

Optical Engineering

SPIEDigitalLibrary.org/oe

Eye region-based fusion technique of thermal and dark visual images for human face recognition

Mrinal Kanti Bhowmik
Debotosh Bhattacharjee
Dipak Kumar Basu
Mita Nasipuri



Eye region-based fusion technique of thermal and dark visual images for human face recognition

Mrinal Kanti Bhowmik

Tripura University (A Central University)
Department of Computer Science and Engineering
Suryamaninagar 799022, Tripura, India
E-mail: mkb_cse@yahoo.co.in

Debotosh Bhattacharjee

Dipak Kumar Basu

Mita Nasipuri

Jadavpur University
Department of Computer Science and Engineering
Kolkata 700032, India

Abstract. We present an approach for human face recognition using eye region extraction/replacement method under low illumination and varying expression conditions. For conducting experiments, two different sets of face images, namely visual and corresponding thermal, are used from Imaging, Robotics, and Intelligent Systems (IRIS) thermal/visual face data. A decomposition and reconstruction technique of Daubechies wavelet co-efficient (db4) is used to generate the fused image by replacing the eye region in the visual image with the same region from the corresponding thermal image. After that, independent component analysis over the natural logarithm domain (Log-ICA) is used for feature extraction/dimensionality reduction, and finally, a classifier is used to classify the fused face images. Two different image sets, i.e., training and test image sets, are mainly prepared using the IRIS thermal/visual face database for finding the accuracy of the proposed system. Experimental results show the proposed method is more efficient than other image fusion techniques which have used region extraction techniques for dark faces. © 2012 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: [10.1117/1.OE.51.7.077205](https://doi.org/10.1117/1.OE.51.7.077205)]

Subject terms: thermal image; wavelet transformation; independent component analysis; multilayer perceptron; face recognition.

Paper 111140 received Sep. 14, 2011; revised manuscript received May 19, 2012; accepted for publication May 29, 2012; published online Jul. 9, 2012.

1 Introduction

In uncontrolled environments, face recognition based on only visible spectrum is still a challenging task. This is due to the large variations in visual images caused by different illumination conditions, pose variations, aging, expressions, and disguises like glasses, facial hair, or cosmetics etc.; the first two, i.e., variations in illumination and pose¹ are unavoidable in different applications, such as surveillance, outdoor access control, etc. Performance of visual face recognition usually degrades a lot when the lighting is too low or when it is not illuminating the face homogeneously. Sometimes, the changes due to illumination variation become more prominent than the differences between the individuals. Many algorithms, e.g., dropping leading eigenfaces, histogram equalization etc., have been studied with partial success to reduce the effects of such variations.² All of these techniques try to reduce the within-class variation caused by illumination changes. A visual image based face recognition system that has been optimized to identify the light-skinned people, may cause higher false alarms while experimenting with the dark-skinned people. Recently, researchers have investigated the use of thermal infrared (IR) face images for person identification to tackle the problems due to illumination variation, presence of facial hair, changes in hairstyle, etc.³⁻⁶; however, thermal images may not perform well for recognition of faces with varying poses, which are very common in this field. Further, IR imagery has few other disadvantages as it is sensitive to temperature changes in the surrounding environment and variations in the heat patterns of the face. In contrast, visual imagery is more

robust in respect to the above factors, but highly sensitive to illumination variation. This suggests that a proper fusion⁷ of the information from both visual and thermal spectra may have the potential to improve face recognition performance.

In the present paper, we have proposed a method toward the partial solution of face recognition under low lighting condition, by eliminating the low illuminated region from the visual image of a face and then replacing it with the corresponding region of the thermal image of the same face; and finally the resultant image is fused with the same thermal image. As, eye is the most important fiducial region on a face and because of the fact that thermal images are less affected due to illumination variation than visual images,⁸ it may be helpful to fuse the visual image, generated after region replacement, with the thermal image. The researchers for human face recognition under low illumination condition have already developed different techniques. Goh et al. in Ref. 9 introduced a wavelet based illumination invariant reflectance model to a solution for visual face image recognition. The method proposed by them intends to remove illumination component by changing the wavelet approximation subband coefficients to zero values. The algorithm was tested on two different databases and achieved the equal error rate (ERR) of 10.83% and 15.37% in two different experiments using the Yale B database, and 14.73% using the CMU PIE database. In Ref. 10, Ekenel and Sankur proposed a method based on subspace projection operation like PCA and ICA and employed multiresolution analysis to decompose the image into its subbands. Their aim was to search for the subband that is insensitive to variations in expression and illumination. This algorithm was tested on face images that differ in expression or illumination separately, from the CMU PIE, FERET and Yale databases, and achieved

91.55% of correct recognition rate. In Ref. 11, Cheng et al. proposed a novel approach for illumination normalization under varying lighting conditions based on a 2D Gaussian illumination model. The algorithm was tested on the Yale B database, and a recognition rate of 88.50% was achieved.

The organization of this paper is as follows. In Sec. 2, the complete overview of the system implementation of this paper is given; in Sec. 3, the experimental results are presented, and in Sec. 4, conclusions and future work of the system are presented.

2 System Overview

The proposed technique for recognition is based on replacement of a low-lighting eye region of a visual face image by the corresponding region from its thermal image, and then fusion of this newly formed face image with the same thermal face image. The steps may be given as follows:

- (1) Creation of a new matrix of dimension equal to that of visual image.
- (2) Detection of low illuminated eye region in the visual image.
- (3) Copying of the pixel values of the visual face image, except the low illuminated region into the newly created matrix.
- (4) Extraction of region, equivalent to the low illuminated region of visual face, from the corresponding thermal face image.
- (5) Copying of the pixel values of the extracted region in step (4) into respective positions in the newly created matrix.
- (6) Fusion of newly created image matrix with the thermal image.

In Fig. 1, sample images from the IRIS database are presented. These sample thermal and visual images are not of the same resolution, but they are of same expression, pose, and illumination. It clearly indicates how a visible image varies with the change of illumination, especially when the images are taken under dark condition, whereas thermal images are unaffected by such changes. Therefore, fusion with thermal image may be the solution to the problem of varying illumination. The block diagram of the entire system is shown in Fig. 2.

2.1 Replacement of Low Lighting Area of Face

This proposed algorithm is an attempt to eliminate the region of low illumination from the visual face with the help of corresponding thermal image. According to the database information of the IRIS thermal/visible face database,¹² it contains 4228 pairs of thermal and corresponding visual images with different expressions, illumination, and poses. Out of these images, only the dark images have been used in this work. These images were captured in a dark room. The presence of natural light may still be noticed in the left side of the face images during photography, but the right side of the images are too dark because of the absence of natural light source; these images with different expressions are mainly used in our proposed work. As the eye is one of the most important fiducial regions of the face, we have replaced only the upper right portion of the visual face image with the corresponding face region of the thermal image. For this, the face images have been divided in such a way, so that the low illuminated area can be extracted from a face.

2.1.1 Division of a face image in multiple parts

The original visual image of the database is resized first so that they are of same resolution. Then the visual and corresponding thermal images are divided into four equal parts. Out of these identified areas of a visual image, the least illuminated area is detected and replaced with the corresponding area of the corresponding thermal image. Figure 1 represents the original thermal and corresponding visual images of the database, and Fig. 3 represents the resized images of the images shown in Fig. 1. Among those four face areas, discussed above, the low illuminated area is only the upper right portion of the visual image, which is to be replaced with the corresponding face area of thermal image. In Fig. 4, the detected and extracted low illuminated area of visual image is shown.

2.1.2 Replacement of the low illuminated area of visual face image

Before replacing the low illuminated area of eye from the visual face, the corresponding thermal image will be divided in the same manner. Finally, the low illuminated eye region of the visual face image is replaced by the corresponding region of thermal face image. In Fig. 5(a), extraction of region from thermal face image corresponding to the low

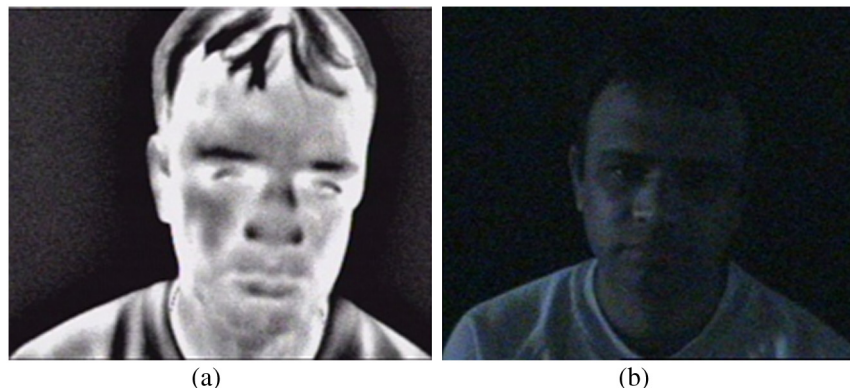


Fig. 1 (a) and (b) are the thermal and corresponding visual images, respectively, of the Imaging, Robotics, and Intelligent Systems (IRIS) database; 320 × 240 pixels with the same expression, pose, and illumination.

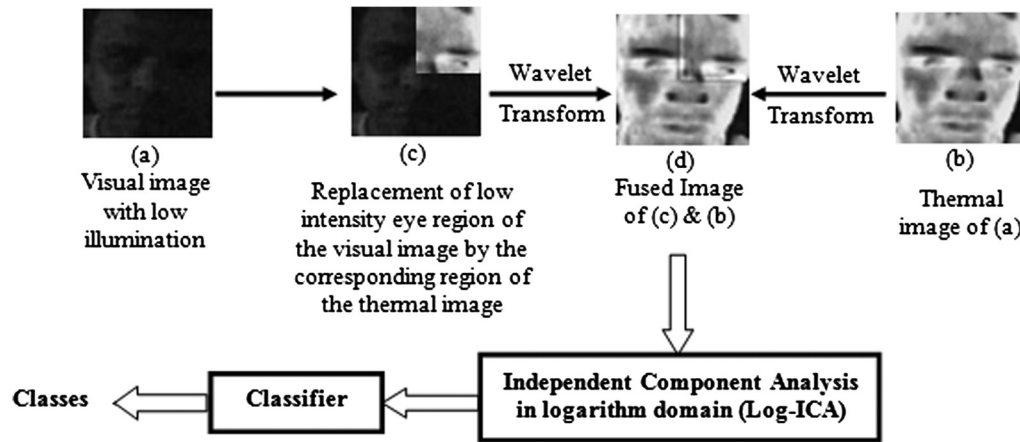


Fig. 2 Block diagram of the proposed system.

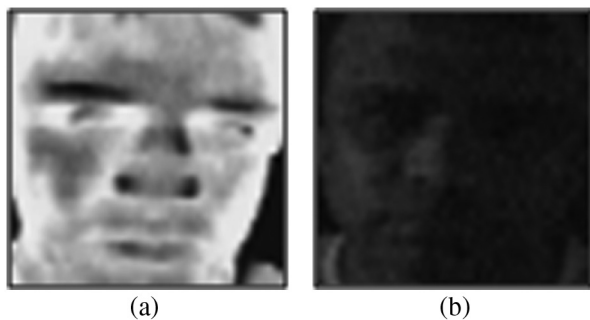
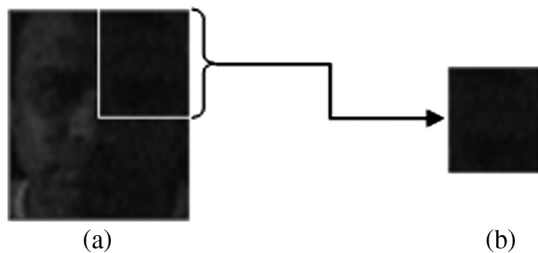
Fig. 3 (a) and (b) are the resized sample images of thermal and corresponding visual images, respectively; 50×50 pixels.

Fig. 4 (a) Detection of low illuminated eye region of the visual face; (b) Extracted low illuminated region of (a).

illuminated visual face image is shown, and in Fig. 5(b), the replaced low illuminated area of visual image is shown.

2.2 Image Fusion in Wavelet Domain

The Daubechies wavelet is used for decomposition and reconstruction of the images during image fusion technique. It belongs to the family of orthogonal wavelets defined as a discrete wavelet transform and is characterized by a maximal number of vanishing moments for some given support. It is more efficient in multi-resolution data fusion using multi-resolution analysis (MRA), which maintains the information of the original data preserving spectral information.¹³

2.2.1 Decomposition of visual and thermal image

The 2-D Daubechies wavelet transform (DWT) aims to decompose the image into approximation coefficient (cA)

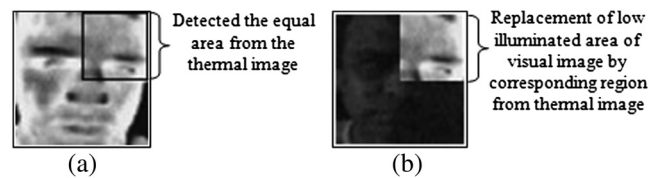


Fig. 5 (a) Extraction of equivalent area in thermal image, which appears dark in visual image; (b) Corresponding visual image after replacement of low illuminated area with the extracted part in (a).

and detailed coefficients: horizontal (cH), vertical (cV), and diagonal (cD), obtained by wavelet decomposition of the input image (x). The corresponding MATLAB instruction is given below:

$$[cA, cH, cV, cD] = \text{dwt2}(x, 'wname').$$

Here, 'wname' is the name of the wavelet used for decomposition and for Daubechies Wavelet coefficient 4 (db4) has been used as 'wname'. The 'dwt2' function performs a single-level, 2-D wavelet decomposition, with respect to the wavelet 'wname'. This kind of 2-D DWT leads to decomposition of approximation coefficients at level j in four components: the approximation coefficients (cA) at level $j + 1$, and the detail coefficients in three orientations: horizontal (cH), vertical (cV), and diagonal (cD).^{9,14,15}

2.2.2 Generating fused images

The main steps of the fusion algorithm are: (1) the two images, to be processed, need to be rescaled to the same size; (2) then both of them are decomposed into the sub-images using forward wavelet transform, which have the same resolution at the same levels and different resolution among different levels; and (3) information fusion is performed based on both the high and low-frequency sub-images of decomposed images; and finally the resulting fused image is obtained using inverse wavelet transform.

Let, $A(x, y)$ and $B(x, y)$ be the two images to be fused, let the decomposed low-frequency sub-images be $lA_j(x, y)$ and $lB_j(x, y)$, respectively, and the corresponding high-frequency sub-images be $hA_j^k(x, y)$ and $hB_j^k(x, y)$, respectively, where j is the parameter of resolution, and $j = 1, 2, \dots, J$. Also for every j , k takes the values 1, 2, and 3 representing the detail coefficients cH , cV , and cD , respectively. Then,

the fused high-frequency sub-images, $F_j^k(x, y)$, may be represented as:

If $hA_j^k(x, y) > hB_j^k(x, y)$, then $F_j^k(x, y) = hA_j^k(x, y)$

If $hA_j^k(x, y) \leq hB_j^k(x, y)$, then $F_j^k(x, y) = hB_j^k(x, y)$.

The fused low-frequency sub images, $F_j(x, y)$, may be defined as follows:

$$F_j(x, y) = k1 \cdot lA_j(x, y) + k2 \cdot lB_j(x, y). \quad (1)$$

In Eq. (1), $k1$ and $k2$ are given parameters, if image B is fused into A , then $k1 > k2$ and vice-versa. In this paper, the values of $k1$ and $k2$ have been picked using the 'max' fusion method. The 'max' method constructs a row vector containing the maximum element from each column of the approximation coefficient matrix. Let us consider two image samples, A and B , as follows:

$$A = \begin{bmatrix} 0.5002 & 0.1340 \\ 0.6897 & 0.0328 \end{bmatrix} \quad B = \begin{bmatrix} 0.4462 & 0.2793 \\ 0.7407 & 0.0051 \end{bmatrix}.$$

The element values for A and B range from zero to one; $\max(A) = [0.6897 \ 0.1340]$ and $\max(B) = [0.7407 \ 0.2793]$. So, when B is fused into A , $k1 = 0.6897$ and $k2 = (1 - 0.6897) = 0.3103$, and when A is fused into B , $k2 = 0.7407$ and $k1 = (1 - 0.7407) = 0.2593$. Now, $F_j(x, y)$ and $F_j^k(x, y)$ are used to reconstruct and generate the fused image $F'(x, y)$ which contains high-frequency and low-frequency information of $A(x, y)$ and $B(x, y)$.¹⁶⁻¹⁹

2.2.3 Reconstruction of decomposed fused image

After generating the decomposed components of fused image, the inverse Daubechies Wavelet (idwt2) is applied to generate synthesized fused image. Repetition of the decomposition scheme increases the concentration of the approximation image in the low frequency energy. The single level wavelet reconstruction is used here to generate the synthesized fused images and the corresponding MATLAB instruction is given below:

$$F = \text{idwt2}(cA, cH, cV, cD, 'db4').$$

Here, idwt2 uses the inverse Daubechies wavelet transform (db4) to compute the single-level reconstruction of fused image F , based on approximation coefficient matrix (cA) and detail coefficient matrices, i.e., horizontal (cH), vertical, (cV) and diagonal (cD), respectively. Finally, the reconstructed image is used as input to feature extraction algorithm for dimensionality reduction. The steps for fusion of thermal and visual image are represented in Fig. 6.

2.3 Dimensionality Reduction by Independent Component Analysis in Natural Logarithm Domain

Face images are very similar, and therefore, highly correlated. It follows that they can be represented in a much lower dimensional feature subspace. For that, independent component analysis (ICA) is implemented over logarithm domain.

ICA is used in many applications like feature extraction, dimensionality reduction, data analysis, and source separation etc. Here, ICA is implemented for dimensionality reduction on face images. It is a technique that is mainly used for subspace projection using projection from high dimensional to low-dimensional space.²⁰⁻²³ ICA is the generalization of the principal component analysis (PCA), which de-correlates the high-order statistics in addition to second-order moments.

To rigorously define ICA, we can use a statistical "latent variables" model.²⁴ Let us assume that, we observe n linear mixtures of n independent components (ICs).

$$x_j = a_{j1}s_1 + a_{j2}s_2 + \dots + a_{jn}s_n \quad \text{for all } j. \quad (2)$$

Instead of summation, it is suitable to use vector-matrix notation in Eq. (2). Let x be a matrix with random vectors x_1, \dots, x_n , as its columns and s be another matrix consisting of random vectors s_1, \dots, s_n as its columns. Let A be a matrix with elements a_{ij} . All vectors are considered here as column vectors; thus, the transpose of x_i , i.e., x_i^T , is a row vector. Using this vector-matrix notation, the above mixing model is written as,

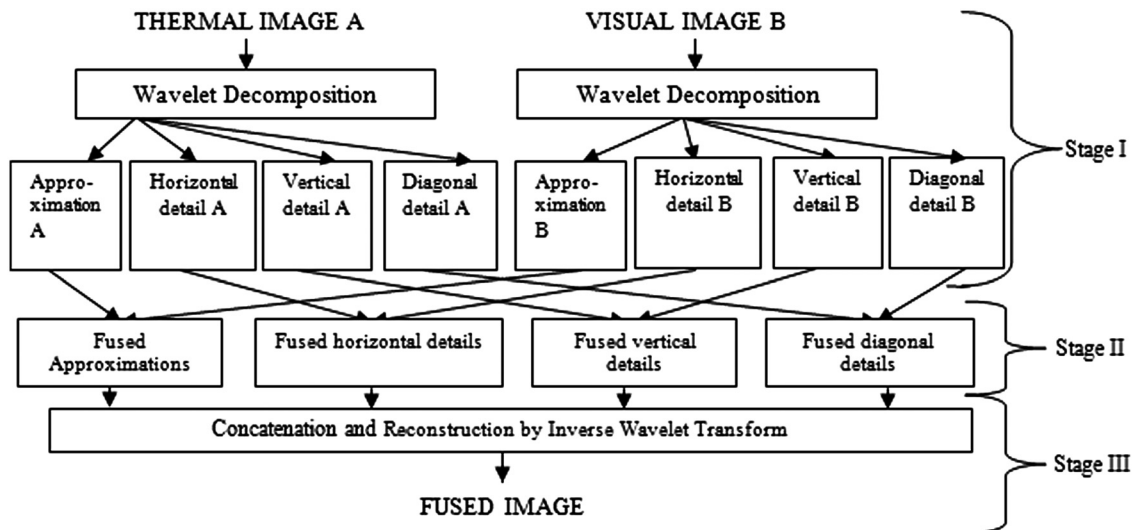


Fig. 6 Block diagram showing the steps for fusion of thermal and visual images.

$$x = As. \quad (3)$$

The statistical model in Eq. (3) is called independent component analysis, or ICA model.

2.3.1 Algorithmic steps for ICA in logarithm domain

With the reference to Fig. 7, the steps followed for dimensionality reduction using ICA are given below:

- Step 1: Represent each fused image as column vector by concatenating the rows of the pixel values in order. Thus, an image of size $M \times N$ pixel will be represented as a column vector with $M \times N$ elements.
- Step 2: Calculate the mean vector of all such column vectors representing fused images.
- Step 3: Subtract this mean vector from each of the column vectors to generate the centered image (I_c).
- Step 4: Apply natural logarithm on each element of the centered image I_c and convert the image to logarithm domain I_{\log} .
- Step 5: Apply whitening on the log-centered image.
- Step 6: Apply further preprocessing steps and implement the FastICA algorithm for dimensionality reduction.

2.3.2 Implementation of ICA in logarithm domain by FastICA

There are different popular ICA algorithms like FastICA,^{20,21} Infomax,²⁵ Common's algorithm,²⁴ Kernel ICA,²⁶ etc. The efficiency of ICA algorithm on a particular dataset may be highly dependent on some of the application dependent pre-processing steps. The FastICA algorithm has many advantages over other ICA algorithms.²¹ When compared with the other existing methods of ICA, the FastICA algorithm and the underlying contrast functions contain a number of desirable properties. It directly calculates ICs. The ICs can be estimated one after another, but it is almost equivalent of doing projection pursuit. This is helpful in investigative data analysis, and reduces the computational load of the system in cases where the need of estimation is restricted to only some of the ICs. The FastICA has most of the advantages of neural algorithms: It is parallel, distributed, and computationally simple and requires little memory space.

Stage 1 (Centering). The first and the most basic as well as essential pre-processing stage is centering the matrix x ,

which consists of fused images represented as its columns. It requires subtracting the mean vector, $m = E\{x\}$, from each column of x , which makes x a zero-mean variable. The elements of the mean vector are computed as the mean of the rows of x matrix. This pre-processing stage is only to make the ICA algorithm simpler. After the estimation of the mixing matrix A with centered data, the mean vector of s is added with the centered estimates of s . The mean vector of s is given by $A^{-1}m$, where m is the subtracted mean during pre-processing.

$$x = \{x_{ij}\} \quad i = 1, 2 \dots q \quad \text{and} \quad j = 1, 2 \dots p, \quad (4)$$

where, i = Row number and j = Column number

Let us consider, x is a matrix of dimension $q \times p$, where $q = p = 3$

$$x = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}_{q \times p}.$$

Mean of x is given as:

$$m_i = \frac{1}{p} \sum_{j=1}^p a_{ij} \quad i = 1, 2 \dots q. \quad (5)$$

Thus, for the above matrix,

$$m = \begin{bmatrix} \frac{(a_{11} + a_{12} + a_{13})}{3} \\ \frac{(a_{21} + a_{22} + a_{23})}{3} \\ \frac{(a_{31} + a_{32} + a_{33})}{3} \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix}$$

$$x_{\text{zeromean}} = x - m = x - E\{x\} = \{a_{ij} - m_i\}. \quad (6)$$

$$i = 1, 2 \dots q; \quad \text{and} \quad j = 1, 2 \dots p$$

$$x_{\text{zeromean}} = \begin{bmatrix} \pm v_{11} & \pm v_{12} & \pm v_{13} \\ \pm v_{21} & \pm v_{22} & \pm v_{23} \\ \pm v_{31} & \pm v_{32} & \pm v_{33} \end{bmatrix}$$

$$= \begin{bmatrix} (a_{11} - k_1) & (a_{12} - k_1) & (a_{13} - k_1) \\ (a_{21} - k_2) & (a_{22} - k_2) & (a_{23} - k_2) \\ (a_{31} - k_3) & (a_{32} - k_3) & (a_{33} - k_3) \end{bmatrix}. \quad (7)$$

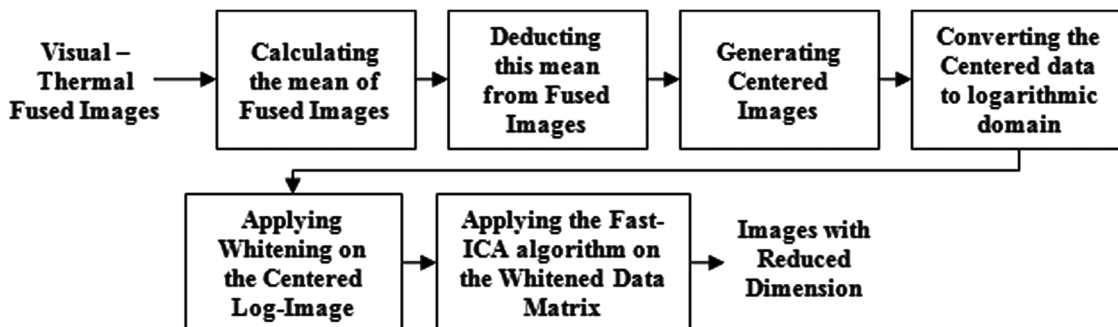


Fig. 7 Block diagram describing the algorithmic steps for independent component analysis (ICA) in logarithm domain.

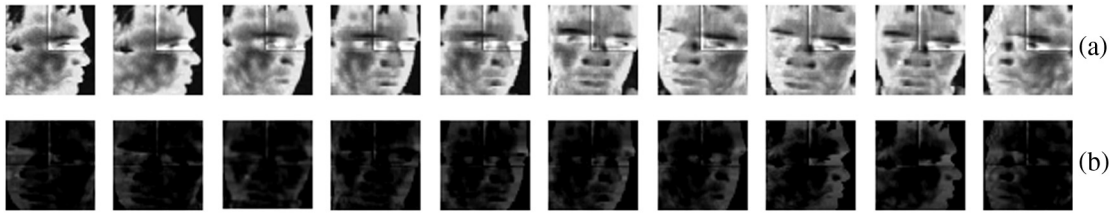


Fig. 8 (a) Eye region replaced fused images; (b) corresponding centered images of (a).

The main advantage of centering is by de-correlating the input vectors, the quantity of redundant information can be reduced. The high dimensional and correlated input vectors can be represented in a lower dimensional space and de-correlated. In Fig. 8, the sample image of a low illuminated eye region replaced fused face images with frontal, mid-left, mid-right, left, and right views along with their corresponding centered image are presented.

Stage 2 (Log-Centered Data). Before employing the algorithm of whitening on the centred data, we first convert the centred image into a log-domain by applying natural logarithm operator. We convert the zero mean variables into logarithm domain.

$$\log(x) = \log(x_{\text{zeromean}}). \quad (8)$$

We know that, the values of zeros and negative numbers are undefined in logarithm domain. After subtracting the mean value from the pixel values, some of these may appear as negative or zero. Therefore, we have to replace the negative values and zeros after subtracting mean values from the original signal. In the present work, all images are grayscale images with pixel values ranging from zero to 255. So, before converting in logarithm domain, the pixel values are represented in the range zero to one by dividing the centered pixel values by 255. Then all the negative values and zero values are replaced by $1/255$ i.e., 0.00392.

After applying centering in fused image, we get two matrices: centered matrix (E_v) and diagonal matrix (E_d). Then, we apply natural logarithm operator on E_v matrix. Observations show that the resultant, log-centered matrix generates more negative pixel values. This fact can be verified from the given example. At first, we have randomly picked some pixels from the centered matrix of 50×50 fused images. Some sample pixel values of centered matrix after replacement (E_v) and log-centered matrix [$\log(E_v)$] are shown below:

$$E_v = \begin{pmatrix} A11 & A12 & A13 & A14 \\ 2.371 \times 10^{-2} & 8.830 \times 10^{-2} & 3.92 \times 10^{-3} & 3.92 \times 10^{-3} \end{pmatrix}$$

$$\text{Log}(E_v) = \begin{pmatrix} A11 & A12 & A13 & A14 \\ -3.74 & -2.43 & -5.54 & -5.54 \end{pmatrix}.$$

Natural logarithm operator enhances the low intensity pixel values, while reducing the actual intensity values into a relatively small pixel range.²⁷ We can call this process of reduction of the pixel values as logarithmic reduction of

pixel matrix. In Fig. 9, some samples of log-centered images are presented.

The main motive of using logarithm is that, in the process of logarithmic reduction of pixel values, original information is almost retained, and there is less chance of information loss.

Stage 3 (Whitening). In the second part of dimensionality reduction using ICA, the task is to find the whitening matrix of the observed matrix. This implies that after centering and before applying ICA algorithm, the observed vector x is transformed linearly, so that a new vector \tilde{x} can be obtained which is white, i.e., the vector contains uncorrelated components, and their variances are equal to unity. In other words, the covariance matrix of \tilde{x} equals to the identity matrix:

$$E\{\tilde{x}\tilde{x}^T\} = I. \quad (9)$$

To calculate the whitening matrix, we use the Eigen values decomposition (EVD) method of the covariance matrix $E\{xx^T\} = EDE^T$, where E is the orthogonal matrix of eigenvectors of $E\{xx^T\}$ and D is the diagonal matrix of its Eigen values, $D = \text{diag}(d_1, \dots, d_n)$. The main advantage of whitening is to reduce the number of parameters to be estimated. Rather than estimating the n^2 parameters, those are the elements of the original matrix (A), only the new orthogonal matrix \tilde{A} is estimated. An orthogonal matrix contains $n(n-1)/2$ degrees of freedom.^{20,21} Thus, half of the problems of ICA are solved by using whitening. As whitening is a standard method and much easier in comparison to other ICA algorithms, it is a good idea to reduce the complexity of the problem using whitening. As we do the whitening, it may also be helpful to decrease the data dimension simultaneously. Often, this has the effect of reducing noise as well.^{20,21}

Algorithmic steps for whitening:

- Step 1: Load log-centered image as a new Vector and original image.
- Step 2: Calculate co-variance for dimensionality reduction by applying PCA.
- Step 3: Calculate the Eigen values and Eigen vectors of the covariance matrix.
- Step 4: Calculate the whitening and de-whitening matrices. These matrices handle dimensionality simultaneously.

Stage 4 (Further Pre-Processing). Some application-dependent, pre-processing steps are needed to be performed to achieve success in ICA for a given dataset. Some band-pass filtering may be very useful, if the data consists of time signals. If we filter linearly the observed signals $x_i(t)$ to



Fig. 9 Sample of centered images in logarithm domain.

Table 1 Database statistics.

Total no. of classes	Image resolution	Image type	No. of expression with types	Total no. of illumination types
28 Individuals	320 × 240	RGB color	Ex1 (Surprise), Exp2 (laughing), Exp3 (Anger)	left light on, right light on, both lights on, dark room, left and right lights off with varying poses like left, right, mid, mid-left, mid-right.

acquire new signals, say $x_i^*(t)$, the ICA model still holds for $X_i^*(t)$, with the same mixing matrix. Now, time filtering of X corresponds to multiplying X from the right by a matrix M . Here, X is a matrix that contains the observations $X(1), \dots, X(T)$ as its columns, and similarly for S .

$$X^* = XM = ASM = AS^*. \quad (10)$$

Equation (10) shows that the ICA model still remains valid.²³ Data can be further pre-processed using different algorithms like FastICA and KernelICA.²⁶ We have implemented fixed-point algorithm for FastICA.

2.4 ANN using Back Propagation with Momentum

Back-propagation²⁸ is a popular scheme for training feed-forward networks. Here, the network consists of a directed acyclic graph which gives the network structure with the activation functions at each unit or node, which relate inputs to activation outputs. A directed acyclic graph is one containing no directed cycles, so the function computed by the network is not computed using any fixed point equations. The hyperbolic tangent sigmoid transfer function has been used to learn the multilayer perceptron network. Inside this transfer function, different weights can be set like input weights and layer weights. In this case, the learning has been performed with the initial weights of zero for both input weights and layer weights. The network deals with real numbers internally, although the inputs may be discrete and represented as -1 and 1 . Denote the input variables to the network as a vector of values x and denote the response variable, which the network is intended to predict as y . In the case of regression, for the given input variables the network output corresponds to the predicted regression for y . This corresponds to the expected value or mean of the real valued variable y conditioned on the values of the input variables, z . In the case of one-of- C classification, the network outputs a conditional probability distribution over C possible values for the discrete variable y , conditioned on the values of the input variables. The output comes from n nodes and corresponds to a vector of real values summing to 1. The i 'th value is the estimated conditional probability that the output variable should have the i 'th discrete value. A multilayer perceptron (MLP) has many hidden layers, so it takes much time to train input layers. The main motive to use the MLP is that, the classes

can be separated via hyper-planes, and depending upon the size of the input data, the number of layers can be changed.

3 Experiments and Discussion

ICA has been implemented in logarithm domain. This ICA algorithm is applied separately on the training and testing images. To evaluate the efficiency of the proposed system, three different data sets were prepared for training and testing of low illuminated faces. For that, the IRIS thermal/visual¹² face database is used, and the face images are of varying expressions, illuminations, and 11 different poses.

3.1 Image Database Description

The IRIS¹² database is one of the most popular thermal/visual face databases, which has been used to demonstrate the effectiveness of the proposed log-ICA algorithm. This is the only freely available thermal/visual face database. Another thermal/visual database is the Equinox database but is not freely available anymore. The IRIS database contains simultaneously acquired unregistered thermal and visible face images under variable illuminations, three different expressions, and 11 pose variations. A detail about the database is given in Table 1. We know there are six human facial expressions like happiness, sadness, fear, anger, surprise, and disguise; however, in this database only happiness, anger, and surprise are available. Therefore, we have not used sadness, fear, and disguise in our experiment.

3.2 Training and Testing using the Proposed Algorithm

For the training set, 10 different classes of dark (low illuminated) faces are picked from the IRIS database; although, this database contains a total of 28 different classes. The main motive of this work is to present a face recognition system for dark human faces by image fusion. In the IRIS database, there are only 10 classes available with dark faces. In these 10 different classes, a total of eight different image sets are available; three sets are based on expression variation, and the other five are based on illumination variation. Among these eight different sets, a single image set of dark faces is taken for training purpose. These dark faces have visual and corresponding thermal faces and these two image sets are used to generate the fused images. Steps for the generation of fused images have already been illustrated in Sec. 2.2. Finally, these three sets of

images, i.e., visual image set, corresponding thermal image set, and the set of fused images, are used for conducting three different experiments.

In case of testing, three other image sets are picked from the expression sets, and these are used at the time of testing in their individuals experiments; i.e., in case of a thermal image experiment, a thermal image set of expression is used to find recognition rates against dark thermal faces. The other two image sets, i.e., visual and fused, are used in their corresponding experiments. Details about image sets are given in Table 2.

3.2.1 Training phase

Three different image sets are prepared for conducting three separate experiments during the training phase. Each set of data contains 110 images of 10 different classes, with each class containing 11 images. In the database, the size of sample images is 320×240 pixels. We have converted the images to 50×50 pixels to remove the background and to maintain a fixed size for all images, and for saving the processing time. The images have been resized using the 'bicubic' interpolation method, where the output pixel value is the weighted average of the nearest 4×4 neighboring pixels values. After resizing, selection of the low illumination area from visual image and replacement of the same with the corresponding area of thermal image is done. After

that, 110 face-eye regions replaced visual images, and the corresponding thermal images are fused together to produce another dataset consisting of 110 fused face images for the third experiment. In Fig. 10, the sample images of dark visual, and thermal face images along with their corresponding eye-region replaced fused images are shown.

After completion of the training set, the FastICA algorithm is used in the logarithm domain for dimensionality reduction on all the training image sets separately, by taking one image at a time. After that, those images of reduced dimension are kept into a single data matrix, i.e., each column represents an image and passed to the network for learning, which took 2 h and 15 min each, for three different experiments.

3.2.2 Testing phase

Three different image sets of expression have been used for testing purpose as compared to dark faces of visual, thermal, and fused images. Each testing dataset, i.e., thermal, visual, and corresponding fused image set, contained 110 images of 10 classes. Some sample images of testing are shown in Fig. 11.

To assess the ability of the neural network, 11 unlearned images with five different poses having 20 deg rotation is given to the network. The evaluation processes were run separately for each class contained in each testing set. The evaluation steps are as follows. First, the log-ICA is

Table 2 Details of training and testing sets used for different experiments.

No. of training images			Testing sets (based on expression)			
Experiment 1(a)	Experiment 2(b)	Experiment 3(c)	Number of training and testing classes	Set-1 (surprise)	Set-2 (laughing)	Set-3 (anger)
110 dark visual faces (11 images per class) images	Corresponding 110 dark thermal faces (11 images per class) images	Face region extracted 110 fused faces of (a) and (b)	10	Testing images are not used in training		
				Total 110 images of 10 different classes	Total 110 images of 10 different classes	Total 110 images of 10 different classes



Fig. 10 (a) Sample of visual dark images; (b) corresponding eye region replaced visual images of (a); (c) corresponding thermal images of (a); (d) fused images of (b) and (c).



Fig. 11 Three different testing sets with varying expressions.

Table 3 Recognition rates of the three expression sets.

Testing set	Set 1 (exp-1) (surprise) (%)	Set 2 (exp-2) (laughing) (%)	Set 3 (exp-3) (anger) (%)
Recognition rates on dark visual image (Experiment 1)	79.57	81.56	81.66
Recognition rates on dark thermal image (Experiment 2)	81.86	82.83	84.84
Recognition rates on face region extracted fused image (Experiment 3)	100	100	100

Table 4 Recognition rates of proposed and existing techniques.

Name of the techniques	Recognition rates (%)
Proposed method	100
Pixel fusion + MLP ²⁹	95.07
Pixel fusion + RBF ²⁹	96.00
Optimum fusion ³⁰	93.00
Wavelet fusion using Haar ³¹	87.00
Wavelet fusion using Daubechies ³¹	91.50
CAQI-other ¹⁷	93.00
CAQI-same ¹⁷	94.00
SQI ¹⁷	70.00

applied to the 11 images. After that, 11 de-correlated images are plotted into a single data matrix where a column of the matrix represents the images. Then, the recognition process is started to match with earlier learned face data matrix using a feed forward neural network. This process is repeated for three different testing sets separately in three different experiments.

3.3 Discussion of Experimental Results

Firstly, an experiment using visual images was conducted. At first, a surprise dataset (Exp-1) was tested against the visual image dataset of dark faces. During the recognition process, 100% recognition rate is achieved for one class only; however, in the case of the laughing/happiness (Exp-2) data, a 100% recognition rate is achieved for two different classes: class-3 and class-9. Finally, the anger dataset (Exp-3) had a better result than the other two expression datasets, and average recognition rate is 81.66%, which is 2% more than the surprise dataset. Nevertheless, in case of thermal image datasets, the recognition process showed better performance over visual images. The surprise dataset of thermal images showed 2% improvement over the corresponding visual image dataset, and for laughing and anger data, the recognition process showed 1% and 3% improvements, respectively, over the visual image dataset. Finally, experiment on fusion image sets has been conducted and all the three expression sets recognized successfully. In Table 3, the recognition rates using three different testing sets on varying expression on the normal lighting condition is presented. From that, it can be justified that, the three different facial expressions are recognized successfully against a single training image set with dark lighting condition by proposed system.

3.4 Comparison of the Present Work with Other Fusion Techniques

The performance of this eye-region-replaced fused image system has been compared with other existing techniques of fused images. In Ref. 29, a pixel fusion of optical and thermal face images was proposed. The fused faces were first projected into Eigen faces, and finally two different classifiers, MLP and RBF, were used separately for classifications. The experimental results show 96% and 95.07% success rates for RBF and MLP classifiers, respectively. In Ref. 30, an optimum level of fusion was proposed to find a fused image of visual and thermal images. In Ref. 30, four different levels of image fusion, i.e., 70% visual and 30% thermal; 60% visual and 40% thermal; 50% visual and 50% thermal; and 40% visual and 60% thermal, have been implemented, and the highest recognition rate obtained

is 93%. In Ref. 31, a comparative study of image fusion in the wavelet domain was presented. Two different wavelets, Haar (db1) and Daubechies (db2), were used to generate the wavelet coefficients of thermal and visual images, and finally, inverse wavelet transform generates the fused images; this system shows 87% and 91.5% recognition rates. So, in comparison to all of these techniques, the proposed system, i.e., eye-region-extracted fused image of dark faces, shows better performance. In Table 4, a comparison between the proposed and existing techniques is given.

4 Conclusions and Future Work

In this paper, a novel approach is presented to recognize the dark human faces. For this study, an IRIS thermal/visible face database is used. To recognize all these faces, ICA was implemented into a logarithm domain (Log-ICA) with a feed forward network. The experimental results show the proposed method can perform the face recognition task for face images with different expressions and pose variations, and the method achieves a 100% success rate. So, it can be said that face images, captured under dark lighting conditions and kept in the gallery images, are able to easily recognize the probe images using the proposed system, even if the probe images were captured in any condition like varying expressions with different lighting conditions. In the future, we plan to study more facial expressions like fear, disguise, etc. using thermal face images. We are also planning to develop a thermal face database which will contain different facial expressions with varying illumination, including different facial action units (AUs), and a real world face recognition system which can recognize different facial AUs like inner brow raised, outer brow raised, brow lowered, upper lid raised, cheek raised, lid lightener, nose wrinkle, lip corner pulled, lip corner depressed, chin raised, mouth stretch, etc. with low illumination conditions.

Acknowledgments

The research was supported by the grant from DIT, MCIT, Government of India, Vide No. 12(2)/2011-ESD, dated March 29, 2011. The first author would also like to thank Barin Kumar De, Dean of Science of Tripura University (A Central University) and Dr. Niharika Nath, PhD of New York Institute of Technology, New York, for their kind support to carry out this research work.

References

1. F. Prokoshi, "History, current status and future of infrared identification," in *Proc. IEEE Workshop Comput. Vis. Beyond Visible Spectrum: Methods Appl.*, pp. 5–14, IEEE Computer Society, Hilton Head, SC (2000).
2. A. Gupta and S. K. Majumdar, "Machine recognition of human face," 2008.
3. B. Fasel and J. Luetten, "Automatic facial expression analysis: a survey," *Pattern Recogn.* **36**(1), 259–275 (2003).
4. J. C. McCall and M. M. Trivedi, "Pose invariant affect analysis using thin-plate splines," in *Proc. 17th Int. Conf. Pattern Recogn. ICPR*, Vol. 3, pp. 958–964, IEEE Computer Society, Cambridge, UK (2004).
5. Y. Kun, Z. Hong, and P. Ying-jie, "Human face detection based on SOFM neural network," in *IEEE Int. Conf. Info. Acquisition*, pp. 1253–1257, IEEE Computer Society, Weihai (2006).
6. Y. Adachi et al., "Extraction of face region by using characteristics of color space and detection of face direction through an Eigenspace," in *4th Int. Conf. Knowledge-Based Intelligent Engineering Systems and Allied Technologies*, Vol. 1, pp. 393–396, IEEE, Brighton (2000).
7. G. Bebis et al., "Face recognition by fusing thermal infrared and visible imagery," *Image Vis. Comput.* **24**(7), 727–742 (2006).

8. M. Hanif and U. Ali, "Optimized visual and thermal image fusion for efficient face recognition," in *IEEE Conf. Info. Fusion, COMSATS Inst. Inf. Technol.*, Abbottabad, pp. 1–6 (2006).
9. Y. Z. Goh, A. B. J. Teoh, and M. K. O. Gog, "Wavelet based illumination invariant preprocessing in face recognition," in *Proc. Int. Congress Image Signal Process*, Vol. 3, pp. 421–425, IEEE Computer Society, Sanya, China (2008).
10. H. K. Ekenel and B. Sankur, "Multiresolution face recognition," *Image Vis. Comput.* **23**(5), 469–477 (2005).
11. Y. Cheng, Z. Jin, and C. Hao, "Illumination normalization based on 2D Gaussian illumination model," in *3rd Int. Conf. Advanced Computer Theory and Engineering*, Vol. 3, pp. 451–455, IEEE, Chengdu (2010).
12. OTCBVS Benchmark Dataset Collection, "Dataset 02: IRIS Thermal/Visible Face Database," IEEE OTCBVS WS Series Bench, <http://www.cse.ohio-state.edu/otcbvs-bench/Data/02/download.html> (2009).
13. K. A. Wahid et al., "An algebraic integer based encoding scheme for implementing daubechies discrete wavelet transforms," in *36th Asilomar Conf. Signals, Systems and Computers*, Vol. 1, pp. 967–971, IEEE, Pacific Grove, CA (2002).
14. I. Daubechies, "Ten lectures on wavelets," in *CBMS-NSF Conf. Ser. Appl. Math. SIAM*, Vol. 61, SIAM, Philadelphia, PA (1992).
15. Y. Meyer, *Wavelets and Operators*, Cambridge University Press, New York, NY (1993).
16. S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern. Anal. Mach. Intell.* **11**(7), 674–693 (1989).
17. C. K. Chui, *An Introduction to Wavelets*, Academic Press Professional, Inc., San Diego, CA (1992).
18. I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.* **41**(7), 909–996 (1988).
19. S. Vekkot and P. Shukla, "A novel architecture for wavelet based image fusion," *World Acad. Sci. Eng. Technol.* **57**, 372–377 (2009).
20. A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, Canada (2001).
21. A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Network* **10**(3), 626–634 (1999).
22. M. S. Barlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Trans. Neural Network* **13**(6), 1450–1464 (2002).
23. A. Hyvarinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Network* **13**(4–5), 411–430 (2000).
24. P. Comon, "Independent component analysis: a new concept," *J. Signal Process.* **36**(3), 287–314 (1994).
25. T. W. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended infomax algorithm for mixed sub-gaussian and super-gaussian sources," *J. Neural Comput.* **11**(2), 417–441 (1999).
26. F. R. Bach and M. I. Jordan, "Kernel independent component analysis," *J. Mach. Learn. Res.* **3**(1), 1–48 (2002).
27. R. Fisher et al., "Logarithm operator," <http://homepages.inf.ed.ac.uk/rbf/HIPR2/pixlog.htm> (2003).
28. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation, parallel distributed processing: explorations in the microstructure of cognition," *MIT Press Computational Models Cognition Perception Series* Vol. 1, pp. 318–362, MIT Press, Cambridge, MA (1986).
29. M. K. Bhowmik et al., "Image pixel fusion for human face recognition," *Int. J. Recent Trends Eng.* **2**, 258–262 (2009).
30. M. K. Bhowmik et al., "Optimum fusion of visual and thermal face images for recognition," in *Proc. 6th Int. Conf. Info. Assurance and Security*, pp. 311–316, IEEE Intelligent Transportation Systems Society, Atlanta, GA (2010).
31. M. K. Bhowmik et al., "Fusion of wavelet coefficients from visual and thermal face images for human face recognition—a comparative study," *Int. J. Image Process.* **4**(1), 12–23 (2010).



Mrinal Kanti Bhowmik received his BE (in CSE) and MTech (in CSE) degrees from Tripura University (A Central University), Suryamaninagar, Tripura, India, in 2004 and 2007, respectively. Since July 2010, he has been working as an assistant professor at Tripura University. He is also pursuing his PhD (Engg.) degree from Jadavpur University, Kolkata, India. His research interests are related to the field of biometric, artificial neural network, information security, etc. He has completed a project on human face recognition as a chief investigator in 2010, and has been involved in another project based on face identification since March, 2011. Both projects are funded by the DIT, MCIT, India. He is also a member of the IEEE (USA).



Debotosh Bhattacharjee received the MCSE and PhD (Engg.) degrees from Jadavpur University, India, in 1997 and 2004, respectively. He was associated with different institutes in various capacities until March 2007. After that, he joined his Alma Mater, Jadavpur University. His research interests pertain to the applications of computational intelligence techniques like fuzzy logic, artificial neural network, genetic algorithm, rough set theory, cellular automata,

etc. in face recognition, Optical Character Recognition (OCR), and information security. He is a life member of Indian Society for Technical Education (ISTE, New Delhi), Indian Unit for Pattern Recognition and Artificial Intelligence (IUPRAI), and senior member of IEEE (USA).



Dipak Kumar Basu received his BETel, METel, and PhD (Engg.) degrees from Jadavpur University (J.U.), in 1964, 1966, and 1969, respectively. He was a faculty member of J.U. from 1968 to 2008. He was an A.I.C.T.E. Emeritus Fellow at the CSE Department of J.U. from 2008 to 2011. His current fields of research interest include pattern recognition, image processing, and multimedia systems. He is a senior member of the IEEE (USA), a Fellow of I.E. (India),

and W.B.A.S.T. (India), and a former fellow of the Alexander von Humboldt Foundation (Germany).



Mita Nasipuri received her BETel, METelE., and PhD (Engg.) degrees from Jadavpur University (J.U.), in 1979, 1981, and 1990, respectively. Nasipuri has been a faculty member of J.U. since 1987. Her current research interest includes image processing, pattern recognition, and multimedia systems. She is a senior member of the IEEE (USA), a 32 Fellow of I.E. (India), and W.B.A.S.T. (India).