}$$

where $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the l_1 and l_2 norms respectively. The underlying assumption of this method is that the error e can be modeled arbitrary but sparse noise, and therefore it can be used to handle partial occlusion. However, in spite of its huge success in handling partial occlusion in tracking and face recognition applications, it suffers from huge computation drawback. Moreover, in this l_1 regularization formulation, the error term is approximated with the assumption that pixels are randomly corrupted. So, for the realistic contiguous block occlusions caused by sunglasses, scarf etc. this framework cannot factorize face and noise effectively.

To model the noise term effectively for contiguous block occlusions with less computation time, fast l_1 regularization is achieved by exploiting PCA subspace in sparse representation framework. Therefore, the face recognition problem can be cast as finding the nearest subspace to the probe image and handling partial occlusion with trivial templates by

$$y = [U \quad I] \begin{bmatrix} \sigma \\ e \end{bmatrix} = Bc \tag{16}$$

where $y \in \mathcal{R}^{d \times 1}$ is an illumination normalized probe vector, $U \in \mathcal{R}^{d \times p}$ represents orthogonal basis matrix (i.e. face templates), $\sigma \in \mathcal{R}^{p \times 1}$ indicates the coefficients of basis vectors, $e \in \mathcal{R}^{d \times 1}$ give the error term, the λ is a l_1 regularization parameter, and $I \in \mathcal{R}^{d \times d}$ indicates an identity matrix for trivial templates. As e is assumed as arbitrary but sparse noise, (15) can be rewritten as,

$$\min_{\sigma, e} \frac{1}{2} \|y - U\sigma - e\|_2^2 + \lambda \|e\|_1 \tag{17}$$

To solve Eq. (17), let the objective function be, $L(\sigma, e) = \frac{1}{2} \|y - U\sigma - e\|_2^2 + \lambda \|e\|_1$, and the optimization problem is

$$\arg \min_{\sigma, e} L(\sigma, e) \quad s.t. \quad U^T U = I \tag{18}$$

As there is no analytical solution for Eq. (18), σ_{opt} and e_{opt} are solved using iterative technique as given in [30]. By using Lemmas 1 and 2, the optimization problem (18), can be solved efficiently as given in algorithm 1.

Lemma 1 Given e_{opt} , the σ_{opt} can be obtained by $\sigma_{opt} = U^T(y - e_{opt})$.

Proof If e_{opt} is given, Eq. (18) becomes the minimization of $J(\sigma)$, where $J(\sigma) = \frac{1}{2} \|(y - e_{opt}) - U\sigma\|_2^2$. The solution for this least squares problem can be obtained via $\sigma_{opt} = U^T(y - e_{opt})$.

Lemma 2 Given σ_{opt} , e_{opt} can be obtained from $e_{opt} = S_\lambda(y - U\sigma_{opt})$, where $S_\lambda(x)$ is a shrinkage operator, and defined as $S_\lambda(x) = \text{sign}(x) \cdot (|x| - \lambda)$. λ is a regularization parameter and it can be a small constant.

The sign function of $x \in \mathfrak{R}$ is defined as

$$\text{sign}(x) := \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases} \quad (19)$$

Proof If σ_{opt} is given, Eq. (18) becomes the minimization of $G(e)$, where $G(e) = \frac{1}{2} \|e - (y - U\sigma_{opt})\|_2^2 + \lambda \|e\|_1$. For this convex optimization problem, global minimum can be obtained by the shrinkage operator, $e_{opt} = S_\lambda(y - U\sigma_{opt})$.

Algorithm 1 Face representation using linear subspace and l_1 regularization

Input: Probe image vector y , orthogonal basis vectors U , total number of iterations N , termination condition T and regularization parameter λ .

```

1: Initialize  $e_0 = 0$ 
2: for  $z = 1 : N$ 
3: Obtain  $e_z$  via  $e_z = U^T (y - e_{z-1})$ 
4: Obtain  $e_z$  via  $e_z = S_\lambda(y - U e_z)$ 
5: if  $|\|y - e_z\|_2^2 - \|y - e_{z-1}\|_2^2| < T$ 
6: break;
6: end
7: Output:  $\sigma_{opt}$  and  $e_{opt}$ 

```

Using the algorithm 1, the given illumination normalized probe image is factorized into target face and sparse noise image. The effectiveness of the proposed algorithm is demonstrated in Fig. 5 using the example images of AR dataset with different kind of illumination variations. In this figure, first column represents probe image, second column represents illumination normalized image, third column represents extracted target face and fourth column represents sparse error image. It can be seen from Fig 5a that when a training image is given as probe image, the target face is exactly reconstructed and intensity value of all the pixels in the sparse error image are zero. Sparse error image generated from the true identity contains only intra-class difference while the ones from incorrect subjects will not only contain the intra-class difference, but also include a lot of detailed information that is the inter-class difference. In the case of Fig 5b, the extreme illumination variation in the probe image is eliminated by illumination normalization step itself. On the other hand, when the given test image is very different from the gallery image, the extracted noise image would contain all the intra class differences like sunglasses, expression variation and shadows as shown in Fig. 5c and d. It is very evident from Fig 5e that the proposed algorithm has handled the partial occlusion and illumination variation simultaneously in very efficient manner.

After probe image factorization, the obtained representation coefficients σ_{opt} and e_{opt} for each subject, are used to calculate the residual error as given below;

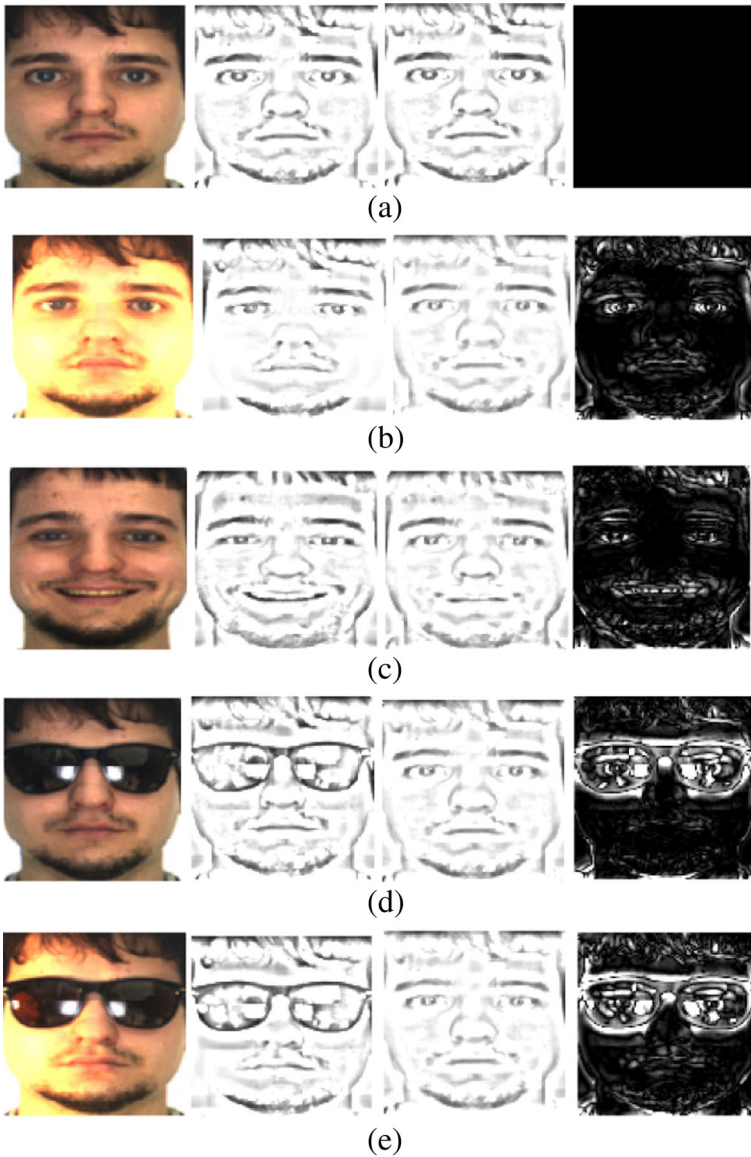


Fig. 5 Illustration of face and noise factorization after illumination normalization using algorithm 1, for (a) training image (b) drastic illumination variation (c) expression variation (d) occlusion, and (e) both illumination variation as well as occlusion of AR dataset

$$r_i(y) = \|y - U_i \sigma_i - e_i\|_2^2 \tag{20}$$

Using r_i , the class of the test face is identified as

$$identity(y) = \arg \min_i r_i(y) \tag{21}$$

Using (20) and (21), the proposed face identification algorithm is given in algorithm 2.

Table 1 Recognition accuracy (%) for the extended Yale –B dataset

Method	Subset			
	2	3	4	5
Eigenfaces + NN	89.91	37.58	5.32	3.08
Fisherfaces + NN	100	97.14	38.59	5.74
LRC [16]	100	100	83.27	33.61
SRC [17]	100	100	67.87	17.51
CRC-RLS[18]	100	100	90.68	45.10
SRPCA-weighted [20]	100	100	95.06	49.38
Proposed method	100	100	82.89	67.36

Algorithm 2 Proposed face identification algorithm

1: **Input:** Subject models $U_i \in \mathcal{R}^{d \times p_i}, i = 1, \dots, N$, illumination normalized probe vector $Y \in \mathcal{R}^{d \times 1}$

2: **for** $i = 1 : N$

3: Calculate e_{opt} and e_{opt} for each U_i using algorithm 1

4: Calculate the reconstruction error $r_i(Y)$ using (20)

5: **end**

6: **Output:** Identity $(y) = \arg \min_i r_i(y)$.

4 Experimental results and discussion

We have validated the proposed method on two popular databases namely Extended Yale B and AR datasets. The experiments are conducted using Matlab 2013 on a computer with Windows XP operating system and Intel core 2 duo processor with 2.53 GHz clock. For each experiment, the recognition rate of the proposed method is compared with that of the other state of the art representative methods.

4.1 Extended Yale B dataset

The Extended Yale B dataset consists of 2414 frontal faces of 38 subjects taken under different lighting conditions. The detailed description about this dataset can be found in [10]. Here, subset 1 is used as training images and subset 2, 3, 4 and 5 are used as test images. The sample images of all the subsets are shown in Fig. 3b. Table 1 shows the recognition rate of various

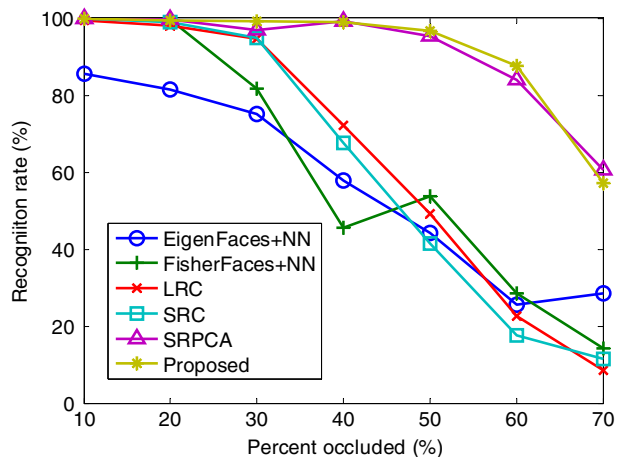
**Fig. 6** Sample images of the 10 to 70 % artificially occluded subset 2 test images

Table 2 Recognition accuracy (%) under different level of contiguous occlusion

Method	10 %	20 %	30 %	40 %	50 %	60 %	70 %
Eigenfaces + NN	85.53	81.58	75.22	57.89	44.3	25.66	28.51
Fisherfaces + NN	100	99.56	81.8	45.61	53.73	28.51	14.25
LRC[16]	99.34	98.03	94.74	72.15	49.12	22.59	8.55
SRC[17]	100	98.9	94.96	67.54	41.45	17.54	11.40
SRPCA-weighted [20]	99.78	99.56	96.92	99.12	95.40	83.99	60.53
Proposed method	99.78	99.34	99.12	98.90	96.71	87.71	57.23

methods for all subsets. As shown in this table, Fisherfaces + NN[13], LRC[16], SRC[17], CRC-RLS[18], SRPCA[20] and proposed method have yielded 100 % recognition accuracy against moderate light variations in subset 2 and subset 3. In case of severe light variations in subset 4, the recognition accuracy of SRC falls to 67.87 % and LRC, CRC-RLS, SRPCA and proposed method achieve good performance with above 80 % recognition accuracy. In subset 4, in most of the images, only half of the face is affected by the illumination and other half is very clear. This fact makes the SRPCA to perform better by relying on one half of the very discriminating information. However, in subset 5, most of the images, entire face is affected by illumination and the SRPCA heavily suffers. The reason behind the lower performance of the proposed method in subset 4 when compared to SRPCA is losing some discriminating face features during illumination normalization. This is the common limitation of most of the reflectance field estimation based illumination preprocessing methods. Interestingly, for extreme light variations in subset 5, the proposed method achieves 67.36 % recognition accuracy and distinguishably outperforms all other representative methods. In this case, the contribution of the DTCWT based reflectance component extraction step in the proposed method is very crucial for its stable performance.

To evaluate the robustness of the proposed method against the contiguous block occlusion, subset 1 is used for training, and subset 2 is used for testing purpose. Like the experimental settings given in [29], test images are simulated with different level of artificial block occlusion from 10 to 70 % by inserting a Baboon image. Moreover, the location of the occlusion is is

Fig. 7 Recognition accuracy plot of artificial occlusion experiment on Extended Yale-B dataset

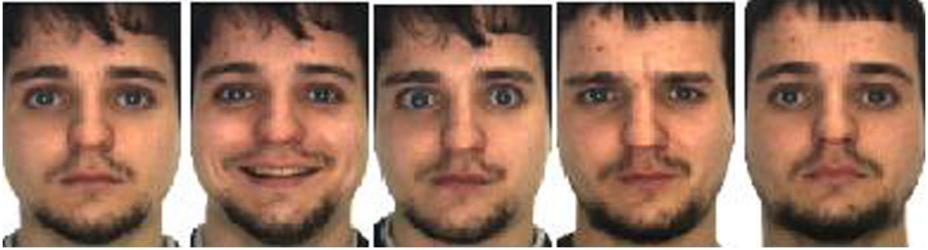


Fig. 8 Training images of a typical subject for experiments using AR dataset

unknown to the algorithms. Figure 6 shows the simulated block occluded test images of subset 2 of an example subject.

Table 2 shows the recognition results of block occluded experiment for all the compared algorithms. Figure 7 shows the recognition performance of block occlusion experiments for all compared algorithms. It can be seen that, the proposed method significantly outperforms all the representative methods in 10 to 60 % occlusion cases. Only in 70 % of block occlusion, the proposed method yields less than 3 % recognition accuracy when compare to the nearest competitor SRPCA-weighted. It is also observed in experiments that the recognition rate varies based on the location of the block occlusion. If the block occlusion is inserted in the lower part of the face, the recognition rate slightly increases, due to the availability of highly discriminant eye features for recognition when compare to the relatively less discriminant mouth and nose features. In our experiment, for an effective evaluation, the block occlusion is inserted in the upper part of the face.

4.2 AR dataset

In this section, 100 subjects (50 male and 50 female) of AR dataset [20] is used to evaluate the proposed method with different kind of experiments. Each subject has 26 images which are captured in two different sessions. The training set for all the experiments conducted in this section is shown in Fig. 8. It can be observed in the training set that only five images are used, in which, images with illumination variations, extreme expressions and partial occlusions are not included. Moreover, our experiments use minimum number of training images compared to other representative methods. First, to individually evaluate the proposed method's ability against illumination variations, the test images are segregated into four different groups like left side light variation (LL), right side light variation (RL), and extreme light variation (EL) as shown in Fig. 9. For each group, 200 test images are used for evaluation. For face recognition, compared to Discrete Wavelet Transform and Gabor wavelet Transform, DTCWT based illumination normalization is shown to be yielding better results [6, 14]. To show the

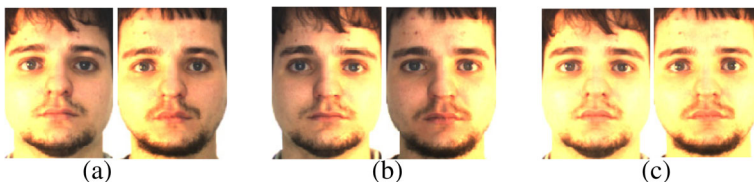


Fig. 9 Testing images of a sample subject in AR dataset with (a) left-side light variation (LL), (b) right side light variation (RL), and (c) extreme light variation (EL)

Table 3 Recognition accuracy of proposed method without and with DCT and DTCWT based illumination normalization techniques on AR dataset with different face disguise

Method	Illumination variation		
	LL	RL	EL
PCA subspace + sparse representation	69.5	80.5	6
DCT	95	97	80.5
DTCWT	98.5	99.5	94.5

robustness of DTCWT against illumination variation, we have compared it with DCT based illumination normalization technique along with the PCA subspace based sparse representation. Table 3 shows the recognition rate of the proposed method against the different test groups and it can be seen that proposed DTCWT based method is obviously superior to the DCT method. Although the proper number of low frequency DCT coefficients has been truncated, its recognition performance is unsatisfactory when testing images are captured under extreme lighting (EL) conditions. This indicates that DCT method is not suitable for dealing with face images with large illumination variations. In the DTCWT based method key facial features are retained using the directional subbands information. Without illumination normalization, PCA subspace learning based sparse representation method performs very poorly against extreme illumination variation. From the results given in Table 1 and 3, it is very clear that the proposed method is robust against illumination variation due to darkness as well as brightness.

To evaluate the robustness against block occlusion in AR dataset, two separate sets of 200 test images are selected. The first set contains the images with nearly 20 % of sunglasses occlusion while the second set contains the images with nearly 40 % scarf occlusion as shown in Fig. 10. Table 4 depicts the recognition rate for this experiment along with all other compared methods. For the sunglasses case, the proposed method achieves 93.3 %, which is considered as very competitive when compare to other results obtained by the representative methods. Moreover, due to the modular approach, LRC could achieve higher recognition rate of 96 %. If the occlusion spreads for more partitions, LRC method would fail as shown in scarves case. In the case of scarf occlusion the proposed method outperforms SRPCA-weighted method by a margin of 40.5 %. In SRPCA [30], classification is relies only on the sparse error images and in case of test images with large contiguous occlusion, discriminative information of the error images used for classification will decrease. From these observations, it is clear that integration of DTCWT based illumination normalization step does not affect the natural ability of the sparse representation framework against the partial occlusion. It also proves that subspace learning based sparse representation can effectively factorize the target



Fig. 10 AR dataset test samples of (a) sunglasses and (b) scarves

Table 4 Recognition accuracy on AR dataset with contiguous occlusion

Method	Sunglasses	Scarves
Eigenfaces + NN	43.5	7.5
Fisherfaces + NN	38.5	67.5
LRC [16]	96	26
SRC [17]	87	59.5
SRPCA-weighted [20]	94	40
Proposed method	93.3	80.5

face and noise than conventional sparse representation method due to the class specific dictionaries and PCA subspace learning based sparse representation.

Finally, to have more comprehensive observation of the ability of these methods against coupling multiple variations simultaneously, a more challenging experiment is conducted. The performance of proposed method is compared with those of LRC, SRC, SRPCA, and PCA subspace based sparse representation based methods. For this experiment another two separate sets of 600 test images are selected. The first set contains all the sunglasses occluded images in which most of them also have illumination variations. In the same way, the second set contains all the scarves occluded images in which most of them also have illumination variations. The example test images of a sample subject for this experiment are shown in Fig. 11. It can be observed that the test images of real disguise are affected with illumination variation as well as large contiguous occlusion.

The results of this experiment are shown in Table 5. It is clear that without illumination normalization, the proposed PCA based sparse representation method outperforms other representative methods. However, illumination variations in the test images affect its recognition accuracy. After DTCWT based illumination normalization, proposed method further improves the recognition accuracy. For the sunglasses case, the proposed method outperforms all other representative methods with significant margin. In the case of scarves, the proposed method achieves 10 % better recognition rate when compare to its nearest competitor SRPCA. Overall, the proposed method shows very robust performance against the real disguise when compare to other representative methods.

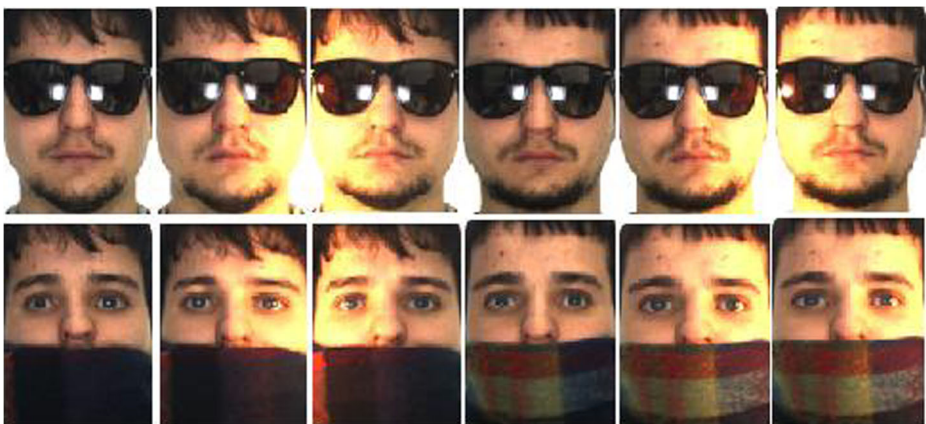
**Fig. 11** Example AR dataset test images with illumination variation as well as occlusion

Table 5 Recognition accuracy of an experiment with illumination variation as well as occlusion

Method	Sunglasses	Scarves
LRC	51	11.2
SRC	69.8	40.8
SRPCA-weighted [20]	73.3	58
PCA subspace + Sparse representation	82	62.3
Proposed method	87	68

The computation time requirement of the proposed and other representative algorithms for face identification is presented in Table 6 using 2.53 GHz Intel Core 2 Duo processor. The computational time required for the training is usually an offline process. The presented computation time is measured for 100 subjects AR dataset with image size of 48×48 and five training images per subject. To solve for l_1 regularization in SRC method, SParse Modeling Software (SPAMS) package [19] and to solve SRPCA, Augmented Lagrange Multiplier (ALM) algorithm software is used [28]. From the table 6, it is very clear that even with DTCWT based preprocessing step, the proposed method takes very minimum computation time compared to other competitive methods. This indicates that the proposed method might be very suitable for time critical face identification applications. However, one drawback of this work is, in all the experiments face misalignment due to pose variation in test images is not considered. In order to pre-align and crop the faces, facial landmarks obtained using Active Shape Models approach can be used [17].

5 Conclusions

An efficient face identification method for simultaneous handling of illumination variation and partial occlusion is presented by using DTCWT and PCA subspace based sparse representation. The good properties of DTCWT like approximate shift invariance, directional selectivity, and fast computation are effectively harnessed to accurately extract the illumination invariant geometrical structure of the face image. Then, partial occlusion is handled via factorizing the probe image into target face and sparse noises by exploiting PCA subspace representation in l_1 regularization with less computation. Experimental results show that structurally stable illumination normalized images do not affect robustness of the sparse representation framework against partial occlusion. Moreover, the proposed method achieves better recognition for illumination variations due to darkness as well as brightness. To test the ability of the proposed method against the real disguise, a challenging experiment is conducted by using the test images affected with partial occlusion as well as illumination variation. Results show that the proposed method significantly outperforms the other representative methods.

Table 6 Average computation time for different methods for an incoming probe image of size 48×48 pixels

Method	LRC	SRC	SRPCA	PCA subspace + Sparse representation	Proposed method
seconds/probe image	0.041	24	18.4	0.11	0.46

References

1. Adini Y, Moses Y, Ullman S (1997) Face recognition: the problem of compensating for changes in illumination variation. *IEEE Trans Pattern Recognit Mach Intell* 19(7):721–732
2. Basri R, Jacobs DW (2003) Lambertian reflectance and linear subspaces. *IEEE Trans Pattern Anal Mach Intell* 25(2):218–233
3. Belhumeur PN, Hespanha JP, Kriegman DJ (1997) Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. *IEEE Trans Pattern Recogn Mach Intell* 19(7):711–720
4. Belhumeur P, Kriegman D (1998) What is the set of images of an object under all possible illumination conditions? *Int J Comput Vis* 28(3):245–260
5. Cai S, Li K. Matlab implementation of wavelet transforms. Polytechnic Univ., Brooklyn, New York. [online] Available: <http://ceweb.poly.edu/iselesni/WaveletSoftware/>
6. Chao-Chun L, Dap-Qing D (2009) Face recognition using dual-tree complex wavelet features. *IEEE Trans Image Process* 18(11):2593–2599
7. Chellappa R, Jie N, Vishal MP (2012) Remote identification of faces: problems, prospects, and progress. *Pattern Recogn Lett* 33(14):1849–1859
8. Chen W, Meng JE, Shiqian W (2006) Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. *IEEE Trans Syst Man Cybern* 36(2):458–466
9. Dong W, Lu H, Yang MH (2013) Online object tracking with sparse prototypes. *IEEE Trans Image Process* 22(1):314–325
10. Georghiades A, Belhumeur P, 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
11. Goh YZ, Teoh A, BJ, Goh MKO (2008) Wavelet based illumination invariant preprocessing in face recognition. *Proceedings of IEEE International Congress on Image and Signal Processing*, 3, pp. 421–425
12. Gonzales RC, Woods RE (1992) *Digital image processing*, 2nd edn. Prentice Hall, Upper Saddle River
13. Gonzalez RC, Woods RE (2009) *Digital Image Processing*. Hall, Prentice
14. Haifeng H (2011) Multiscale illumination normalization for face recognition using dual-tree complex wavelet transform in logarithm domain. *Comput Vis Image Underst* 115(10):1384–1394
15. Horn BKP (1986) *Robot vision*. MIT Press
16. Joni-Kristian K, Ville K, Heikki K (2006) Invariance properties of Gabor filter based features- overview and applications. *IEEE Trans Image Process* 15(5):1088–1099
17. Kathryn B, Brendan FK, Anil KJ (2012) Component based representation in automated face recognition. *IEEE Trans Inform Forensics and Secur* 8(1):239–253
18. Kingsbury NG (2000) A dual-tree complex wavelet transform with improved orthogonality and symmetry. *Proceedings of IEEE international conference on image processing, Canada*, vol. 2, pp. 375–378
19. Mairal J, Bach F, Ponce J, Sapiro G (2010) Online learning for matrix factorization and sparse coding. *J Mach Learn Res* 11:19–60
20. Martinez A, Benavente R (1998) The AR face database, CVC technical report
21. Mei X, Ling H (2009) Robust object tracking using l_1 minimization. *Proceedings of international conference on computer vision, Kyoto*, pp. 1436–1443
22. Nabatchian A, Abdel-Raheem E, Ahmadi M (2011) Illumination invariant feature extraction and mutual information based local matching for face recognition under illumination variation and occlusion. *Pattern Recogn* 44(10-11):2576–2587
23. Naseem I, Togneri R, Bennamou M (2010) Linear regression for face recognition. *IEEE Trans Pattern Recogn Mach Intell* 32(11):2106–2112
24. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Recogn Mach Intell* 24(7):971–987
25. Selesnick IW, Baraniuk RG, Kingsbury NC (2005) The dual-tree complex wavelet transform. *IEEE Signal Process Mag* 22(6):123–151
26. Taiping Z, Yuan YT, Bin F, Zhaowei S, Xiaoyu L (2009) Face recognition under varying illumination using gradientfaces. *IEEE Trans Image Process* 18(11):2599–2606
27. Wanger A, Wright J, Ganesh A, Zhou Z, Mobahi H, Ma Y (2012) Toward a practical face recognition system: robust alignment and illumination by sparse representation. *IEEE Trans Pattern Recogn Mach Intell* 34(2):372–386
28. Wright J, Ganesh A, Shankar R, Yigang, P, Ma Y (2009) Robust principal component analysis: exact recovery of corrupted low-rank matrices by convex optimization, *advances in neural information systems*, pp. 2080–2088
29. Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y (2009) Robust face recognition via sparse representation. *IEEE Trans Pattern Anal Mach Intell* 31(2):210–227

30. Xiao L, Bin F, Linghui L, Weibin Y, Jiye Q (2014) Extracting sparse error of robust PCA for face recognition in the presence of varying illumination and occlusion. *Pattern Recogn* 47(2):495–508
31. Zhang D, Meng Y, Xiangchu F (2011) Sparse representation or collaborative representation: Which helps face recognition? *Proceedings of IEEE international conference on computer vision, Barcelona*, pp. 471–478



K. Selvakumar received his PhD degree in electrical engineering from Anna University, Chennai, India in 2015. His research interests include Signal Processing, Computer Vision and Embedded Vision Systems.



Jovitha Jerome received her BE (EEE) degree from College of Engineering, Guindy, Chennai. She received her ME and PhD degrees in Electrical Engineering from Government College of Technology, Coimbatore, India and Asian Institute of Technology, Bangkok, Thailand respectively. She has published numerous papers in international journals and conferences. Currently she is working as Professor at PSG College of Technology, Coimbatore, India. Her research interests include Signal Processing and Machine Learning.



Kumar Rajamani received his MSc (Mathematics) and MTech (CSc) from Sri Sathya Sai Institute of Higher Learning. He received his PhD degree in Biomedical Engineering from University of Bern, Switzerland. He has three patents to his credit and has published extensively in international journals and conferences. He is currently working as architect at Robert Bosch Engineering and Business Solutions Limited, Bangalore, India. His areas of interests include image processing, medical image analysis and pattern recognition.