Optimal fusion aided face recognition from visible and thermal face images

Madheswari Kanmani[1] · Venkateswaran Narasimhan[2]

© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

This paper proposes a novel Eigen face recognition that is aided by fusion of visible and thermal face images to improve the face recognition accuracy. We adopt three different fusion schemes where in the face information is fused by the optimal weights obtained by different optimization algorithms. The first two fusion approaches operate in the dual tree discrete wavelet transform (DT-DWT), while the third one operates in the Curvelet transform (CT) domain. We employ particle swarm optimization (PSO), self-tuning particle swarm optimization (STPSO) and brain storm optimization algorithm (BSO) to find optimal fusion coefficients. The proposed fusion aided face recognition approaches are evaluated through extensive experiments using OCTVBS benchmark face database and the Eigen face detection methodology. Simulation results show that proposed face recognition techniques have significant performance improvement in recognition accuracy suggesting fusion aided face recognition approach that deserves further study and consideration whenever high recognition accuracy is desired.

Keywords Visible Image · Thermal Image · Image fusion · Multi-resolution · PSO STPSO · Curvelet transform · DT-DWT · Face recognition

1 Introduction

Face recognition is one of the most successful and widely applied biometric traits for security purposes [50]. Face recognition based on the visible spectrum has shown good performance when the face images are captured in a controlled environment [20, 42]. But, the performance

✉ Madheswari Kanmani
   madheswarik@ssn.edu.in

   Venkateswaran Narasimhan
   venkateswarann@ssn.edu.in

1   Computer Science and Engineering, SSN College of Engineering, Chennai, Tamil Nadu, India

2   Electronics and Communication Engineering, SSN College of Engineering, Chennai, Tamil Nadu, India

of such face recognition systems degrades significantly in uncontrolled illumination conditions [21, 32]. The accuracy of face recognition degrades quickly when the lighting is dim or when the face is not uniformly illuminated [27]. Thus, face recognition with visible face images in an uncontrolled lighting environment is challenging task. Face recognition using thermal infrared imaging (IR) sensors has become an area of growing interest [18]. The use of thermal images in face recognition systems has been shown to improve the recognition accuracy and be robust in uncontrolled illumination conditions [48]. However, in such systems thermal face image is sensitive to the ambient temperature changes that lead to misidentification [26].

Infrared images can distinguish targets from their backgrounds based on the radiation difference, which works well in all-weather and all-day/night conditions. By contrast, visible images can provide texture details with high spatial resolution and definition in a manner consistent with the human visual system [36, 37]. Therefore, it is desirable to fuse these two types of images, which can combine the advantages of thermal radiation information in infrared images and detailed texture information in visible images.

In our proposed work, we have developed three multi-resolution based fusion schemes to enhance the face recognition performance. In the first proposed scheme, the source images are decomposed into high and low frequency coefficients through DT-DWT. The reason to choose multi resolution approach is that the high frequencies are relatively independent of global changes in the illumination, while the low frequencies take into account the spatial relationships among the pixels and less sensitive to noise and small changes (e.g. Facial expression). Fusion in the multi resolution domain involves combining the coefficients of the visible and thermal images. The fused image is obtained by applying the inverse transform on the combined coefficients.

The rest of the paper is organized as follows. Section 2 describes about related work in face recognition. The proposed image fusion schemes are described in section 3. Experimental results and discussions are given in section 4. Finally, conclusions are given in section 5.

2 Related work

Face recognition is one of the most efficient and broadly used biometric modality in today's scenario [9]. Face recognition methods can be classified into two main categories: holistic and texture-based methods [58, 59, 64]. In the holistic approach, all the pixels in the entire face image are taken as a single signal, and processed to extract the relevant features for classification [12].

Holistic or appearance-based approaches to face recognition involve encoding the entire facial image in a high-dimensional space [29, 64]. It is assumed that all faces are constrained to particular positions, orientations, and scales. The most widely used holistic approaches are the principal component analysis (PCA) [6], linear discriminant analysis (LDA) [55] and a blind source separation technique, called independent component analysis (ICA) [4].

Principal Component Analysis was used for face recognition by Turk and Pentland [56]. Later, Principal Component Analysis was compared with Linear Discriminant Analysis in [39]. Gabor based Kernel PCA with fractional power polynomial model were used by Liu in [33]. Yang et.al proposed two dimensional PCA for face recognition [61]. In 2005, Locally Linear Discriminant Analysis (LLDA) was used for face recognition [25]. Texture-based approaches rely on the detection of individual facial characteristics and their geometric relationships prior to performing face recognition [40, 51, 64]. Apart from these approaches, face recognition can also be performed by using different local regions of face images [5, 11].

Jiayi Ma et al. [15] proposed a novel fusion algorithm, named Gradient Transfer Fusion (GTF), based on gradient transfer and total variation (TV) minimization. The authors formulated the fusion problem as an ι1-TV minimization problem, where the data fidelity term keeps the main intensity distribution in the infrared image, and the regularization term preserves the gradient variation in the visible image. It can simultaneously keep the thermal radiation information in the infrared image and preserve appearance information in the visible image. The fusion results look like high-resolution infrared images with clear highlighted targets and hence, it will be beneficial for fusion-based target detection and recognition systems.

Recently, the Sparse Representation based Classification (SRC) method, has received a lot of attention for face recognition [15]. In SRC, a sparse coefficient vector was introduced in order to represent the test image by a small number of training images. Then the SRC model was formulated by jointly minimizing the reconstruction error and the 1-norm on the sparse coefficient vector. The main advantages of SRC have been pointed out in [15]: i) it is simple to use without carefully crafted feature extraction, and ii) it is robust to occlusion and corruption.

Jiayi Ma et al. [15] address the problem of face recognition when there is only few, or even only a single, labelled examples of the face that we wish to recognize. Moreover, these examples are typically corrupted by nuisance variables, both linear (i.e., additive nuisance variables, such as bad lighting and wearing of glasses) and non-linear (i.e., non-additive pixel-wise nuisance variables, such as expression changes). The small number of labelled examples means that it is hard to remove these nuisance variables between the training and testing faces to obtain good recognition performance. To address the problem, the authors proposed a method called semi-supervised sparse representation-based classification.

Some other methods proposed for face recognition are: Bayesian inference [41], Elastic Bunch Graph Matching (EBGM) [30], Support Vector Machines (SVM) [43], Linear Discriminant Analysis (LDA) [17], Kernel Methods [49], Neural Networks [31], Local Feature Analysis (LFA) [45]. Parkhi et al. [44] used deep convolutional neural networks for face recognition and this approach achieved results comparable to the state of the art.

Face recognition system also uses transform domain techniques to achieve challenges like illumination compensation and normalisation [10]. Discrete Cosine transform (DCT) minimizes illumination variations and is robust and can be implemented in real time [7]. High speed face recognition can be implemented combining DCT and Fisher Linear Discriminant (FLD) and Radial Basis Function (RBF) neural networks. The proposed system achieves excellent performance with training and high-speed recognition, high recognition rates and illumination challenges [14]. 3D Discrete Wavelet Transform (DWT) is employed for feature extraction of hyper-spectral facial analysis and achieved accuracy proves that 3D DWT method is superior to spatio-spectral classification [16]. The authors [2] used Multi-Resolution transform such as, Gabor Wavelet Transform (GWT) is used for recognizing facial images, collected from benchmark Yale Database. Alaa Eleyan et al. [13] combined wavelet with PCA to improve face recognition accuracy. Hafiz Imtiaz et al. [24] proposed face recognition approach based on two-dimensional discrete wavelet transform (2D-DWT), which efficiently exploits the local spatial variations in a face image. Other formal algorithms for feature extraction with multivariate statistical techniques in complex domain are fused with deep learning and results show advancement as compared to state of art methods in computer vision and pattern recognition [54].

But the existing above-mentioned face recognition approaches either works on visible image or thermal image. It is well-known that, face recognition based on the visible image has shown good performance when the face images are captured in a controlled environment.

But, the performance of such face recognition systems degrades significantly in uncontrolled illumination conditions. The accuracy of face recognition degrades quickly when the lighting is dim or when the face is not uniformly illuminated. The use of thermal images in face recognition systems has been shown to improve the recognition accuracy and be robust in uncontrolled illumination conditions. However, in such systems thermal face image is sensitive to the ambient temperature changes that lead to misidentification. Hence, considering the complementary information contained in visible and thermal face images, the fusion can be used to improve the accuracy of face recognition task [1, 8, 60].

Image fusion literature related to visible and thermal images show that multi-resolution approach (MRA) fusion is efficient and it is possible to integrate information at different level of decomposition [47]. Multiresoultion methods provide powerful signal analysis, which are widely used in feature extraction. Wavelet transform techniques achieve optimal decomposition without affecting much the image quality. Because of the property of shift-invariance, it is well known that wavelet-based approach is one of the most robust feature extraction schemes, even under variable illumination. Some of the most popular multi-resolution approaches include the Laplacian pyramid (LAP) [52], Gradian pyramid (GRAD), ratio of Laplacian pyramid (ROLP) [53], Contourlet transform, non sub-sampled Contourlet transform (NSCT) [28], discrete wavelet transform (DWT), shift invariant discrete wavelet transform (SIDWT) [62], dual tree discrete wavelet transform (DT-DWT) [38] and Curvelet transform (CT) .

In this paper, we propose, three optimization based fusion methods that aids face recognition problem. The ultimate goal of the paper is to enhance the face recognition performance by various optimizations based fusion methods. In the first proposed scheme, the source images are decomposed into high and low frequency coefficients through DT-DWT. Particle swarm Optimization (PSO) is used to find the optimal weights to combine face information from thermal and visible images. Then the fused images are recognized using Eigen face approach for the purpose of demonstrating the benefits of fusion.

In the second proposed scheme, the source images are decomposed into high and low frequency coefficients through DT-DWT. Self Tunning Particle swarm Optimization (ST-PSO) is used to find the optimal weights to combine face information from thermal and visible images. Then the fused images are recognized using Eigen face approach for the purpose of demonstrating the benefits of fusion.

In the third scheme, curvelet transform is applied for image decomposition that preserves the edges along the curves. Again, further to improve the searching of optimal weight coefficients a Brain storm optimization algorithm is used for optimization. Then the fused images are recognized using Eigen face approach. In our work, we have used OTCBVS [22] face database for carrying experiments using proposed fusion methods.

3 Preliminaries

3.1 Eigen face detection methodology

In Eigen face detection methodology [57] PCA is applied to the task of face recognition. The PCA converts the pixels of face image into a number of Eigen feature vectors. These Eigen Feature vectors used to measure the similarity between the two face images. The mean vector of the training face images is calculated. Let the training set of face images can be I_1, I_2, I_M.

Every training image (I_i) is represented as a vector Γ_i and mean face vector (ψ) is computed as follows:

$$\psi = \frac{1}{S} \sum_{i=1}^{S} \Gamma_i \tag{1}$$

Where S is the to be number of faces in the training set. Subtracting the mean from the training images gives mean shifted images vector (ϕ_i) as

$$\phi_i = \Gamma_i - \psi_i \tag{2}$$

The Eigen vectors and Eigen values of the mean shifted images are computed from covariance matrix (c).

$$c = \frac{1}{S} \sum_{i=1}^{S} \phi_i \phi_i^T = AA^T \tag{3}$$

Where $A = [\phi_1, \phi_2, \phi_3, ..., \phi_M]$. The Eigen face can be defined by computing the Eigen face vector μ_i of c.

$$\mu_i = \sum_{j=1}^{S} V_{ij} \phi_j, \tag{4}$$

Where $j = 1, 2, 3, ..., S$. V_{ij} is the Eigen vectors of $A^T A$. The Eigen vectors are ordered in descending order by its corresponding Eigen values. The Eigen vectors having the largest Eigen values could be retained and projected into the Eigen face.

The last step in this method is to classify a given face image. To perform face recognition the similarity score is calculated between the test image and each of the training images. The given new image F, transforms into its Eigen face components (projected into face space), subtract the mean ($\phi = \Gamma - \psi$) and compute the projection

$$\phi = \sum_{i=1}^{S_k} w_i \mu_i \tag{5}$$

Where $w_i = \mu_i^T \Gamma$ are the coefficients of the projection and (w_i) referred as Eigen features. The matched image is the one with the highest similarity score.

3.2 Particle swarm optimization algorithm (PSO)

PSO is a population-based optimization technique that finds an optimal solution to the problem in a feasible solution space. PSO is initialized with a population of random solutions called particles that distributed over the search space. The moment of each particle is updated based on the two factors: (i) personal best ($P_i(t)$) position that the ith particle has found (ii) global best position ($P_g(t)$) found by the whole swarm. Each particle updates its velocity ($V_i(t)$) and position ($X_i(t)$) as follows

$$V_i(t+1) = \omega V_i(t) + c_1 r_1 (P_i(t) - X_i(t)) + c_2 r_2 (P_g(t) - X_i(t)) \tag{6}$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \tag{7}$$

Where ω is the inertia weight that controls the convergence of PSO. The parameters r_1 and r_2 controls the diversity of the population. The acceleration coefficients c_1 and c_2 take fixed value (i.e., $c_1=c_2=2$) that stabilizes the influence of the $P_i(t)$ and $P_g(t)$. The current global best value replaces the previous iteration $P_g(t)$ value if it has a better fitness value. The same process is repeated until the maximum number of iterations is reached.

3.3 Self-tuning particle swarm optimization (ST-PSO)

The PSO can easily locate nearly optimal solutions with fast convergence speed, but fails to adjust the acceleration coefficients leading to premature convergence. The PSO algorithm tends to give poor performance when the acceleration coefficients (c_1 and c_2) are fixed. In the proposed work [38], dynamically varying acceleration coefficients have been introduced to improve search ability and premature convergence. The modified acceleration coefficients are represented as follows

$$c_1 = \left(c_{1fv} - c_{1iv}\right) \times \left(\frac{P_i(t)}{\sum\limits_{i=1}^{t} P_i(t)/t}\right) + c_{1iv} \tag{8}$$

$$c_2 = \left(c_{2fv} - c_{2iv}\right) \times \left(\frac{P_g(t)}{\sum\limits_{t=1}^{max-iter} P_g(t)/t}\right) + c_{2iv} \tag{9}$$

where $c_{1iv}, c_{1fv}, c_{2fv}$ and c_{2iv} are constants fall in the range of [2.5,0.5] and [0.5, 2.5] respectively. The objective of ST-PSO is to avoid premature convergence.

3.4 Brain storm optimization algorithm (BSO)

Shi proposed BSO [19] in 2011 by modelling the human brain storm process by creatively mapping it to the optimization field. In BSO, each position within the solution space is called an idea that is randomly initialized in the solution space. During each generation, the ideas are grouped into cluster using k-means clustering and the idea with best fitness is selected as cluster center. To avoid from premature convergence and improve the search efficiency, a randomly selected cluster center could be replaced by a newly generated individual with a probability of p_r. To generate new idea, one cluster or two clusters are randomly selected with pre-determined probability (p_1, p_2). If a new idea is generated using one existing idea, it can be produced by Eq. (6).

$$X_{new}^t = X_{old}^t + \xi N(\mu, \sigma) \tag{10}$$

Where X_{new}^t and X_{old}^t are the tth dimension of X_{new} and X_{old}, respectively. $N(\mu, \sigma)$ represents the Gaussian distribution with mean μ and variance σ and ξ is the regulatory factor which controls the convergence speed, is defined as

$$\xi = logsig\left(\frac{N_{\max}/2 - N_m}{K}\right) \times rand \qquad (11)$$

where $N_{\max}$ is the maximum number of iterations, N_m is the current number of iteration, K is the scale factor ($K = 20$). If the new idea is generated by two existing ideas, it can be defined as

$$X^t_{new} = X^t_{old} + \xi N(\mu, \sigma) \qquad (12)$$

$$X^t_{old} = \left(w_1 \times X^t_{old1}\right) + \left((1 - w_1) \times X^t_{old2}\right) \qquad (13)$$

where w_1 is the weight of selected idea. After the new idea is generated, the quality of the new idea is evaluated by fitness function, if the new idea is better than old one, it replaced by new idea. The above process is repeated for all ideas and until the maximum number of iterations is reached. Then output result as the optimal solution to the problem.

4 Proposed optimized image fusion framework for face recognition

In our proposed work, we have developed three multi-resolution based fusion schemes to enhance the face recognition performance. Visible (V) and thermal (T) face images are taken using different cameras so that the images have different fields of view and spatial resolutions. Thermal images are registered using affine transform by considering the visible image as base image [34, 35]. After registration, the source images are decomposed using transform and fused using optimal weights obtained by optimization algorithm. Fused images are trained and recognized using Eigen face detection methodology.

In the first proposed scheme, the source images are decomposed into high and low frequency coefficients through DT-DWT. The reason to choose multi resolution approach is that the high frequencies are relatively independent of global changes in the illumination, while the low frequencies take into account the spatial relationships among the pixels and less sensitive to noise and small changes (e.g. Facial expression). Fusion in the multi resolution domain involves combining the coefficients of the visible and thermal images. The fused image is obtained by applying the inverse transform on the combined coefficients.

The key question in implementing this idea is how optimally combining the coefficients from each spectrum. Using unweighted average is not appropriate since it assumes that the two spectra are equally important (weight = 0.5 for both the images). George Bebis et al. [3] employed genetic algorithm (GA) to find an optimal fusion strategy to combine information from thermal and visible images. Gabriel et al. [19] also used genetic algorithm to choose optimal face areas where one spectrum is more representative than other. Genetic algorithm requires additional operations such as cross over and mutation which is a time-consuming process. Moreover, genetic algorithms struck with local optima.

In the first scheme, PSO is used to find the optimal weights to combine face information from thermal and visible images. The reason for the choice of PSO is that, it has less time complexity compared with GA free from complex crossover and mutation operations. The advantages of PSO is that it can locate nearly optimal solutions with a fast convergence, but usually fails to adjust acceleration coefficients which often leads to premature convergence [38].

In the second scheme, we have modified the PSO by introducing dynamically varying acceleration coefficients to improve the global search ability and to avoid premature convergence. The modified version of PSO is named as self-tuning particle swarm optimization (ST-PSO). ST-PSO is employed to find optimal weights and used to combine information from thermal and visible face images.

Edges in the face images need to be properly synthesized in the fused image in order to improve face recognition accuracy. But DT-DWT fails to preserve edges along the curves. Therefore, to improve the presence of edges further in the third scheme, Curvelet transform is applied for image decomposition that preserves the edges along the curves. Again, further to improve the searching of optimal weight coefficients a Brain storm optimization algorithm is used for optimization. The Fig. 1 illustrates the steps involved in the proposed work.

The fused images are recognized using Eigen face approach for the purpose of demonstrating the benefits of fusion. The projected Eigen face space is constructed from training face images. Similarity score between the test image and each of training images is calculated. The matched image is the one with the highest similarity score. Recognition performance is computed by finding the percentage of images in the test set, for which the top match is an image of the same person from the training images. Experimental results show that the proposed image fusion image fusion scheme is a viable approach for enhancing face recognition performance.

4.1 Algorithm 1: Image fusion through DT-DWT and PSO for face recognition

Step 1:　The visible and thermal images are resampled to common size (m x n). The reason is that DT-DWT operates only on images size of power two. Hence, we generated 128×128 size images using bi-cubic interpolation.
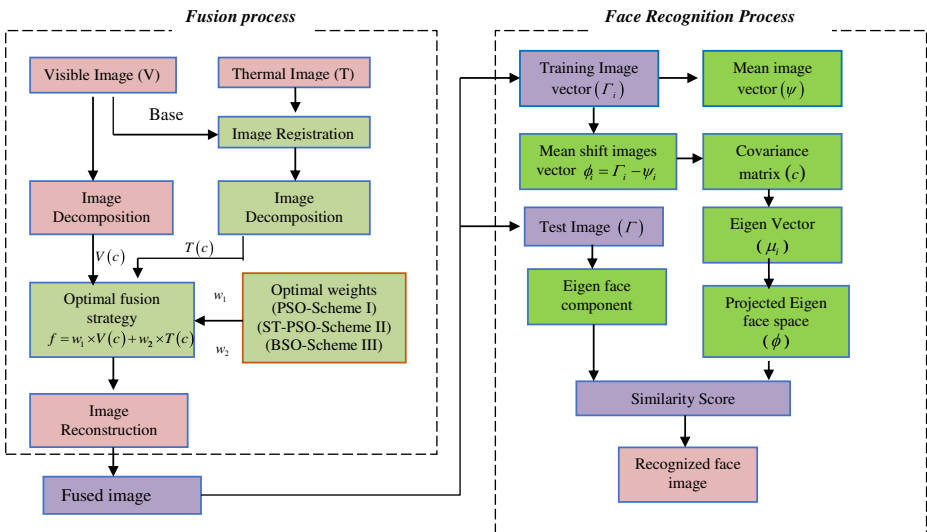

Fig. 1 Block diagram of the proposed image fusion scheme-based face recognition

Step 2:   The images T and V are registered using affine transformation in order to spatially align the images.

Step 3:   The images T and V are decomposed into low and high frequency components using DT-DWT.

$$f = Tf(T, V) \qquad (14)$$

Step 4:   The coefficients of the T and V face images are combined using the fusion rule (FR)

$$FR = w_1 \times V(c) + w_2 \times T(c) \qquad (15)$$

Here w_1 and w_2 determine the percentage of each image coefficient in the fused image. Here, PSO is used to obtain optimal weights that maximize the entropy and minimize the root mean square error. The procedure for obtaining optimal weights is given below.

In our work, image fusion is formulated as an optimization problem. The set of solution is defined as a set of N particles (weights)

$$w = \left\{ \begin{array}{c} w_{11}, w_{12}, ..., w_{1N} \\ w_{21}, w_{22}, ..., w_{2N} \end{array} \right\}$$

where, $w = (w_1, w_2)^T \in A$, which maximizes the entropy (E) of the fused image.

$$f_1(x) = E = - \sum_{j=0}^{255} p(j) \times \log_2(p(j)) \qquad (16)$$

where $p(j)$ is the probability of the occurrence of jth intensity of the fused image. The solution set (w) also minimize the objective function (RMSE)

$$f_2(x) = RMSE$$

$$= \frac{1}{2} \left[\sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} [F(i,j) - V(i,j)]^2} + \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} [F(i,j) - T(i,j)]^2} \right] \qquad (17)$$

The overall objective function is defined as follows

$$f(x) = \alpha_1 \times f_1(x) + \alpha_2 \times f_2(x) \qquad (18)$$

Here α_1 and α_2 are constants whose value indicate the relative significance of the objective function. In this work, we choose $\alpha_1 = \alpha_2 = 0.5$. The solution set which gives maximum entropy value and minimal RMSE value can be taken as global best. After the maximum number of iterations is reached the global best value is used to get the final fused image. The Fig. 1 illustrates the steps involved in the proposed work.

The reason to choose entropy and RMSE as objective function is to maximize the information content (entropy) thus indicates the quality of the fused image.

Step 5:   The fused image (F) is obtained by taking inverse DT-DWT to the fused coefficients.

$$F = T^{-1}(T(V), T(B)) \tag{19}$$

Where T^{-1} is the inverse DT-DWT.

Step 6:   The fused face images are recognized using Eigen face detection methodology. As shown in Fig. 1 the mean vector (ψ) of the training face images and mean shifted images vector (ϕ_i) Calculated using Eq. (9) and (10).

Step 7:   Calculate the Eigen vectors and Eigen values of the mean shifted images from covariance matrix (c). The Eigen vectors are ordered in descending order by their corresponding Eigen values. The Eigen vectors having the largest Eigen values could be retained and projected into the Eigen face

Step 8:   The last step is to classify a face image. To perform face recognition the similarity score is calculated between the test image and each of the training images. The matched image is the one with the highest similarity score.

4.2 Algorithm 2: Image fusion through DT-DWT and ST-PSO for face recognition

Step 9:   Thermal and visible images are resample using bi-cubic interpolation to the size of 128×128.

Step 10:   Images are registered using affine transform to spatially align the images.

Step 11:   The registered images are decomposed using DT-DWT and fused using optimal weights obtained from ST-PSO using the step 4 mentioned in section 3.1. The optimal weights are obtained by substituting solution set values in Eq. (10) and the solution which gives maximum fitness value at the end of maximum iterations will be considered as optimal value.

Step 12:   As in Inverse DT-DWT is applied on the fused coefficients to get the final fused image.

Step 13:   To perform face recognition on the fused images the steps 6 to 9 mentioned in section 4.1 has followed.

4.3 Algorithm 3: Image fusion through Curvelet and BSO for face recognition

Step 1:   As in Scheme I and II, the thermal and visible images are registered using bi-cubic interpolation to the size of 128×128. Curvelet transform can operate in any size of image. Here the reason for resampling is that to use consistent image size for all the proposed fusion approaches.

Step 2:   After image resampling, image registration is performed. The DT-DWT has better reconstruction and shift invariant property. But edges are not effectively handled in DT-DWT. Curvelet transform effectively captures the edges along the curves can improve the face recognition accuracy.

Step 3: Optimal weights are obtained using BSO formulation by substituting idea set in Eq. (15). The idea set which gives the maximum fitness value at the end of the maximum iteration is taken as global best value.

Step 4: The fused image (F) is obtained by taking inverse curvelet transform to the fused coefficients.

$$F = T^{-1}(T(V), T(B)) \qquad (20)$$

Where T^{-1} is the inverse Curvelet transform.

Step 5: To perform face recognition on the fused images the steps 6 to 9 mentioned in section 4.1 has followed.

5 Experimental results and discussion

5.1 OTCBVS-dataset

In our experiments, we used the face database called OTCBVS which is a standard bench mark of thermal and visible images for face recognition techniques [23]. OTCBVS consists of 700 visible and 700 thermal images of 16 persons. The images were taken at different times that contain variability in illumination, facial expression (open /closed eyes, smiling/ non smiling), various poses like upright, frontal position and facial details (glasses / without glasses. Out of 700 thermal images only 400 images of 10 persons are taken out of which 200 are thermal images and 200 are visible images. 20% images are used as training set and 80% images are used as testing set.

5.2 Parameter settings

The proposed image fusion based face recognition techniques are compared with several image fusion based faced recognition techniques namely, Laplacian pyramid (LAP) [52], Ratio-of-laplacian pyramid (ROLP) [53], Gradian pyramid (GRAD) [47], shift invariant discrete wavelet transform (SIDWT) [62], Non sub-sampled contourlet transform (NSCT) [28]. All techniques are implemented using Matlab R2015. In this work, DT-DWT and curvelet transform is used for image decomposition. The number decomposition level is set to 6 for DT-DWT, 5 for curvelet transform, 4 for NSCT, 2 for LAP, ROLP, GRAD, and SIDWT. The high frequency components are fused using maximum selection rule whereas the low frequency components are fused using average fusion rule for state-of-art-methods. The parameters selected for PSO, ST-PSO and BSO are listed in Tables 1 and 2.

Table 1 Parameters of PSO and ST-PSO

PSO and ST-PSO parameters	Values
Particle size (N)	50
Dimension (D) and iteration	2,80
Inertia (ω)	ω -initial =0.9,ω -final =0.4
Learning Factors (c_1,c_2) for PSO	$c_1 = 2, c_2 = 2$
Learning Factors (c_1,c_2) for STPSO	As defined in Eq. (3) and (4)

Table 2 Parameters of BSO

BSO parameters	Values
Idea size (N)	50
Dimension (D) and iteration	2,80
Number of clusters (m)	5
K (slope factor)	20
Mean (μ)	20
Variance (σ)	1
p, p_1, p_2	0.2, 0.8, 0.4

5.3 Face recognition accuracy

The visible and thermal face images are pre-processed prior to recognition. The visible and thermal images are converted into grayscale images. Thermal images are registered based on the visible images. The images are fused using proposed image fusion schemes and state-of-art image fusion techniques. The fused face images are recognized using Eigen face Recognition methodology. The projected Eigen face space is constructed from training face images. Similarity score between the test image and each of training image is calculated. The matched image is the one with the highest similarity score. Recognition performance is computed by finding the percentage of images in the test set, for which the top match is an image of the same person from the training images. The recognition ratio (R) is computed as follows

$$R = \frac{1}{N} \sum_{i=1}^{N} f_i \tag{21}$$

where N is the number of images in the test set. Here f_i=1, if the top most match from the training set belongs to the same object and f_i=0 otherwise.

5.4 Evaluation metrics

Evaluating the quality of the fused image is a challenging task as the reference image is not available to compare to the fusion results. Researchers have proposed several quality metrics to assess the quality of such an image. Zheng Liu et al. [63] classified the twelve-quality metrics as being grouped into four categories in which mutual information (MI), Petrovic metrics, and spatial frequency (SF) are considered in the performance analysis of the proposed work. The first two metrics come under information theory as the image fusion aims to combine information content and does not require a reference image. The other quality metrics, such as fusion symmetry (FS) and correlation coefficients (CC), are also used in the proposed system. The metrics are defined and computed as follows:

1) **Mutual information**: MI quantifies the mutual dependence between the source and fused image, which is given by

$$MI = MI_{AF} + MI_{BF}$$

$$= \sum_{m,n} \left(h_{AF}(m,n) \log_2 \frac{h_{AF}(m,n)}{h_A(m)h_F(n)} + h_{BF}(m,n) \log_2 \frac{h_{BF}(m,n)}{h_B(m)h_F(n)} \right) \tag{22}$$

where MI_{AF} and MI_{BF} are the mutual information between the source images A, B and the fused image. $h_{AF}(m, n)$ is the joint probability distribution function of A and F, and $h_A(m)$ and $h_F(n)$ are the marginal probability distribution functions of A and F, respectively.

2) **Petrovic metrics:** $Q^{AB/F}$ computes the amount of edge information transferred from the source image to the fused image. $L^{AB/F}$ computes the loss of information and $N^{AB/F}$ computes the artefacts (noise) in the fused image due to the fusion process. The procedure for computing $Q^{AB/F}$, $L^{AB/F}$ and $N^{AB/F}$ given in [46] is adapted in our work to compute the petrovic metric.

3) **Fusion symmetry** (FS) defines the symmetry of the fused image with respect to the source images and is computed by

$$FS = 2 - \left| \frac{MI_{AF}}{MI} - 0.5 \right| \tag{23}$$

A higher value of FS denotes better performance of the fusion system.

4) **Spatial Frequency** (SF) is used to measure the action level in an image. A large value of SF represents the clarity of the image. The spatial frequency is computed as

$$SF = \sqrt{(RF)^2 + (CF)^2} \tag{24}$$

where RF is the row frequency and given by

$$RF = \sqrt{\frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=2}^{n} [F(i,j) - F(i,j-1)]^2} \tag{25}$$

and CF is the column frequency denoted by

$$CF = \sqrt{\frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=2}^{n} [F(i,j) - F(i-1,j)]^2} \tag{26}$$

5) **The correlation coefficient (CC)** computes the relevance of the fused image to the source image

Fig. 2 Set of Visible images taken for overall test (experiment I)

Fig. 3 Set of thermal images taken for overall test (experiment I)

and is defined by

$$CC = {(R_{AF} + R_{BF})}\big/_{2},\tag{27}$$

where

$$R_{AF} = \frac{\sum\limits_{i,j}\left(a(i,j)-\overline{A}\right)\left(f(i,j)-\overline{F}\right)}{\sqrt{\left(\sum\limits_{i,j}\left(a(i,j)-\overline{A}\right)^{2}\right)\left(\sum\limits_{i,j}\left(f(i,j)-\overline{F}\right)^{2}\right)}}\tag{28}$$

$$R_{BF} = \frac{\sum\limits_{i,j}\left(b(i,j)-\overline{B}\right)\left(f(i,j)-\overline{F}\right)}{\sqrt{\left(\sum\limits_{i,j}\left(b(i,j)-\overline{B}\right)^{2}\right)\left(\sum\limits_{i,j}\left(f(i,j)-\overline{F}\right)^{2}\right)}}\tag{29}$$

where $\overline{A}$, $\overline{B}$ and $\overline{F}$ are the average pixel intensity of the source images and fused image that measure an index of contrast. $f(i,j)$, $a(i,j)$ and $b(i,j)$ represent the pixel intensity at (i,j) for the fused and source images, respectively.

5.5 Performance of various image fusion algorithms

We have conducted the experiments in four ways. (i) The first experiment includes the training data set using all type of images (varying facial expression, facial illumination and eye glasses). (ii) The second experiment includes only the images which are having varying facial expression as training dataset. (iii) Third experiment includes the face images from varying illumination condition; (iv) The last experiment contains the face images with eye glasses.

The sample input visible and thermal images for over all test and detailed test are given in Figs. 2, 3, 4, 5, 6 and 7. Fusion results of various image fusion algorithms are given in Figs. 8 and 9. The LAP and ROLP fusion methods not effectively captured the eye part of the face. Moreover, the features like nose, ears not effectively synthesized.

The edges of the nose and ears are not completely reproduced compared with GRAD based fusion approaches. Compared with LAP and ROLP, GRAD based fusion method have better eye feature. But, the teeth in the face image lost it contrast. The reason is that visible image having clear representation about eye and teeth under bright illumination conditions. Thermal

Fig. 4 Set of Visible images taken for various facial expression test (experiment II)

Fig. 5 Set of thermal images taken for various facial expression test (experiment II)

images don't have these sharp features. But thermal images have same features if images are taken during presence of light or absence of light.

The visibility of the eye feature and sharpness of the nose are better in SIDWT. But the contrast of teeth is not preserved in SIDWT. The results of NSCT are better than SIDWT. Compared with pyramid-based approach, wavelet based fusion approach gives better fusion result. The proposed image fusion scheme I based on dual tree discrete wavelet transform and PSO effectively captures the eye and teeth features from visible images under the well-illuminated condition and effectively captures from thermal images under the absence of light.

Compared with LAP, ROLP, GRAD and NSCT, the proposed scheme-I effectively synthesized the facial features from thermal and visible images. The factor is that the proposed algorithm is based multi resolution approach. So, the details that are missing at one level can be easily acquired at another level. The coarse details of image are effectively fused by DT-DWT. The decomposed coefficients of thermal and visible images are fused using the optimal weights determined from PSO that improves quality of image interpretation. Though, in Fig. 9a, the last face image, the person with eyeglass is not effectively fused in the resultant image of proposed scheme-I. From that figure we can observe that left eye inside the eye glass is not visible. That feature is effectively captured in proposed image fusion scheme-II.

The Curvelet and brain storm optimization based image fusion results gives better results compared with other methods. The fused images are highly contrasted image and having sharp edge and nose features. Proposed approach gives better result than all other methods. The reason is that the Curvelet transform preserved edges along the curves. Naturally face images contain more curves that is effectively synthesized in the fused images which gives better image representation. From Figs. 10, 11, 12 and 13 we can observe that the proposed image fusion algorithms effectively combine the information from thermal and images under various illumination condition, expression and eyeglass test.

The quantitative analysis, based on various image fusion quality metrics for the fused images by applying the fusion algorithms, is presented in Table 3. Among the quality metrics, a high value for SF, MIF, FS, CC, and $Q^{AB/F}$ and a low value for the RMSE, $L^{AB/F}$ and $N^{AB/F}$ indicate the good quality of the fused image. The high value of $Q^{AB/F}$ for the proposed algorithm indicates that more edge information has been transferred to the fused image. It can also be observed that the low frequency value of $L^{AB/F}$ indicates only a minimal information loss compared with other metrics. The low value of $N^{AB/F}$ indicates that the proposed method introduces minimal artefacts in the fused image, whereas LAP introduces more artefacts compared with all of the other methods.

Fig. 6 Set of Visible images taken for various facial illumination test (experiment III)

Fig. 7 Set of thermal images taken for various facial illumination test (experiment III)
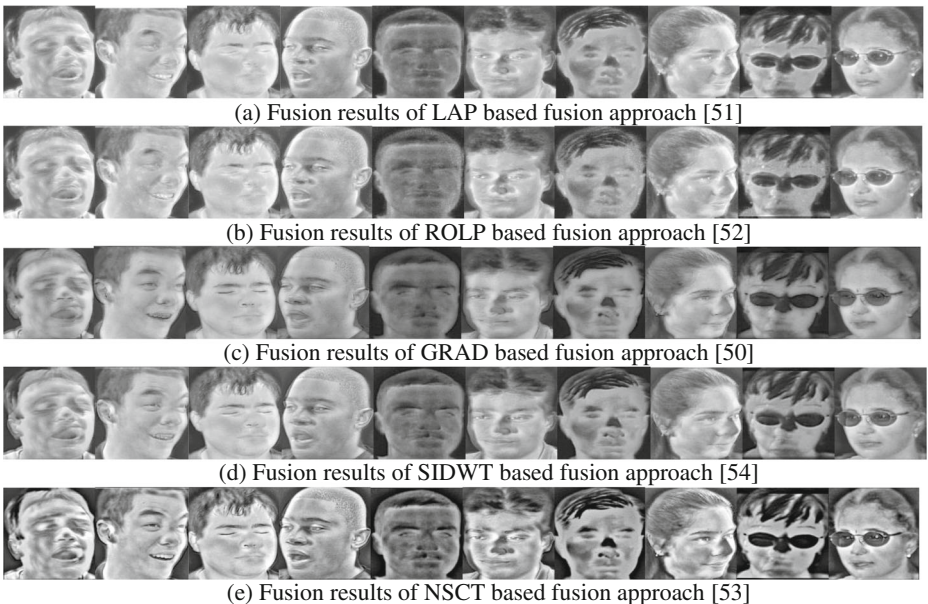
5.6 Face recognition performance of fused images for overall dataset

The overall dataset tests had varying success as shown in Table 4, Figs. 8, 9 and 10. Face recognition using visible image gives the recognition accuracy as 80.00%. In general, fusion led to improved recognition performance compared to visible images.

The face recognition accuracy (Table 4) using fusion methods such as LAP, ROLP, GRAD, SIDWT, NSCT, Jiayi Ma et al. [15] and Parkhi et al. [44] are, 90.50%, 90.21% 90.80%, 92.32%, 93.54%, 94.00% and 93.98% respectively. The authors thank of Jiayi Ma et al. [15] for posting the code in github (https://github.com/jiayi-ma/S3RC). The authors thank Parkhi et al. [44] for posting the code in website (http://www.robots.ox.ac.uk/~vgg/software/vgg_face/). Comparing Proposed fusion methods with other methods (face recognition accuracy 94.17, 94.50, 96.00), proposed method better than other methods and vice versa.
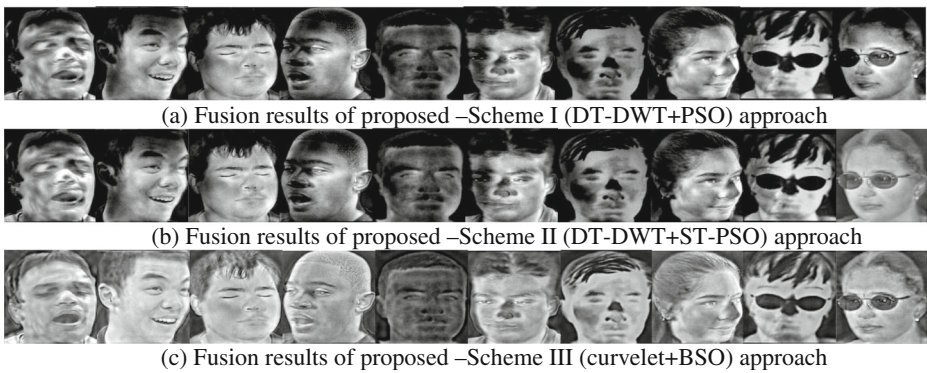
5.7 Face recognition performance based on different facial expressions

The facial expression tests had varying success as shown in Table 4, Figs. 10, 11 and 13. Face recognition using visible image gives the recognition accuracy as 85.32%. In general, fusion led to improved recognition performance compared to recognition in visible spectrum.

(a) Fusion results of LAP based fusion approach [51]

(b) Fusion results of ROLP based fusion approach [52]

(c) Fusion results of GRAD based fusion approach [50]

(d) Fusion results of SIDWT based fusion approach [54]

(e) Fusion results of NSCT based fusion approach [53]

Fig. 8 Fusion results for overall test (combination of facial expression, illumination and eyeglasses) using various image fusion algorithms. **a** Fusion results of LAP based fusion approach [52]. **b** Fusion results of ROLP based fusion approach [53]. **c** Fusion results of GRAD based fusion approach [47]. **d** Fusion results of SIDWT based fusion approach [62]. **e** Fusion results of NSCT based fusion approach [28]

(a) Fusion results of proposed –Scheme I (DT-DWT+PSO) approach

(b) Fusion results of proposed –Scheme II (DT-DWT+ST-PSO) approach

(c) Fusion results of proposed –Scheme III (curvelet+BSO) approach

Fig. 9 Fusion results for overall test (combination of facial expression, illumination and eyeglasses) using proposed image fusion algorithms. **a** Fusion results of proposed –Scheme I (DT-DWT+PSO) approach. **b** Fusion results of proposed –Scheme II (DT-DWT+ST-PSO) approach. **c** Fusion results of proposed –Scheme III (curvelet+BSO) approach
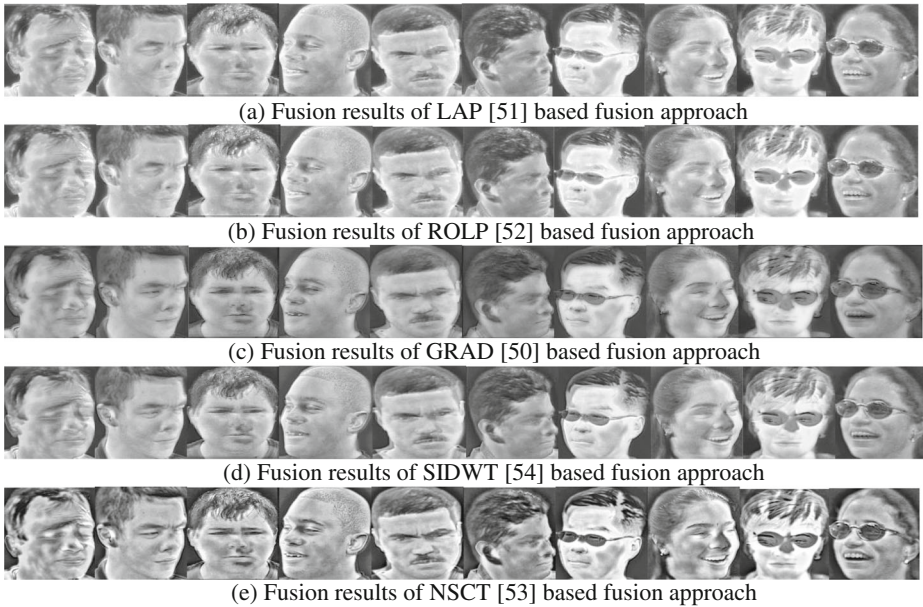
Comparing thermal images with fusion, sometimes thermal images performed better than fusion and vice versa. The reason is that the presence of undesired illumination effect of visible images taken account into the fused image. Among all the methods the proposed method based on curvelet and BSO based fusion approach gives better recognition accuracy (90.90%).

5.8 Face recognition performance under varying illumination conditions

The facial expression tests had varying success as shown in Table 4, Figs. 14, 15, 16 and 17. Face recognition using visible image gives the recognition accuracy as 85.32%. In general, fusion led to improved recognition performance compared to recognition in visible spectrum.
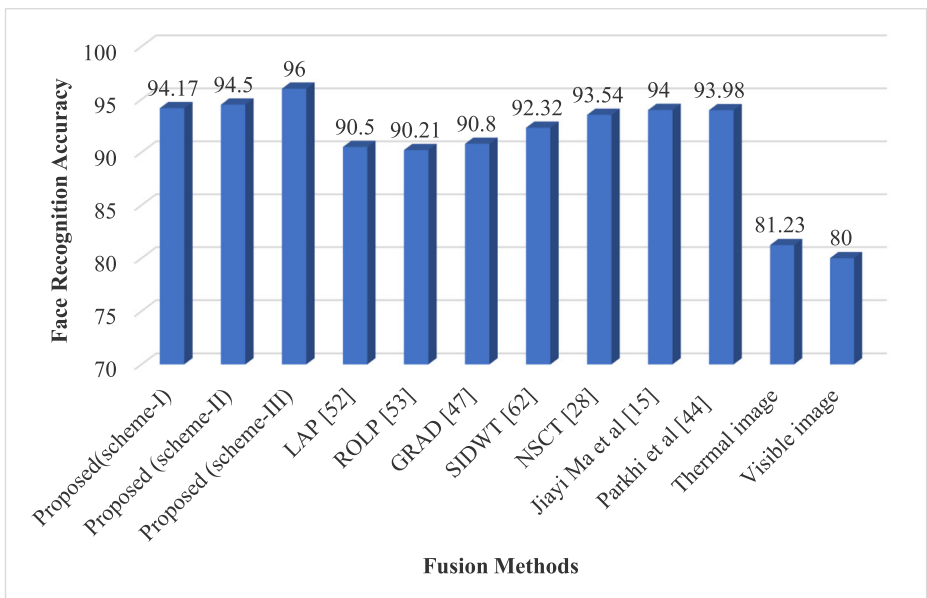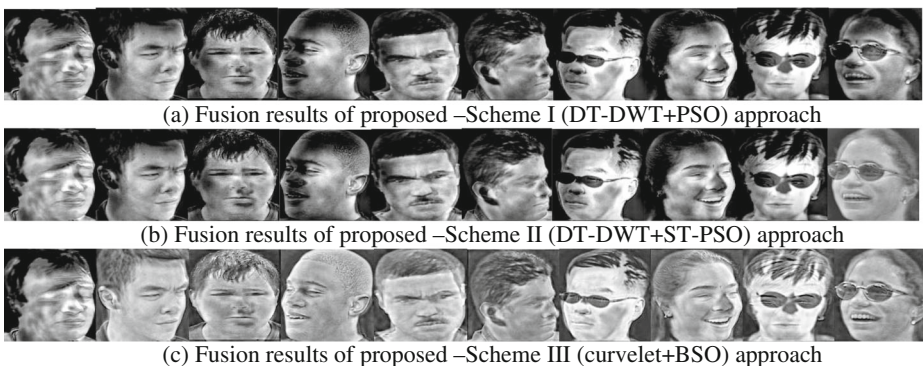

Fig 10 Graphical representation of Face Recognition performance for overall dataset

(a) Fusion results of LAP [51] based fusion approach

(b) Fusion results of ROLP [52] based fusion approach

(c) Fusion results of GRAD [50] based fusion approach

(d) Fusion results of SIDWT [54] based fusion approach

(e) Fusion results of NSCT [53] based fusion approach

Fig. 11 Fusion results for various facial expression using various image fusion algorithms. **a** Fusion results of LAP [52] based fusion approach. **b** Fusion results of ROLP [53] based fusion approach. **c** Fusion results of GRAD [47] based fusion approach. **d** Fusion results of SIDWT [62] based fusion approach. **e** Fusion results of NSCT [28] based fusion approach

Comparing thermal images with fusion, sometimes thermal images performed better than fusion and vice versa. The reason is that the presence of undesired illumination effect of visible images taken account into the fused image.

Among all the methods the proposed method based on Curvelet and BSO based fusion approach gives better recognition accuracy (90.90). From Fig. 17 we can observe that under varying illumination condition fused images give better recognition accuracy (LAP - 85.30%, ROLP – 86.74%, GRAD-86.52%, SIDWT-87.88% and NSCT-84.00%) compared with recognition in visible spectrum. Recognition in the visible spectrum was not satisfactory while recognition using proposed fused images had comparable performance to that in the thermal

(a) Fusion results of proposed –Scheme I (DT-DWT+PSO) approach

(b) Fusion results of proposed –Scheme II (DT-DWT+ST-PSO) approach

(c) Fusion results of proposed –Scheme III (curvelet+BSO) approach

Fig. 12 Fusion results for various facial expression using proposed image fusion algorithms
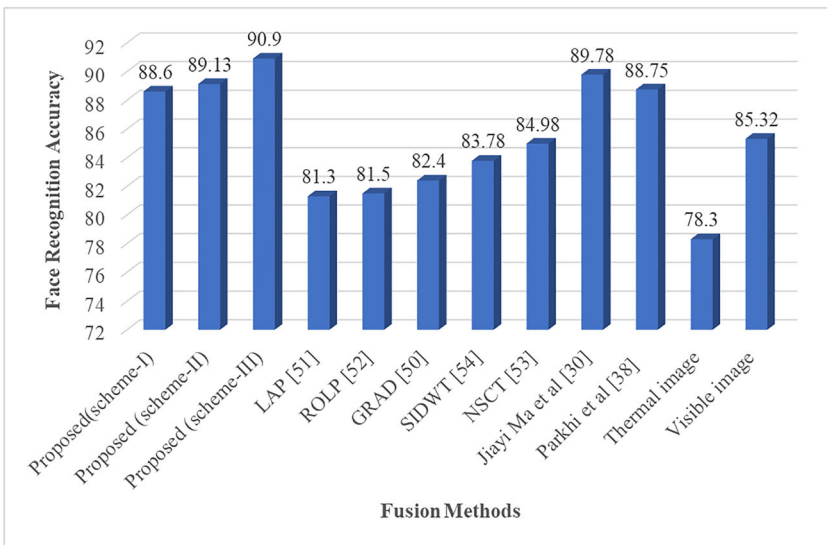
Fig 13 Graphical representation of Face Recognition performance on different facial expressions

spectrum. The recognition accuracy using proposed fused images are 90.65%, 91.23% and 90.34% respectively.

5.9 Face recognition performance with eyeglasses

Face recognition accuracy for images with eyeglasses using IR images give poor recognition performance (62.67%). The reason is that eyeglasses are not sensitive to heat. The part under the eyeglasses cannot effectively capture in thermal images. Our experimental results illustrate clearly that IR is robust to illumination changes but perform poorly when glasses are present in the face image (Fig. 17). From Table 4 and Fig. 12 we can observe that considerable improvement is achieved in this case by fusing IR with visible images in curvelet and dual tree discrete wavelet transform domain.

We have also attempted to analyse PSO, ST-PSO and BSO solutions in order to understand what part of the face are encoded by IR features and what parts are encoded by visible features. Eye and teeth were optimally combined mostly using features from visible spectrum. Head parts of the face were optimally combined using features from thermal spectrum.

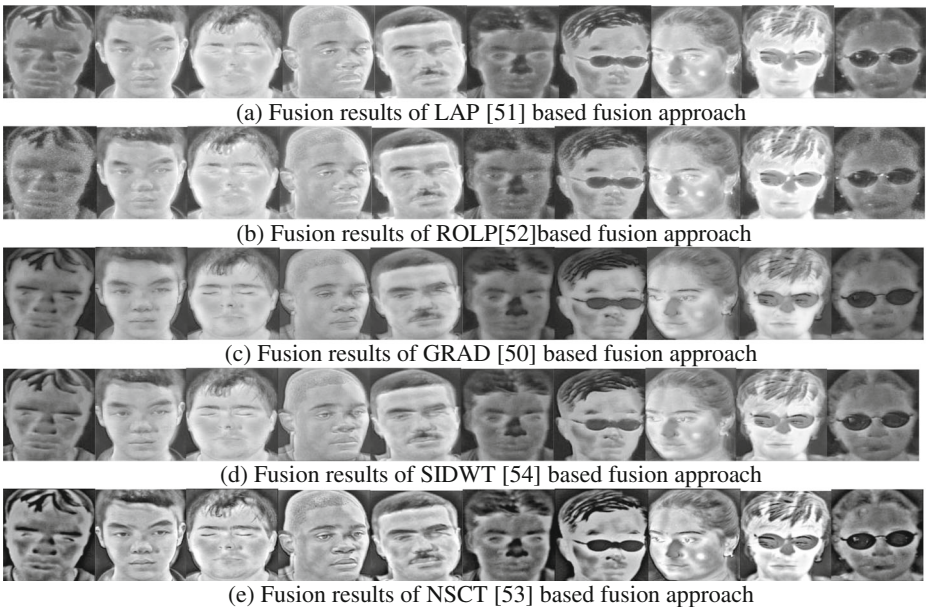**Table 3** Quantitative Analysis of the image quality metrics

Technique/Measure	Entropy	RMSE	SF	MIF	FS	CC	$Q^{AB/F}$	$L^{AB/F}$	$N^{AB/F}$
Proposed (scheme-I)	7.6619	7.7457	38.8010	3.6446	1.9797	0.8803	0.8061	0.1687	0.0028
Proposed (scheme-II)	7.7435	7.6450	38.9012	3.7894	1.9823	0.8908	0.8123	0.1542	0.0025
Proposed (scheme-III)	7.7578	7.5013	39.1239	3.8010	1.9734	0.9012	0.8231	0.1432	0.0022
LAP [52]	7.5485	10.9867	27.2795	3.0479	1.9837	0.8549	0.6918	0.4033	0.0059
ROLP [53]	7.4770	9.8190	26.5332	3.0251	1.9848	0.8641	0.7071	0.3898	0.0051
GRAD [47]	7.6431	13.9366	33.6090	2.8272	1.9413	0.8644	0.7873	0.3918	0.0040
SIDWT [62]	7.5342	8.0123	34.7689	2.8970	1.8341	0.8691	0.7918	0.3033	0.0039
NSCT [28]	7.5450	9.9234	33.7845	2.9076	1.9231	0.8637	0.7871	0.2898	0.0037

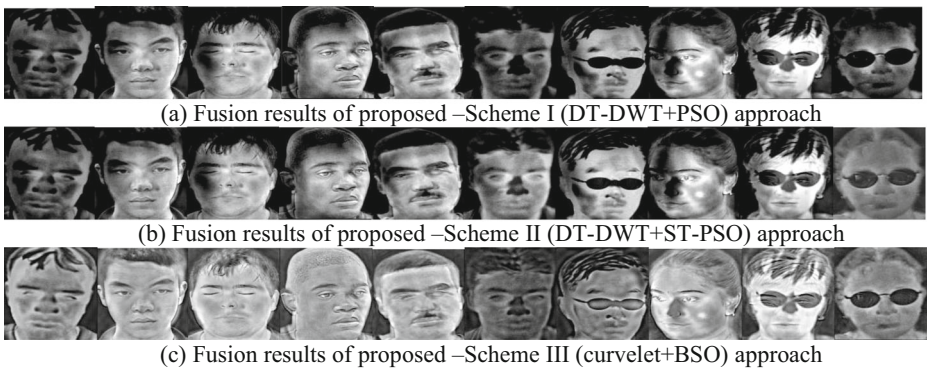**Table 4** Face recognition performance of thermal, visible and fused images

Methods	Recognition Accuracy (overall test) in %	Recognition Accuracy (detailed test in %) with varying facial		
		Expression	Illumination	Eyeglasses
Proposed(scheme-I)	94.17	88.60	90.65	97.34
Proposed (scheme-II)	94.50	89.13	91.23	98.12
Proposed (scheme-III)	96.00	90.90	90.34	98.50
LAP [52]	90.50	81.30	85.30	88.89
ROLP [53]	90.21	81.50	86.74	90.98
GRAD [47]	90.80	82.40	86.52	91.24
SIDWT [62]	92.32	83.78	87.88	92.90
NSCT [28]	93.54	84.98	84.00	95.19
Jiayi Ma et al [15]	94.00	89.78	90.20	96.98
Parkhi et al [44]	93.98	88.75	90.45	96.50
Thermal image	81.23	78.30	98.00	62.67
Visible image	80.00	85.32	58.50	85.50

5.10 Processing time

The computational efficiency of different fusion methods is compared here. In our experiments, all the five test methods are implemented in MATLAB R2015 on a computer with a 3.0 GHz CPU and 4 GB RAM. The average running time of different fusion methods are listed in Table 5. The GRAD method has a high computational efficiency whereas LAP and SIDWT take 7 s to fuse source images. The proposed method takes 100 s to complete the fusion

(a) Fusion results of LAP [51] based fusion approach

(b) Fusion results of ROLP[52]based fusion approach

(c) Fusion results of GRAD [50] based fusion approach

(d) Fusion results of SIDWT [54] based fusion approach

(e) Fusion results of NSCT [53] based fusion approach

Fig. 14 Fusion results for various illumination condition using various image fusion algorithms. **a** Fusion results of LAP [52] based fusion approach. **b** Fusion results of ROLP [53] based fusion approach. **c** Fusion results of GRAD [47] based fusion approach. **d** Fusion results of SIDWT [62] based fusion approach. **e** Fusion results of NSCT [28] based fusion approach

(a) Fusion results of proposed –Scheme I (DT-DWT+PSO) approach

(b) Fusion results of proposed –Scheme II (DT-DWT+ST-PSO) approach

(c) Fusion results of proposed –Scheme III (curvelet+BSO) approach

Fig. 15 Fusion results for various illumination condition using proposed image fusion algorithms. **a** Fusion results of proposed –Scheme I (DT-DWT+PSO) approach. **b** Fusion results of proposed –Scheme II (DT-DWT+ST-PSO) approach. **c** Fusion results of proposed –Scheme III (curvelet+BSO) approach

process. We believe that with a more efficient implementation approach such as C++, the running time can be easily reduced.

6 Conclusion

We presented and compared three different fusion schemes for combining thermal and visible imagery for the purpose of face recognition. IR images are more robust to varying illumination conditions, but gives performance when eyeglasses are present in the face images. The proposed swarm intelligence based fusion methodology is general enough and can be applied
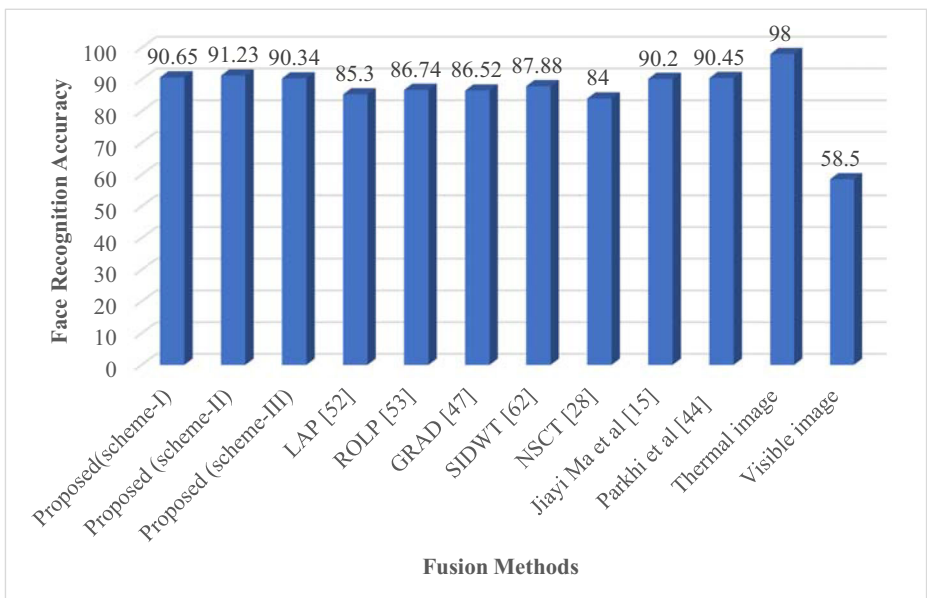
Fig. 16 Graphical representation of Face Recognition performance on varying illumination conditions
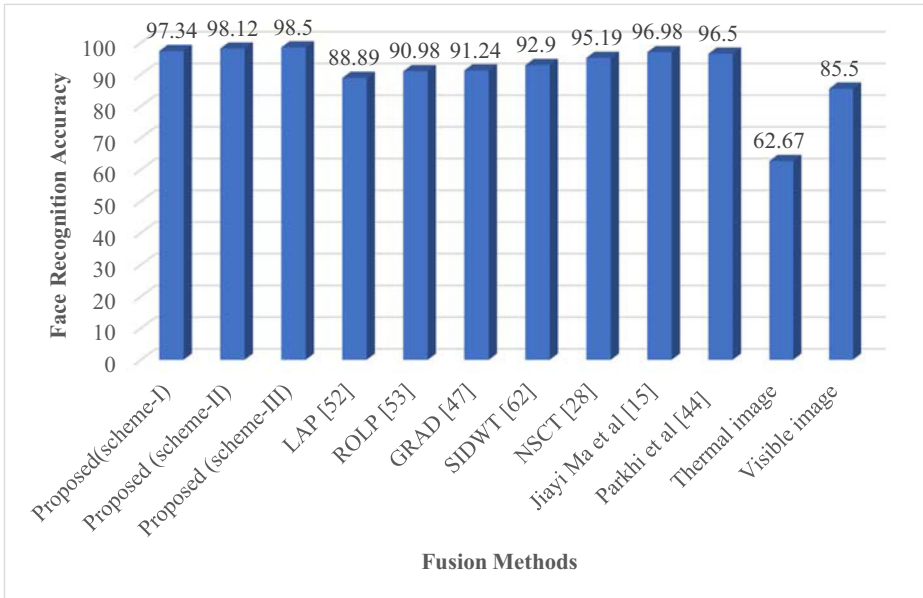
Fig. 17 Graphical representation of Face Recognition performance with eye glasses

in these cases as well as to improve recognition performance when thermal and visible images are not very reliable.

Several interesting conclusions can be drawn by considering these results.

(i)   As expected, face recognition in the thermal images is not influenced by the illumination changes. However, thermal image yielded very low success when eyeglasses were present in the face images.

(ii)   Illumination changes had an important influence on the success of face recognition in the visible domain. Illumination changes also affect the fused images. The fact is that the fusion was not able to completely discard undesired illumination effects present in the visible images.

(iii) Success of face recognition using fused images implies that fusion was able to become less sensitive both to eyeglasses and illumination changes

(iv)   Between the three proposed fusion schemes tested, fusion in the Curvelet domain yield higher recognition performance over all.

Table 5 Average running time for various image fusion methods

Method	Time /Seconds
Proposed (scheme-I)	90.56
Proposed (scheme-II)	92.56
Proposed (scheme-III)	100.45
LAP [52]	7.26
ROLP [53]	8.50
GRAD [47]	6.50
SIDWT [62]	7.05
NSCT [28]	45.03

References

1. Ahmad A, Riaz MM, Ghafoor A, Zaidi T (2016) Noise resistant fusion for multi-exposure sensors. IEEE Sensors J 16(13):5123–5124
2. Arivazhagan S, Mumtaj J (2007) Face Recognition using Multi-Resolution Transform. International Conference on Computational Intelligence and Multimedia Applications, 301–306
3. Bebis G, Gyaourova A, Singh S, Pavlidis I (2006) Face recognition by fusing thermal infrared and visible imagery. Image Vis Comput 24:727–742
4. Belhumeur PN, Hespanha JP, Kriegman DJ (1997) Eigenfaces vs.fisherfaces: recognition using class specific linear projection. IEEE Trans Pattern Anal Mach Intell 19(7):711–720
5. BenAbdelkader C, Griffin P (2005) A local region-based approach to gender classification from face images. Proc. IEEE Comp. Society Conf. Computer Vision and Pattern Recognition, pages 52–57
6. Brunelli R, Poggio T (1993) Face recognition: features versus templates. IEEE Trans Pattern Anal Mach Intell 15(10):1042–1052
7. Chen W, Er MJ, Wu S (2006) "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain" IEEE transactions on systems, man, and cybernetics. Part B (Cybernetics) 36(2):458–466
8. Chen YL, Jahanshahi MR, Manjunatha P, Gan WP, Abdelbarr M, Masri SF, Becerik-Gerber B, Caffrey JP (2016) Inexpensive multimodal sensor fusion system for autonomous data acquisition of road surface conditions. IEEE Sensors J 16(21):7731–7743
9. Choudhary A, Vig R (2018) Face Recognition Using Multiresolution Hybrid Kekre-DCT Wavelet Transform Features with Multiclass ECOC Framework. Proc Comp Sci 132:1781–1787
10. Choudhary A, Vig R (2018) Face recognition using multiresolution hybrid Kekre-DCT wavelet transform features with multiclass ECOC framework. Proc Comp Sci 132:1781–1787
11. Dakin SC, Watt RJ (2009) Biological `bar codes' in human faces. World Acad Sci Eng Technol 9:1–10
12. Ekenel HK, Sankur B (2005) Multiresolution face recognition. Image Vis Comput 23:469–477
13. Eleyan A, Demirel H (2007) Face Recognition using Multiresolution PCA. IEEE International Symposium on Signal Processing and Information Technology, 52–55.
14. Er MJ, Chen W, Wu S (2005) High-speed face recognition based on discrete cosine transform and RBF neural networks. IEEE Trans Neural Netw 16(3):679–691
15. Gao Y, Ma J (2017) Semi-Supervised Sparse Representation Based Classification for Face Recognition with Insufficient Labeled Samples. IEEE Transactions on Image Processing
16. Ghasemzadeh A, Demirel H (2017) 3D discrete wavelet transform-based feature extraction for hyperspectral face recognition. IET Biometrics 7(1):49–55
17. Grotschel M, Lovász L (1993) Combinatorial optimization: a survey, DIMACS technical report 93–29, Princeton University. In Internet
18. Guzman AM, Goryawala M, Wang J, Barreto A, Andrian J, Rishe N, Adjouadi M (2013) Thermal imaging as a biometrics approach to facial signature authentication. IEEE J Biomed Health Inf 17(1):214–222
19. Hermosilla G, Gallardo F, Farias G, San Martin C (2015) Fusion of visible and thermal descriptors using genetic algorithms for face recognition systems. Sensors 15:17944–17962. https://doi.org/10.3390/s150817944
20. Hizem W, Allano L, Mellakh A, Dorizzi B (2009) Face recognition from synchronised visible and near-infrared images. IET Signal Process 3(4):282–288
21. Hollingsworth KP, Darnell SS, Miller PE, Woodard DL, Bowyer KW, Flynn PJ (2012) Human and Machine Performance on Periocular Biometrics Under Near-Infrared Light and Visible Light. IEEE Transact Inf Foren Sec 7(2):588–601
22. IEEE OTCBVS WS Series Bench (n.d.); DOE University Research Program in Robotics under grant DOE-DE-FG02-86NE37968; DOD/TACOM/NAC/ARC Program under grant R01–1344-18; FAA/NSSA grant R01–1344-48/49; Office of Naval Research under grant #N000143010022
23. IEEE OTCBVS WS Series Bench (n.d.); DOE University Research Program in Robotics under grant DOE-DE-FG02-86NE37968; DOD/TACOM/NAC/ARC Program under grant R01–1344-18; FAA/NSSA grant R01–1344-48/49; Office of Naval Research under grant #N000143010022. http://vcipl-okstate.org/pbvs/bench/Data/02/download.html
24. Imtiaz H, Fattah SA (2011) A face recognition scheme using waveletbased dominant features. Signal & Image Processing : An International Journal (SIPIJ). 2(3)
25. Kim T-K, Kittler J (2005) Locally Linear Discriminant Analysis for Multi-modally Distributed Classes for Face Recognition with a Single Model Image. IEEE Trans Pattern Anal Mach Intell 27(3):318–327
26. Klare BF, Jain AK Heterogeneous Face Recognition Using Kernel Prototype Similarities. IEEE Trans Pattern Anal Mach Intell 35(6):1410–1422

27. Kong SG, Heo J, Boughorbel F, Zheng Y, Abidi BR, Koschan A, Yi M, Abidi MA (2007) Multiscale Fusion of Visible and Thermal IR Images for Illumination-Invariant Face Recognition. Int J Comput Vis 71(2):215–233

28. Kong WW, Lei YJ, Lei Y, Zhang J (2010) Technique for image fusion based on non-subsampled contourlet transform domain improved NMF. SCIENCE CHINA Inf Sci 53(12):2429–2440

29. Kumar A, Zhang D (2006) Personal recognition using hand shape and texture. IEEE Trans Image Process 15(8):2454–2461

30. Laurenz W, Fellous JM, Krüger N, von der Malsburg C (1997) Face recognition by elastic bunch graph matching. 19(7): 775–779

31. Lawrence S, Giles CL, Tsoi A, Back AD (1998) Face recognition: a convolutional neural network approach. IEEE Trans Neural Netw 8(1):98–113

32. Li SZ, Chu RF, Liao SC, Zhang L (2007) Illumination invariant face recognition using near-infrared images. IEEE Trans Pattern Anal Mach Intell 29(4):627–639

33. Liu C (2004) Gabor-Based Kernel PCA with Fractional Power Polynomial Models for Face Recognition. IEEE Trans Pattern Anal Mach Intell 26(5):572–581

34. Ma J, Zhao J, Ma Y, Tian J (2015) Non-rigid visible and infrared face registration via regularized Gaussian fields criterion. Pattern Recogn 48:772–784

35. Ma J, Qiu W, Zhao J, Ma Y, Yuille AL (2015) Robust L2E Estimation of Transformation for Non-Rigid Registration, IEEE Transactions on Signal Processing. IEEE Transactions on Signal Processing, VOL. 63, NO. 5, MARCH 1, 2015, 1115–1129

36. Ma J, Chen C, Li C, Huang J (2016) Infrared and visible image fusion via gradient transfer and total variation minimization. Info Fusion 31:100–109

37. Ma J, Ma Y, Li C (2019) Infrared and visible image fusion methods and applications: a survey. Info Fusion 45:153–178

38. Madheswari K, Venkateswaran N (2016) Swarm intelligence based optimisation in thermal image fusion using dual tree discrete wavelet transform. Quant Infra Thermograph J, First Online, 1–20

39. Martõ Ânez AM, Kak AC (2001) PCA versus LDA. IEEE Trans Pattern Anal Mach Intell 23(2):228–233

40. Meraoumia A, Chitroub S, Bouridane A (2010) Gaussian modeling and Discrete Cosine Transform for efficient and automatic palmprint identification. Int Conf Mach Web Intell (ICMWI):121–125

41. Moghaddam B, Pentland A (1997) Probabilistic visual learning for object representation. IEEE Trans Pattern Analysis Mach Intell 18:696–710

42. Nicolò F, Schmid NA (2012) Long range cross-spectral face recognition: matching SWIR against visible light images. IEEE Transact Inf Foren Sec 7:1717–1725

43. Osuna E, Freund R, Girosi F (1997) Training support vector machines: an application to face detection

44. Parkhi OM, Vedaldi A, Zisserman A (2015) Deep Face Recognition British Machine Vision Conference

45. Penev P, Atick J (1996) Local Feature Analysis: A general statistical theory for object representation

46. Petrovic V (2007) Subjective tests for image fusion evaluation and objective metric validation. Info Fusion 8(2):2018–2216

47. Raghavendra R, Dorizzi B, Rao A, Kumar GH (2011) Particle swarm optimization based fusion of near infrared and visible images for improved face verification. Pattern Recogn 44:401–411

48. Rajoub BA, Zwiggelaar R (2014) Thermal facial analysis for deception detection. IEEE Transact Inf Foren Sec 9(6):1015–1023

49. Schoelkopf B, Smola A, Muller KR (1997) Kernel principal components analysis", artificial neural networks, ICANN97

50. Seal A, Bhattacharjee D, Nasipuri M (2016) Human face recognition using random forest based fusion of à-trous wavelet transform coefficients from thermal and visible images. Int J Electron Commun (AEÜ) 70: 1041–1049

51. Shen LL, Bai L (2004) Gabor feature based face recognition using kernal methods. Proc IEEE Int Conf Auto Face Gest Recog:386–389

52. Tan H, Huang X, Tan H, He C Pixel-level Image Fusion Algorithm Based on Maximum Likelihood and Laplacian Pyramid Transformation. J Comput Inf Syst 9(1):327–334

53. Toet A (1989) Image fusion by a ratio of low-pass pyramid. Pattern Recogn Lett 9:245–253

54. Tripathi BK (2017) On the complex domain deep machine learning for face recognition. Appl Intell 47(2): 382–396

55. Turk M, Pentland A (1991) Eigenfaces for recognition. J Cogn Sci:71–86

56. Turk M, Pentland A (1991) Eigen faces for recognition. IEEE Transact Cogn Neurosci 13(1):71–86

57. Turk MA, Pentland AP (1991) Face recognition using eigenfaces. Computer Vision and Pattern Recognition. Proceedings {CVPR'91.}, {IEEE} Computer Society Conference on 1991

58. Villegas-Quezada C, Climent J (2008) Holistic face recognition using multivariate approximation,genetic algorithms and AdaBoost classifier: preliminary results. World Acad Sci Eng Technol 44:802–806

59. Xiaozheng Z, Gao Y (2009) Face recognition across pose: A review. Pattern Recogn 42:2876–2896
60. Yang Y, Que Y (2016) Shuying Huang, and Pan Lin, 'multimodal sensor medical image fusion based on Type-2 fuzzy logic in NSCT domain'. IEEE Sensors J 16(10):3735–3745
61. Yang J, Zhang D, Frangi AF, Yang J-y (2004) Two-Dimensional PCA : A New Approach to Appearance-Based Face Representation and Recognition. IEEE Trans Pattern Anal Mach Intell 26(1):131–137
62. Zheng L, Blasch E, Xue Z, Zhao J, Laganiere R, Wei W (2012) Objective Assessment of Multi-resolution Image Fusion algorithms for Context enhancement in Night Vision: A Comparative Study. IEEE Trans Pattern Anal Mach Intell 34(1):94–109
63. Zheng L, Blasch E, ZhiyunXue JZ, Laganiere R, Wu W (2012) Objective assessment of multi-resolution image fusion algorithms for context enhancement in night vision: a comparative study. IEEE Trans Pattern Anal Mach Intell 34(1)
64. Zhou M, Wei H (2006) Face verification using Gabor wavelets and adaboost. Proc Int Conf Pattern Recog: 404–407

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

Dr. K.Madheswari, Associate Professor in the Department of Computer Science and Engineering has 7 years of teaching experience. She received her PhD in the area of image fusion from Anna University, Chennai. She received her Bachelor's degree in Computer Science and Engineering (First Class with Distinction) from Hindusthan College of Engineering and Technology, Coimbatore. M.E. Software Engineering (First class with Distinction) from Anna University. She has done SCJA (Sun certified Java Associate) certification in Java. Her area of interest is on image processing, Optimization techniques, Software Engineering, Software Testing and Wireless Networks. She is a member of professional bodies such as ACM, IEEE, Indian Society for Technical Education (ISTE). She has published 12 papers in Journals, 12 papers in International conferences. She has reviewed research articles for IEEE Geo science and Remote Sensing Letters journal . She also reviewed research articles for Springer Journal "Pattern Analysis and Applications", "Wireless Personal communication" journal and "Multimedia Tools and Applications".

Venkateswaran. N received his PhD from Anna University, India in Signal and Image Processing in 2008. He also received hisM. Tech degree in Communication Systems from Pondicherry Engineering College, Pondicherry University in 1995. He was Professor at Sri Venkateswara College of Engineering, India until 2010. Later, he joined as professor in the Department of Electronics and Communication Engineering at SSN college of Engineering, India during June 2010. He has more than 23 years of teaching, research and industry experience. He has published more than 50 technical papers in various journals and conferences. His research interest includes Signal and Image processing, Wireless Communications and its applications. He is member of IEEE and Fellow of IETE (India).