IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 19 June 2022, accepted 27 June 2022, date of publication 5 July 2022, date of current version 20 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3188717

RESEARCH ARTICLE

Face Recognition System Based on Four State Hidden Markov Model

DANISH ALI, IMRAN TOUQIR, ADIL MASOOD SIDDIQUI, JABEEN MALIK, AND MUHAMMAD IMRAN

Department of Electrical Engineering, National University of Sciences and Technology, Islamabad 46000, Pakistan

Corresponding author: Imran Touqir (imrantqr@mcs.edu.pk)

This work was supported by the National University of Sciences and Technology (NUST), Islamabad, Pakistan.

ABSTRACT Computational complexity is a matter of great concern in real time face recognition systems. In this paper, four state hidden Markov model for face recognition has been presented whereby coefficients of feature vectors have been curtailed. Face images have been divided into a sequence of overlapping blocks. An observation sequence containing coefficients of eigen values and eigenvectors of these blocks have been used to train the model and each subject is associated with a separate hidden Markov model. The computational complexity of the proposed model has been minimized by employing discrete wavelet transform in the preprocessing stage. Furthermore, singular value decomposition has been employed on face images and a threshold singular value is determined empirically to reject or accept test images. Principal component analysis is used for feature extraction. Accepted test images are classified based on the majority vote criteria using different observation sequences of image features. Experimental findings on Yale and ORL databases in noisy such as Salt and Pepper and noise free environments reveal that the recognition accuracy of the proposed model is comparable to the existing techniques with reduced computational cost.

INDEX TERMS Face recognition, four state hidden Markov model, principal component analysis, singular value decomposition, discrete wavelet transform.

I. INTRODUCTION

Face recognition (FR) has become challenging and fast-growing research topic. FR is a passive method for biometric identification and avoids direct personal interaction. Researchers have developed a variety of methods in ideal and noisy environments. Most of the FR techniques focus on improving the recognition rate and pay little attention to decreasing the computational complexity. In five state hidden Markov model (HMM) the image has been divided into five face regions; hair, forehead, eyes, nose and mouth [1]. In seven state HMM two more face regions; eyebrows and chin have been included [2]. Three state HMM is based on three different facial poses as states [3]. We have eliminated the hair region from five state hmm to make it as four state that reduces computational complexity which is proportional to the number of states.

The associate editor coordinating the review of this manuscript and approving it for publication was Zahid Akhtar.

Feature extraction and classification are the two key steps in automatic face recognition (AFR) [4], [5]. There are two types of features extracted from image contents: global and local features. Typical global feature extraction techniques are eigen faces [6] and Schur faces [7]. However, the performance of these methods degrades under varying facial expressions. Principal component analysis (PCA) is generally employed for extraction of image features [8]. Support Vector Machine (SVM) realizes the partitioning of classes with an optimal hyperplane [9]. Kim *et al.* [10] presented a singular value decomposition (SVD) face which is an efficient feature descriptor under varying illumination conditions. Jian *et al.* [11] proposed that SVD can be used for the reconstruction of high-resolution (HR) images from low-resolution images using a mapping function. These HR images effectively improve the recognition accuracy. Local features are extracted to illustrate the variations in the local patches of an image in more details. However, execution time is the limitation of local methods in real time applications [12], [13]. Local Binary Pattern (LBP)

extracts local patterns of images based on pixel's intensity comparison values that serve as a basis for remote learning transformation. Gabor filter extracts texture pattern and edges of faces but it increases redundancy [14]. Linear discriminant analysis (LDA) reduces number of features to a more manageable number prior to classification [15]. Independent component analysis (ICA) extracts the hidden features of image and defines a generative model in FR [16], [17].

The second main problem in FR is to classify the test images of trained databases based on minimum distance or maximum likelihood criteria [18], [19]. Classification of face images under an uncontrolled environment requires a robust algorithm. However, the computational cost of these algorithms is a serious challenge that hampers its application in real-time AFR. It necessitates to reduce the image dimensions using a transformation approach that retains significant image features. Numerous transformations have been used for dimensionality reduction in the preprocessing stage, however, in this paper, discrete wavelet transform has been envisaged for the subject before FR. The computational cost of an algorithm depends upon observation vector length and classification model used. Therefore, selection of classifier has a major impact on the operation of FR system. Numerous other classification models have been developed such as linear regression classifier [20], minimum distance classifier [21], neural networks (NN) [22] and HMM [23]. These distance measures misclassify test images of untrained databases to one of the trained database images. The architecture of the convolutional neural networks has been widely used but requires more time for training [24], [25]. SVM has already been used in binary classification of data points. Nebti *et al.* used SVM for FR in combination with HMM [26].

Nonlinear minimum static [27] bilateral and trained filters based on non-local mean (NLM) [28], [29] are used to denoise images, same have been used in processing stage of the proposed algorithm. In real time applications, the noise is a complex phenomenon [30]. In [31] the proposed architecture significantly outperforms the classical K-SVD algorithm while preserves the original K-SVD essence [32] and approaches the latest state-of-the-art learnable denoising methods. Ren *et al.* [33] used non-local self-similarity technique using Haar transform and Wiener filter based on blind denoising method.

Feature extraction and classification of FR have been discussed in section 1. Rest of the paper is structured such that section 2 describes the related work on the existing FR techniques. Section 3 describes the proposed algorithm for FR. Experimental results have been demonstrated in section 4, followed by comparative study and conclusion in section 5 and section 6, respectively.

II. RELATED WORK

Mittal *et al.* [19] suggested a hybrid feature extraction approach that was centered on the blend of SVD and LBP features. Query image was classified based on the minimum distance between observation vectors of query and database images. Alobaidi *et al.* [34] used two frequency domain features; discrete cosine transforms (DCT) and DWT. Majority voting scheme was employed to classify face images. Abuzneid *et al.* [35] used local binary pattern histogram (LBPH) for dimensionality reduction. Five different distance methods were employed to form a feature vector. Back propagation neural network (BPNN) was exploited for recognition of the test images. However, training process of BPNN is computationally expensive. HMM has been used in AFR due to its computational efficiency in training and evaluation of images. The number of states have a major impact on the efficiency of HMM based AFR. Samaria *et al.* [36] proposed five state HMM to classify face images using pixel intensity values as features. Nefian and Hayes [1] employed 2D-DCT to reduce complexity of FR system. Shen *et al.* [38] presented DWT based HMM to efficiently recognize the high dimensional images. DWT was used to replace Discrete Cosine Transform (DCT) to achieve better accuracy of FR algorithm based on HMM for extraction of observation sequence. Wang *et al.* [39] proposed an assessment function to evaluate the performance of five state HMM for different overlap ratio between image blocks. In [40] two more facial regions (eyebrows and chin) were incorporated and seven state HMM was employed to evaluate observation vectors formed using singular values and coefficients of singular vectors. This results in a high recognition accuracy but increase computational cost. Arindam *et al.* [41] proposed local modified Zernike moment per unit mass LMZMPM for facial recognition which is quite robust. For each pixel in a neighborhood, the proposed LMZMPM computed and then the complicated tuple that encompasses both the phase and magnitude coefficients of LMZMPM as the extracted features are considered only. Edges and structural data have additional information about the image such as, both the phase and the magnitude elements of the complex feature. However, it is computationally expensive. Wan *et al.* [42], [43] used maximum margin criteria to classify face images. To obtain the finest class projection matrix, maximum inter-class distance and minimum intra class distance was calculated. They built the multiple manifolds inter-class and intra-class scatter matrix. Still determination of classification threshold parameter is a major issue that affects the recognition accuracy of the model. Lahaw *et al.* [44] proposed SVM model to classify face images based on ICA, PCA and LDA features. However, it is inadequate for pose variant expressions.

HMM is the most successful statistical classifier that isolates credible features from image. However, it depends on each individual state and its respective observation sequence. It is employed for each subject separately from a set of training images. Recognition accuracy is proportional to number of training images. Training images can be excluded but the parameters of the model are stored. In the recognition process, the maximum probability of unknown face image with the trained models leads to the unknown person.

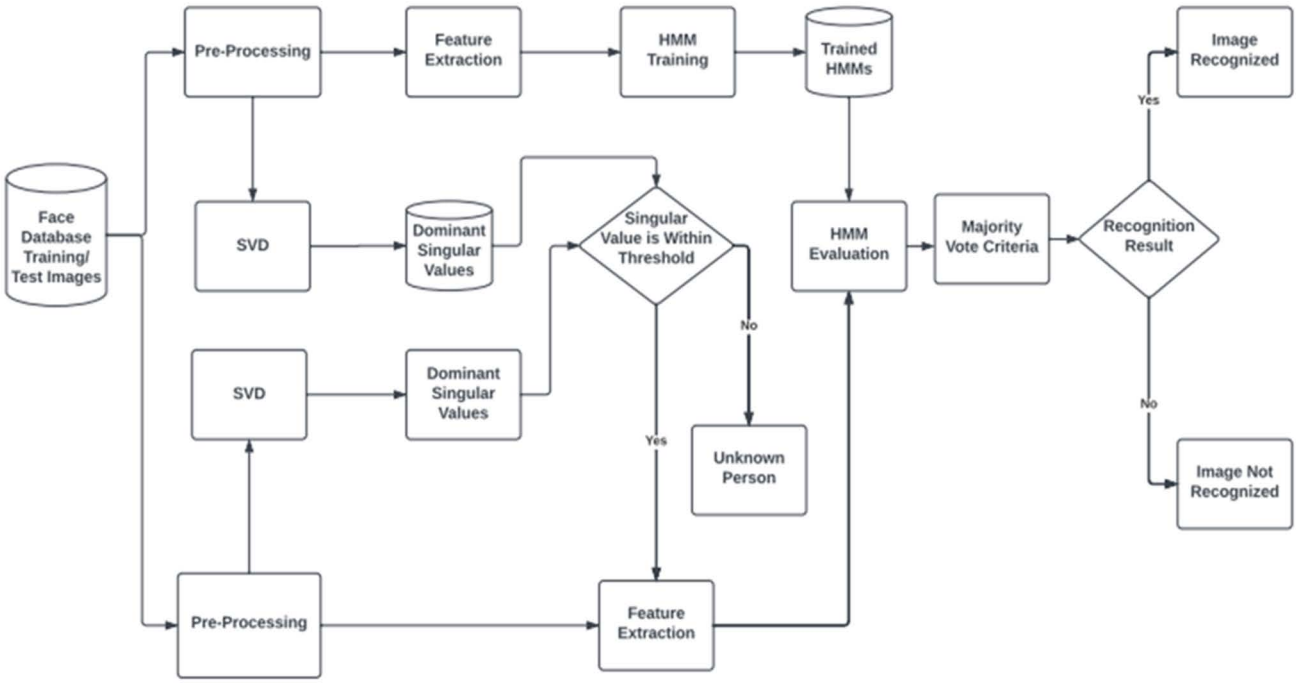


FIGURE 1. Proposed model for face recognition.

Jameel *et al.* [45] proposed seven state HMM using DCT followed by PCA as feature extraction technique that is computationally expensive.

III. THE PROPOSED WORK

This paper focuses on efficient FR system using PCA for feature extraction and four state HMM for classification. In this paper, we proposed an efficient four-state HMM to replace higher order HMMs for classification. The proposed model employs singular value threshold and majority vote criteria to minimize the misclassification rate. It successfully rejects test images of untrained databases using distinct threshold singular value prior to feature extraction.

Preprocessing stage includes noise removal and sparse representation of image using Haar. Eigenvalue decomposition is used to extract informative features from face images. The proposed framework is shown in Fig. 1. Face images are divided into four facial regions that are modelled by the hidden states of HMM as shown in Fig. 2. The feature coefficients of different facial regions represent visible symbols of HMM.

A. PRE-PROCESSING

A non-linear minimum static filter of order three is used to suppress undesired distortions and enhances image characteristics. Furthermore, image dimensions are reduced using the Haar wavelet. Figure 3 shows an example explaining the working of a minimum order-static filter. Table 1 (a) and (b) shows a minimum filter that uses a 3 × 3 window that works

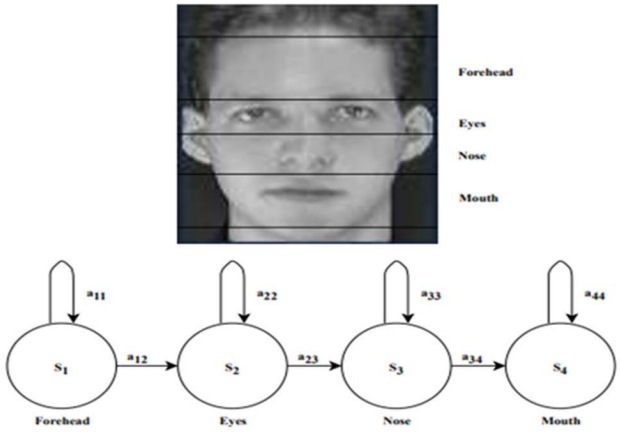


FIGURE 2. Four state HMM for human face image.

on a 6 × 6 region of an image. DWT has exceptional localization properties that makes it computationally effective.

SVD is used to factorize face images obtained after preprocessing step using (1). The peculiarity of SVD is that it can be performed on any arbitrary (i × j) matrix.

$$I = MUN^T \tag{1}$$

where I represent face image, U is diagonal matrix that contains singular values of image, M and N are orthogonal matrices containing orthonormal arbitrary vectors. Principal/ Dominant singular values are determined for all training images and difference between the normalized singular

(a) (b)

FIGURE 3. An example of operation of the order static filter (a) Before filtering (b) After filtering.

TABLE 1. An example of the operation of minimum order static filter.

200	198	195	191	201	190
193	194	198	201	197	201
200	199	196	200	198	193
190	201	197	194	196	197
194	202	200	199	193	194
191	196	198	195	199	191

(a) Before filtering

0	0	0	0	0	0
0	193	191	195	190	0
0	190	194	194	193	0
0	190	194	193	193	0
0	190	194	193	191	0
0	0	0	0	0	0

(b) After filtering

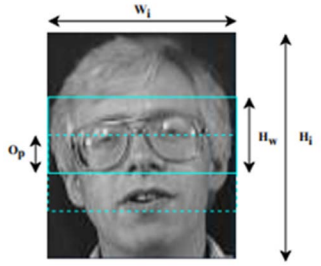


FIGURE 4. Overlapping blocks.

values is computed. Empirically a distinct threshold singular value is determined for each face database to reject test images of untrained databases prior to feature extraction. Thereafter, significant features of accepted test images are extracted. The strength of proposed approach is to employ PCA on local blocks of images for feature extraction.

B. FEATURE EXTRACTION

Face images are scanned using a sampling window from top to bottom approach and overlapping blocks are extracted as shown in Fig. 4. Number of blocks obtained for each face

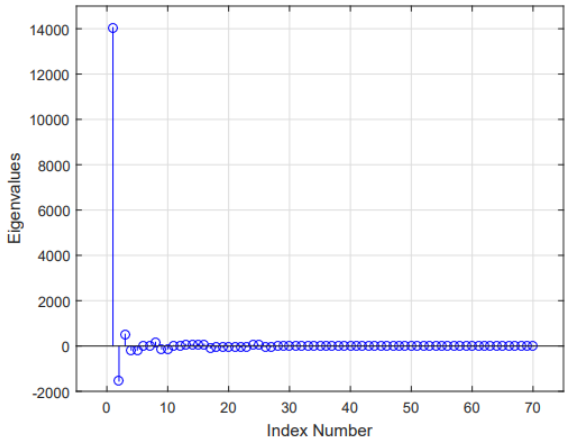


FIGURE 5. Eigen values for a face image.

image are determined using (2).

$$N_B = \left(\frac{H_i - H_w}{H_w - O_p}\right) + 1 \tag{2}$$

where $N_B = 50$ is number of blocks extracted using $H_i = 53$, $H_w = 4$ and $O_p = 3$ that represent image height after dimensionality reduction, sampling window height and overlap size respectively. Covariance matrix of each block is factorized using eigen decomposition.

$$C_i V_i = U_i V_i \tag{3}$$

where U_i and V_i are matrices containing eigen values and eigenvectors of the covariance matrix respectively. Then select only those eigen values and coefficients of principal components for which classification accuracy of the proposed model is maximum at reduced computational cost. Fig. 5 shows eigen values of a face image as a function of its index numbers. It is clear from this figure that eigenvalues corresponding to first two principal components carry most of the information about face image while remaining eigenvalues are negligible. Multiple combinations of these eigen values and coefficients of principal components are tested to evaluate recognition accuracy. Desired results are obtained when we used two eigen values (U_{pc1}, U_{pc2}) corresponding to first two principal components and first coefficient of principal component $V_{pc1}(1, 1)$. These three coefficients are used as features describing image block. These block features are quantized to specific discrete levels to reduce the computational cost using (4) and (5).

$$f_j^{qn} = \frac{f_j - f_j^{min}}{\Delta_j} \tag{4}$$

$$\Delta_j = \frac{f_j^{max} - f_j^{min}}{L_j} \tag{5}$$

where f_j^{min}, f_j^{max} and f_j^{qn} are minimum, maximum and quantized values of feature coefficients respectively. Δj is the difference between two successive quantization levels and

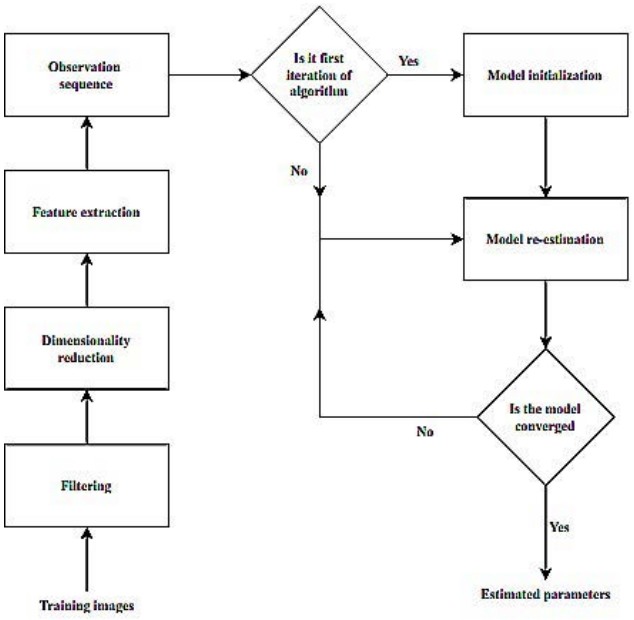


FIGURE 6. Training process of HMM.

L_j is the number of quantization levels. Quantized features are assigned labels using (6)

$$Labels = q_n(1) \times 8 + q_n(2) \times 60 + q_n(3) \times 2 + 6 \quad (6)$$

where $q_n(1)$, $q_n(2)$ and $q_n(3)$ represent quantization levels of $V_{pc1}(1, 1)$, U_{pc1} and U_{pc2} respectively. Empirically good values of $q_n(1)$, $q_n(2)$ and $q_n(3)$ are found to be 10, 10 and 7 respectively. Thus, for each block there are 700 possible training feature values. An observation sequence is formed for each training image which contains feature labels of all image blocks. This observation sequence is further used to train HMM.

C. HMM TRAINING PROCESS

Human facial regions: forehead, eyes, nose and mouth have been modelled by four states of HMM. Training process of HMM is shown in Fig. 6. Feature coefficients of face image are the visible symbols of model. Two major steps are involved in HMM training. In the first step model parameters {A, B, Π} are initialized where A and B are transition and emission probability matrices respectively and Π is initial probability vector. In proposed model adequate results have been obtained using following initial parameters.

$$\begin{aligned} N &= 4 \\ |O| &= 700 \\ \Pi &= [1000] \\ a_{r,r} &= 0.85 \quad 1 \le r \le N-1, \\ a_{r,r+1} &= 0.15 \\ a_{4,4} &= 1 \end{aligned}$$

TABLE 2. Transition matrix A.

	Forehead	Eyes	Nose	Mouth
Forehead	0.85	0.15	0	0
Eyes	0	0.85	0.15	0
Nose	0	0	0.85	0.15
Mouth	0	0	0	1

$$B = \left(\frac{1}{O}\right) \times ones(N, O) \quad (7)$$

where $a_{r,r}$ and $a_{r,r+1}$ are transition probabilities from state r to r and r to r+1 respectively. Number of states and visible symbols are represented by N and O. Table 2. shows the initial transition probability matrix of four states HMM for FR.

Probability of transition for a particular visible symbol (b_d) is determined using (8).

$$P(S_r \xrightarrow{b_d} S_{r+1}) = \frac{C(S_r \xrightarrow{b_d} S_{r+1})}{\sum_{l=1}^{N}\sum_{k=1}^{T} C(S_r \xrightarrow{b_k} S_l)} \quad (8)$$

where $C(\cdot)$, a_r, a_l and b_i represent number of counts of transition, specific hidden state, all possible hidden states and visible symbols respectively. Number of counts is determined using (9).

$$C(S_r \xrightarrow{b_d} S_{r+1}) = \sum_{i=1}^{i_{max}} \left(P\left(\frac{S_i^T}{b^T}\right)\right) \times \left(S_r \xrightarrow{b_d} S_{r+1}\right) \quad (9)$$

where a_i^T and T represent sequence of hidden states and sequence length of visible symbols respectively. After each iteration of Baum Welch algorithm model parameters are updated and these parameters are used as initial parameters in next iteration. Updated transition and emission probabilities are determined using (10) and (11) respectively.

$$a_{r,r+1} = \frac{\sum_k P(S_r \xrightarrow{b_k} S_{r+1})}{\sum_l \sum_k P(S_r \xrightarrow{b_k} S_l)} \quad (10)$$

$$b_{r+1,d} = \frac{\sum_l P(S_l \xrightarrow{b_d} S_{r+1})}{\sum_l \sum_k P(S_l \xrightarrow{b_k} S_{r+1})} \quad (11)$$

where $b_{r+1,d}$ represents emission probability of visible symbol b_d that is emitted by hidden state S_{r+1}. This process continues until model is converged i-e variation of probability values in two consecutive iterations is within a specified threshold which is 0.09 in proposed technique.

D. CLASSIFICATION PROCESS

After training each person is associated to a separate HMM. Each test image is represented by two different observation sequences containing coefficients of eigen values

and eigenvectors that are slightly different from each other as given below.

$$b_1^T = [V_{pc1}(1,1), U_{pc1}, U_{pc2}]$$
$$b_2^T = [V_{pc1}(1,1) \times 1.01, U_{pc1}, U_{pc1} \times 1.04]$$

Probability of each observation sequence of a test image is to be determined against all trained HMMs.

$$P\left(\frac{b_u^T}{\theta_n}\right) = ? \quad u = 1, 2 \tag{12}$$

where θ_n represents the trained HMMs. To determine this probability, we find out all possible sequences of hidden states/facial regions that may generate the given observation sequence. $P(b^T/\theta)$ is calculated by taking summation of probabilities of all these hidden state sequences using (13).

$$P\left(\frac{b_u^T}{\theta n}\right) = \sum_{i=1}^{i_{max}} P(b_u^T / S_i^T) P(S_i^T), \quad i_{max} = N_s^T \tag{13}$$

where i_{max} represents the maximum number of possible paths of hidden states through which model can make transitions while generating b_u^T. Hidden sequence S_1^T represents one of those possible hidden sequences S_i^T of length T that has generated b_u^T.

$$S_1^T = \{S_1, S_2, S_3, , \ldots, S_T\} \tag{14}$$

Probability of a particular hidden sequence is given in (15) that is the product of transition probabilities at different time instances. Similarly, probability of observation sequence for a known hidden sequence is given in (16) that is the product of emission probabilities at different time steps.

$$P\left(S_1^T\right) = \prod_{t=1}^{T} P(S_t, / S_{1,-1}) \tag{15}$$
$$P(b_u^T / S_i^T) = \prod_{t=1}^{T} P(b_u(t) / S_t,) \tag{16}$$

For time instant $t = 1$, $P(b_u^T / S_i^T)$ is the probability that model is at first state of a particular hidden sequence, and it has generated first visible symbol of known observation sequence of test image. By putting the values of (15) and (16) in (13) we get

$$P(b_u^T/\theta) = \sum_{i=1}^{i_{max}} \prod_{t=1}^{T} P\left(\frac{b_u^T}{S_i(t)}\right) P\left(\frac{S_i(t)}{S_i(t-1)}\right) \tag{17}$$

Test image is classified based on the majority vote rule of evaluation probabilities for known observation sequences as given in (18).

$$\mathrm{I_{test}} = \begin{cases} I_k \\ \text{unknown} \end{cases} \quad \left(\begin{array}{c} if\ P\left(\frac{b_u^T}{\theta_k}\right) = \max_{\mathrm{n}} P\left(\frac{b_u^T}{\theta \mathrm{n}}\right) \\ \text{otherwise} \end{array} \right) \tag{18}$$

where $\mathrm{P}\left(\frac{b_u^T}{\theta_k}\right)$ is evaluation probability of observation sequence b_u^T. Test image I_{test} is classified to $\mathrm{k^{th}}$ person in the database if evaluation probability for both observation sequences is maximum against trained HMM; θ_k.

E. COMPLEXITY ANALYSIS

The computational complexity and features required for the characterization of face in a system is associated with N, which is number of states. Five state [1] and seven state [2] HMM used for FR achieved better recognition rate but their computational complexity increases significantly. Comparatively, for three state HMM [3] the proposed method achieved better recognition rate. Reduction in the number of states leads to minimal memory occupation but on the same time efficiency compromises. Thus, decreasing HMM dimensions decrease computational complexity significantly. Complexity of HMM training process is proportional to complexity of Baum Welch algorithm. Baum Welch algorithm has time complexity of $\mathrm{O}(N_s^2 T)$ per iteration [46]. Observation sequence length T in the proposed technique is calculated using (19).

$$\mathrm{T} = \mathrm{N_I} \times \mathrm{N_B} \times \mathrm{N_F} \tag{19}$$

where N_I is number of images used in training, N_B is number of blocks per training image and N_F is number of feature coefficients per block. Overall complexity of HMM training for all individuals in the face database is calculated using (20).

$$\text{Overall Complexity} = \mathrm{N_p} \times \mathrm{N_{it}} \times \mathrm{N_s^2} \times \mathrm{T} \tag{20}$$

where N_p is total number of persons and N_{it} is number of iterations. Forward algorithm is used for testing or evaluation of an observation sequence. Complexity of forward algorithm is $O(N_s^2 T)$ [47]. Overall complexity of HMM evaluation for all persons in the database is $N_p N_s^2 T$. DWT is employed prior to FR that reduces computational complexity to its one quarter.

F. ASSESSMENT FUNCTION

In previous techniques, performance of FR system was evaluated based on recognition accuracy. In the proposed technique an assessment function is used in which error rate and computational complexity have been considered. Assessment process is depicted in Fig. 7. Error rate and complexity of the algorithm have been calculated for different number of states for the model.

$$\text{Error Rate} = 100 - \text{Recognition Rate}(\%) \tag{21}$$

Error rate and computational complexity are normalized using Min Max data normalization method as given in (22) and (23) respectively.

$$\mathrm{N}(E_m) = \frac{E_m - \min(E_m)}{\max(E_m) - \min(E_m)} \tag{22}$$
$$\mathrm{N}(C_m) = \frac{C_m - \min(C_m)}{\max(C_m) - \min(C_m)}, \quad \text{where } m = 1, 2, \ldots 8 \tag{23}$$

where $N(E_m)$ and $N(C_m)$ represent normalized values of error rate and computational complexity for eight different states of model respectively. Assessment function is given in (24).

$$\mathrm{AF} = w_1 \times \mathrm{N}(E_m) + w_2 \times \mathrm{N(C_m)} \tag{24}$$

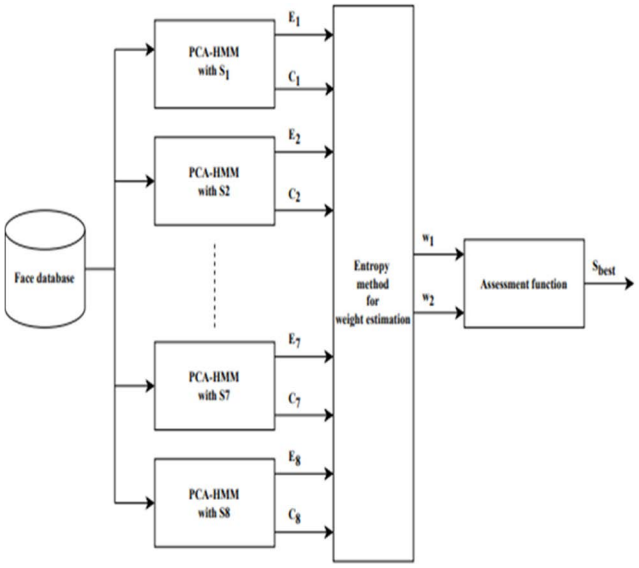


FIGURE 7. Assessment process.

where w_1 is the weight of error rate and w_2 is the weight of computational complexity and their weights are assigned based on their entropy.

$$w_i = \frac{1 - \text{entropy(i)}}{\sum_i (1 - \text{entropy(i)})}, \quad i = 1, 2. \tag{25}$$

Parameters of HMM with minimum value of assessment function are considered as favorable parameters.

IV. EXPERIMENTAL RESULTS

Implementation of the proposed method has been executed using C++ with MATLAB 2022a software on a system with an 11th Generation Intel Core i7-1165G7processor @ 2.8 GHz, 16.00 GB RAM, 64- bit. ORL and Yale databases have been used to verify the effectiveness of the proposed algorithm in recognizing the face. Petya Dinkova et al [41] proposed 7-state HMM for FR. Our proposed methodology increase efficiency by reducing the number of HMM states and achieves satisfactory recognition rate.

ORL database contains 400 face images of 40 individuals with different face orientations and facial expressions [48]. The original image size is 112 × 92 pixels. The images have been captured at various moments, varying illumination, facial expressions (open/closed eyes, happy/unhappy) and facial characteristics. Yale database contains 165 face images of 15 individuals with different illumination conditions and face behavior [49]. There are 11 images per subject with different facial expressions having original image size of 231 × 95 pixels. Local images have been used to augment the performance of proposed algorithm. To expedite the algorithm time and to reduce memory utilization along with using only three SVD coefficients. We reduced the formatted images of the given databases to 53 × 44 jpeg images. Image resizing causes information losing which tends to reduce the recognition rate. However, we have achieved 98.5% classification accuracy on ORL database while significantly speeding up the proposed algorithm. In Fig. 8(a) to 8(c) sample images of these databases are shown. Crossed images represent misclassification.

TABLE 3. Recognition accuracy for different training images.

Number of Training Images	Recognition Rate (%)	
	ORL database	YALE database
Ntr = 5	98.5	94.44
Ntr = 6	99.375	97.33
Ntr = 7	100	98.33
Ntr = 8	100	100
Ntr = 9	100	100

Two different experiments are performed on these databases. In first experiment, images are grouped into training and testing pairs. In the training phase, HMM model parameters have been initialized for each training face. Baum welch algorithm has been used to associate separate HMM θ to each face. Subsequently, recognition accuracy of proposed model is calculated using 5, 6, 7, 8 and 9 image of each subject for training the model and remaining images for testing as shown in Table 3.

In second experiment, images are classified based on the maximum evaluation probability without considering the majority vote rule of classification. In the testing phase, evaluation probability of observation sequence has been calculated and compared with trained HMM model θ. Test face image is classified to k^{th} person in the database if evaluation probability for both observation sequences is maximum against trained HMM θ_k. It is also worth noting that recognition accuracy of four state HMM is comparable to higher order HMMs for feature coefficients of principal component. As every image block is transformed into a covariance matrix of size 4 × 4. Eigen decomposition of each of these covariance matrices results in two matrices of 4 × 4 dimensions. Fig. 9(a) to Fig. 9(f) show the eigen values for sample images from ORL, Yale and Local databases. Table 3 is depicting the objective analysis of the proposed model.

Diversity exists in two principal eigen values of image blocks while remaining eigen values are negligible. So, we used these eigen values and changed the coefficients of eigenvectors and computed recognition accuracy for all coefficients of eigenvectors. Fig 10(a) and 10(b) show the recognition accuracy of proposed algorithm for 8 different states ($S_1, S_2, \ldots, S_8$) of HMM on ORL and Yale databases respectively. $PC_{ij}(i, j = 1, 2, 3, 4)$ in these graphs represent the j^{th} coefficient of i^{th} principal component for all image blocks. It is evident from the graphs that maximum recognition accuracy is achieved for coefficients of principal component. It gradually decreases as the feature coefficient is changed from PC_1 to PC_4.

FIGURE 8. Example images (a) ORL database (b) Yale database (c) Local database.

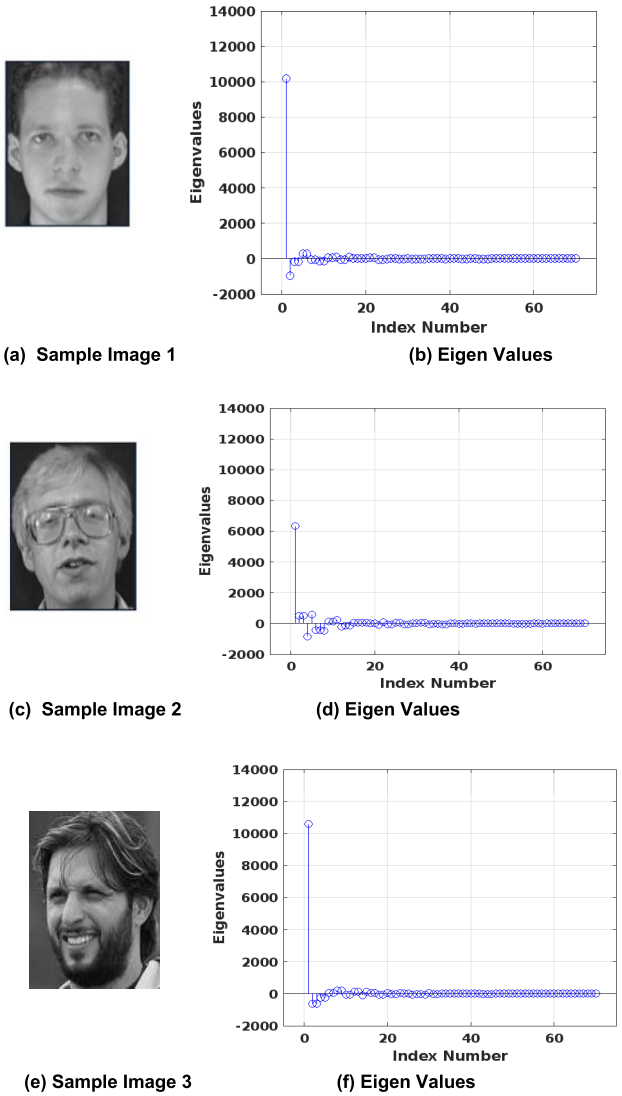
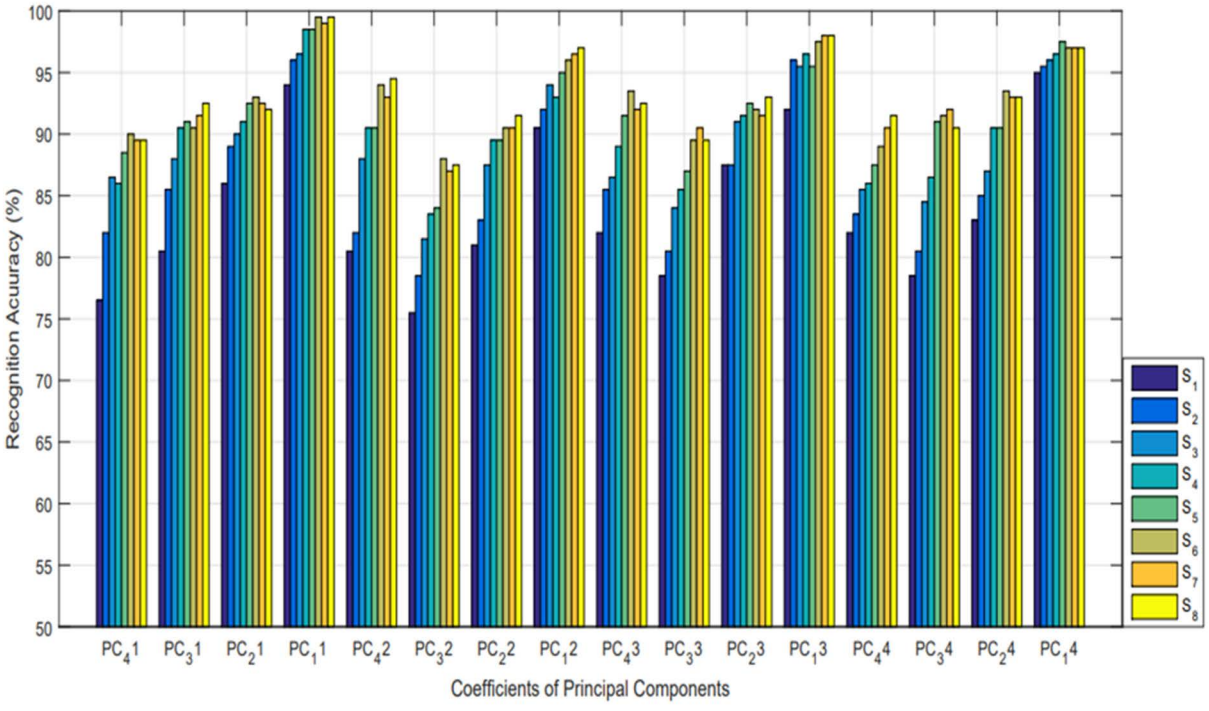
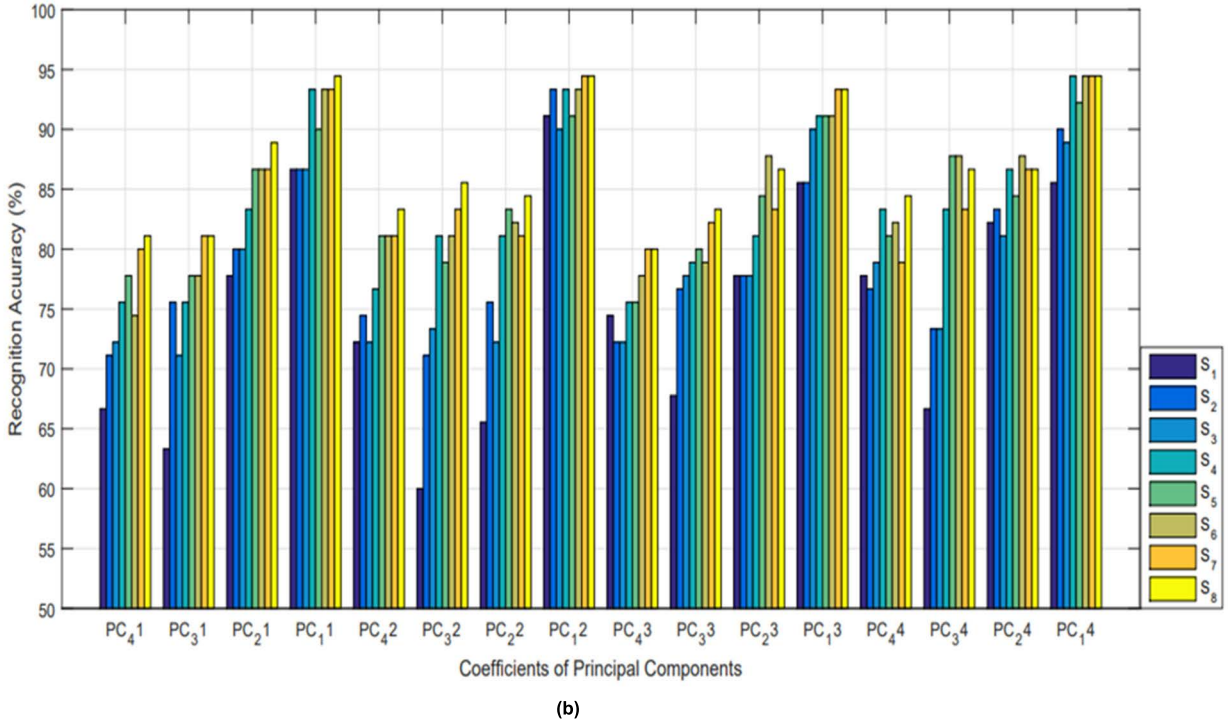


FIGURE 9. Eigenvalues for sample images for ORL, Yale and Local Databases, respectively.

It is also worth noting that recognition accuracy of four state HMM is comparable to higher dimensional HMMs for feature coefficients of principal component. Proposed 4-state HMM model gives acceptable results. Moreover, Assessment function index values, ROC curve represented by Fig. 11 and Fig. 12 respectively, are other assessment methods used to show the effectiveness of model. Assessment function index values are also determined for 8 different states of HMM using their computational costs and error rates. Fig. 11 illustrates that assessment function index value is smaller for four states HMM that gradually raises by increase in the states of the model. This is due to the reason that complexity boosts by number of states of the model while variation in the error rate is minimum for higher states. In second experiment, test face images are categorized into known and unknown images. Known images are those test images that are associated with trained database whereas unknown test images belong to untrained database. In the evaluation of ORL database 150 images of Yale database are used as unknown face images and vice versa.



(a)



(b)

FIGURE 10. Recognition accuracy of proposed algorithm for different states of HMM, (a) ORL database, (b) Yale database.

Fifty local images are also used as unknown images for these databases. In this case, singular value threshold and majority vote rule are also incorporated to recognize known images and successfully reject the unknown images. Recognition and rejection rates of proposed technique for known and unknown images are listed in Table 4.

Recognition accuracy for trained database images is dropped down by incorporating face rejection. This is due to

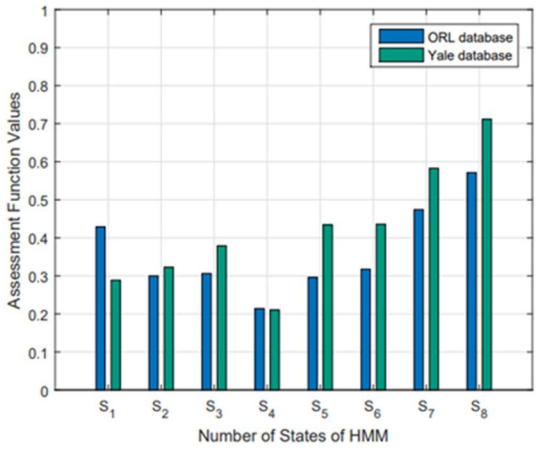


FIGURE 11. Assessment function values for ORL and Yale databases.

TABLE 4. Recognition and rejection rates using five training images.

Face databases	Recognition Rate (%)	Rejection Rate (%)
ORL with SVD threshold	97	73
ORL with SVD and majority vote	94.4	82
Yale with SVD threshold	92.22	84
Yale with SVD and majority vote	85.56	86

the mismatch in threshold singular value and classification results against two slightly different observation sequences of some known images. Therefore, these images are also rejected by the model along with unknown images. Parameters used for proposed model on Yale and ORL database are specified in Table 5.

V. COMPARATIVE STUDY

Although, HMM has its own demerits and not a novel method but it provides flexibility to manipulate the training and verification processes. It has strong statistical foundation with efficient learning algorithms where learning can take place directly from raw sequence data. In image recognition, we need such methods that should be able to outperform with any type of input image (different illuminations, noise, orientation etc.) Due to these main characteristics, HMM is good choice as compared with other state of the art FR techniques [37]. Table 6 summarizes the recognition precision of state-of-the-art FR techniques and proposed model on ORL and Yale data sets. The proposed model performance is estimated for three separate observations. In first observation, face images are classified based on the maximum evaluation probability in which comparable recognition accuracy is achieved but no unknown test image is rejected by the model. In second observation, we only used singular value threshold to reject unknown test images in which six and seven images of ORL and Yale databases are misclassified.

TABLE 5. Parameters design for ORL and Yale database.

Parameters	ORL Dataset	Yale Dataset
Block Height	5	5
Block overlap	75%	75%
Quantization level 1	10	10
Quantization level 2	10	10
Quantization level 3	7	7
Number of states	4	4
Index of training images	[5 6 7 8 9]	[5 6 7 8 9]
Index of testing images	[1 2 3 4 10]	[1 2 3 4 10 11]
Selected features	U_{pc1}, U_{pc2},V_{pc} (1,1)	U_{pc1}, U_{pc2},V_{pc} (1,1)
Image size	53x44	53x44
Remarks	Accuracy: 98.5% Total of 200 testing images for 40 individuals	Accuracy: 94.44% Total of 90 testing images for 15 individuals

TABLE 6. Recognition rate comparison on ORL and Yale databases.

Method	Ref	Recognition Rate (%)	
		ORL database	Yale database
Schur faces	[7]	85	86
SVM-HMM	[26]	100	-
SVD-HMM	[18]	96.6	82.7
BPNN	[35]	98	97.7
DCT-HMM	[39]	85.5	82.23
SVD-HMM	[40]	96.6	82.6
PCA-DCT	[45]	95.122	-
LGE/MMC	[42]	95.62	94.90
MLGE/MMC	[43]	97.64	96.94
DWT-PCA	[44]	96	-
Proposed without threshold	-	98.5	94.4
Proposed with SVD threshold	-	97	92.22
Proposed with SVD and majority vote	-	94.5	85.56

The proposed model also rejected 73% and 84% unknown images for trained databases. In third observation, we incorporated majority vote rule along with singular value threshold in which recognition accuracy for known images is dropped down slightly as depicted in Table 6. The ability of proposed model to successfully reject 82% and 86% unknown images is advantageous for trained ORL and Yale databases respectively. It is evident from these observations that proposed algorithm is effective in classifying face images captured at varying illumination conditions, different facial expressions, and behavior. Table 6 shows that for ORL database, 4-state HMM has better recognition rate than other techniques proposed in [7], [18], [35], [39], [40], [42]–[44]

TABLE 7. Computational complexity comparison (in terms of counting elementary operations) on ORL and Yale databases.

Method	Ref	Complexity per image			
		ORL database		Yale database	
		NB	Complexity	NB	Complexity
7 State HMM	[26]	60	8820	60	8820
7 State HMM	[18]	60	8820	111	16317
7 State HMM	[40]	52	7644	52	7644
Proposed without threshold	-	50	2400	50	2400
Proposed with majority vote	-	50	4800	50	4800

and [45] except [26] in which 100% recognition rate was claimed. However, computational complexity of [26] is much higher than our proposed HMM model for FR. Table 7 depicts that complexity regarding recognition process, and it varies quadratically with several states of model and linearly with the observation sequence length.

In state-of-the-art systems seven state, HMM has been used more frequently with the aim of maximizing recognition rate but its computational cost is very high. We tested several HMM states to reduce the complexity. Comparable results at reduced computational cost are obtained using four states of model.

Computational time per image recognition observed is 50 ms for proposed model shown in Table 8. Computational cost of proposed model for face rejection is twice the forward algorithm as in this case it runs for two different observation sequences. Number of states determines the number of features used to characterize the face besides computational complexity of the system. In three state HMM [3], the computational complexity decreases with decrease in recognition rate to 82.76%. On the other hand, in 7-state [2] by using five training image (ORL and Yale Database) the predicted HMM model gives recognition rate of 97.78 % however, its computational complexity increases significantly. Whereas in 5-state [1] MIT database 9 images were used in the training set and achieved 90 % recognition. Although its recognition rate decreased as compared to seven state, but its computational complexity is less than seven state HMM as it takes less sequential information. Reduction in the number of states leads to minimal memory occupation is obvious but on the same time accuracy compromises. Therefore, it is evident that it is a tradeoff between recognition accuracy and computational cost. 4-state HMM is better as we are achieving satisfactory recognition rate with minimized computational complexity.

Confusion matrix shows the complete picture of face image classification i.e. face image is correctly classified or otherwise. True positive rate and true negative rate is computed for each class using the coefficients of confusion

TABLE 8. Recognition time comparison for ORL dataset.

Image size	64×64 (7-state HMM)	32×32 (7-state HMM)	53x44 (Proposed 4-state HMM)
Number of training image(s)	5	5	5
Recognition time per image(second)	0.28	0.15	0.05
Number of Symbols	1260	960	700
Recognition rate	99	92.5	98.5

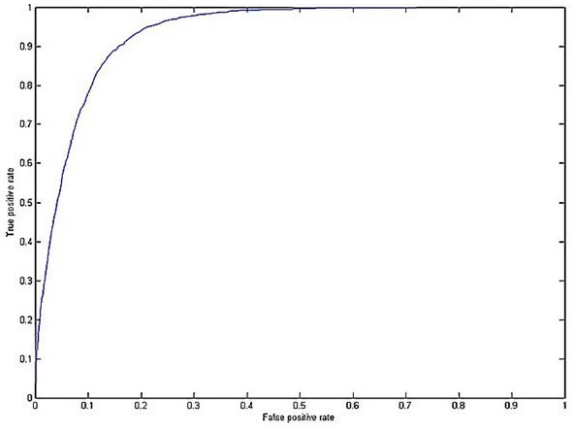


FIGURE 12. ROC curve between True positive and False positive rate.

matrix as shown in Fig. 12. The ROC curve has outperformed the average ROC curve and nearly converges for values above 0.6 false positive. As compared with the results of [3] using three state HMM, the proposed 4-state HMM achieved more recognition accuracy.

VI. CONCLUSION

An efficient four-state HMM has been envisaged for FR. Recognition accuracy and computational complexity of the proposed model are calculated, and it is compared with previous approaches. The proposed work provides satisfactory recognition rate but significantly reduces computational cost. The two standard face image databases; ORL and YALE are used for evaluations and comparisons. Better accuracy rates have been achieved by using a minimal number of features that describes blocks and resizing images to smaller size. An assessment function is employed for the evaluation of the proposed model performance by using error rate, and complexity simultaneously. The result after experiments validates that the recognition accuracy of the proposed model is comparable with the existing techniques for trained databases. Rejection rates of 82% and 86% are also achieved by the proposed model for unidentified images of trained ORL, and Yale databases correspondingly.

The algorithm does not perform well for low-intensity images and is restricted to frontal and upright faces. The research work can be enhanced to cater for intensity and pose variant facial images.

REFERENCES

[1] A. V. Nefian and M. H. Hayes, "Hidden Markov models for face recognition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process., (ICASSP)*, May 1998, pp. 2721–2724, doi: 10.1109/ICASSP.1998.678085.

[2] H. Miarnaeimi and P. Davari, "A new fast and efficient HMM-based face recognition system using a 7-state HMM along with SVD coefficients," *Iranian J. Elect. Electron. Eng.*, vol. 4, nos. 1–2, pp. 46–57, 2008.

[3] K. R. Singh, M. A. Zaveri, and M. M. Raghuwanshi, "Recognizing faces under varying poses with three states hidden Markov model," in *Proc. IEEE Int. Conf. Comput. Sci. Automat. Eng. (CSAE)*, vol. 2, May 2012, pp. 359–363.

[4] H. Tan, Z. Ma, and B. Yang, "Face recognition based on the fusion of global and local HOG features of face images," *IET Comput. Vis.*, vol. 8, no. 3, pp. 224–234, Oct. 2013, doi: 10.1049/IET-CVI.2012.0302.

[5] Z. Mahmood, T. Ali, and S. U. Khan, "Effects of pose and image resolution on automatic face recognition," *IET Biometrics*, vol. 5, no. 2, pp. 111–119, Jun. 2016, doi: 10.1049/IET-BMT.2015.0008.

[6] L. C. Paul and A. A. Sumam, "Face recognition using principal component analysis method," *Int. J. Adv. Res. Comput. Eng. Tech.*, vol. 1, no. 9, pp. 135–139, 2012.

[7] G. Ghinea, R. Kannan, and S. Kannaiyan, "Gradient-orientation-based PCA subspace for novel face recognition," *IEEE Access*, vol. 2, pp. 914–920, 2014, doi: 10.1109/ACCESS.2014.2348018.

[8] M. K. Halidu, P. Bagheri-Zadeh, A. Sheikh-Akbari, and R. Behringer, "PCA in the context of face recognition with the image enlargement techniques," in *Proc. 8th Medit. Conf. Embedded Comput. (MECO)*, Jun. 2019, pp. 1–5.

[9] H. I. Dino and M. B. Abdulrazzaq, "Facial expression classification based on SVM, KNN and MLP classifiers," in *Proc. Int. Conf. Adv. Sci. Eng. (ICOASE)*, Apr. 2019, pp. 70–75.

[10] W. Kim, S. Suh, W. Hwang, and J.-J. Han, "SVD face: Illumination-invariant face representation," *IEEE Signal Process. Lett.*, vol. 21, no. 11, pp. 1336–1340, Nov. 2014, doi: 10.1109/LSP.2014.2334656.

[11] M. Jian and K.-M. Lam, "Simultaneous hallucination and recognition of low-resolution faces based on singular value decomposition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 11, pp. 1761–1772, Nov. 2015, doi: 10.1109/TCSVT.2015.2400772.

[12] D. A. Zebari, A. R. Abrahim, D. A. Ibrahim, G. M. Othman, and F. Y. H. Ahmed, "Analysis of dense descriptors in 3D face recognition," in *Proc. IEEE 11th Int. Conf. Syst. Eng. Technol. (ICSET)*, Nov. 2021, pp. 171–176.

[13] B. T. Chinimilli, A. T., A. Kotturi, V. R. Kaipu, and J. V. Mandapati, "Face recognition based attendance system using Haar cascade and local binary pattern histogram algorithm," in *Proc. 4th Int. Conf. Trends Electron. Informat. (ICOEI)*, Jun. 2020, pp. 701–704.

[14] G. Mattela and S. K. Gupta, "Facial expression recognition using Gabor-mean-DWT feature extraction technique," in *Proc. 5th Int. Conf. Signal Process. Integr. Netw. (SPIN)*, Feb. 2018, pp. 575–580.

[15] Y. Ren, Z. Song, S. Sun, J. Liu, and G. Feng, "Outsourcing LDA-based face recognition to an untrusted cloud," *IEEE Trans. Depend. Secure Comput.*, early access, May 3, 2022, doi: 10.1109/TDSC.2022.3172143.

[16] R. Naik, D. Pratap Singh, and J. Chaudhary, "A survey on comparative analysis of different ICA based face recognition technologies," in *Proc. 2nd Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Mar. 2018, pp. 1913–1918.

[17] M. Rahul, R. Shukla, P. K. Goyal, Z. A. Siddiqui, and V. Yadav, "Gabor filter and ICA-based facial expression recognition using two-layered hidden Markov model," in *Advances in Computational Intelligence and Communication Technology*. Singapore: Springer, Jun. 2020, pp. 511–518.

[18] P. Dinkova, P. Georgieva, and M. Milanova, "Face recognition using singular value decomposition and hidden Markov model," in *Proc. WSEAS Int. Conf. Math. Methods, Comput. Techn. Intell. Syst. (MAMECTIS)*, Lisbon, Portugal, Oct. 2014, pp. 144–149.

[19] R. KumarMittal and A. Garg, "Face recognition through combined SVD and LBP features," *Int. J. Comput. Appl.*, vol. 88, no. 9, pp. 23–27, Feb. 2014.

[20] Y. Peng, J. Ke, and S. Liu, "An improvement to linear regression classification for face recognition," *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 9, pp. 2229–2243, 2019.

[21] E. I. Abbas, M. E. Safi, and K. S. Rijab, "Face recognition rate using different classifier methods based on PCA," in *Proc. Int. Conf. Current Res. Comput. Sci. Inf. Technol. (ICCIT)*, Apr. 2017, pp. 37–40, doi: 10.1109/CRCSIT.2017.7965559.

[22] S.-H. Lin, S.-Y. Kung, and L.-J. Lin, "Face recognition/detection by probabilistic decision-based neural network," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 114–132, Jan. 1997, doi: 10.1109/72.554196.

[23] M. M. M. Farag, T. Elghazaly, and H. A. Hefny, "Face recognition system using HMM-PSO for feature selection," in *Proc. 12th Int. Comput. Eng. Conf. (ICENCO)*, Dec. 2016, pp. 105–110, doi: 10.1109/ICENCO.2016.7856453.

[24] Z. Deng, X. Peng, Z. Li, and Y. Qiao, "Mutual component convolutional neural networks for heterogeneous face recognition," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 3102–3114, Jun. 2019.

[25] M. Z. Khan, S. Harous, S. U. Hassan, M. U. G. Khan, R. Iqbal, and S. Mumtaz, "Deep unified model for face recognition based on convolution neural network and edge computing," *IEEE Access*, vol. 7, pp. 72622–72633, 2019, doi: 10.1109/ACCESS.2019.2918275.

[26] S. Nebti and B. Fadila, "Combining classifiers for enhanced face recognition," in *Proc. Int. Conf. Adv. Inf. Sci. Comput. Eng.* Dordrecht, The Netherlands, 2015, pp. 339–345.

[27] L. Lazeta, I. Markovic, and V. Simovic, "IIR filters designed for comparison and minimum-order design exploration using MATLAB," in *Proc. 44th Int. Conv. Inf., Commun. Electron. Technol. (MIPRO)*, Sep. 2021, pp. 875–879, doi: 10.23919/MIPRO52101.2021.9596760.

[28] A. Khmag, A. R. Ramli, S. J. bin Hashim, and S. A. R. Al-Haddad, "Additive noise reduction in natural images using second-generation wavelet transform hidden Markov models," *IEEJ Trans. Electr. Electron. Eng.*, vol. 11, no. 3, pp. 339–347, Jan. 2016.

[29] A. Khmag, S. A. R. Al Haddad, R. A. Ramlee, N. Kamarudin, and F. L. Malallah, "Natural image noise removal using nonlocal means and hidden Markov models in transform domain," *Vis. Comput.*, vol. 34, no. 12, pp. 1661–1675, Sep. 2017.

[30] G. Chen, F. Zhu, and P. A. Heng, "An efficient statistical method for image noise level estimation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 477–485.

[31] M. Golla and S. Rudra, "A novel approach of k-SVD-based algorithm for image denoising," in *Histopathological Image Analysis in Medical Decision Making*. Hershey, PA, USA: IGI Global, 2019, pp. 154–180.

[32] M. Scetbon, M. Elad, and P. Milanfar, "Deep K-SVD denoising," *IEEE Trans. Image Process.*, vol. 30, pp. 5944–5955, 2021, doi: 10.1109/TIP.2021.3090531.

[33] X. Ren, W. Yang, W.-H. Cheng, and J. Liu, "LR3M: Robust low-light enhancement via low-rank regularized retinex model," *IEEE Trans. Image Process.*, vol. 29, pp. 5862–5876, 2020, doi: 10.1109/TIP.2020.2984098.

[34] T. Alobaidi and W. B. Mikhael, "Face recognition system based on features extracted from two domains," in *Proc. IEEE 60th Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2017, pp. 977–980, doi: 10.1109/MWSCAS.2017.8053089.

[35] M. A. Abuzneid and A. Mahmood, "Enhanced human face recognition using LBPH Descriptor, multi-KNN, and back-propagation neural network," *IEEE Access*, vol. 6, pp. 20641–20651, 2018, doi: 10.1109/ACCESS.2018.2825310.

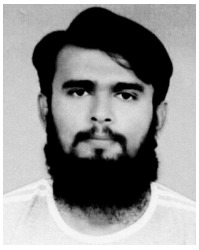
[36] F. Samaria and S. Young, "HMM-based architecture for face identification," *Image Vis. Comput.*, vol. 12, no. 8, pp. 537–543, Oct. 1994, doi: 10.1016/0262-8856(94)90007-8.

[37] Q. Deng and D. Soffker, "A review of HMM-based approaches of driving behaviors recognition and prediction," *IEEE Trans. Intell. Vehicles*, vol. 7, no. 1, pp. 21–31, Mar. 2022.

[38] L. Shen, Z. Ji, L. Bai, and C. Xu, "DWT based HMM for face recognition," *J. Electron. China*, vol. 24, no. 6, pp. 835–837, Nov. 2007.

[39] X. Wang, Y. Cai, and M. Abdulghafour, "A comprehensive assessment system to optimize the overlap in DCT-HMM for face recognition," in *Proc. 11th Int. Conf. Innov. Inf. Technol. (IIT)*, Nov. 2015, pp. 290–295, doi: 10.1109/INNOVATIONS.2015.7381556.

[40] P. Dinkova, P. Georgieva, A. Manolova, and M. Milanova, "Face recognition based on subject dependent hidden Markov models," in *Proc. IEEE Int. Black Sea Conf. Commun. Netw. (BlackSeaCom)*, Jun. 2016, pp. 1–5, doi: 10.1109/BlackSeaCom.2016.7901570.
[41] A. Kar, S. Pramanik, A. Chakraborty, D. Bhattacharjee, E. S. L. Ho, and H. P. H. Shum, "LMZMPM: Local modified Zernike moment per-unit mass for robust human face recognition," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 495–509, 2021, doi: 10.1109/TIFS.2020.3015552.
[42] M. Wan, S. Gai, and J. Shao, "Local graph embedding based on maximum margin criterion (LGE/MMC) for face recognition," *Informatica*, vol. 36, no. 1, pp. 103–112, 2012.
[43] M. Wan and Z. Lai, "Multi-manifold locality graph embedding based on the maximum margin criterion (MLGE/MMC) for face recognition," *IEEE Access*, vol. 5, pp. 9823–9830, 2017, doi: 10.1109/ACCESS.2017.2706525.
[44] Z. B. Lahaw, D. Essaidani, and H. Seddik, "Robust face recognition approaches using PCA, ICA, LDA based on DWT, and SVM algorithms," in *Proc. 41st Int. Conf. Telecommun. Signal Process. (TSP)*, Jul. 2018, pp. 1–5, doi: 10.1109/TSP.2018.8441452.
[45] S. Jameel, "Face recognition system using PCA and DCT in HMM," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 4, no. 1, pp. 13–18, Jan. 2015, doi: 10.17148/IJARCCE.2015.4103.
[46] W. Khreich, E. Granger, A. Miri, and R. Sabourin, "On the memory complexity of the forward–backward algorithm," *Pattern Recognit. Lett.*, vol. 31, no. 2, pp. 91–99, Jan. 2010, doi: 10.1016/J.PATREC.2009.09.023.
[47] M. T. Johnson, "Capacity and complexity of HMM duration modeling techniques," *IEEE Signal Process. Lett.*, vol. 12, no. 5, pp. 407–410, May 2005, doi: 10.1109/LSP.2005.845598.
[48] ORL Cambridge. *The ORL Database of Faces*. Accessed: Feb. 20, 2019. [Online]. Available: https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
[49] TYV Group. *Yale Faces*. Accessed: Feb. 20, 2019. [Online]. Available: https://vismod.media.mit.edu/vismod/classes/mas622-00/datasets/



DANISH ALI received the bachelor's degree in electrical engineering from the Government College University, Faisalabad, in 2017, and the master's degree in electrical engineering from NUST, Islamabad, Pakistan. His research interests include image processing and computer vision.

IMRAN TOUQIR received the bachelor's, master's, and Ph.D. degrees in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, in 1994, 2005, and 2009, respectively. He has been with the Faculty of Electrical Engineering, NUST, since 2009. His research interests include image processing, time frequency analysis, and probabilistic models.

ADIL MASOOD SIDDIQUI received the bachelor's, master's, and Ph.D. degrees in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, in 1994, 2005, and 2009, respectively. He has been with the Faculty of Electrical Engineering, NUST, since 2009. His research interest includes signal and image processing.

JABEEN MALIK received the bachelor's degree in electrical engineering from the University of Engineering and Technology, Taxila, in 2018. She is currently pursuing the master's degree in electrical engineering with NUST, Islamabad, Pakistan.

MUHAMMAD IMRAN received the bachelor's, master's, and Ph.D. degrees in electrical engineering from the National University of Sciences and Technology, Islamabad, Pakistan, in 2007, 2011, and 2014, respectively. He is currently with the Faculty of Electrical Engineering, NUST.