PCA+LDA Method for Face Recognition using Neural Network

Dhiren Pandit, Jayesh Dhodiya

Abstract*— *Face recognition plays important role in many applications like video surveillance, retrieval of an identity from a database for criminal investigations and forensic applications. The face is considered as good bio-metric for many reasons: the acquisition process is non-intrusive and does not require collaboration of the subject to be recognized. The acquisition process of a face from a scene is simpler and cheaper than the acquisition of other bio-metrics as the iris and the fingerprint. On the other hand, many problems arise, because of the variability of many parameters like face expression, pose, scale, lighting, and other environmental parameters. Face recognition involved in application like problem of recognition of an identity in a scene. A system that automatically recognizes a face in a scene first detects it and normalizes it with respect to the pose, lighting and scale. Then, the system tries to associate the face to one or more faces stored in its database, and gives the set of faces that are considered as nearest to the detected face. This requires more computational resources and very robust algorithms for detection, normalization and recognition. In this paper we have implement different face recognition methods like Principle component analysis, Linear Discriminant Analysis and Fusion of PCA and LDA for face recognition. And better recognition rate is achieved by implementing neural network for classification.

Index Terms*—*PCA, LDA, FFNN, MLP, PCA-NN, LDA-NN

I. INTRODUCTION

So many algorithms have been proposed during the last decades for research in face recognition [3]. Bledsoe [2] use semi-automated face recognition with a human-computer system that classified faces on the basis of marks entered on photographs by hand. The principal components of the distribution of faces, or the eigenvectors of the covariance matrix of a face images, treating an image as point in a very high dimensional space is sought. The eigenvectors are ordered and each one accounting for a different amount of the variation among the face images. These eigenvectors are set of features that together characterize the variation between face images. Principal Component Analysis (PCA) method [7] is widely used for dimensionality reduction and recorded a great performance in face recognition. PCA based approaches can be divided in two phases: training and classification. In the training phase, an eigenspace is constructed from the training samples using PCA method. In classification phase, an input face is projected to the same eigenspace and classified by an appropriate classifier such as Euclidean distance [12] or Bayesian [9]. PCA encodes information in an orthogonal linear space while the Linear Discriminant Analysis (LDA) method encodes discriminatory information in a linear separable space of which bases are not necessarily orthogonal. Researchers have demonstrated that the LDA based algorithms outperform the PCA algorithm for many different tasks [1]. In this paper, the PCA and LDA methods are used for dimensionality reduction and feedforward neural network (FFNN) classifier is used for classification of faces. The proposed methods are called PCA-NN and LDA-NN respectively. The methods consist of two phases which are the PCA or LDA preprocessing phase, and the neural network classification phase. The proposed systems show improvement on the recognition rates over the conventional LDA and PCA face recognition systems that use Euclidean Distance based classifier.

Manuscript Received on April 2015.

Prof. Dhiren Pandit, Science & Humanities Dept., LDRP-ITR, Gandhinagar, India.

Dr. Jayesh Dhodiya, AHMD, SVNIT, Surat, India.

II. PCA AND LDA METHODS FOR FACE RECOGNITION

Let X be a d-dimensional feature vector. Here d is equal to the number of pixel of each face image. Therefore methods for reducing the dimensionality of such image space are required. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are used for dimensionality reduction. Principal Component Analysis [12] [4] is defined by the transformation:

$$y_i = W^T x_i \tag{1}$$

Where $x_i \in X \subseteq R^d$, i=1,2,3,…,n (n samples). W is a d-dimensional transformation matrix whose columns are the eigenvectors related to the eigenvalues computed according to the formula:

$$\lambda e_i = S e_i \tag{2}$$

S is covariance matrix defined by

$$S = \sum_{i=1}^{n}(x_i - m)(x_i - m)^T , m = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3}$$

This transformation is called Karuhnen-Loeve transform. It defines the d-dimensional space in which the co-variance among the components is zero. Hence it is possible to consider a less number of "principal" components exhibiting the highest variance. In the face space, the eigenvectors related to the most expressive features are called "eigenfaces". The Linear Discriminant Analysis is defined by the transformation:

$$y_i = W^T x_i \tag{4}$$

The columns of W are the eigenvectors of $S_W^{-1} S_b$, where S_W is the within-class scatter matrix, and S_b is the between-class scatter matrix. It is possible to show that this choice maximizes the ratio $\frac{\det(S_b)}{\det(S_W)}$. These matrices are computed as follows:

$$S_W = \sum_{j=1}^{c}\sum_{i=1}^{n_j}\left(x_i^j - \mu_j\right)\left(x_i^j - \mu_j\right)^T ;\ \mu_j = \frac{1}{n_j}\sum_{i=1}^{n_j} x_i^j \tag{5}$$

Where x_i^j is i^{th} pattern of j^{th} class and n_j is number of patterns for the j^{th} class.

$$S_b = \sum_{j=1}^{c}(\mu_j - \mu)(\mu_j - \mu)^T ;\ \mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{6}$$

LDA transformation is strongly dependent on the number of classes (c), the number of samples (n), and the original space dimensionality (d). It can be shown that there are atmost $c - 1$ nonzero eigenvectors. $c - 1$ being the upper bound of the discriminant space dimensionality. We need $d + c$ samples atleast to have a nonsingular S_W.

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

Consequently, an intermediate transformation is applied to reduce the dimensionality of the image space. For this we used the PCA transform.

III. NEURAL NETWORK

Neural Networks are widely used in pattern recognition because of their ability to generalize and to respond well to novel patterns. During training phase neurons are taught to identify specific patterns. If a novel pattern is obtained each neuron selects the output that corresponds to the training pattern that is slight different from the input. According to Haykin [6][99], a Neural Network is a massively parallel distributed processor that has a natural prosperity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:

1) Knowledge is acquired by the network through a learning process.

2) Inter-neuron connection strengths (synaptic weights) are used to store the knowledge.

The required parameters to train neural network are as follows.

A. Multilayer Perceptrons (MLPs)

MLP is one of the most popular neural network models for solving pattern classification and image classification problems. Because of their ability to learn complex decision boundaries, MLPs are mostly used in practical applications involving classification. When connection weights in a MLPs have been learn, the network can be used repeatedly for classification of new input patterns.

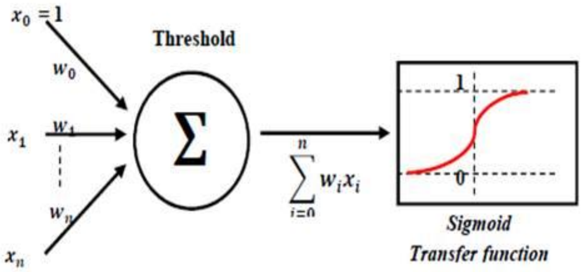

Fig. 1 Sigmoid function for MLP

We imply feed-forward networks and Back-propagation algorithm while referring MLPs. A typical topology of a fully connected feed-forward network is shown in Fig. 2. Backpropagation algorithm is a variation of Delta rule. When we pass input to ANN forwardly, the word Back in Backpropagation algorithm refers to the direction to which the error is propogated.

Stochastic Gradient Descent version of BACKPROPOGATION Algorithm **[8]**

• Initialize all network weights to small random numbers.

• Untill the termination condition do:

Propagate the input forward to the network and compute the observed outputs.

Propagate the errors backward as follows:

For each network output unit k calculate its error term

$$\delta_k = O_k(1 - O_k)(t_k - O_k)$$

For each hidden unit calculate its error term

$$\delta_h = O_h(1 - O_h) \sum_{k\ outputs} W_{kh}\delta_k$$

Finally update each weight $W_{ji} = W_{ji} + \Delta W_{ji}$, where $\Delta W_{ji} = -\eta \delta_j x_{ji}$.

Algorithm gives basic steps for the stochastic gradient descent version of BACKPROPAGATION algorithm [8].

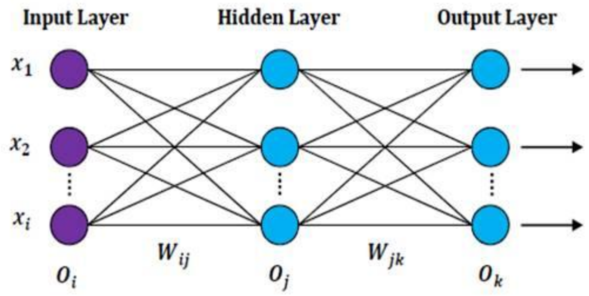

Fig. 2 Fully connected, Feed-Forward MLP network

A training sample $X = (x_1, x_2, \ldots, x_i)$, is pass to input layer. Weighted connections exits between each layer, where W_{ij} denotes the weight form a unit j in one layer to a unit i in the previous layer. We have an input layer (i) consisting of input nodes and an output layer (k) consisting of output nodes. The input nodes are connected to the output nodes via one or more hidden layers (j). The nodes in the network are connected together, and each of the links has a weight associated with it. The output value from a node is a weighted sum of all the input values to the node. By changing the different weights of the input values we can adjust the influence from different input nodes. For facial expression recognition the input nodes will typically correspond to image pixel values from the face image.

The output layer will correspond to classes or individuals in the database. Each unit in the output layer can be trained to respond with + 1 for a matching class and 0 for all others. In practice real outputs are not exactly + 1 or 0, but vary in the range between these values. Classification is done by finding the output neuron with the maximal value. Then a threshold algorithm can be applied to reject or confirm the decision.

B. Error Function

The total error, which in standard gradient descent version of BACKPROPAGATION is the SSE (Sum of Squared Errors):

$$E(\bar{W}) = \sum_{n\ pattern} \sum_{k\ output} (t_{n,k} - O_{n,k})^2 \qquad (7)$$

Since this E is actually a function of the network's weight vector, we conclude that E is multidimensional parabola. Gradient descent starts with an arbitrary weight vector and tries to minimize E at each step.

C. C. Hidden Layer

To determine the best number of hidden units we need to train several networks and find the generalization error of each.

If units are very less then there will be high training error and high generalization error due to under fitting and high statistical bias. On the other hand, the training error can be made as small as desired by adding more neurons, but generally each additional unit will produce less and less benefit. We should take into account the cost in processing time and storage requirements for each extra unit. Beside a relatively large number of neurons in the hidden layer can give high generalization error due to over fitting and high variance.

D. BACKPROPAGATION Parameters

Learning rate and the Momentum term are main parameters of BACKPROPAGATION. The learning rate, W_{ji} is a scaling factor of the gradient of the error function. The momentum term is an extra factor,

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

which is added to the term and makes it more or less dependent to the weight update of the previous step in the algorithm.

1) Learning Rate:

In BACKPROPAGATION algorithm weights are updated by

$$\Delta W_{ji} = -\eta \delta_j x_{ji} \tag{8}$$

This is called learning rate of the BACKPROPAGATION algorithm. With standard steepest descent, the learning rate is held constant throughout training. If the learning rate is set too high, the algorithm may oscillate and become unstable. If the learning rate is too small, the algorithm will take too long to converge. It is not practical to determine the optimal setting for the learning rate before training. This is most likely obtained by trial and error.

2) Momentum Term :

A common modification, of the basic weight update rule is the addition of a momentum term. By adding this term to the formula of the final step in BACKPROPAGATION, we obtain the following update rule:

$$\Delta \mathrm{W}_{\mathrm{ji}} - \eta \delta_j x_{ji} + \alpha \Delta \mathrm{W}_{\mathrm{ji}}(\mathrm{n}-1) \tag{9}$$

Therefore the update in iteration is affected by the update in nth iteration multiplied by a factor α, called momentum. Momentum takes values in the range $0 \leq \alpha \leq 1$. Empirical evidence shows that the use of a momentum in the BACKPROPAGATION algorithm can be helpful in speeding the convergence and avoiding local minima in the error surface.

E. Input Standardization and Weights Initialization

The main emphasis in the NN on initial values has been on the avoidance of saturation, hence the need to use small random values. Symmetry breaking in the weight space is needed in order to make neurons compute different functions. If all nodes have identical weights then they would respond identically. Therefore the gradient, which updates the weights, would be the same for each neuron. This way the weights would remain identical even after the update and this means no learning. Small weights (as well as small inputs) are needed to avoid immediate saturation because large weights could amplify a moderate input to produce an extremely large weighted sum at the inputs of the next layer. This would put the nodes into the flat regions of their nonlinearities for sigmoid saturation and learning would be very slow because of the very small derivatives.

F. Training Stopping Criteria

Following are four basic termination conditions when training an ANN:

- ***Fixed number of iterations.*** Iterations refer to the number of times the total training set is being presented in the Neural Network.
- ***Use threshold for the error.*** Empirically estimate a certain value for the error, which considered being acceptable.
- *Use threshold for the error gradient.* Usually we have to restrict training to steps which error gradient is larger than a fixed value. Small changes in error gradient mean that training reached a minimum (local or global) and it would be wise to stop without delay.
- *Early stopping.* Divide the available data into training and validation sets. Commonly use a large number of hidden units and very small initial values. Compute the validation error rate periodically during training. Finally, stop training when the validation errors rate 'start to go up'.

We can combine stopping criteria when constructing Neural Networks.

G. Generalization

Generalization is the ability of capturing the underlying function [6], [8] [10], during the training phase, and hence producing correct outputs in response to novel (patterns patterns that has not seen before). A system then is said to generalize well. From a statistic perspective the generalization error can be considered as the summation of a variance and a bias term:

$$E_{gen} = Variance + Bias^2 \tag{10}$$

Minimizing the generalization error is not equivalent to selecting a model where the bias is zero. This is because the model variance penalty may be too high. This is called the bias/variance trade-off.

There are a few conditions that are typically necessary although not sufficient -for good generalization:

- In order to generalize well, a system needs to be sufficiently powerful to approximate the target function. If it is too simple to fit even the training data then generalization to new data is also likely to be poor.
- The inputs contain sufficient information pertaining to the target, so that really exists a concept that relates inputs with correct outputs. We cannot expect a network to learn a nonexistent function or a non-existed classification rule.
- In general, the training set must be a representative subset of the theoretical population. A poor set of training data may contain misleading regularities not found in the underlying function/classifier.

IV. FACE RECOGNITION USING PCA+LDA FUSION METHOD WITH NEURAL NETWORK

The Face Recognition using PCA+LDA fusion method with Neural network shown in Fig.4 contains five different levels of recognition process which are as follows

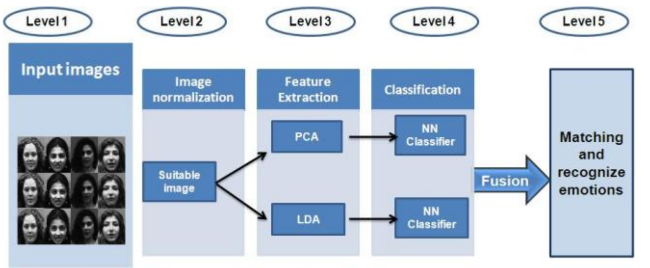

Fig. 4 Recognition process

At the initial stage multiple face images of different persons are entered in proposed system at Level 1. This set are normalized at Level 2 by applying different normalization method like, Gamma Correction (Balancing lightness and darkness)[5], Dimension reduction (RGB to Gray Scale), Histogram Equalization (enhancement of contrast) and Resizing Image [5]. At Level 3 suitable feature vectors called Eigen images are extracted parallel by the methods Principle Component Analysis (PCA)[11] and Linear Discriminant Analysis (LDA).

At the Level 4 the eigen images of both PCA and LDA are classified by Neural Network algorithm especially Multilayer feed forward and Back propagation algorithm of artificial neural network. At this level the images are classified according to the different face images and it is called as training set. Here the newly entered face image is classified on the basis of this training set for respective emotion. At Level 5 the output (Classified Images of emotion) by neural network are combined and system matches the final output with stored image sets and will exposed the suitable emotion.

A. Face images of different persons(Level -1)

This is Level 1 of proposed system in which face images of different persons are given as input images.

B. Image Normalization (Level -2)

This elaborates the techniques to normalize the converted face images in order to increase the efficiency of feature extraction and classification of face image. To normalize the face image the following set of mathematical techniques are adopted.

1) Gamma correction
2) Dimension reduction - RGB to gray
3) Histogram equalization
4) Image Resizing

1) Gamma Correction (on RGB Image): Gamma correction balance darkness and lightness of an image, which improve the visibility of an image. Especially in facial feature extraction it is necessary that the prime feature of face area must be visible separately from the other part of the image.

Gamma function: If X is an original value of an image then the new value of X is obtained as;

$$X_{NEW} = X^{Gamma} \quad (11)$$

Fig. 5 shows several gamma curves demonstrating the effect that the gamma value has on the shape of the gamma curve

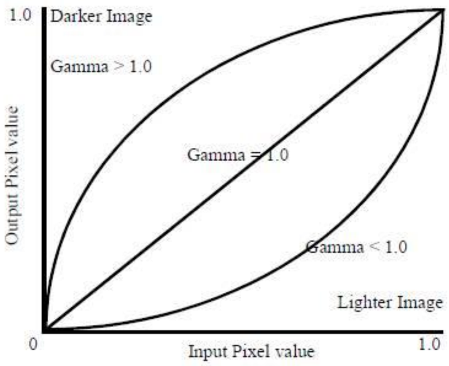

Fig. 5 Gamma Curve

2) Dimension Reduction: *(RGB to Gray):* Basically the colored image has three dimensional array pixel value $M * N * 3$ of Red, Green and Blue (RGB)[5]. It is quite difficult to perform image processing technique on three dimensional matrix in MATLAB, so for convenient and smoother processing the image is converted in gray format. A gray image has two dimensional $M * N$ pixel value which lies between [0,255] as per the gray value of an image pixel.

I = rgb2gray(RGB) converts the true color image RGB to the grayscale image I.

The MATLAB function to perform this operation is IamgeGray = rgb2gray(ImageGamma);

3) Histogram Equalization: In a feature extraction it is necessary that the gray image must has high contrast so that the gray values of each pixel become significantly differ from each other. The better variation in range [0, 255] makes the identification feature better. The Histogram Equalization [5] enhances the contrast of images by transforming the values in an intensity image so that the histogram of the output image is approximately flat.

The MATLAB function to perform this operation is Image-Histogram = histeq(ImageGray, hgram);

4) Image Resizing: Since different face images are considered, there is a possibility that the different face images are extracted in different sizes, and also for image processing in MATLAB the dimensions of each images are required to be same. To make the image process smoothly it is necessary to make all the expression images same sized, as per desired but not so small not so big and in sufficient sized that never affect the quality of result [5].The MATLAB function to perform this operation is.

ImageResize=imresize(ImageHistogram,[numrowsnumcols]

C. Feature Extraction (Level -3)

Features of face images are considered as Eigen images that represents the whole face with faces of different persons. Different mathematical techniques are available for feature extraction, here we used (I) Principle Component Analysis and (II) Linear discriminant Analysis approaches on same face images.

D. Classification (Level -4)

At Level 3 eigen features of normalized face images have been obtained by PCA and LDA which are classified for face recognition at this level using neural network (NN) individually. At Level 4 neural network is trained to classify PCA and LDA eigen images by using Multilayer Perceptron (MLP) with Back Propagation classification algorithm. In this level the NN parameters like, Error surface, number of hidden layers, learning rate, momentum, input standardization, eight initialization, training and stopping criteria, generalization of NN is decided. We adopt following terminologies of NN in Fusion method for face recognition

1) Multilayer Perceptrons *(MLPs)* To train NN for face recognition from face images we have used Multilayer Perceptorns of NN which includes one hidden layer as in pattern recognition it is universal approximation and four output layers as basic four faces of four different persons are considered. Log-sigmoid transfer function is used at both the layers hidden as well as output layer to restrict the output in [0, 1]. Since we have taken outputs range of 0.1 to 0.9 (0.9 for confident, 0.1 for not confident) log- sigmoid is the ideal for system.

2) Weight initialization In the training of NN a problem is to initialize weights at hidden and output layer. There are three ways for that (1) hidden-zero, output-zero, then no update occurs, which means that weights remain zero and performance constant so it has not been chosen (2) hidden random, output-random, then the output of face images are noisy and the face can be visible.(3) hidden zero, output random, then the output as clear appearance of face images. Thus, in the training of NN the third option that hidden layer is initialized with zero weight and output layer initialized with random weight is chosen.

3) Error in Classification In the classification of face the two processes are performed by NN on dataset, training and testing.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

During this procedure NN finds the percentage of accuracy for classification of face from eigen images in training and testing process and finds the related error in both the process in the term of Mean Square Error (MSE). The MSE shows the mean of squared error in the different between output and target, smaller MSE means the output of the system is closer to the target function to be minimized.

4) Number of neurons in hidden layer (size of hidden layer) To choose optimum number of neurons in hidden layer prior training or testing is still an issue of research, it is totally based on experiment and on the choice of other parameters of NN. In the implementation of system for face recognition from expressions of a person on the trial and error base we have chosen 16 neurons in hidden layer, at which the accuracy of classification of face is higher and the test MSE is lower.

5) Learning rate Back propagation parameter In the training of NN for classification of face image the value of learning rate decides how quick output converges to target. Experiment shows that the larger value of learning rate increase MSE in classification, while at low learning rate NN training process long time as because update in weight too small. Thus the moderate learning rate between 0 to 1 is most preferable. Here we have chosen 0.9 learning rate.

6) Momentum Back propagation parameter Momentum makes the training process smoothly by speeding the convergence and avoiding local minima in the error surface but it only works when learning rate is quite small. The very small and very large value of momentum increase MSE in training, thus by experimental results we have chosen 0.6 a moderate value of momentum for NN classification.

7) Input Standardization The contribution of an input will depend heavily on its variability relative to other inputs. If for example one of the inputs has range of 0 to 1 and another has a range of 0 to 1000, then the contribution of the first input will be swamped by the second input. So it is essential to rescale the inputs such that their variability reflects their importance. Here, in the training of NN we have used PCA or LDA feature vector as input, here PCA is centralized with zero mean and 1 standard deviation before feeding in NN while is LDA already rescaled in image data. So both the inputs are rescaled before trained in NN.

E. Fusion of PCA-NN and LDA-NN (Level -5)

After Level 4 the PCA-NN and LDA-NN classified face images are fused in this Level 5 using score based strategy discussed as under. If the score functions are directly comparable or if there exists at least one acceptable transformation scheme to make the involved classifiers comparable, score based strategies are good ways for decision process. In this work, NN is used as a classifier for both systems (PCA and LDA), hence outputs of both systems are in same format, so we select score based strategy (SBS) as combiner

Algorithm of Score Based Strategy (SBS)

```
step 1   Assemble Label, PCA-NN and LCA-NN Score(S) of
           faces
step 2   Set threshold value(TV) for PCA-NN and LDA-NN
step 3   If PCA-NN(S) > PCA-NN(TV) and
         LDA-NN(S) > LDA-NN(TV)
         If PCA-NN(S) > LDA-NN(S)
         PCA-NN(S)
         Else
         LDA-NN(S)
         End
         Else Go to Step 4
         End
step 4   If PCA-NN(S) > PCA-NN(TV) and LDA-NN(S) <
         LDA-NN(TV)
         PCA-NN(S)
         Else if PCA-NN(S) < PCA-NN(TV) and DA-NN(S) >
         LDA-NN(TV)
         LDA-NN(S)
         Else PCA-NN(S) < PCA-NN(TV) and
         LCA-NN(S) < LCA-NN(TV)
         Access Denied
         End
```

Where, TV = Threshold Value , S = Score

Neural Network gives the individual PCA and LDA accuracy score of classification of a face in percentage which are used as PCA-NN and LDA-NN scores at Step 1 of algorithm.

The threshold value is consider as expected accuracy we want from the NN classification as per our expectation that can be set manually, for both PCA-NN and LDA-NN classifier at Step2 which can be same or differ.

In Step 3 conditional approach is used for both PCA-NN and LDA-NN, if both classification scores crossed their limit of threshold value then the higher classified score is accepted otherwise in failure situation it indicates not to reach the expected level of accuracy. In winning situation if one of the classifier crosses the threshold limit at Step 4 then its accuracy score will be considered as final matching score.

F. Face Recognition (Level-5)

After fusion of PCA-NN and LDA-NN, to identify a face, system needs the training set of a face images of individual persons.

At this level-5 we consider face images according to the class of different persons (E1,E2, ...,Ek) by NN with same feature extraction techniques (PCA and LDA) is called as “Training Set” which have been already stored in our database. The more customized face images makes the database stronger for highly accurate classification, so the update in training set is much essential for face recognition. When the new face image or sequence of images are entered as test image, at that time proposed method performs all processes of Levels 1 to 4 on it and obtained an output which is matched with all the stored face class(E1,E2, ...,Ek) of an individual. At this level system finds the association between output face and stored classes of categorized faces then identifies the correct face.

V. RESULTS AND DISCUSSION

Here we have considered the face images of different persons for recognition, which we call as train images. From these images our algorithm is trained for different class of faces. Then we normalize sequence of images. Here we did classification using Neural network.

A. Flow of Process

Following is the flow of entire process of Classification of nature of the video. Entire process is divided in five different levels.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

- ***Level 1 Video to Image sequence***
  At first level sequence of face images are given as input. Here in train images we have consider all face images with different persons and test images for which we are seeking recognition of person.
- ***Level 2 Image Normalization***
  In second level normalization of face image is done. Here face image is crop according to face portion of a person in that image. After these images are converted to grey scale images for ease of calculation.
- ***Level 3 Feature Extraction***
  We use Principle component analysis and Linear discriminant analysis for feature extraction. At this level eigen faces and fisher faces which contains important characteristic of images are extracted from both train and test images.
- ***Level 4 Classification***
  At level 4 image classification is done. For classification purpose we have use Neural network with back propogation algorithm.We use same classifiers in PCA and LDA both. We call it as PCA+NN and LDA+NN.
- **Level 5 Recognition**
  In level 5 matching is done by combining both technique PCA+NN and LDA+NN. We have use fusion of PCA and LDA technique for this purpose. And with this classification we obtain recognized face image. And with the help of no of such recognized images we classify number of images.

B. Results and Discussion

- For experiment we have taken 200 different images of 4 persons for train purpose. All these images will train our algorithm for recognition.
- Once algorithm is trained we have taken test images.
- Figure 6 are few images extracted from set of train Images.

Fig. 6 Set of Train Images

- We have passed 40 test images of a person as input and obtained the following results.
- All 39 images are correctly classified by PCA+LDA fusion method using neural network.
- Hence we get 97.5% classification rate for this set of train and test images.

VI. CONCLUSION

We used PCA+LDA method with classification using Neural network to recognize face of individual person. PCA+LDA method with classification using neural network gives very good recognition rate. Hence PCA+LDA fusion method with classification using neural network is good method for face recognition over PCA and LDA.

REFERENCES

1. Belhumeur P., Hespanha J., Kriegman D. Using discriminant eigenfeatures for image retrieval. PAMI, 19(7):711720, 1997.
2. Bledsoe W. Man-Machine Facial Recognition, Panoramic Research Inc.Palo Alto, CA, (1966) Rep. PRI:22.
3. Chellappa R., Wilson C.L.. Human and machine recognition of faces:A survey, Proc, IEEE 83Vol. 5 (1995) 705-741.
4. Duda R.O., Hart P.E., Stork D.G., Pattern Classification, John Wiley Sons, USA (2001)
5. Gonzalez R. Digital Image Processing Using MATLAB, Publisher-Tata McGraw-Hill Education, ISBN 0070702624, 9780070702622, 2013.
6. Haykin S. Neural Networks: A Comprehensive Foundation , Prentice Hall Inc.,New Jersey, 1999.
7. Kirby M., Sirovich L. Sirovich, Application of the Karhunen-Loeve Procedure for the Charac terization of Human Faces, IEEE PAMI, Vol. 12 (1990) 103-108.
8. Mitchell T. Machine Learning, , McGraw-Hill, Editions International,New York, 1997***
9. Moghaddam B., Pentland A. Probabilistic visual learning for object recognition. PAMI, 9(7):696710, 1997.
10. Russell D. , Robert J. Neural Smithing, Supervised Learning in Feed forward Artificial Neural Networks (A Bradford Book), The MIT Press, 1999
11. Tunglin D. Facial Expression Classification Using PCA and Hierarchical Radial Basis Function Network", Journal of Information Science And Engineering 22, pp. 1033-1046, Taiwan, 2006.
12. Turk M. , Pentland A. Eigenfaces for Recognition, Journal of Cognitive Neuroscience, Vol. 3, (1991) 71-86.

AUTHORS PROFILE

Prof. Dhiren Pandit- Education: Research Scholar, SVNIT Surat, Gujarat, M. Phil. From Veer Narmad South Gujarat University, Surat, M.Sc. from The M.S. University, Baroda, Gujarat, **Current affiliation**:-Pursuing Ph.D. in the area of Image processing from SVNIT Surat, Gujarat, India. Assistant professor at LDRP-ITR Gandhinagar, Gujarat (9 Years Teaching experience till date) **Publications**: Publications in International Journals/Proceedings :2 Presentation: Various paper presentation at national and international conferences

Dr. Jayesh Dhodiya, Education: Ph. D, M.Sc.(Gold Medalist), G.P.S.C(2010), CCC+ **Teaching Experience:** Total=11 Years P.G=6.3 Years U.G=3 Years **Current affiliation**: Assistant Professor at Sardar Vallabhbhai National Institute of Technology, Surat, Gujarat, India **Research Projects:** 04(On Going) **Expert Talk:** 12 **Publications**: Publications in International Journals : 12 Publications in Proceedings of International Conferences: 03Publications in National Journals: 05 Publications in Proceedings of National Conferences: 06 **Attended number of various national and International conferences / workshops/ Seminars in various area of Mathematics**

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication