IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 20, Issue 1, Ver. III (Jan.- Feb. 2018), PP 08-15
www.iosrjournals.org

Face Recognition Using Coupled the Hidden Markov Model With An Artificial Neural Network

Sandeep Saxena[1], Dr. Mansaf Alam[2], Dr. B. Mohd. Jabarullah[3], Dr. Rakesh Kumar Yadav[4]

[1]*(Research Scholar, Computer Application, IFTM University, India)*
[2]*(Computer Science, JMI University, India)*
[3]*(Computer Application, PIT, India)*
[4]*(Computer Science & Engg., IFTM University, India)*
Corresponding Author: Sandeep Saxena1

Abstract: *Face recognition is being used in a variety of fields because of its advantages such as a non-contact process, fast and accurate results, reliable matching, and diverse applications. Machine recognition of faces can be classified into two types: still image recognition and video image recognition of a person from video image face databases. Nowadays, face recognition or verification systems have a wide range of commercial and law enforcement applications. To achieve good recognition results, this paper proposes a coupled the Hidden Markov Model (HMM) with an Artificial Neural Network (ANN) to recognize the face image. The proposed system detects the facial region and recognizes the faces using the existing video face databases and finally, the system is experimentally analyzed. The proposed method is ascertained that it has small observation vector set, reduced number of transitions among states, low computing time, and most promising recognition accuracy.*

Keyword :– *accuracy, Artificial Neural Network, face recognition, Hidden Markov Model*

Date of Submission: 16-01-2018 Date of acceptance: 17-02-2018

I. Introduction

A face recognition system from video is a computer application having the ability of identifying or verifying a person from a video frame of a video. As shown in Figure 1, this system has three basic steps, namely face detection, feature extraction, and recognition. One of the ways of performing face recognition is by comparing selected facial features from the created image to those from the video database.

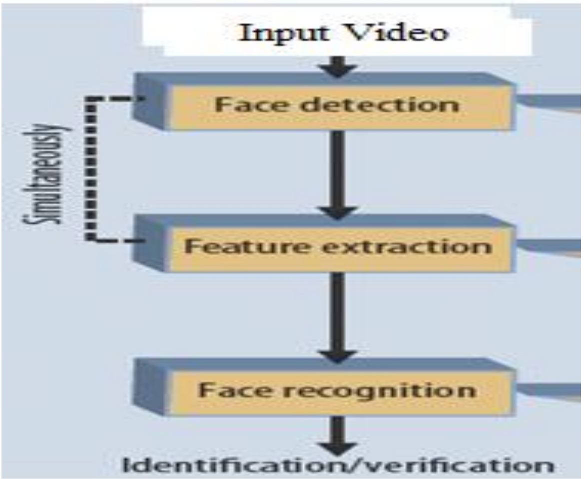

Fig. 1: Steps of video face recognition

In recent years, video face recognition has been used on a larger scale, with much emphasis on further research and innovation such as video surveillance and access control. The face recognition process [1-2] has two basic steps: face detection and classification. Several algorithms for detecting and matching faces with those in the database have been developed. An Artificial Neural Network (ANN) is successfully applied for face detection [3-6] with applied high detection rate. Several studies have proposed different models of ANN. A challenge is to identify the most appropriate neural network model, which can work reliably for solving realistic problems. Forface recognition, a Hidden Markov Model (HMM) with feature classification function is developed [7-10].

The proposed system comprises of two main steps. The first step is identifying the facial region of a raw image (face detection) and the second step is face recognition to identify an individual. Localization and normalization (face detection and alignment) are processing steps that are performed before face recognition (facial feature extraction and matching). However, HMM along with ANN is used for detection in combination with face recognition because of its potential to yield better results. The method developed in this study is thus warranted. The detailed descriptions are given Literature Review in Section 2. In Section 3, elaborates the proposed method. Section 4 briefly discusses Experimental Discussion. Conclusion is presented in Section 6.

II. Literature Review

For face detection, researchers have proposed different models using ANNs. However, identifying the most appropriate neural network model that can work reliably and produce optimal result is challenging. Yang et al. [3] attempted to determine the effective parameters that influence the performance of neural network face recognition systems. They concluded that a robust face detection system can be affected by factors such as lighting conditions, orientation, pose, partial occlusion, facial expression, and presence of glasses, facial hair, and a variety of hair styles.Rowleyet.al. [4] presented a Neural Network-based algorithm to detect frontal views of faces in grayscale images. The authors presented a straightforward procedure for aligning positive face examples for training. To collect negative examples, they used a bootstrap algorithm for training the networks that add false detections into the training set as training progresses. This eliminates the difficult task of manually selecting non face training examples that must be selected to span the entire space of non-face images. The proposed system has better performance in terms of face detection and false positive rates compared with other face detection systems.Pankaj et al. [5] proposed a method for fuzzy neural network classification based on Integrated Adaptive Fuzzy Clustering (IAFC). IAFC forms cluster boundaries by using a combined similarity measure and by integrating the advantages of the fuzzy c-means, adaptive resonance theory, and a Fuzzified Kohonen-type learning rule. The system achieved a higher recognition rate for the AT&T and Yale databases, respectively, higher than that of the back propagation neural network system. Considering the rejection rate for the no registrants, the system achieved an equal error rate of 3.7% and 1.3% for the AT&T and Yale databases, respectively. Thus, the proposed extension of ANN is better than conventional ANN. Sinha.et. al [6] presented a method for detecting the location of frontal faces in a small amount of time. The method involved creating a dataset from a database of faces that has localized parts of facial images along with the corresponding scores of the facial regions being detected. Using the obtained scores, further processing is performed in the corresponding facial area to determine whether a frontal face is present in the given region. The detection can be achieved in a small amount of time because the neural network can identify parts of a face from only one kernel. Nowadays, various decision-making systems are utilized for implementing face recognition systems. HMM plays an important role in resolving many of the related issues in face recognition because HMM, after proper training, can function as a powerful tool for facer ecognition. A major challenge faced by any face recognition system is its ability to identify images. Therefore, it is necessity to review HMM.The mathematical theory of HMMs is originally described in 1970s. The most notable initial efforts for face recognition are made by Samaria & Young (1994). HMM have five states, each stage modeling a specific area of a facial image. The HMM is a promising method that attention as a complex task because of noticeable changes in the appearance caused by illumination, lighting, variations in facial expression, size, orientation, and other external factors. Samaria et al. [7] investigated the factors that affect the choice of the model type and parameters. The authors obtained successful results for different facial expressions, lighting condition, and orientation changes. Experiments are performed to evaluate the performance of the HMM-based approach and the results are better than those with the Eigen face method. Nefian.et.al [8] presented a method that calculates the observation vectors used to characterize the states of HMM. The HMM is developed using the coefficients of the Karhunen Loeve transform. The developed system reduced the computational complexity slightly improving the recognition rate.Phad.et al. [9] proposed a method that uses both feature and score levels. Embedded HMM (E-HHM) is used in this research. The performance of the proposed method depends on the choice of the model parameters. A discriminating set of multiple E-HMMs based on the face recognition algorithms proposed. Experimental results revealed higher recognition accuracy and higher generalization ability.Chihaaoui.et.al. [10] presented a HMM-Local Binary Pattern (LBP) model that used an LBP tool for classification of a 2D face image through feature extraction and used an HMM as a classifier. The method comprised four steps. First, the face image is decomposed into blocks. Then, the image features are extracted using LBP. The probabilities are subsequently calculated. Finally, the maximum probability is selected. The results of the developed method are better than those of the other methods.Various extensions of ANN and HMM have been used in face recognition as face detectors and classifiers, respectively, yielding satisfactory results. HMM is a strong statistical foundation for developing efficient learning algorithms. The performance of a face recognition system combining ANN and HMM depends on choosing the optimal model of face detection by ANN and optimized recognition by using HMM.

III. Proposed Method

Several studies have attempted to improve the accuracy and speed of video face detection and recognition. However, no ideal method that can yield satisfactory results exists. Therefore, in this study, a coupled HMM with ANN for face recognition is proposed, which is presented in Figures2-4. The proposed work flows in manner that the detection is accomplished by using ANN and recognizing by HMM.

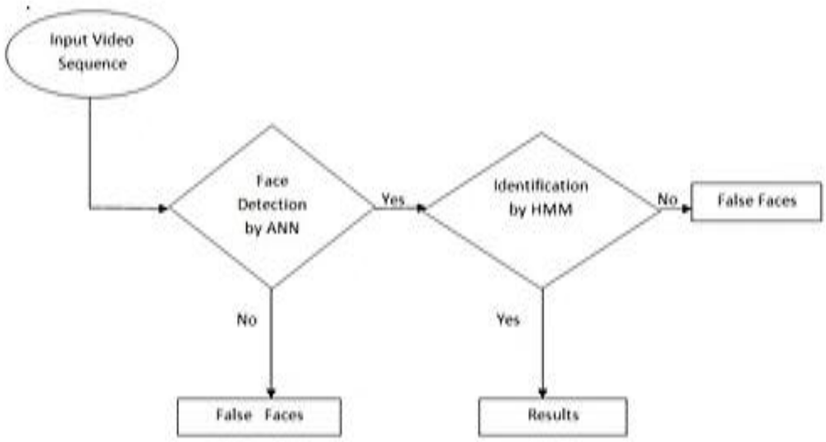

Fig. 2: Proposed Model of ANN-HMM

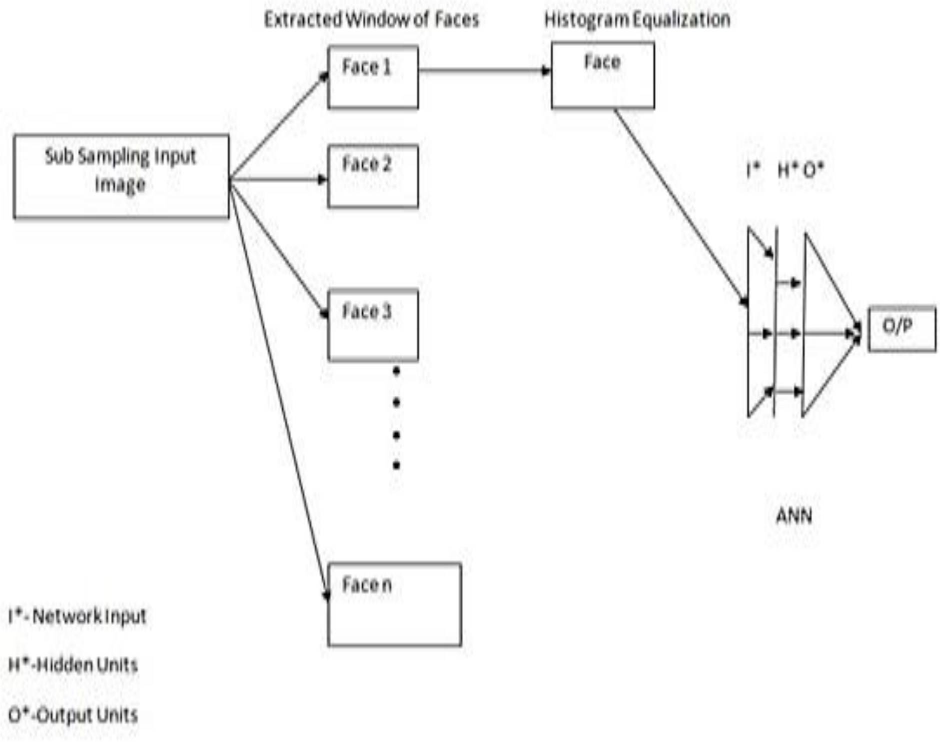

Fig. 3: Steps of Face Detection Using ANN

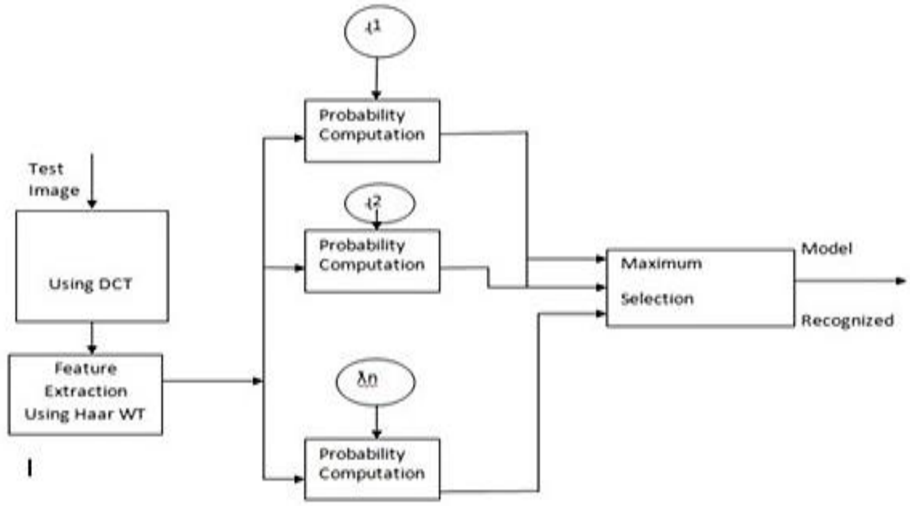

Fig. 4: Recognition Model Using HMM

3.1: Methodology of ANN for face detection:

The Gaussian pyramid method is used for extracting the window of faces. Furthermore, a histogram equalization method is used for equalizing the distribution of various intensities. For identifying the frontal-view of a face, an input is given to the ANN, as shown in Figure 3. Here, ANN is used to solve the following three fundamental problems: (1) learning effective features from a large feature set; (2) constructing weak classifiers, each of which is based on one of the selected features; (3) boosting the weak classifiers for constructing a strong classifier.The selected neural network shown here [3] is a 3-layer Feed Forward Neural Network with a back propagation algorithm. In this algorithm, given set of input output pair, provides a procedure for changing the weights in a Network to classify the given input patterns correctly, where changing the weight means the error is propagated back to hidden unit. The purpose of the neural network is to train the network to get a balance between the responses of network ability. In this, the number of input neurons is equivalent to the length of the extracted feature vector, and the number of output neurons is just one. The result is true if the image contains a human face and false otherwise. The number of hidden neurons is 2 × n + 1, where n is the number of input neurons. The system uses hidden layer with neuron to represent local features that characterize faces by using activation function that is sigmoid function. Sigmoid function is used in making the relation between the value of the function at a point and the value of the derivative at that point which reduces computational time during training:

$$f(x) = \frac{1}{1+\exp(-\lambda x)} \qquad (1)$$

Where λ is the steepness parameter.

The derivative of this function is

$$f'(x)=\lambda f(x)[1 - f(x)] \qquad (2)$$

3.2 Methodology Of Hmm For Face Recognition:

The output of the detection is a cropped image containing only the face. The HMM is used for reorganization, and each individual in the database is represented by a HMM face model. A set of images representing different instances of the same face are used to train each HMM. A set of 2D DCT coefficients is obtained from each block which is used to form the observation vectors. The observation vectors are effectively used in the testing and the Haar wavelet is used for further reduction of the extracted features, resulting in the problems of overhead and memory.HMMs represent a popular statistical pattern recognition technique and can be considered as a state-of-the-art face recognition technique. Consider 30 × 32 pixel facial region with a scanning window of 3 × 3 with 2 pixels overlap. The observation vector can be calculated as follows:

$$T = \frac{(P-R)}{(R-S)} + 1 \qquad (3)$$

$$T = \frac{(32-3)}{(3-2)} + 1 = 30 \qquad (4)$$

Where T = observation vector, P = height of the face, Q = width of the face, R = height of the overlapped block, and S = width of the overlapped block.Thirty observation vectors are obtained per scanning row both horizontally and vertically, and thus, 900 observation vectors are obtained in total. The observation vectors are divided between 5 states—30 observation vectors per row/5 states, resulting in 6 observation vectors per row in each state, that is, 6 × 30 observation vectors per state. Corresponding mixture of Gaussians defined the observation probability density function get through observation vectors which are extracted from each state. In the next step, the model parameters are re-estimated using the expected maximization procedure to maximize P (O|ƛ), where P is the probability of the model, O is the observation, and ƛ is the whole model. The iterations are stopped after the model convergence is achieved, that is, the difference between the model probability between consecutive iterations (k and k+1) is smaller than a threshold C.

$$|P(O|\lambda^{k+1}) - P(O|\lambda^{k})| < C \quad (5)$$

The probability of the observation vector for each HMM face model is computed. A face is recognized by obtaining the maximum probability of the model.

IV. Experimental Discussion

In the experimental discussion, Matlab 2014 with Windows 7 operating system on a Core 2 Duo 1.8-GHzPC with 4 GB RAMS is used. The experiments are performed using FJU and VIDTIMIT database obtained from the web. The FJU video database has been created for experimental usage for face recognition and tracking. The challenges of these videos include 9 poses, 10 illumination conditions, and glass effects. Our system comprised two phases: training and testing.

4.1 Detection by using ANN

Images are created by extracting video frame from FJU video database. In training, 14 set of face examples are created from each original each of 30 × 32 windows in the set is then preprocessed by applying lighting correction and histogram equalization. Larger variations in the translation and scale are managed by applying the filter at every pixel position in an image pyramid, in which the images are scaled by a factor of 1.2. The resulting image (30 × 32 pixels) is inputted to the ANN. The output of ANN is a real value between 0 (false) and 1 (true). The original image is decomposed into a pyramid of images using the Gaussian pyramid method as 4 blocks of 20 × 1 2 pixels, 16 blocks of 10 × 6 pixels, and 5 overlapping blocks of 24 × 12 pixels. Thus, the ANN has 4 + 16 + 5 = 25 input neuron. Every block has passed through hidden units for the purpose of detection of faces through local features. The objective of this technique is to find the important facial features. The system used hidden layer with 51 neurons to represent local features that characterize faces. For training the Neural Network, learning rate is initially defined 0.25 while momentum factor is less than o.9 which is in such a way that it generates 1 for face and other for non-face.

4.2 Recognition By Hmm

After detection, the faces are segmented using DCT and the extracted features are further reduced through Haar Wavelet transform. The HMM is used as a recognizer, as shown in Figure 3. Consequently, the HMM is used for preparing a lookup table for storing previous history to match data during execution. This technique reduces a large number of features into small number of features for frame selection. Therefore, this technique able to reduce the complexity of the system compared with the other techniques.For training the HMM, the no. of state five is defined. Observation vectors of each of the state are calculated and observation probabilities Corresponding mixture of Gaussians are obtained through observation vectors from each of the state. Now initial probability is set to 1 and remaining probabilities of states is yet to start that is transition probability which are dependent on occurrences of the transition. In this way, initial and transition probabilities are calculated by using viterbi algorithm which substitute the segmentation. These processes are continued until consecutive iteration is less than threshold.

4.3 Results Analysis on FJU Database:

In the first experiment, we used two sets of images for face detection. Each set of image had 14 images; therefore, we trained 14 × 2=28 images. Two images from each set are used for testing the system. That is, we used 2 × 2=4 images as test images. Table 1 Comparison of results of Conventional HMM algorithm and our proposed method.In the second experiment, we used 5 sets of images for face detection. Each set of images had 14 images; thus, we trained 14 × 5=70 images. Two images from each set are used for testing the system. That is, we used 2 × 5=10 images as test images. Table 1 Comparison of results of Conventional HMM algorithm and our proposed method.In the third experiment, we used ten of images for face detection. Each set of image had 14 images; therefore, we trained 14 × 10 = 140 images. Two images from each set are used for testing the system. That is, we used 10 × 2 = 20 images as test images. Table 1 Comparison of results of Conventional HMM algorithm and our proposed method.In the fourth experiment, we used 20 sets of images for face detection. Each set of images had 14 images; thus, we trained 14 × 20 = 280 images. Two images from each set are used for testing the system. That is, we used 2 ×20 = 40 images as test images. Table 1 Comparison of results of Conventional HMM algorithm and our proposed method.In the fifth experiment, we used twenty five sets of images for face detection. Each set of image had 14 images; therefore, we trained 14 × 25 =350 images. Two images from each set are used for testing the system. That is, we used 2 × 25 = 50 images as test images. Table 1 Comparison of results of Conventional HMM algorithm and our proposed method.

4.4 Comparative analysis of results

The proposed method is performing efficiently for different datasets of faces. This method requires rigorous training, which requires considerable time. However, once learned, it can efficiently recognize faces and improve itself. The tradeoff between time and accuracy is clear from Table 1.The data shown in Table 1 is based on the running the corresponding algorithms. The time given in the Table 1 is based on the training and testing of only one type of image, and accuracy depends on number of images. Each time an image is tested, the trained network is corrected, and thus, the accuracy depends as the number of tests performed. The Conventional HMM algorithm requires less time than our proposed method; it has an accuracy of only 94.5% and is highly dependent on the environment. Moreover, it fails in the cases of low light conditions and tilted faces. The comparison of recognition rate of the proposed method and the Conventional HMM face-based algorithm is defined. The proposed method recognition reaches up to 99.72% as compared with 94.5% of the conventional HMM face-based algorithm.

Table 1: Comparison of results of Conventional HMM algorithm and our Proposed Method

	Conventional HMM	PROPOSED METHOD				
Parameters		Experiment 1	Experiment 2	Experiment3	Experiment 4	Experiment 5
Video Input	25 fps, 360 X 240 pixels	25 fps, 360 X 240 pixels	25 fps, 360 X 240 pixels	25 fps, 360 X 240 pixels	25 fps, 360 X 240 pixels	25 fps, 360 X 240 pixels
Number of faces detected per frame	1	1	1	1	1	1
No. of Haar coefficients	--	2	2	2	2	2
Accuracy for same test dataset	94.5%	99.72%	99.68%	99.64%	99.59%	99.56%
Number of trained images	--	28	70	140	280	350
Number testing images	--	4	10	20	40	50
Running time	--	10.100 s (for training and testing)	23.172s	46.05s	92.168s	102.72S
Accuracy for different dataset for testing	84%	99.68%	99.65%	99.62%	99.58%	99.53%

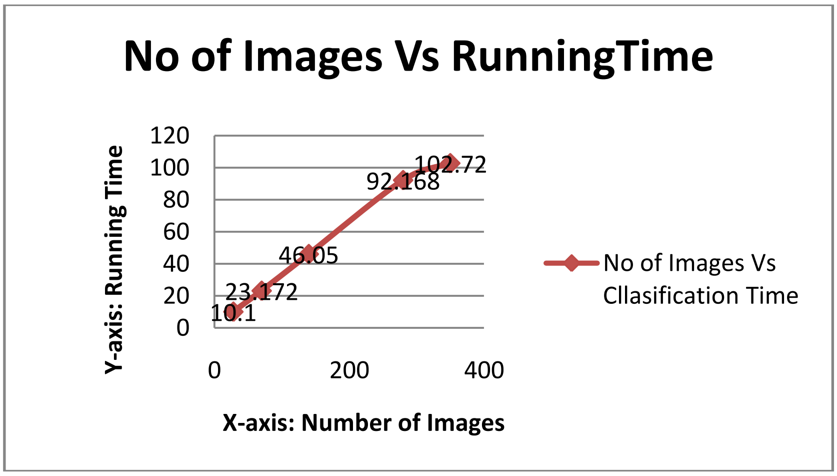

Fig. 5: Comparison of running time required by the algorithms for a set of images.

Fig. 5 shows the running time required by the algorithms for a set of images. Proposed method has every set has 14 images for training purpose. The running time required by algorithm increases as the sets of images increases.

4.5 Results Analysis On Vidtimit Database:

In the first experiment, we used two sets of images for face detection. Each set of image had 14 images; therefore, we trained 14 × 2 = 28 images. Two images from each set are used for testing the system. That is, we used 2 × 2 = 4 images as test images. The given below Table 2 shows Comparative analysis of accuracy on VIDTIMIT database.In the second experiment, we used 5 sets of images for face detection. Each set of images had 14 images; thus, we trained 14 × 5 = 70 images. Two images from each set are used for testing the system. That is, we used 2 × 5 = 10 images as test images. The given Table 2 shows the Comparative analysis of accuracy on VIDTIMIT database.

4.6 Comparative analysis of results:

Table 2: Comparative analysis of accuracy on VIDTIMIT database [11][12]

	HMM+2-D DCT METHOD	PROPOSED METHOD	
Parameters		Experiment 1	Experiment 2
Video Input	25 fps, 320x240 pixels	20 fps, 160x120 pixels	20 fps, 160x120 pixels
Number of faces detected per frame	1	1	1
No. of Haar coefficients	-	4	4
Accuracy for same test dataset	95%	98.75%	97.6%
Number of trained images	-	28	70
Number testing images	-	4	10
Running time	-	10.172s (for training & testing)	11.100s
Accuracy for different dataset	-	97.8%	97.0%

The table 1 clearly shows the measurement in the first experiment, used two sets of images, each set of image had 14 images; therefore, trained 14 × 2 = 28 images. Two images from each set are used for testing the system. That is, used 2 × 2 = 4 images as test images. The second experiment is used 5 sets of images. Each set of images had 14 images; thus, trained 14 × 5 = 70 images. Two images from each set are used for testing the system. That is, we used 2 × 5 = 10 images as test images. Accuracy of first experiment on the same dataset is 98.75% while in the second experiment it is 97.6%. It means as the no. of trained images and testing images increases, it slightly reduces the accuracy. But experiment done on VIDTIMIT Database [11][12] with the same no. of trained and tested images on the accuracy comes around 95%.

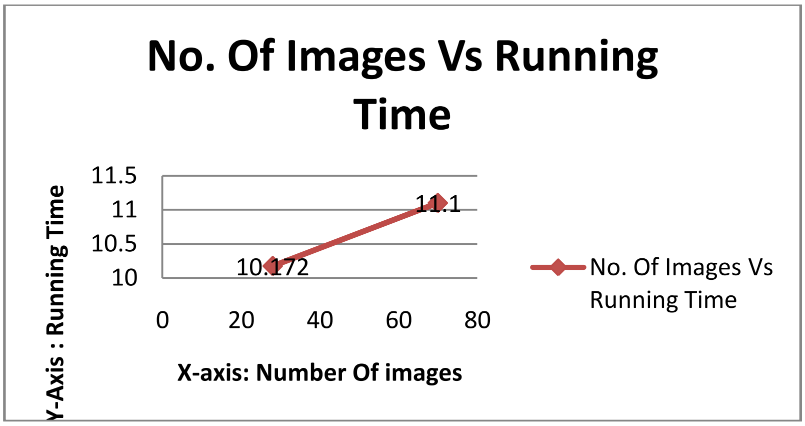

Fig. 6: Running time required by the algorithms for a set of images.

Figure 6 clearly shows that the no. of trained images and testing images increases, it slightly increases the running time.In the proposed method, when different no. of experiments performed on the same dataset and

different dataset, accuracy comes around 99.72% and 99.68% respectively while in the Conventional HMM algorithm, the recognition of comes around 94.5% on the same dataset and it differs very much such as it comes around 84% on different datasetOur experiment performed on 28, 70, 140, 280,350 no of trained images and 4, 10,20,40,50 no. of testing images with the same and different dataset. In both cases On FJU video database, running time increases when the trained and testing images for the different experiment increases respectively.
The basis of experiments performed on FJU video database, proposed method gives more than 99.5% accuracy on the same dataset and different dataset while in different dataset by using Conventional HMM method it gives only 84%.On the basis of experiments performed VIDTIMIT Database, our proposed method gives result 97.6% in the first experiment.So, it is conclude that our proposed method is better than other methods it gives promising result in terms of accuracy by using same dataset and different dataset.

V. Conclusions

Here, the Coupled HMM-ANN model for face recognition is proposed. All experiments show improved performance over the conventional HMM model. The proposed method has experimentally testes on FJU and VIDTIMIT video databases, successfully detected the facial region and recognized the faces. The experimental analysis is performed on both still images and video frame images. Then the result is compared with conventional HMM. The accuracy of recognition is derived for the same test dataset and different datasets. The experimental result itself is evident that the proposed method has higher accuracy that the conventional HMM. Therefore the proposed method is better than the conventional method with the challenging parameters low-lighting condition, titled faces. The proposed method may also be used for face recognition in different areas such as information security, law enforcement and surveillance, smart environment, commercial, and government applications.In future development of this work; the input images may be further enhanced to high quality of the images, the ANN method may be improved in order to minimize the time needed for adjustment of errors during back propagation.

References

[1]. W. ZhaoSarnoff Corporation, Princeton, NJR. Chellappa," Face recognition: A literature survey", ACM computing surveys (CSUR) ,2003,Volume 35 ,pp.399-458.
[2]. ElhamBagherian, RahmitaWirza O.K. Rahmit," Facial feature extraction for face recognition: a review ", IEEE, 2008, pp: 1-9.
[3]. Yang, M.H., Kriegman, D.J., Ahuja, N, "Detecting Faces in Images: A Survey",IEEE TransPAMI, 2002, Volume 24 , Issue 1, pp.34–58 .
[4]. Rowley, Baluja, Kanade, " Neural Network-Based Face Detection", IEEE Transaction on Pattern Analysis and Machine Intelligence, 1998 , Volume 20 , Issue 1, pp.23–38 .
[5]. Pankaj D.S., Wilscy M., "Face Recognition Using Fuzzy Neural Network Classifier", Parallel Distributed Computing. Communications in Computer and Information Science, Springer, 2011, Volume 203. pp. 53-62.
[6]. Sinha A., "Efficient Face Detection Using Neural Networks", Advances in Intelligent Systems and Computing, Springer, 2017, Volume 509, pp. 279-285.
[7]. Samaria, F., Young, S.,"HMM-based architecture for face identification ", Image and Vision Computing, Elsevier, 1994, Volume 12, Issue 8 , pp.537–543.
[8]. Nefian, A.V., Hayes III, and M.H, "Face detection and recognition using hidden markov models ", Image Processing conference, IEEE, 1998, Volume 1, pp 141–145 .
[9]. VitthalSuryakantPhad, Prakash S. Nalwade, Prashant M. Suryavanshi "A face recognition system by hiddenmarkov model and descrimining set approches," I.J. Modern Education and Computer Science,2014, Volume 6, Number. 7,pp; 25-30.
[10]. Chihaoui M., Bellil W., Elkefi A., Amar C.B. , "Face Recognition Using HMM-LBP ", Hybrid Intelligent Systems , Springer, 2016 , Volume 420,pp 249-258.
[11]. BeldiMakrem, LachiriZied,"Video-Based Face Recognition Using Hidden Markov Models and 2D Discrete Cosine Transform:Application to (VIDTIMIT database)", 3rd International conference on Automation Control,Engineering and computer science(ACES'16),proceedings of Engineering & Technology,pp. 5-10.
[12]. C. Sanderson and B.C. Lovell Multi-Region Probabilistic Histograms for Robust and Scalable Identity Inference. Lecture Notes in Computer Science (LNCS), Vol. 5558, pp.199-208, 2009.

Sandeep Saxena "Face Recognition Using Coupled the Hidden Markov Model With An Artificial Neural Network." IOSR Journal of Computer Engineering (IOSR-JCE) 20.1 (2018): 11-15.