Multimed Tools Appl (2016) 75:16781–16807
DOI 10.1007/s11042-015-2945-2

Expressions Recognition of North-East Indian (NEI) Faces

Priya Saha[1] · **Mrinal Kanti Bhowmik**[1] ·
Debotosh Bhattacharjee[3] · **Barin Kumar De**[2] ·
Mita Nasipuri[3]

Received: 17 June 2013 / Revised: 21 July 2015 / Accepted: 8 September 2015 /
Published online: 29 September 2015

Abstract Facial expression is one of the major distracting factors for face recognition performance. Pose and illumination variations on face images also influence the performance of face recognition systems. The combination of three variations (facial expression, pose and illumination) seriously degrades the recognition accuracy. In this paper, three experimental protocols are designed in such a way that the successive performance degradation due to the increasing variations (expressions, expressions with illumination effect and expressions with illumination and pose effect) on face images can be examined. The whole experiment is carried out using North-East Indian (NEI) face images with the help of four well-known classification algorithms namely Linear Discriminant Analysis (LDA), K-Nearest Neighbor algorithm (KNN), combination of Principal Component Analysis and Linear Discriminant Analysis (PCA + LDA), combination of Principal Component Analysis and K-Nearest Neighbor algorithm (PCA + KNN). The experimental observations are analyzed through confusion matrices and graphs. This paper also describes the creation of NEI facial expression database, which contains visual static face images of different ethnic groups of the North-East states. The database is useful for future researchers in the area of forensic science, medical applications, affective computing, intelligent environments, lie detection, psychiatry, anthropology, etc.

Keywords Visual face image · Facial expressions · Pose and illumination variations · NEI facial expression database · Baseline algorithms

✉ Mrinal Kanti Bhowmik
mkb_cse@yahoo.co.in

[1] Department of Computer Science and Engineering, Tripura University (A Central University), Suryamaninagar 799022 Tripura, India

[2] Department of Physics, Tripura University (A Central University), Suryamaninagar 799022 Tripura, India

[3] Jadavpur University, Kolkata 700032 West Bengal, India

1 Introduction

Facial expressions are the fundamental means of conveying social information among individuals. A face is more emotion expressive medium than verbalization of one's feeling. A small and the slightest change in any of the facial features like eyebrow, lip may alter the entire meaning of the body language. Facial expressions are one of the three main factors that significantly influence the human faces. The other two factors are face pose and illumination. Face recognition technology becomes more complicated when all these three factors are present simultaneously. Development of a robust face recognition algorithm requires a face image database of sufficient size containing visual face images captured with controlled variations in pose, illumination, expression, occlusion, etc. [6]. In the recent years, the development of different facial expression databases containing variations in pose, illumination and other factors has been getting importance from the research community. These face databases can be utilized for the improvement and assessment of different face recognition algorithms.

In the Japanese Female Facial Expression Database (JAFFE) [11], total 213 images of 10 subjects with six basic facial expressions (happiness, sadness, surprise, anger, disgust, fear) have been captured. The AR Database [13] contains total 3,288 images of 116 individuals, where face images are captured under four different illumination conditions and expressions (neutral, smile, anger and scream). The Radboud Faces Database [8] contains 8040 number of images with different expressions (happy, angry, sad, contemptuous, disgusted, neutral, fearful, and surprised) where each expression consists of three different gaze directions (looking left, looking frontal, looking right). The Korean Face Database [15] contains overall 52,000 images of 1000 subjects collected under varying illumination condition, varying poses and facial expressions (neutral, happy, surprise, anger, and blink). The CAS-PEAL (pose, expression, accessory, and lighting) Database [2] contains 1040 individuals with a different expression, lighting, pose and accessory. The database contains six different expressions (neutral, laughing, frowning, surprise, eyes closed, mouth open). The CMU Pose, Illumination and Expression (PIE) database [16] has been collected at the Carnegie Mellon University in 2000. It contains face images with variations in pose, illumination and expression. However, due to some limitations, they have created another database, called the Multi-PIE database [7], with more than 750,000 images of 337 people using four recording sessions over the period of 5 months. 15 different poses and 19 illumination conditions were used for imaging the subjects with a wider range of facial expressions. The MMI facial expression database [14] consists of both still images and image sequences. It contains two types of pose variations with six basic expressions. The Extended Cohn-Kanade Dataset (CK+) database [10] contains 593 image sequences of 210 subjects. The database has expression and pose variations. The detailed description of the databases mentioned above is given in Table 1.

In this paper, a brief study of North-East Indian (NEI) face database, which is a subset of DeitY-TU face database [1], along with baseline evaluations with four popular classification algorithms on this database, has been presented. The paper also discusses the image acquisition, design, and variations of the NEI database. The captured images of NEI facial expression database are all static images with variations in pose, illumination and expression. NEI database consists of visual face images of 320 individuals from three different North-Eastern states of India namely, Assam, Mizoram and Nagaland. These individuals belong to different ethnic groups of Mongoloid races. Depending upon the cultural heritage of the ethnic groups, specific facial expressions may have variations in the intensity of expression or the degree of

Table 1 Description of different facial expression based databases

Name of Database	Number of Images	Resolution	Number of Subjects	Expressions	Pose and Illumination variation	Year of release
JAFFE	213	256×256	10	Happiness, sadness, surprise, anger, disgust, fear	No variation	1998
AR database	3288	768×567	116	Neutral, smile, anger, scream	Frontal view; 3 lighting conditions	Year of release not available publicly
Radboud Face Database	8040	1024×681	67	happy, angry, sad, contemptuous, disgusted, neutral, fearful, and surprised	5 poses and 3 gaze directions; No illumination variation	2010
Korean Face Database (KFDB)	52,000	640×480	1000	Neutral, happy, surprise, anger, smile, and blink	7 poses; 2 illumination colours	2003
CAS-PEAL Database	30900	360×480	1040	Neutral, laughing, frowning, surprise, eye closed, mouth open	27 poses; 15 lighting conditions	2003
MultiPIE	More than 7,50,000	3072×2048	337	Neutral, smile, surprise, squint, disgust, scream	16 poses; 19 illumination conditions	2010
MMI Facial Expression Database	800+ sequences, 200+ images	720×576	52	Anger, happy, sad, surprise, fear, disgust	2 poses; No illumination variation	2005
The Extended Cohn-Kanade Dataset (CK+)	593 image sequences	640×490	210	Anger, contempt, disgust, fear, happy, sadness, surprise	2 poses; No illumination variation	2010

movement of action units [4, 5]. However, in NEI database, we have observed that captured facial expressions of different states are almost universal. Most of the people of NEI belong to Mongolian races, and this might be the reason of the universality of facial expressions for peoples of all these states. Because of excessive storage requirement, often it becomes difficult to capture the images of every person under all variations. Sometimes the participants are non-cooperative during image capturing session. Still our database contains the face images having almost all variations for every single person. Evaluations of baseline recognition performance on this face database have been carried out on the face images containing all three types of variations, namely expression, pose, and illumination. Four different classification algorithms, namely, Linear Discriminant Analysis (LDA), K-Nearest Neighbor algorithm (KNN), combination of Principal Component Analysis and Linear Discriminant Analysis (PCA + LDA), combination of Principal Component Analysis and K-Nearest Neighbor algorithm (PCA + KNN) have been experimented in three different ways on three different datasets excluding face images with closed eye, glasses and neutral expression. Observations reveal that more face variations in the experiment cause more performance degradation of the classification algorithms. Here, a brief description of the NEI facial expression database is given below:

i) Face images have been captured from different ethnic groups of North-East India with six different expressions, four illumination conditions and five different poses;
ii) Neutral face images containing five different poses and four different illumination conditions are available in this database;
iii) Neutral face images with glasses/spectacles are also available in this database;
iv) Another variation i.e., closed eye is captured in full and half illumination in all five pose variations;
v) In total, each person has 60 expressive face images and 35 neutral face images;
vi) Database contains 30400 static visual face images of 320 people of three NE states of India;

The remaining part of the paper is organized as follows. The studio setup for face image capture is illustrated in Section 2. The development of the database is detailed in Section 3. Section 4 contains the thorough discussion of the baseline evaluation of NEI database. Section 5 reports the baseline evaluation results of other databases and finally conclusions are drawn in Section 6.

2 Studio setup

NEI facial expression database has been developed at the Biometrics Laboratory in the Department of Computer Science & Engg. of Tripura University, Tripura, India. The face images are captured with varying poses, expressions and lighting conditions in a special studio room of size 12 ft×10 ft×10 ft. The camera and lighting setup in the photographic room are described below.

2.1 Camera setup

To capture the face images, total four numbers of Nikon D5100 cameras with Nikkor 18–55 mm lens and one Canon EOS 1000D camera with Canon EF-S 18–55 mm lens are being used. Five different visual images of each subject have been captured from five different angles in a single shot using a remote sensor called Nikon ML-L3. This infrared remote can work within the range of 16 ft from the camera. The cameras are placed and fixed in front of the subject, at a 4.5 ft distance from the subject. The five cameras (cam) namely Cam1, Cam2, Cam3, Cam4, and Cam5 are placed at +50°, +25°, 0°, −25°, and -50° respectively with respect to the subject. Heights of the cameras are adjusted according to the subject's head position. The resolution of two cameras (cam1 and cam5) is 3696×2448. The resolution of other two cameras (cam2 and cam4) is 2464×1632, and the last one i.e., cam3 possesses a resolution of 1936×1288. The subject is asked to look directly at cam3. The side view and top view of the camera setup are shown in Figs. 1 and 2 respectively.

2.2 Lighting/Illumination setup

This setup regulates the indoor lighting conditions using different directions of the multiple lamps as shown in Figs. 1 and 2. All doors, windows and curtains are kept closed at the time of capturing the face images in the photographic room. We have used three numbers of Simpex photo light systems (Simpex Pro 23) bulbs for controlling the illumination conditions. The two light systems labeled as L2 and L3, are set at angles $+60^0$ and -60^0 respectively with respect to the subject. The third one L1 is set from the top of the subject to reduce the shadow effects formed by the side lights. The distance of the side lights from the subject is 4.5 ft, and the top light is 7 ft above the ground. The illumination conditions are controlled by three modelling bulbs of 100 W for capturing the face images of each image. The simpex lighting system can produce half and full power output.

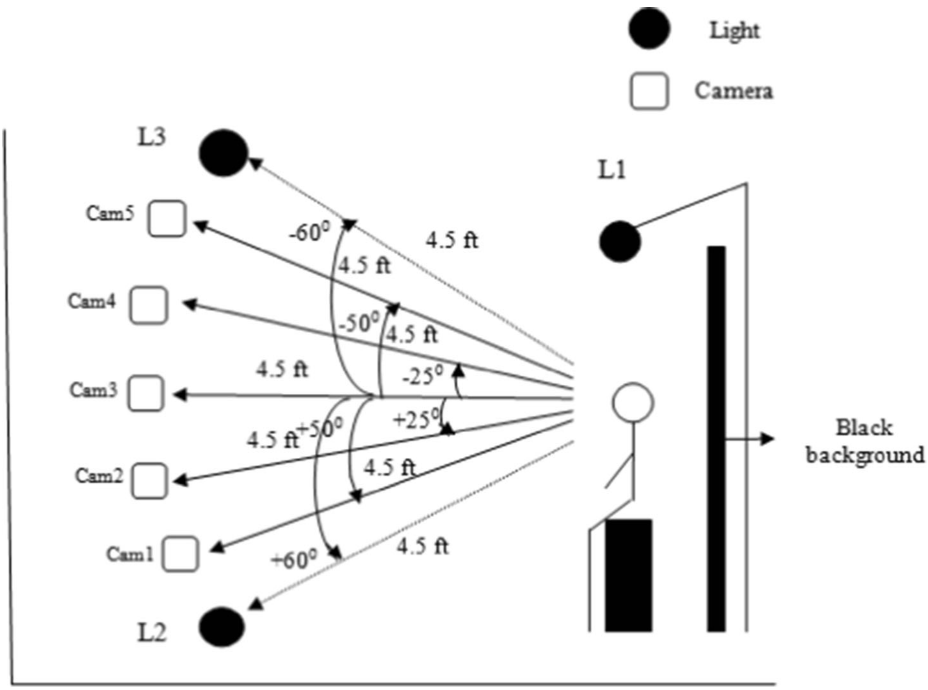


Fig. 1 Side view of the camera and lighting setup

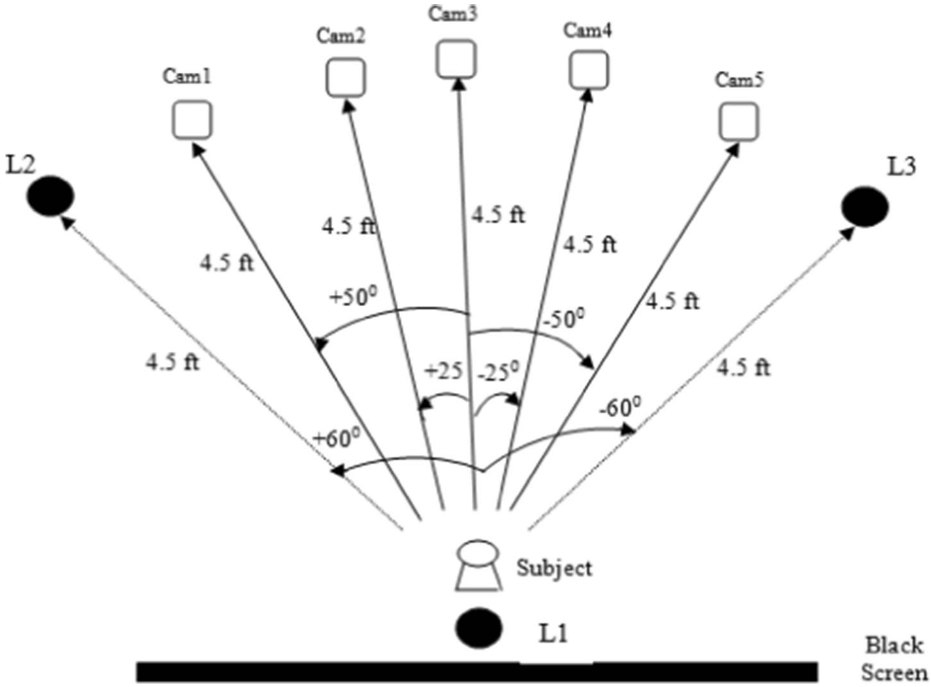


Fig. 2 Top view of the camera and lighting setup

2.3 Accessories and background

All the images are captured against a homogeneous black background of 8.7×6.5 ft dimension to prevent light reflection. As we are capturing face images in an indoor environment, the uniformly coloured background is taken. A glass with the thin black frame is taken as an accessory when capturing the images, but face images with glasses are captured only in a neutral expression. The accessory is added to database to increase the variety of the database.

3 Database development

The NEI facial expression database has been organised as a combination of personal information and captured face images of each subject participated in the development of this database.

3.1 Subjects

Currently, NEI facial expression database consists of visual face images of three North-Eastern states. All raw images are captured in JPEG format. The total storage required to accumulate all the captured face images till date is approximately 61 GB. 140 MB disk space is allotted for storing each person's face image (95 images per person). The research team has captured total 10,640 face images of 112 individuals from Mizoram. There are 62 males and 50 females. The subjects have been categorized with respect to age. There are 42 persons less than or equal to the age of 20. 68 persons are there in between 21 and 30. In the age of 31 to 50, there are 5 persons. However, no person is found above the age of 50. From Assam, the team has collected 10,165 images of 107 individuals where the numbers of males and females are 46 and 61 respectively. Total 54 persons are there whose age is less than or equal to 20. The number of persons belonging to the age of 21 to 30 and 31 to 50 is 48 and 4 respectively. Here is also no person found above the age of 50. A total number of subjects in Nagaland is 101 where the number of male and female is 57 and 44 respectively. The subjects from Nagaland consist of 16 persons whose age is less than or equal to 20; 79 persons whose age is in the range of 21 and 30; 4 persons in between 31 and 50. Lastly, two persons' face images are captured whose age is above 50.

3.2 Expression variations

In addition to neutral and closed eye expression, six basic facial expressions i.e., anger, happy, sad, surprise, fear and disgust have been captured for each person. Face images with glasses are the part of the neutral expression. A total number of face images for each expression is listed in a tabular format that is shown below in Table 2. Sample face images with expressions are shown in Fig. 3.

3.3 Illumination variations

Illumination variation affects the appearance of a face. Using the illumination setup described in Section 2.2, we have controlled four types of illumination conditions for each person. These are "half illumination", "full illumination", "left light on" and "right light on". To create different illumination conditions, we have controlled two side lights (L2 and L3), keeping the top light (L1) on. In "full illumination", we set both L2 and L3 to full power. For "half

Table 2 Number of face images per expression

Expressions	Types of variations	Number of face images per illumination and face images with glass	Total face images per person
Neutral expression	With half illumination	5	25
	With all pose variations		
	With full illumination	5	
	With all pose variations		
	With Left light on	5	
	With all pose variations		
	With Right light on	5	
	With all pose variations		
	With Glass	5	
	With all pose variations		
Anger, happy, sad, surprise, fear, disgust, closed eye expressions	With half illumination	5	10
	With all pose variations		
	With full illumination	5	
	With all pose variations		

illumination”, both the lights are set to glow at half power. In case of “left light on”, L2 is set at full power, and the right light (L3) is set off. Similarly for “right light on”, we set the L3 to full power and switch off L2. Four types of illuminations have been applied only to a neutral expression. The six expressions are captured only in half and full illumination. Some sample face images captured with varying illuminations are shown in Fig. 4.

3.4 The combination of expression, illumination and pose variation

Figure 5 shows the face images that contain three types of variations. Five cameras capture five different poses in one shot. The camera description with their respective angles is already discussed in Section 2.

Each person has six facial expressions, and every facial expression contains five different poses. Facial expressions are captured in full and half illumination conditions. It means that each person has 30 images in full illumination condition and 30 images in half illumination condition for various facial expressions. Other than six facial expressions, neutral face images are considered with “full illumination”, “half illumination”, “left light on”, “right light on”, and glasses along with five poses i.e., total 25 neutral face images. Another variation i.e., closed eye is captured in the both full and half illumination levels for all five pose variations and thus the database contains 10 face images in closed eye position. So, each person contains total 95 face images including neutral and expressive face images.

3.5 Image naming convention

The file naming rule is designed for ease of understanding the contents of the image. In Table 3, we represent the naming rules of face images. Figure 6 shows the naming hierarchy of the image files. First of all, we consider the state code; under the state code, the individual

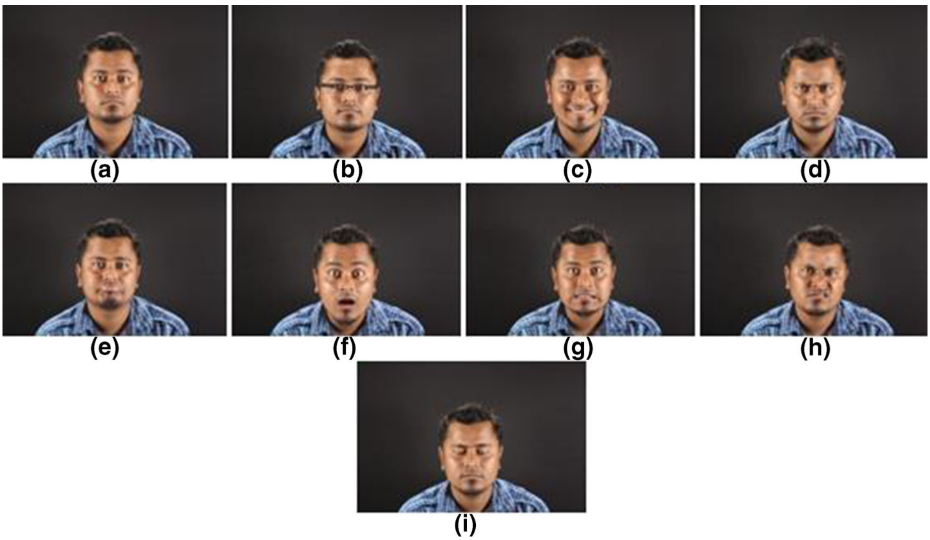


Fig. 3 Facial expressions **a** neutral face **b** neutral face with glasses **c** happy **d** anger **e** sad **f** surprise **g** fear **h** disgust **i** closed eye

person is considered; each face image of that person has specific expression, illumination and pose with or without glass. The naming format is:

State code_Person number_Expression code_Illumination code_Glass code_Pose code.jpg

Suppose, we have taken a face image 'AS_009_E4_I1_0_P3.jpg', it means that person number 9 is from Assam and the image is captured with 'sad' expression, in 'full illumination', and the person is not wearing any glass, and finally, the image has been captured from 0^0 angle, i.e., this image is a frontal image.

4 Evaluation and analysis of NEI facial expression database

In this section, baseline evaluations of NEI facial expression database, using four well-known classification algorithms have been performed to assess the complexity of the database in context to expression recognition. Future researchers can utilize the reference evaluation results of expression recognition. Four classification algorithms namely: Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA) + LDA, K-Nearest Neighbor (KNN), KNN + PCA have been selected for this purpose. The expression recognition results are presented through confusion matrices. A brief introduction of the above-mentioned algorithms has been presented in the following subsection.

4.1 Selected classification algorithms

4.1.1 Linear discriminant analysis (LDA)

Linear discriminant analysis (LDA) discovers a linear combination of features which divides two or more classes of objects. This method maximizes the ratio of between-class variance to

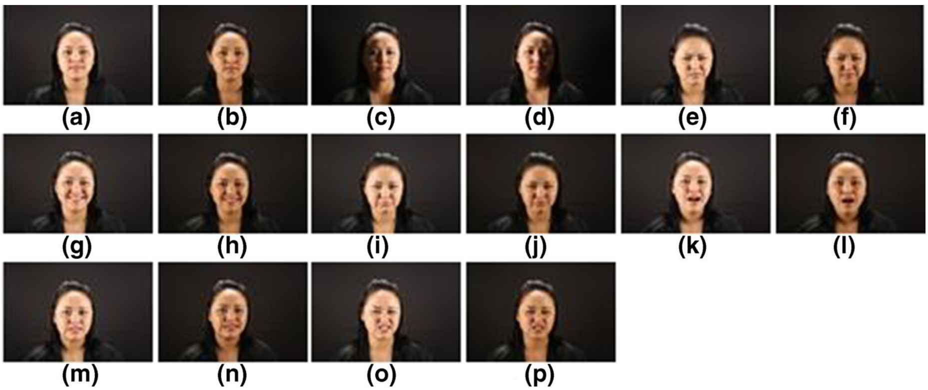


Fig. 4 Combination of expression and illumination **a** neutral face with full illumination **b** neutral face with half illumination **c** neutral face with left light on illumination **d** neutral face with right light on illumination **e** anger expression with full illumination **f** anger expression with half illumination **g** happy expression with full illumination **h** happy expression with half illumination **i** sad expression with full illumination, **j** sad expression with half illumination, **k** surprise expression with full illumination, **l** Surprise expression with half illumination, **m** fear expression with full illumination, **n** fear expression with half illumination, **o** disgust expression with full illumination, **p** disgust expression with half illumination

Fig. 5 Face images captured with combination of expression, illumination and pose

the within-class variance in any particular data set thereby guaranteeing maximal separability. For Multiclass LDA, the intra-class matrix of multivariate observations x is defined as

$$\hat{\sum}_w = \sum_{i=1}^{n} \sum_{x \in c_i} \left(x - \overline{x}_i\right)\left(x - \overline{x}_i\right)' \tag{1}$$

The inter-class scatter matrices are given as

$$\hat{\sum}_b = \sum_{i=1}^{n} m_i \left(\overline{x}_i - \overline{x}\right)\left(\overline{x}_i - \overline{x}\right)' \tag{2}$$

Where m_i is the number of training samples for each class, $\overline{x}_i$ is the mean for each class and $\overline{x}$ is the total mean vector given by $\overline{x} = 1/m \sum_{i=1}^{n} m_i \overline{x}_i$. The transformation ϕ can be obtained by solving generalized eigen value problem

$$\hat{\sum}_b \phi = \lambda \hat{\sum}_w \phi \tag{3}$$

After obtaining the transformation function ϕ, the classification is performed in the transformed space based on some distance metric. When a new instance z arrives, it is classified to

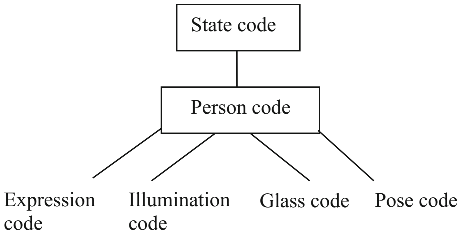
$$\underset{k}{\operatorname{argmin}}\, d\left(z\phi, \overline{x}_k \phi\right),$$

Where $\overline{x}_k$ is the centroid of kth class [9].

Table 3 Symbols and their meaning for image data format

Symbol	Meaning
AS, MZ, NL	State code of Assam, Mizoram and Nagaland
001, 002, 003…	Person Number under particular state
E1, E2, E3, E4, E5, E6, E7, E8	Expression code E1: Neutral; E2: Anger; E3: Happy; E4: Sad; E5: Surprise; E6: Fear; E7: Disgust; E8: Closed eye
I1, I2, I3, I4	Illumination code I1: Full Illumination; I2: Half Illumination; I3: Left Light On; I4: Right Light On
0, 1	Glass code 0: Absence of glass 1: Presence of glass
P1, P2, P3, P4, P5	Pose code P1: $+50^0$ P2:$+25^0$ P3:0^0 P4:-25^0 P5:-50^0

Fig. 6 Image naming hierarchy



4.1.2 Combination of principal component analysis and linear discriminant analysis (PCA + LDA)

In PCA [17], eigenfaces are comprised of eigenvectors of the covariance matrix of the face images. Only the eigenvectors that possess significantly large eigenvalues are selected in dimension reduction process. The principal component w_1 of a dataset X can be defined as:

$$w_1 = \arg \max_{\|w\|=1} E\left\{ \left(W^T X \right)^2 \right\} \tag{4}$$

$W = \{W[p,q]\}$ is the matrix of basis vectors, one vector per column, where each basis vector is one of the eigenvectors of the covariance matrix. Vectors constituting W are a subset of those in V, where V is the matrix consisting of the set of all eigenvectors of the covariance matrix. LDA is closely associated with PCA. LDA clearly gives the effort to model the dissimilarity between the classes of data, but PCA does not take into account.

4.1.3 K-nearest neighbor algorithm (KNN)

The K-nearest neighbor algorithm classifies objects based on K number of closest training samples in the feature space. The similarity between two example vectors is measured using the concept of Euclidean distance between them. The smaller distance is regarded as more similar vectors. Let, an example vector be represented by a_1 (x), a_2 (x),…a_n (x). The Euclidean distance between two example vectors is given in Eq. (5)

$$d\left(x_i, x_j\right) = \sqrt{\sum_{r=1}^{n} \left(a_r(x_i) - a_r\left(x_j\right)\right)^2} \tag{5}$$

After sorting the distances of the test sample from each of the training samples, K number of nearest neighbors are taken. The test sample is assigned to the class in which majority of these K samples belong to [3].

4.1.4 Combination of principal component analysis and K-nearest neighbor algorithm (PCA + KNN)

The purpose of the PCA is to reduce the large dimension of the data to a relatively small dimension feature vector. Keeping the major independent variables, PCA reduces the number of features and thus reduces the dimension of the feature vector. The independent components of a lower dimensional space are called principal components. KNN classifies the reduced numbers of features.

4.2 Pre-processing of face images

Before applying the classification algorithms, all the face images are pre-processed. The first pre-processing operation is registration of face images. Registration has been performed to provide proper alignment of two images, which are captured in two different illumination conditions i.e., full and half. The research team have taken two face images of a same person with the same expression and pose but different illumination. Keeping fully illuminated face image as a base image, registration has been done on that person's half illuminated face image. Registration has been done using affine transformation [12]. Properties like parallelism, the ratio of lengths of collinear or parallel segments, the ratio of areas and a linear combination of vectors are invariant under affine transformation. Completing the registration, the face images are cropped manually to remove the background. The cropped face includes only the face region with little hair excluding ear. The cropped face images are then resized into particular width×height rectangle to reduce the computational complexity without losing much information. In our experiment, we have taken width=height=50 pixels. Then, the cropped images have been turned to greyscale. Sample cropped face images of individuals from three different states are shown in Fig. 7.

4.3 Experimental design

The total experimentation is carried out using the face images of three states i.e., Mizoram, Nagaland, and Assam. After pre-processing, all the face images have been grouped into training and testing sets. Three independent experiments have been done in three states. We have performed facial expression recognition experiments on both interstate (training and

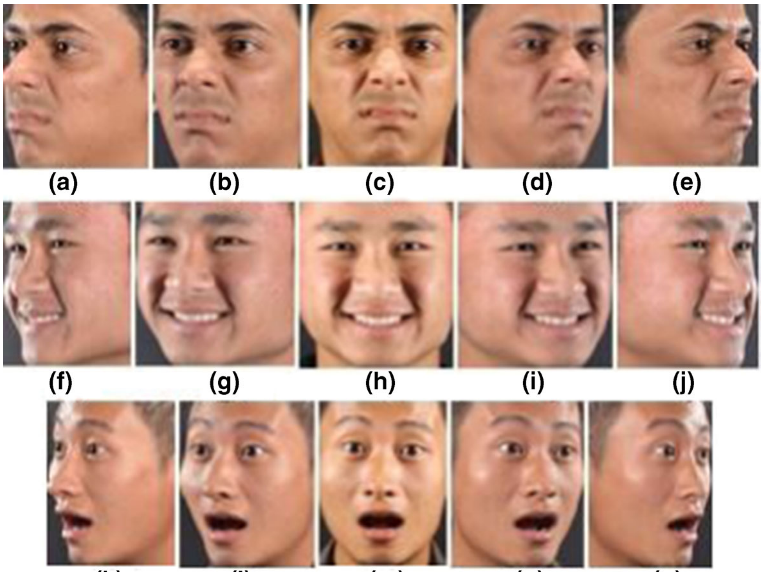


Fig. 7 Sample registered and cropped face images **a**–**e** Assam face images with full illumination, disgust expression containing pose variations, **f**–**j** Mizoram face images with full illumination, happy expression containing pose variations, **k**–**o** Nagaland face images with full illumination, surprise expression containing pose variations

testing based on two different states) and intrastate (training and testing within the same state). Experiment results have shown that both types of experiments produce an almost similar result for a particular experimental design. We have stated previously that most of the people of NEI belong to Mongolian races, and this may be the reason of finding no major difference in interstate and intrastate expression recognition accuracy. So, we have kept the intrastate performance evaluation in the manuscript. The design of experimental intrastate data has been described below.

4.3.1 First experimental design

The first experiment has been conducted using the face images of Mizoram people. Out of 112 subjects, face images of 50 subjects have been taken for training and rest 62 subjects have been taken for the testing purpose. Here, only frontal face images with six expressions have been considered. The training set contains total 300 face images (50 subjects×6 expressions), and the testing set contains 372 face images (62 subjects×6 expressions).

4.3.2 Second experimental design

The second experiment is done on Nagaland face images. In this case, in addition to expressions all poses have been considered i.e., five poses have been taken per expression for the experiment. Training set consists of 300 half illuminated face images of 10 subjects (10 subjects×6 expressions×5 poses). The testing set contains 600 full illuminated face images of 20 subjects (20 subjects×6 expressions×5 poses).

4.3.3 Third experimental design

The last experiment contains four different parts. Here, only frontal face images with full illumination have been taken for training. 180 face images (30 subjects×6 expressions×1 frontal face) have been taken for training purpose. The test set consists of four different parts. Each part contains 70 subjects with six expressions and one pose ($+25^0$ or $+50^0$ or -25^0 or -50^0) with half illumination. So, each test part contains total 420 face images (70 subjects×6 expressions×1 pose).

4.4 Evaluation results

Three independent experiments have been conducted using above said four classification algorithms. Image projection of LDA is the feature vector of LDA and PCA+ LDA. Pixel values of the cropped gray scale face image are used as feature vector for the classification algorithms LDA and KNN. For PCA + LDA and PCA + KNN, we have chosen first m principal components from the list sorted in a decreasing order of the magnitudes of the corresponding eigenvalues. The value of m is so chosen that the ratio between the sum of magnitudes of top m eigenvalues in the sorted list and the sum magnitudes of all eigenvalues is in the range 0.85-0.86. In KNN algorithm, the value of k has been chosen through experimentation and it is found that k=1 gives best recognition results. The following tables show the performance of each of these algorithms in each experimental setup. The experiment results are represented as confusion matrices (Tables 4, 5, 6, 7, 8, 9).

Table 4 Confusion matrices of expression recognition of Mizoram People

	Anger	Happy	Sad	Surprise	Fear	Disgust
	LDA					
Anger	26	12	10	3	1	10
Happy	5	49	0	0	2	6
Sad	13	6	30	2	3	8
Surprise	3	5	0	48	3	3
Fear	5	10	3	2	36	6
Disgust	10	11	3	0	0	38
	PCA + LDA					
Anger	0	6	6	10	14	26
Happy	0	8	2	13	15	24
Sad	2	11	3	11	16	19
Surprise	1	8	0	6	21	26
Fear	1	3	1	12	19	26
Disgust	0	12	2	13	10	25
	KNN					
Anger	23	7	10	2	10	10
Happy	6	35	6	1	8	6
Sad	17	8	18	1	10	8
Surprise	8	4	1	43	2	4
Fear	10	6	13	2	25	6
Disgust	10	17	9	1	4	21
	PCA + KNN					
Anger	9	10	0	2	0	41
Happy	6	18	0	5	0	33
Sad	6	9	0	1	0	46
Surprise	5	10	0	0	0	47
Fear	5	6	0	0	0	51
Disgust	7	10	0	1	1	43

The following observations are being made from the tables:

1. The experimental setup of Mizoram dataset is already illustrated in the previous section. The training and testing dataset does not contain so much variation i.e., no variations in pose, illumination present in this experimental dataset. Here, the happy expression has acquired highest recognition rate using LDA. The happy expression is the most easily identified expression than other expressions because the facial feature changes caused by the happy expression are almost similar for all participants. The highest average recognition rate is 61.02 % that is achieved through LDA. Other algorithms perform very poorly. In case of PCA + LDA and PCA + KNN, disgust expression is mostly recognized.
2. In the second experiment, the highest average recognition rate has been degraded by 8.19 % in comparison to the result obtained in the first experiment. Here, also LDA generates a better result than other classification algorithms. Unlike the first experiment, here variations lie in pose and illumination in addition to expressions. Illumination and

Table 5 Confusion matrices of expression recognition of Nagaland People

	Anger	Happy	Sad	Surprise	Fear	Disgust
	LDA					
Anger	69	0	25	2	1	3
Happy	10	49	38	0	3	3
Sad	19	0	76	1	4	0
Surprise	7	4	1	85	3	0
Fear	15	0	39	0	33	13
Disgust	28	9	51	1	3	8
	PCA + LDA					
Anger	7	9	27	2	44	11
Happy	6	5	27	2	50	10
Sad	8	3	31	1	43	14
Surprise	5	4	30	3	44	14
Fear	7	2	32	3	40	16
Disgust	9	3	28	0	54	6
	KNN					
Anger	45	0	36	4	3	12
Happy	18	31	22	2	13	14
Sad	28	0	49	0	16	7
Surprise	7	0	4	78	5	6
Fear	13	1	34	1	39	12
Disgust	29	10	31	3	5	22
	PCA + KNN					
Anger	13	27	9	18	19	14
Happy	8	29	3	28	20	12
Sad	10	34	10	16	18	12
Surprise	17	20	14	23	16	10
Fear	14	37	10	7	23	9
Disgust	16	19	14	18	21	12

pose variation greatly influence the recognition performance, and that may be the cause of performance degradation. Surprise expression secures maximum recognition rate using LDA and KNN. It is observed from the confusion matrices that when PCA is combined with LDA or KNN, recognition accuracy degrades seriously. The reason may be laid in sensitivity of PCA due to large variations in illumination, facial expressions.

3. According to the third experiment, four different test cases are being considered. In this experiment, complexity in training and test set is more increased than the previous two datasets. Here, training set contains only frontal faces with full illumination and testing set contains four types of poses ($+50^0$, $+25^0$, -50^0 and -25^0) in four independent sets with half illumination with all expressions. Among the four cases, the highest average recognition rate is achieved in $+25^0$ pose experiment using LDA. However, recognition accuracy is immensely degraded than previous two observations and it is obvious because training images contain frontal faces whereas testing set consists of posed face images with different illumination

Table 6 Confusion matrices of expression recognition of Assam People (−50^0 pose)

	Anger	Happy	Sad	Surprise	Fear	Disgust
	LDA					
Anger	5	33	0	0	19	13
Happy	7	19	2	1	14	27
Sad	3	21	4	0	10	32
Surprise	1	41	0	2	23	3
Fear	3	23	0	0	28	16
Disgust	6	31	0	0	10	23
	PCA + LDA					
Anger	0	36	6	1	18	9
Happy	0	29	5	1	21	14
Sad	2	34	8	0	12	14
Surprise	1	47	2	1	15	4
Fear	0	37	6	2	11	14
Disgust	2	28	8	4	16	12
	KNN					
Anger	12	0	7	8	21	22
Happy	9	0	3	5	21	32
Sad	4	0	5	2	28	31
Surprise	5	0	5	7	23	30
Fear	3	0	4	3	38	22
Disgust	10	0	5	2	18	35
	PCA + KNN					
Anger	3	15	8	8	20	16
Happy	7	22	6	4	18	13
Sad	5	20	9	11	16	9
Surprise	3	16	5	9	21	16
Fear	2	12	7	8	31	10
Disgust	7	11	11	10	19	12

condition. Out of 16 different experiments, 15 experiments result below 30 %. The instances of happy and fear expression are most correctly recognized ones in these 16 experiments. The highest average recognition rate is 34.29 % that clearly proves the difficulty of our database in the combined variations in pose, expressions and illuminations. Here, PCA + LDA algorithm is the poorest performer than other three algorithms.

From the above three observations, the following conclusions can be made

- The facial expressions of emotions, namely, happy and fear are two mostly recognized expressions throughout the whole experiment,
- Anger is the lowest recognized expression,
- In most of the cases, anger is misclassified with sad and disgust,
- Likewise, disgust is confused mostly with facial expression anger and sad,

Table 7 Confusion matrices of expression recognition of Assam People (-25^0 pose)

	Anger	Happy	Sad	Surprise	Fear	Disgust
	LDA					
Anger	15	3	17	8	1	26
Happy	4	5	5	4	4	48
Sad	5	0	26	3	2	34
Surprise	4	5	4	35	9	13
Fear	9	2	14	4	17	24
Disgust	11	2	20	7	7	23
	PCA + LDA					
Anger	10	38	3	3	2	14
Happy	6	50	5	2	3	4
Sad	9	48	3	5	3	2
Surprise	9	40	4	5	4	8
Fear	13	41	1	6	3	6
Disgust	4	50	3	3	5	5
	KNN					
Anger	3	0	20	6	8	33
Happy	1	3	5	1	6	54
Sad	4	0	1	3	13	39
Surprise	4	0	10	10	11	35
Fear	1	0	7	1	16	45
Disgust	3	0	7	6	5	49
	PCA + KNN					
Anger	10	30	7	3	10	10
Happy	3	38	9	3	4	13
Sad	5	39	8	5	6	7
Surprise	2	29	4	7	13	15
Fear	2	29	5	9	4	21
Disgust	7	34	8	2	8	11

- The confusion among the three expressions (anger, disgust, sad) reveals that all three expressions have almost similar facial feature movements like lips are tightened, brows are lowered,
- The overall recognition rate of surprise expression is also low. In many cases of the entire experiment, surprise achieves the lowest recognition rate,
- It has been noticed that surprise expression is misclassified with happy and fear. The reason behind this may be the openness of the mouth and rise of inner brow,
- The performance of LDA is better than other three baseline algorithms,
- As the variation increases in the experimental dataset, the performance degradation is noticed,
- The performance of PCA + LDA is very poor throughout the experiment because relevant information may be lost due to dimension reduction.

The expression recognition rates are graphically shown in Fig. 8a–c. In these graphs, the expressions are numbered as 1, 2, 3, 4, 5 and 6 corresponding to the expressions anger, happy,

Table 8 Confusion matrices of expression recognition of Assam People (+25^0 pose)

	Anger	Happy	Sad	Surprise	Fear	Disgust
	LDA					
Anger	35	2	14	2	11	6
Happy	11	12	17	1	20	9
Sad	9	3	34	1	18	5
Surprise	21	10	2	15	19	3
Fear	10	2	10	3	36	9
Disgust	21	6	13	2	16	12
	PCA + LDA					
Anger	0	23	7	1	32	7
Happy	0	32	7	1	25	5
Sad	0	20	8	1	31	10
Surprise	0	19	6	0	43	2
Fear	0	20	4	0	44	2
Disgust	0	21	5	2	32	10
	KNN					
Anger	2	0	51	2	0	15
Happy	2	1	54	4	2	7
Sad	1	0	61	0	0	8
Surprise	2	0	47	4	4	13
Fear	0	0	59	1	4	6
Disgust	0	0	52	0	0	18
	PCA + KNN					
Anger	4	17	5	2	24	18
Happy	3	15	5	2	28	17
Sad	3	9	13	1	28	16
Surprise	1	15	4	0	35	15
Fear	3	17	8	1	19	22
Disgust	1	18	8	2	23	18

sad, surprise, fear and disgust respectively. The 8 (a) graph illustrates the 1st experiment where LDA shows better accuracy in case of happy expression than other three algorithms. The accuracies obtained using PCA + LDA and PCA + KNN are very low. The graphical illustration of performances obtained in the 2nd experiment, as shown in Fig. 8b, clearly shows that the performance of LDA is poorer than the previous experiment. The happy expression has secured less accuracy than the previous but surprise expression has got maximum accuracy among the six expressions. Figure 8c is the graph of average expression recognition rates obtained by four different algorithms for four different pose based experiments. Expression 2 i.e., happy expression has secured the highest recognition rate in -50 degree posed experiment. The maximum average expression recognition has been generated through LDA. In graph (a), it has been noticed that LDA produces maximum accuracy for happy expression (expression 2), but in graph (b), it has been clearly observed that accuracy of happy expression reduces and this is because of illumination variation. Face in happy expression consists of three important facial movements, the first two are lip corner puller and lips part, where teeth are visible (in most of the cases), and another is cheek

Table 9 Confusion matrices of expression recognition of Assam People (+50^0 pose)

	Anger	Happy	Sad	Surprise	Fear	Disgust
	LDA					
Anger	22	4	8	2	27	7
Happy	14	4	4	3	29	16
Sad	9	5	7	3	38	8
Surprise	22	13	0	2	32	1
Fear	4	6	2	2	50	6
Disgust	25	6	0	1	21	17
	PCA + LDA					
Anger	0	64	4	0	2	0
Happy	1	65	4	0	0	0
Sad	0	63	6	0	1	0
Surprise	0	67	3	0	0	0
Fear	0	67	3	0	0	0
Disgust	0	67	2	0	1	0
	KNN					
Anger	3	0	50	2	6	9
Happy	2	0	40	3	7	18
Sad	0	1	55	1	4	9
Surprise	1	0	44	2	8	10
Fear	1	0	46	2	12	9
Disgust	0	2	46	0	7	15
	PCA + KNN					
Anger	4	40	6	2	17	1
Happy	2	44	7	0	15	2
Sad	2	39	5	1	21	2
Surprise	4	42	5	0	17	2
Fear	4	45	4	0	16	1
Disgust	5	42	3	2	18	0

raiser. These two features are more brightened than other facial parts. In half illumination condition, these facial parts have become low illuminated, and this causes degradation in expression recognition performance. The other thing is that LDA is not performed well in pose and illumination variant condition. Likewise, in pose based experiment also, LDA cannot produce good accuracy for happy expression. In pose based experiment, happy is mostly confused with fear and disgust. Happy and fear have two common facial movements i.e., lips part and lips stretched; whereas in disgust, sometimes lips are parted in our dataset. So, confusion causes between those expressions and generates low recognition accuracy. Expression 3 i.e., sad is also achieved low recognition accuracy because of confusion with disgust expression.

Figure 9 summarizes the performance of four baseline algorithms as a bar graph. Mizoram dataset experiment contains fewer variations than others. So, among the six experiments, better accuracy (61.02 %) has been achieved in Mizoram dataset experiment. The combination of expression, pose and illumination with complete variations in training and test cases causes serious performance degradation in Assam dataset experiments.

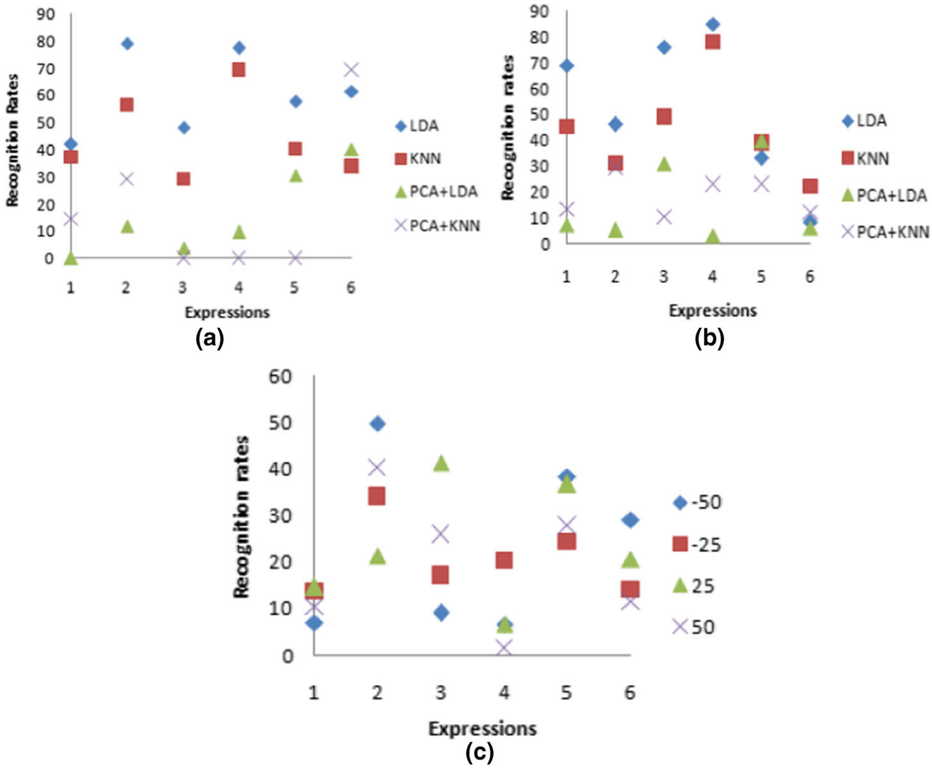


Fig. 8 Graphical representation of results **a** first experiment, **b** second experiment, **c** third experiment

5 Baseline evaluations of other databases

Different researchers have developed different face databases and performed baseline evaluations. Table 10 illustrates baseline evaluations of different databases.

In Korean face database [15], three types of variations have been introduced. However, experiments have been conducted using only one variation. Performance is measured graphically using Correlation Matching (CM) and PCA. Radboud Face Database [8] contains pose and expression variation. Experiments have been conducted using frontal face images, and happy

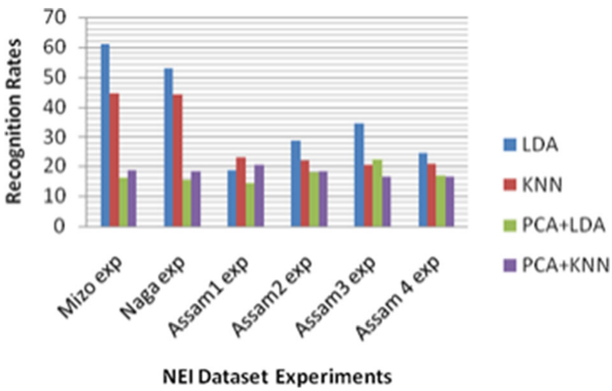


Fig. 9 Average recognition rates of four classification algorithms for six experiments

Table 10 Summarization of baseline evaluations of different databases

Database reference	Classification algorithm used	Experiments conducted	Reported performance
Korean Face Database [15]	Correlation Matching (CM) and PCA	Pose Varying Experiment (No expression variation) Expression varying experiment (No pose variation) Illumination varying experiment (No expression and pose variation)	Performance is reported using graphical representation
Radboud Face Database [8]	Analysis of Variance (ANOVA)	Expression varying experiment (No pose and illumination variation)	98 %
CMU PIE [16]	PCA and LDA	Experiment across illumination	PCA-36.6 % LDA-95.0 %
CMU Multi-PIE [7]	PCA and LDA	Experiment across illumination	PCA-35.4 % LDA-71.3 %
		Experiment across expression and illumination	PCA-18.5 % LDA-41.4 %
CAS-PEAL Database [2]	PCA, PCA + LDA, PCA + LDA (HE), PCAvLDA (GIC), PCA + LDA (RHE), PCA + LDA (RGIC), G PCA + LDA, G PCA + LDA (HE), G PCA + LDA (GIC), G PCA + LDA (RHE), PCA + LDA (RGIC)	Using frontal probe sets	PCA-28.2 % PCA+LDA- 42.2 % PCA + LDA (HE)- 48.4 % PCA + LDA (GIC)- 45.4 % PCA + LDA (RHE)- 47.8 % PCA + LDA (RGIC)- 45.5 % G PCA + LDA- 57.4 % G PCA + LDA (HE)- 58.3 % G PCA + LDA (GIC)- 57.8 % G PCA + LDA (RHE)- 53.7 % G PCA + LDA (RGIC)- 56.8 %
	PCA + LDA, PCA + LDA (RGIC), G PCA + LDA, PCA + LDA (RGIC)	Using pose probe sets	PCA + LDA- 19.0 % PCA + LDA (RGIC)- 23.3 % G PCA + LDA- 28.7 %

Table 10 (continued)

Database reference	Classification algorithm used	Experiments conducted	Reported performance	
			G PCA + LDA (RGIC)- 32.8 %	
JAFFE Database [11]	Gabor coding and semantic similarity and Geometric and semantic similarity	Expression varying experiment (No pose and illumination variation)	56.8 % and 36.6 %	
NEI facial expression database (Present Database)	LDA, PCA + LDA, KNN, PCA + KNN	Expression varying experiment (No pose and illumination variation)	LDA-61.02 % PCA + LDA-16.1 % KNN-44.35 % PCA + KNN-18.82 %	
		Expression and illumination varying experiment	LDA-52.83 % PCA + LDA-15.33 % KNN-44 % PCA + KNN-18.33 %	
		Expression, illumination and Pose varying experiment (training images are frontal)	-50^0 posed testing face image	LDA-18.83 % PCA + LDA-14.52 % KNN-23 % PCA + KNN-20.48 %
			-25^0 posed testing face image	LDA-28.67 % PCA + LDA-18.1 % KNN-21.91 % PCA + KNN-18.41 %
			$+25^0$ posed testing face image	LDA-34.29 % PCA + LDA-22.38 % KNN-20.48 % PCA + KNN-16.43 %
			$+50^0$ posed testing face image	LDA-24.29 % PCA + LDA-16.91 % KNN-20.72 % PCA + KNN-16.43 %

expression is mostly recognized. Experiments have been carried out using illumination varying face images in CMU PIE database [16]. PCA and LDA have been used as baseline algorithms, and experimental results show that LDA performs better than PCA. CMU MultiPIE [7] extends the PIE database and experiments have been performed using the same algorithms. Here experiments have been conducted in two ways. The first experiment is done using only illumination varying face images and the second experiment is conducted using the combination of expression and illumination varying face images. The performance of both algorithms degrades in case of second experiment. In the baseline evaluations of CAS-PEAL Database [2], the authors have used three algorithms namely Principle Component Analysis (PCA), a combined Principle Component Analysis and Linear Discriminant Analysis (PCA + LDA), and PCA + LDA algorithm based on Gabor features (G PCA + LDA). Four face image preprocessing methods are combined, with those algorithms, to test the performance of the algorithms. Experiments have been conducted using frontal face images and posed face images. The recognition rate is higher in case of frontal face images than posed ones. The JAFFE Database [11] contains variations in only expressions. Two algorithms namely Gabor coding and semantic similarity and Geometric and semantic similarity are used as algorithms. Gabor coding and the semantic similarity has obtained higher recognition accuracy than the other. Three types of experiments have been conducted using NEI facial expression images. The combination of illumination, expression and pose face images are used in the experiments. The three successive experiment results reveal that the increasing number of variations in training and testing phases causes performance degradation. All variations have been included in the experiments to test the efficiency of the baseline algorithms.

6 Conclusions

Presently, NEI facial expression database contains 30,400 static face images of 320 tribe and non-tribe individuals containing variations in pose, lighting and expression. Cameras with different resolution have been utilized for image capturing purpose. Each person contains 95 face images that are the most important feature of this database. The four well-known classification algorithms are employed in this paper for the baseline evaluation of the database. The observations resulted from the conducted experiments using the combination of variations reveal the fact that illumination, pose and expression greatly influence the face recognition performance. The database is designed to assess the robustness of the face recognition algorithms and also to explore the difficulties that are not yet resolved in face recognition domain.

Acknowledgment The work presented here is being conducted in the Biometrics Laboratory of Tripura University, under the research project supported by the Grant No. 12(2)/2011-ESD, dated 29/03/2011, from DeitY, MCIT, Government of India. The first author is grateful to Department of Science and Technology (DST), Government of India for providing her Junior Research Fellowship-Professional (JRF-Professional) under DST INSPIRE fellowship program (No. IF131067). The authors would like to thank anonymous reviewers for their comments/suggestions to improve the quality of the paper.

References

1. Bhowmik MK, Saha K, Saha P, Bhattacharjee D (2014) DeitY-TU face database: its design, multiple cameras capturing, characteristics, and evaluation. Opt Eng 53(10):102106-1–102106-24. doi:10.1117/1.OE.53.10.102106

2. Cao B, Shan S, Zhang X, Gao W (2004) Baseline evaluations on the CAS-PEAL-R1 face database. In: Proc. 5th Chinese Conf on Biometric Recognition, Guangzhou, China, pp. 370–378. doi: 10.1007/978-3-540-30548-4_42
3. Dasarathy BV (1990) Nearest neighbor (NN) norms: NN pattern classification techniques. IEEE Computer Society Press, Los Alamitos
4. Ekman P (1972) Universals and cultural differences in facial expressions of emotion. In: Cole J (ed) Nebraska sym. on motivation. University of Nebraska Press, Lincoln, pp 207–283
5. Ekman P (1993) Facial expression and emotion. Am Psychol 48:384–392
6. Gross R (2005) Face databases. In: Li SZ, Jain A (eds) Handbook of face recognition. Springer, New York, pp 301–328
7. Gross R, Matthews I, Cohn J, Kanade T, Baker S (2010) Multi-PIE. Image Vis Comput 28(5):807–813. doi: 10.1016/j.imavis.2009.08.002
8. Langner O, Dotsch R, Bijlstra G, Wigboldus DHJ, Hawk ST, van Knippenberg A (2010) Presentation and validation ofthe Radboud faces database. Cognit Emot 24(8):1377–1388. doi:10.1080/02699930903485076
9. Li T, Zhu S, Ogihara M (2006) Using discriminant analysis for multi-class classification: an experimental investigation. Knowl Inf Syst 10(4):453–472. doi:10.1007/s10115-006-0013-y
10. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar J, Matthews I (2010) The extended Cohn-Kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression. Proc. IEEE Int Conf on Computer Vision and Pattern Recognition Workshop, San Francisco
11. Lyons MJ, Akamatsu S, Kamachi M, Gyoban J (1998) Coding facial expressions with Gabor wavelets. Proc. 3rd IEEE Int Conf on Automatic Face and Gesture Recognition, Nara, pp 200–205
12. Majumder G, Debnath R, Bhowmik MK, Bhattacharjee D, Nasipuri M (2012) Image registration of North-Eastern Indian (NEI) face database. Proc 1st Int Conf on Intelligent Infrastructure, Kolkata, pp 286–290
13. Martinez AM, Benavente R (1998) The AR face database. CVC Technical Report #24. Computer Vision Center, Barcelona, Spain
14. Pantic M, Valstar M, Rademaker R, Maat L (2005) Web-based database for facial expression analysis. Proc. IEEE Int Conf on Multimedia and Expo (ICME05), Amsterdam, pp 317–321
15. Roh M-C, Lee S—W (2007) Performance analysis of face recognition algorithms on Korean face database. Int J Pattern Recognit Artif Intell 21(6):1017–1033. doi:10.1142/S0218001407005818
16. Sim T, Baker S, Bsat M (2003) The CMU pose, illumination, and expression database. IEEE Trans Pattern Anal Mach Intell 25(12):1615–1618. doi:10.1109/TPAMI.2003.1251154
17. Turk M, Pentland A (1991) Eigenfaces for recognition. J Cogn Neurosci 3(1):71–86

Priya Saha received her B.E. (CSE) degree from National Institute of Technology, Agartala, Tripura, India in 2010 and completed her M. Tech (CSE) from Tripura University (A Central University), Suryamaninagar, Tripura, India in 2012. She was worked as a research assistant in a DeitY funded project in Tripura University from 2012 to 2013. She is currently pursuing Ph.D as a DST-INSPIRE Fellow, Govt. of India in Tripura University. Her topics of interests are related to the field of Facial Expressions and Emotions, Image Fusion, Face Recognition etc.

Mrinal Kanti Bhowmik received his BE, M. Tech (CSE), from Tripura University, India, in 2004 and 2007, respectively; and received his PhD (engineering) degree from Jadavpur University, India, in 2014. From July 2010 onward, he is working as an assistant professor at Tripura University. His research interests are related to the field of biometric, artificial neural network, information security, etc. He is also a member of the IEEE (USA).

Debotosh Bhattacharjee received the MCSE and PhD (engineering) degrees from Jadavpur University, India, in 1997 and 2004, respectively. He was associated with different institutes until March 2007. After that, he joined his alma mater, Jadavpur University. His research interests pertain to 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 (New Delhi), Indian Unit for Pattern Recognition and Artificial Intelligence, and a member of IEEE (USA).

Barin Kumar De received his B.Sc and M.Sc in Physics Honors from Calcutta University, India, in 1975 and 1977. He has completed his Ph.D in Atmospheric Science and Radio Propagation from Calcutta University. His research interests are related to field of Atmospheric Science and Radio Propagation. He is the former Dean of Science of Tripura University. He is a member of International Academy of Physical Science, Allahbad, Institute of Electronics and Tele Communication Engineering (IETE), Physics Academy of North-East (North-East India), SK Mitra Centre for Research in Space Environment, Calcutta University, Kolkata.

Mita Nasipuri received her BETel., METelE., and PhD (Engg.) degrees from Jadavpur University (J.U.), in 1979, 1981 and 1990 respectively. She 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).