Facial Expression Invariant Person Recognition using Feature Level Fusion of Visual and Thermal Images

Mrinal Kanti Bhowmik

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

Debotosh Bhattacharjee, Dipak Kumar Basu*,
Mita Nasipuri

Department of Computer Science and Engineering
Jadavpur University
Kolkata - 700032, India
*Former Professor
E-mail: debotosh@indiatimes.com,
dipakkbasu@gmail.com, mitanasipuri@yahoo.com

Abstract—In this paper feature level fusion scheme of visual and thermal images in wavelet domain is being analysed. Here, Daubechies wavelet transform, termed as db4, coefficients from visual and corresponding coefficients computed in the same manner from thermal images are combined to get fused coefficients. After decomposition up to fifth level (Level 5) fusion of coefficients is done. At the time of generating fused images of coefficients, two different approaches (absolute maximum and absolute minimum) have been followed. Inverse Daubechies wavelet transform of those coefficients gives us fused face images. The main advantage of using wavelet transform is that it is well suited to manage different image resolutions and after decomposition, the high and low frequency sub-band images can be separated. Therefore, the reconstructed fused images have the information of low and high frequencies. These fused images have been passed logarithmic domain (log-ICA) for reduction of dimensions, which is suitable for face images varying with expressions. Finally, those reduced fused images are classified using a multi-layer perceptron. For experiments, IRIS Thermal/Visual Face Database was used. Experimental results show that the performance of the approach presented here gives an average success rate of 94.81%.

Keywords-fused image; daubechies wavelet transform; principal component analysis (PCA); independent component analysis in logarithmic domain (Log-ICA); multi-layer perceptron; classification; IRIS database

I. INTRODUCTION

The researchers for face recognition have developed a large number of methods. However, these methods give poor performance under variable illumination, pose and facial expression. Visual image-based face recognition has shown difficulties in performing consistently under uncontrolled operating conditions. Under the condition of uncontrolled illumination face, recognition accuracy degrades significantly. The use of thermal infrared (IR) images can improve the performance of face recognition under uncontrolled illumination conditions [16], [17], [18]. Thermal infrared (IR) images have several advantages over conventional visible images. Most of the light in the mid-to-long wave IR is emitted rather than reflected [19], [20]. Thus, thermal imaging can work well under low illumination conditions or even in complete darkness, where visual face recognition techniques fail. However, face recognition using thermal images includes different challenging problems like; thermal images change when temperature varies due to physical exercise or ambient temperature. IR imagery has other limitations, as it is obscure to glass. Recently the analysis of facial expressions in computer vision systems is an important research area for many applications. A facial expression recognition system includes several difficulties in different areas like, lighting conditions, variations in the expressions, head pose and head movement etc [14], [15]. Lighting variations along with pose, exaggerates the complicacy, as expressions can greatly change the way in which the face is lightened and it becomes severe when there are shadows. This creates problems by producing shadows that shift and unclear important features useful in identifying facial expressions. The fusion of visual and thermal IR images is presented to enhance the robustness of face recognition. Two different information fusion schemes have been used to generate fused images for the two different methods discussed in the section II. In method-I, fused images have been generated by taking maximum level for approximate coefficient and minimum level for details coefficients and in case of method-II, maximum level for both the frequencies subbands have been taken into account. A new approach is introduced here to find a solution of the problem like expression in faces and independent component analysis in logarithmic domain, called log-ICA, for this purpose. Again, the dimension of the fused images was reduced using method-I (i.e. Principal Component Analysis (PCA)) in [6] and method-II (Independent Component Analysis in logarithm domain, log-ICA). Finally, Multi-layer Perceptron has been used as a classifier for both the methods.

The paper is organized as follows: system overview has been given in Section II, which includes the description of information fusion in wavelet domain, different algorithms used for dimension reduction, difference between ICA and ICA in logarithm domain (log-ICA) and Artificial Neural Network. Section III shows and analyses the experimental results using IRIS database. Finally, the conclusion is made in section IV.

II. SYSTEM OVERVIEW

The complete system implementation in this work is described with a block diagram, which is shown in Fig. 1. At first, 2-D discrete wavelet transform of thermal and visual images has been completed at level five. After that, the fusion of coefficients (i.e. approximate and details) of thermal and visual images have been generated and for this two different fusion methods have been used.

A. Feature Level Fusion in wavelet domain

Wavelet transformations are multi-resolution image decomposition tool that provides a variety of channels representing the image feature by different frequency sub-bands at multi-scale.

It is a famous technique for analyzing signals. When decomposition is performed, the approximation and detail component can be separated [13]. The Daubechies wavelet (db4) decomposition has been done upto level 5 for image fusion, which is the highest level of decomposition. The higher the decomposition level, the more is the number of subbands, which ensures that more fusion of subbands will be generated [13]. After the single level reconstruction of subband fused images, the generated synthesized fused image is of higher quality than the other decomposition levels.

1) Decomposition of thermal and visual images:
In this technique, first decomposition of the thermal and visual images has been done by 2-D discrete wavelet transformation $(dwt2)$ up to 5th level using Daubechies Wavelet (db4). Both the thermal and visual images are in the pixel resolution of 40 x 50 within same pose, illumination and expression.

Daubechies wavelets are a family of orthogonal wavelets defining a discrete wavelet transformation and characterized by a maximal number of vanishing moments for some given support. This kind of wavelet transform aims to decompose the image into approximation coefficients (cA) and detailed coefficients cH (horizontal), cV (vertical) and cD (diagonal) obtained by wavelet decomposition of the input image.

2) Generating Fused Images of different coefficients:
In this paper, the process of fusion algorithm is as follows: (a)

two images to be processed and resampled with the same dimension; and (b) they are respectively decomposed into the sub-images using forward wavelet transform $(dwt2)$, which have the same resolution at the same levels; and (c) information fusion is performed based on the high-frequency sub-images of decomposed images.

Let $A(x, y)$ and $B(x, y)$ be the images to be fused, the decomposed low frequency sub-images of $A(x, y)$ and $B(x, y)$ be respectively $lA_J(x, y)$ and $lB_J(x, y)$ (J is the parameter of resolution). The decomposed high frequency sub-images of $A(x, y)$ and $B(x, y)$ be respectively $hA_j^k(x, y)$ and $hB_j^k(x, y)$ (j is the parameter of resolution and $j = 1, 2, \ldots, J$ for every j, k=1, 2, 3, 4, 5). Then the fused high-frequency sub-images $F_j^k(x, y)$ are if $hA_j^k(x, y) > hB_j^k(x, y)$ (absolute maximum) then $F_j^k(x, y) = hA_j^k(x, y)$ and if $hA_j^k(x, y) < hB_j^k(x, y)$ (absolute minimum) then $F_j^k(x, y) = hB_j^k(x, y)$; and the fused low-frequency sub images $F_j(x, y)$ are as follows.

$$F_J(x, y) = k1 \cdot lA_J(x, y) + k2 \cdot lB_J(x, y). \quad (1)$$

In (1), $k1$ and $k2$ are given parameters, if image B is fused into A, then $k1 > k2$. Now, $F_J(x, y)$ and $F_{j,}^k(x, y)$ are use to reconstruct and generate the fused image $F'(x, y)$ which contains high frequency as well as low frequency information of $A(x, y)$ and $B(x, y)$ [21], [22], [23], [24].

3) Reconstruction of Decomposed Fused Image:
After generating decomposed co-efficient fused image the single level inverse Daubechies Wavelet transform $(idwt2)$ was used to generate the synthesized fused image. Finally, the reconstructed image is used as an input to the next module reducing the dimension.

B. Algorithm used for Dimension reductions

Face images are very much alike, and therefore they are highly correlated. It follows than that which can be represented in a much lower dimensional feature subspace. For this study, two different dimensionality reduction methods were applied here, principal component analysis
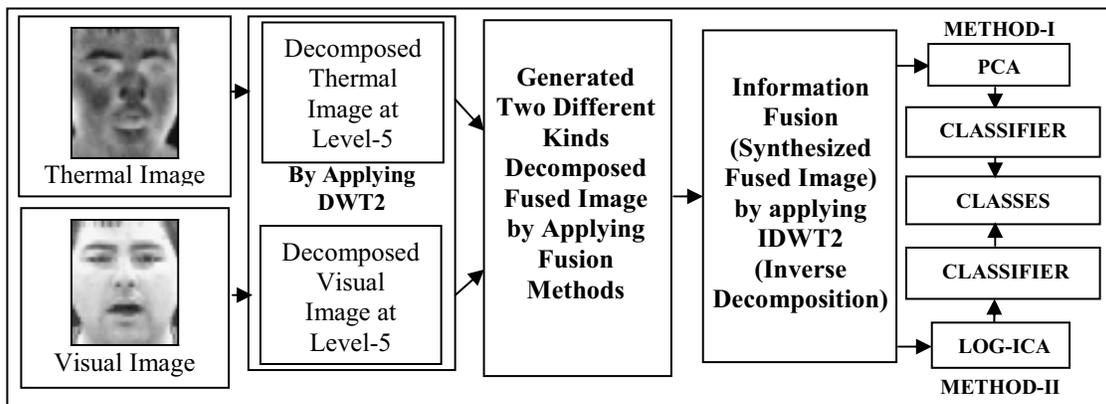

Figure 1. Flow of the system.

(PCA) and independent component analysis in logarithmic domain (log-ICA).

1) Principal Component Analysis (PCA): The Principal component analysis (PCA) was used for subspace projection technique for method-I [6], [8]. It is based on the second-order statistics of the input image, which tries to attain an optimal representation that minimizes the reconstruction error in a least-squares sense. Eigenvectors of the covariance matrix in the face images constitute the eigenfaces. The dimensionality of the face feature space is reduced by selecting only the eigenvectors possessing significantly large eigenvalues. Once the new face space is constructed, resulting in arrival of a test image, it is projected onto this face space to yield the feature vector— the representation coefficients in the constructed face space. The classifier decides for the identity of the individual, according to similarities scored among the test images [3], [7].

2) Independent Component Analysis (ICA): In this paper, independent component analysis (ICA) was used for dimensionality reduction of face images. Although it has other applications like data analysis and source separation. Here, Independent Component Analysis was implemented in natural logarithm domain (this implementation was represented in this paper as a log-ICA) to reduce the parameters. This log-ICA projects data from high dimensional to low-dimensional space [1], [25], [27], [28]. Independent Component Analysis (ICA) is the generalization of the Principal component analysis (PCA), which decorrelates the high-order statistics in addition to second-order moments.

To define rigorously, ICA (Jutten and Herault, 1991; Comon, 1994) [26], we can use a statistical "latent variables" model. Assume that, we observe n linear mixtures $x_1, ..., x_n$ of n independent components.

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

It is convenient to use vector-matrix notation instead of the sums as given in (2). Let us denote by x, the random vector whose elements are the mixtures $x_1, ..., x_n$, and likewise by s the random vector with elements $s_1, \cdots, s_n$. Let us denote by A the matrix with elements a_{ij}. All vectors are understood as column vectors; thus x^T, or the transpose of x, is a row vector. Using this vector-matrix notation, the above mixing model is written as

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

Sometimes we need the columns of matrix A; denoting them by a_j the model can also be written as,

$$x = \sum_{i=1}^{n} a_i s_i. \quad (4)$$

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

a) Algorithm for Independent Component Analysis in natural logarithm domain (log-ICA):

Step 1: Input an M x N Thermal image I.

Step 2: Calculate and remove the mean of thermal image and generate the centered image I_c.

Step 3: Apply natural logarithm of the each element of centered image I_c and convert the image to logarithm domain I_{log}.

Step 4: Applying whitening on the centered-log image.

Step 5: Apply the FastICA algorithm for dimensionality reduction.

Step 6: Plot the image in a data matrix.

b) Implementing log-ICA: There exist a number of popular ICA algorithms. These include FastICA [27], [28] Infomax [24], [29], Comon's algorithm [26], and KernalICA [30]. The success of ICA for a given set of data may depend crucially on performing some application-dependent pre-processing steps. The FastICA algorithm has many advantages over other independent component analysis (ICA) algorithms. The FastICA algorithm and the underlying contrast functions have a number of desirable properties when compared with existing methods for ICA. The FastICA algorithm directly calculates independent components (ICs) of an image. The ICs can be estimated one by one, which is roughly equivalent to doing projection pursuit. This is useful in exploratory data analysis, and decreases the computational load of the method in cases where only some of the independent components need to be estimated. The FastICA has most of the advantages of neural algorithms: it is parallel, distributed, computationally simple and requires little memory space.

Stage 1 (Centering): The first is the most basic and necessary preprocessing is to center x, i.e. subtract its mean vector $m = E\{x\}$ so as to make x a zero-mean variable. This implies that it is zero-mean as well, as can be seen by taking expectations on both sides of (3). This preprocessing is made solely to simplify the ICA algorithm. It does not mean that the mean could not be estimated. After estimating the mixing matrix A with centered data, we can complete the estimation by adding the mean vector of s back to the centered estimates of s. The mean vector of s is given by $A^{-1}m$, where m is the mean that has been subtracted in the preprocessing.

Algorithmic Steps for Centering:

Input: Original Image I

Step 1: Create an array with all zeros of size I

      newVectores = zeros (size (I));

Step 2: Find the mean value of loaded image I

Step 3: Subtract the mean value from the I

newVectores = I - meanValue;

The main advantage of centering is that it decreases the amount of redundant information by decorrelating the input vectors. The input vectors, with high dimension and correlated, can be represented in a lower dimension space and de-correlated.

Stage 2 (Log-Centered Data): Before applying the algorithm of whitening on the centered data we first convert the centered image into a log-domain by applying natural logarithm to the elements of I_c. At the time of converting the image we first check whether the pixel value tends to 'zero' and those pixels values are replaced with a minimum value, which is 1/512. After replacing those pixels values, we apply log on generated image as pixel-by-pixel.

$$I_{log}(x, y) = \log(I(x, y)). \qquad (5)$$

The primary reason to apply the natural log over to centered image is to protect partially from the loss of some information during the conversion from one to another class of image type that uses less number of bits to represent pixel values. For example, a class of image type uint16 (unsigned integer 8) grayscale image is capable of storing up to 65,536 distinct shades of gray, but a class of image type uint8 (unsigned integer 8) grayscale image can store only 256 distinct shades of gray. When we convert a uint16 grayscale image to a uint8 grayscale image, im2uint8 function (of image processing toolbox of matlab) quantizes the gray shades in the original image. In other words, all values from 0 to 127 in the original image become 0 in the uint8 gray scale image, values from 128 to 385 all become 1, and so on. Therefore, without converting such image classes, if natural log is applied which reduces the pixel values of a uint8 class image. However, in the case of color images more time is required for processing, so finally all the pixel values of images after converting the gray scale image to logarithm domain have been reduced and less time is required for process as compare to color image.

After subtracting the mean from original image, some pixel values may be found as negative or zero. Therefore, before applying natural log over centered image, there is a need to do some changes of the image pixel values. We consider that, pixel values of centered image may have the conditions given below:

Case 1 (When pixel values are negative): In this case, after finding the image type, the highest pixel range is 256. Therefore, during the conversion in logarithm domain, we first eliminate all the negative values with highest minimum (1/256) values, as natural log does not accept negative values. The image intensity does not change too much in this case.

Case 2 (When pixel values are 0): In this case, if natural log is applied over the zero pixel values of centered image then that pixel value becomes undefined. Therefore, this is also another reason for changing the zero pixel values to a minimum by 1/image class type.

Stage 3 (Whitening): Second part of dimensionality reduction using independent component analysis is to find whitening of the observed matrix. This means that before the application of the ICA algorithm and after centering, we transform the observed vector $\tilde{x}$ that is white, i.e. its components are uncorrelated and their variances are of equal unity. In other words, the covariance matrix of $\tilde{x}$ equal s the identity matrix:

$$E\{\tilde{x}\tilde{x}^T\} = I. \qquad (6)$$

To calculate the whitening matrix we use the eigenvalues 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 eigenvalues, $D = diag(d_1, \ldots d_n)$.

The main advantage of whitening is to reduce the number of parameters to be estimated. Instead of estimating the n^2 parameters that are the elements of the original matrix A, we only need to estimate the new, orthogonal matrix $\tilde{A}$. An orthogonal matrix contains $n(n-1)/2$ degrees of freedom [1], [28]. Thus, whitening solves half of the problem of ICA. Because whitening is a very simple and standard procedure, in fact much simpler than any ICA algorithms, it is a good idea to reduce the complexity of the problem using whitening. It may also be quite useful to reduce the dimension of the data at the same time as we do the whitening. It often involves the effect of reducing noise [1], [28].

Algorithmic Steps for Whitening:

Input: Load centered image as a newVector and original image.

Step1: Calculate co-variance for dimensionality reduction by applying PCA.

Step2: Calculate the eigenvalues and eigenvectors of covariance matrix.

Step3: Calculate the whitening and de-whitening matrices (these handle dimensionality simultaneously).

C. Difference Between ICA and Log-ICA

The FastICA algorithm is firstly developed by A. Hyvarinen [26]; in this case, the ICA algorithm was implemented by fixed point algorithm. This ICA algorithm has many stages to calculate the dimension reduction of a face image; (i) calculating centered data (to make correlate to un-correlate data); (ii) apply the algorithm of whitening (for reducing the number of parameters) and, (iii) third and final stage for dimension reduction. In this paper, an efficient

approach called ICA in logarithm domain (log-ICA) was introduced to solve the problem like expression in face recognition. The stages for log-ICA as follows: (i) calculating centered data (to make correlate to un-correlate data); (ii) convert the centered image to logarithm domain by applying element wise natural log on the face image (which reduces the changes of facial structure due to expressions); (iii) whitening; and (iv) dimensionality reduction [1], [28].

D. PCA vs ICA

ICA is more advantageous than PCA for the following reasons: i) It has better architecture than PCA, for which it can easily understand where the data is mainly concentrated in n-dimensional space; ii) ICA uniquely identifies the mixing matrix; iii) In case of PCA it reconstructs the data on the orthogonal basis, but on the other hand it is not necessary for ICA; iv) ICA can easily reconstruct the data even in the presence of the noise; v) It is sensitive to high-order statistics in the data, but not just the covariance matrix [1], [28].

E. ANN using Back propagation with Momentum

Either humans or other computer techniques can use neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, to extract patterns and detect trends that are too complex to be noticed. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyze. It has reawakened the scientific and engineering community to the modelling and processing of many quantitative phenomena using neural networks. This learning algorithm is applied to multilayer feed forward networks consisting of processing elements with continuous differentiable activation functions. Such networks associated with the backpropagation learning algorithm are also called backpropagation networks [9], [10], [11], [12].

III. EXPERIMENTAL DISCUSSION

Three different experiments were conducted here, to find a solution of the problem of facial expression for face recognition. For this study, a newly developed feature extraction algorithm called ICA in logarithm domain (log-ICA) has been introduced here. For these experiments three different face images (visual, thermal and fusion) datasets were prepared and for that IRIS thermal / visible [2] face database was used and the face images are of varying expressions with illuminations and total of 11 different poses. This work has been simulated using MATLAB 7.9.0.529 (R2009b) in a machine of configuration: Intel Xeon Quad Core Processor with 16GB Physical Memory.

A. Image Database Description

One of popular thermal/visible face database was used to demonstrate the effectiveness of the proposed log-ICA algorithm. The IRIS [2] database simultaneously acquired unregistered thermal and visible face images under variable illuminations, expressions, and poses. It contains 30 individuals of 320 x 240 bitmap RGB image format. Each individual has the images of three different expressions Exp1 (Surprise), Exp2 (laughter), Exp3 (Anger) and five different

illuminations Lon (left light on), Ron (right light on), 2on (both lights on), Dark (dark room), Off (left and right lights off) with varying poses like (left, right, mid, mid-left, mid-right). Two different sensors are used to capture this database. One is Thermal - Raytheon Palm-IR-Pro and another is Visible - Panasonic WV-CP234.

B. Training and Testing

1) Details of Training Set: The detailed overview of the training and testing process has been presented here. To start the face recognition experiments three different datasets ('v' for visual, 't' for thermal, and 'f' for fusion of 'v' and 't') was prepared on the basis of different types of image available in the database. 10 individuals were randomly picked up out of 30 for the preparation of the dataset. Each dataset contains 330 images (i.e. 33 images per class) separately of 3 different expressions. Besides these 330 images of each dataset are grouped into 3 categories (anger, surprise and laughter) on the basis of expressions.

Out of 330 images of each set (i.e. v, t, f), 110 images from 10 different classes were randomly picked up as a training data set for each of three different experiment.

2) Training Phase: Three different experiments were performed separately to learn the three different data sets. The fusion of visual and thermal images was completed with the help of Daubechies wavelet coefficient (db4). The whole process of generation of fused image was presented in section II. *A*. First: the feature selection algorithms i.e. independent component analysis in logarithmic domain (log-ICA) was applied on the face images of for three different datasets. Second: all the features of selected images were kept on a different data matrix.The same training processes were followed for three different datasets.

3) Details of Testing Set: Three different testing sets were prepared on the basis of three different expression details (i.e. anger, surprise and laughter). Each training set contains total 110 images of 10 different classes (i.e. 11 images per class). These same testing sets were used for three different experiments to test the efficiency of the system. During testing, 33 images have been used for each class and out of 33 images, 11 images each for three expression s namely laughter, anger, and surprise. This same approach was followed for preparation of testing sets for three different experiments.

C. Results of Experiments (Method-I)

In method-I, total 100 fused images of 10 different classes (i.e. 10 images per class) have been used, which does not include the left-to-right, angle i.e. 90^0 rotation images of every class. Using method-I, the highest recognition rate was achieved for four different classes; class-3, 6, 8 and for class-10. At the time of training we used exp-1 (i.e. Surprise) for first 7 classes and exp-2 (i.e. happiness/laughter) for class-9 and we used exp-3 (i.e.

anger) is used for class-8 and class-10 and in the testing set the illuminated faces (i.e. both light on left and right) with normal facial expression was picked for 6 classes (class-1 to class-4, class-6 and 7) from IRIS database. In addition, for the other class's data of expression were chosen for testing. The testing processes were conducted separately for all the classes and 85% was the average recognition rate of all classes using method-I.

D. Result of all Experiments using the Images with Different Types of Expressions using Method-II

Depending upon the expression sets (i.e. anger, laughter, and surprise); all of these image sets were recognised separately. It means that against the same learned images different types of expression image sets were tested. From Table I, it was clearly shown that for a particular expression set testing data was not available because these sets of images were used in the training set. So, three separate testing of method-II using three different image sets were not repeated those image were used in their corresponding training set. Because, in the IRIS database only 11 images are available of each expression for each class. The recognition rates on three different experiments were illustrated below:

1) Results of Experiments using Visual Image: Three different image sets on the basis of expression was examined against the same training data set of 10 different classes which have been presented in Table I. Let us discuss the recognition performance of three different expressions image sets. In case of laughter / happiness and anger image sets, when these sets were used as a training set of classes, the images from other expression sets were successfully recognised. It means when laughter set was used as training data for any class among 10, other two facial expression sets (i.e. anger and surprise) were recognised and also for anger image sets. But when surprise dataset was used for training data of any class, other expressions of same class were not

successfully recognised. So, 90.91% is the recognition rate for laughter and anger dataset when surprise was used for training and the highest success rate was achieved for surprise when laughter and anger was used for training.

2) Results of Experiments using Thermal Image: The same approach (like visual face) was followed during the testing phase of three different facial expression sets of thermal image. From Table I, it is clear that data of anger (in case of thermal) were never used in the training set for any class. So, in the testing set of anger expression it has all 10 classes and using the anger expression in thermal experiment it has 93.637% as average recognition rate and 81.82% is the lowest success rate of class-5, when data of training were from laughter set. On the other hand surprise data have 93.94% as 6% less than that of average recognition rate which is surprise data of visual images.

3) Results of Experiments using Fused Image: Third and final experiment was conducted on the fused images of corresponding thermal and visual images which were used in previous two experiments. Generation of error-free fused images was not a very easy task to do. Because of position of the facial images (i.e. thermal and visual) are not in the same place due to difference in expressions in the IRIS database. So, recognition rates degrade in few cases as compared to the recognition rates of visual and thermal images. 94.81% is the average recognition rate, when the testing set is laughter and surprise. So, from these recognition rates it was clearly shown that surprise fused data were more recognisable than surprise thermal images but not from surprise visual data set and 95.46% is the average recognition rate of anger dataset and in this case this recognition rate is more than the corresponding visual and thermal image set.

A comparative study has been presented here in Table II.

IV. CONCLUSION AND FUTURE WORK

In this paper, we present a novel approach to recognize the faces of varying expressions like anger, happiness / laughter and surprise using fusion. For this study, we used IRIS thermal/visible face database and it contains the faces

TABLE I.   DETAILS OF TRAINING SET

Class Number	Visual (v)	Thermal (t)	Fusion (f)
Class-1	Surprise	Surprise	Visual Anger and Thermal Anger
Class-2	Surprise	Surprise	Visual Anger and Thermal Anger
Class-3	Laughter	Laughter	Visual Surprise and Thermal Surprise
Class-4	Anger	Surprise	Visual Laughter and Thermal Laughter
Class-5	Laughter	Laughter	Visual Anger and Thermal Anger
Class-6	Surprise	Surprise	Visual Surprise and Thermal Surprise
Class-7	Surprise	Surprise	Visual Anger and Thermal Anger
Class-8	Surprise	Surprise	Visual Laughter and Thermal Laughter
Class-9	Laughter	Laughter	Visual Surprise and Thermal Surprise
Class-10	Anger	Surprise	Visual Laughter and Thermal Laughter

TABLE II.   RECOGNITION RATES OF DIFFERENT METHODS

Method	Recognition Rate
Present Method-I	85.00%
Present Method-II (for visual image)	90.91%
Present Method-II (for thermal image)	93.637%
Present Method-II (for fused image)	94.81% 95.46% (maximum)
PCA [4]	83.4%
PCA for Visual indoor Probes[5]	81.54%
Wavelet Subband + Kernel associative Memory with XM2VTS database [31]	84%
Face Recognition using Gabor Filter [32]	85%

having different expressions with illumination. The visual and thermal face has own advantages and disadvantages. Therefore, we try to combine the advantages from visual and thermal image from the IRIS database. The experimental results show that our method i.e. ICA in logarithmic domain can significantly perform the face recognition task for the face images with different expressions and rotation and achieved maximum success rate by further reducing the parameters. We are planning to develop our own thermal and corresponding visual face database which contain different facial expression with varying illumination with different facial action units (AUs) like inner brow raiser, outer brow raiser, brow lowerer, upper lid raiser, cheek raiser, lid lightener, nose wrinkle, lip corner puller, lip corner depressor, chin raiser, mouth stretch etc.

REFERENCES

[1] A. Hyvarinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis", IEEE Trans. on Neural Networks, Vol. 10, No. 3, May 1999.

[2] http://www.cse.ohio-state.edu/otcbvs-bench/Data/02/download.html.

[3] I. Daubechies, Ten lectures on wavelets, CBMS-NSF conf. series in applied mathematics. SIAM Ed. 1992.

[4] L. Shen and L. Bai, "Information Theory for Gabor Feature Selection for face recognition", Hindawi Publishing Corporation EURASIP Journal On Applied Signal Processing 2006, Article ID 30274, pp. 1-11.

[5] A. Socolinsky and A. Seinger, "Thermal face recognition in an operational scenario". Proc. of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition workshops (CVPRW'04).

[6] M. K. Bhowmik, D. Bhattacharjee, M. Nasipuri, D. K. Basu and M. Kundu, "Fusion of Daubechies Wavelet Coefficients for Human Face Recognition", International Journal of Recent Trends in Engineering (ISSN: 1797 -9617), by the Academy Publisher, Finland, Vol. 3, No. 2, May 2010.

[7] H. K. Ekenel and B. Sankur, "Multiresolution face recognition", Image and Vision Computing, 2005. Vol. 23, pp. 469 – 477.

[8] M. K. Bhowmik, D. Bhattacharjee, M. Nasipuri, D. K. Basu and M. Kundu, "Fusion of Wavelet Coefficients from Visual and Thermal Face Images for Human Face Recognition – A Comparative Study", International Journal of Image Processing (IJIP) [ISSN 1985-2304], published by CSC Press, Computer Science Journals, Kuala Lumpur, Malaysia, Vol. 4, Issue 1, March 2010.

[9] M. K. Bhowmik, D. Bhattacharjee, M. Nasipuri, D. K. Basu and M. Kundu, "Classification of Log Polar Visual Eigenfaces using Multilayer Perceptron", Proc. of The 2nd International Conf. on Soft computing (ICSC-2008), IET, Alwar, Rajasthan, India, published by Norosa Publishers, pp. 107-123, Nov. 8–10, 2008.

[10] M. K. Bhowmik, D. Bhattacharjee, M. Nasipuri, D. K. Basu, M. Kundu, "Human Face Recognition using Line Features", Proc. of National Seminar on Recent Advances on Information Technology (RAIT-2009) which will be held on Feb 6-7, 2009, Indian School of Mines University, Dhanbad. pp: 385.

[11] D. Bhattacharjee, M. K. Bhowmik, M. Nasipuri, D. K. Basu and M. Kundu, "Classification of Fused Face Images using Multilayer Perceptron Neural Network", Proc. of International Conf. on Rough sets, Fuzzy sets and Soft Computing, organized by Deptt. of Mathematics, Tripura University, pp. 289 – 300, Nov. 5–7, 2009.

[12] M. K. Bhowmik, "Artificial Neural Network as a Soft Computing Tool – A case study", Proc. of National Seminar on Fuzzy Math. & its application, Tripura University, November 25-26, 2006, pp: 31-46.

[13] S. Mallat, "A theory for Multiresolution signal decomposition: the wavelet representation", IEEE PAMI, 1989. vol. 11, no. 7, pp. 674-693.

[14] B. Fasel, J. Luettin, "Automatic facial expression analysis: a survey", Pattern Recognition, Vol. 36, pp. 259-275, 2003.

[15] J. C. McCall and M. M. Trivedi, "Pose Invariant Affect Analysis using Thin-Plate Splines", Int. Conf. on Pattern Recognition, Cambridge, UK, August 2004.

[16] F. Prokoski, "History, current status, and future of infrared identification," Proc. IEEE Workshop on Computer Vision beyond the Visible Spectrum: Methods and Applications, pp. 5-14, 2000.

[17] D. A. Socolinsky, A. Selinger, and J. D. Neuheisel, "Face recognition with visible and thermal infrared imagery," Computer Vision and Image Understanding, Vol. 91, pp. 72-114, 2003.

[18] J. Heo, S. G. Kong, B. R. Abidi, and M. A. Abidi, "Fusion of Visual and Thermal Signatures with Eyeglass Removal for Robust Face Recognition", Proc. of the 2004 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW'04).

[19] L. Wolff, D. Socolinsky, and C. Eveland, "Quantitative measurement of illumination invariance for face recognition using thermal infrared imagery," Proc. CVPR Workshop on Computer Vision Beyond the Visible Spectrum, December 2001.

[20] C. K. Eveland, Diego A. Socolinsky and L. B. Wolff, "Tracking Human Faces In Infrared Video".

[21] P. Ekman and W. Friesen, "Facial Action Coding System (FACS): A Technique for the Measurement of Facial Action", Palo Alto, Calif.: Consulting Psychologists Press, 1978.

[22] C. K. Chui, "An Introduction to Wavelets", Academic Press, Inc., America. 1992.

[23] I. Daubechies, "Orthonormal bases of compactly supported wavelets, Communication on Pure and Applied Mathematics", vol. XLI, pp.909-996, 1988.

[24] 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," Neural Computation, vol. 11, no. 2, pp. 417-441, 1999.

[25] M. S. Barlett, J. R. Movellan and T. J. Sejnowski, "Face Recognition by Independent Component Analysis", IEEE Trans. Neural Networks, vol. 13, no. 6, pp. 1450-1464, 2002.

[26] P. Comon, "Independent Component Analysis, A New Concept", Signal Processing, vol. 36, no. 3, pp. 287-314, 1994.

[27] A. Hyvarinen and E. Oja, "Independent Component Analysis: Algorithms and Applications". Neural Networks Research Centre, Helsinki University of Technology.

[28] A. Hyvarinen, J. Karhunen and E. Oja, "Independent Component Analysis", John Wiley & Sons, 2001.

[29] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind decondolution", Neural Compute, vol. 7, no. 6, pp. 1129-1159, 1995.

[30] F. R. Bach and M. I. Jordan, "Kernel Independent Component Analysis," J. Machine Learning Research, vol. 3, pp. 1-48, Mar.2002.

[31] M. Marchand, M. Golea and P. Rujan, "A convergence theorem for sequential learning in two-layer perceptrons", Europhys. Lett. vol. 11: 487-492, (1990).

[32] Y. Moses, Y. Adini and S. Ullman, "Face Recognition: the problem of compensating for changes in illumination direction", Lecture notes in Computer Science, Vol. 800, Oxford University, 1994.

Mrinal Kanti Bhowmik received his B.E. (in CSE) and M.Tech (in CSE) degrees from Tripura University (A Central University), Suryamaninagar, Tripura, India, in 2004 and 2007 respectively. From July, 2010 to till date, he is working as an Assistant Professor in Tripura University. He is also pursuing his Ph.D. 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 the year of 2010 and another project based on Face Identification is going on since March, 2011. Both the projects are funded by the DIT, MCIT, India. He is also a member of the IEEE (USA).

Debotosh Bhattacharjee received the MCSE and Ph. D.(Eng.) 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, 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 member of IEEE (USA).

Dipak Kumar Basu received his B.E.Tel.E., M.E.Tel., and Ph.D. (Engg.) degrees from Jadavpur University, in 1964, 1966 and 1969 respectively. Prof. Basu has been a faculty member of J.U from 1968 to January 2008. He is presently an A.I.C.T.E. Emiretus Fellow at the CSE Department of J.U. His current fields of research interest include pattern recognition, image processing, and multimedia systems. He is a senior member of the IEEE, U.S.A., Fellow of I.E. (India) and W.B.A.S.T., Kolkata, India and a former Fellow, Alexander von Humboldt Foundation, Germany.

Mita Nasipuri received her B.E.Tel.E., M.E.Tel.E., and Ph.D. (Engg.) degrees from Jadavpur University, in 1979, 1981 and 1990, respectively. Prof. 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, U.S.A., Fellow of I.E (India) and W.B.A.S.T, Kolkata, India.