Face Recognition using Maximum Variance and SVD of Order Statistics with only Three States of Hidden Markov Model

Hameed R. Farhan
Dept. of Electrical and Electronic Engineering.
college of Engineering
University of Kerbala
Kerbala, Iraq

Mahmuod H. Al-Muifraje
Assistant Professor
FPGA System Design and Pattern Recognition Group
Dept. of Electrical Engineering.
University of Technology
Baghdad, Iraq

Thamir R. Saeed
Assistant Professor
FPGA System Design and Pattern Recognition Group
Dept. of Electrical Engineering.
University of Technology
Baghdad, Iraq

ABSTRACT
This paper presents a fast face recognition (FR) method using only three states of Hidden Markov Model (HMM), where the number of states is a major effective factor in computational complexity. Most of the researchers believe that each state represents one facial region, so they used five states or more according to the number of facial regions. In this work, a different idea has been proven, where the number of states is independent of the number of facial regions. The image is resized to 56x56, and order-statistic filters are used to improve the preprocessing operations and thereby reducing the influence of the illumination and noise. Up to three coefficients of Singular Value Decomposition (SVD) are utilized to describe overlapped blocks of size 5x56. Experimental results show that the proposed work manages to achieve 100% recognition rate on ORL face database using the maximum variance and two coefficients of SVD and can, therefore, be considered as the fastest face recognition type.

Keywords
Face recognition, Hidden Markov Model, Order Statistic Filter, Number of states of HMM, Singular Value Decomposition.

1. INTRODUCTION
Face recognition (FR) is a type of technical mechanism to identify people based on facial characteristics as vital elements. The research activities in the field of FR have recently witnessed considerable development in view of the many important applications in several disciplines and areas, such as access control, surveillance and security, criminal identification and credit card verification.

During the last 20 years, research on FR has received increasing attention and different approaches have been proposed in this field which varied in aspects including memory size and time elapsed, as well as, the recognition rate. The early approaches used a comparison of geometric features [1], and later the trend was to focus on reducing the system's complexity. Principle Component Analysis (PCA) is a dimensionality reduction method used to generate eigen faces. Orthogonal eigenvectors are obtained using normalization. The low-order eigenvectors are used, because they encode a wide variation in the training space, while the high-order eigenvectors represent noise that can be excluded [2]. Linear Discriminant Analysis (LDA) or fisher-face is another FR technique. LDA projects the training images onto a fisher-space specified by the fisher-faces. The test image is projected onto the fisher-space and an algorithm, such as the Euclidean distance, is applied for verification [3]. Independent Component Analysis (ICA) is a FR method, similar to PCA, but it finds a linear representation of face image, in which the components are statistically independent [4]. Another method is Support Vector Machine (SVM), where it constructs a hyper-plane to separate classes on both sides of the plane and the distance between each class and the hyper-plane is maximized [5]. Artificial Neural Network (ANN) is also used as a classifier of faces, where different methods of ANNs are used including either supervised or unsupervised learning models. ANNs have the ability of adapting, learning and clustering of data in a parallel processing operation. The task of ANNs is to train or cluster some face images for each person in the database within the hidden layers and obtain the best match class to the test image [6]. Recently, Hidden Markov Model (HMM) [7] is considered as one of the successful face classifiers. For each person, one HMM is generated to model features from a particular set of training images. The recognition step depends on the probability of the feature vectors of the test image with the trained HMMs, and the image can be considered representative of a person who gives a high probability. Different from the previous HMM methods which use five states or more, the proposed work presents a fast and robust FR method based on only three states and achieves 100% recognition rate with much lower computational complexity.

2. RELATED WORK
Many surveys on HMM have been performed regarding FR techniques. The major differences between these efforts are restricted in the nature of feature extraction methods and the types of HMM. Some of these are briefly described below.

The first attempt of FR system based on HMM was reported by F. Samaria and F. Fallside in 1993 [8]. They used an 8-state ergodic (fully connected) HMM and 5-state left-to-right HMM. A given image matrix of pixel intensity values is scanned in overlapping blocks from top to bottom to form the observation vectors used in HMM. The pixel intensities are more complicated in calculations due to the large memory they occupy as well as the processing time. Thereafter, the trend is to reduce the size of the observation vectors by using dimensionality reduction methods instead of pixel intensities. In [9, 10, 11, 12, 13, 14], Discrete Cosine Transform (DCT) was used to extract features from images. The Wavelet Transform (WT) is also used either in continuous or discrete form as in [15, 16, 17, 18, 19, 20, 21, 22, 23, 24]. In [15, 17];

the authors made a comparison between DCT and DWT and the results proved that the DWT is the best. Many other methods have been used, such as ANN [25], Local Binary Pattern (LBP) [26, 27], Singular Value Decomposition (SVD) [28], Fast Fourier Transform (FFT) and the Partial Least Squares (PLS) [29], Canonical Correlation Analysis (CCA) [30], Active Appearance Model and Shape Model [31], and Separable Lattice [32,33].

Most of the researchers used 5-state HMM [8, 9, 10, 11, 12, 13, 16, 18, 20, 21, 25, 26, 27, 30] and others used 7-state HMM [22, 23, 28]. In [19], 64 states were used, while in [24], the authors used 9-state HMM for Yale database and 15-state HMM for ORL database. Some researchers used variable number of states; (5-25) states and the best is 12 states [31], (8-64) states and the best is 40 [32], (4-12) states and the best is 8 [33]. The first effort of using 3-state HMM in FR is carried out in [14]. It achieves a poor recognition rate and needs a specific modification to improve the performance.

3. BACKGROUND

3.1 Nonlinear Spatial Filters

Nonlinear spatial filtering, as the name implies, is a nonlinear filtering operation which encompasses the surrounding of the pixels by sliding the center point of a filter window through an image. The Order Statistic Filter is a kind of nonlinear spatial filter where the center pixel of the sliding window is replaced by the ordering (ranking) value [34].

3.1.1 Minimum Order Statistic Filter

In order to improve the performance of the FR system, Minimum Order Statistic Filter (MOSF) is used as follows: The face image is scanned from top to bottom and left to right in steps of one pixel using a mask window. The minimum value in the window replaces the centered of that window after excluded all zeros in the window.

3.1.2 Median Filter

A median filter is a special kind of Order Statistic Filter which gives better results in eliminating noise from the image especially salt and pepper noise. The operation of the filter is the same as that discussed in MOSF, except the center point is replaced by the median value instead of the minimum value. The median value is obtained by arranging all the values in ascending order from low to high, and then taking the center value. If there are two values in the center, their average is taken. The 2D block image is converted to 1D vector of length n and arranged in ascending order as illustrated below [36]:

Let $X_1, X_2, \dots\dots, X_n$ be pixel values of the vector (X_i) after sorting in ascending order, where X_1 and X_n are minimum and maximum values respectively. The median of (X_i) is given by

$$median(X_i) = \begin{cases} X_{(v+1)} & n = odd \\ \dfrac{X_{(v)} + X_{(v+1)}}{2} & n = even \end{cases} \qquad (1)$$

where v is the integer value of $n/2$, or in Matlab statement

$\{\ v = floor(n/2)\}$. For simplicity, filter windows having odd length are used in most cases.

3.2 Singular Value Decomposition

The SVD is a mathematical method widely used in signal and image processing. Let X be a real matrix with rank of k, SVD of X consists of three matrices U, Σ and V, and defined as;

$$X = U\Sigma V^T \qquad (2)$$

U and V are unitary matrices with the properties $U^T U = I$ and $V^T V = I$, where I is the identity matrix. The non-zero elements of the diagonal matrix Σ are arranged in descending order $(\Sigma_{11} \geq \Sigma_{22} \geq \ldots\ldots \geq \Sigma_{kk})$ which represent the singular values of X. The columns of Umatrix are the left singular vectors for corresponding singular values. The columns of V are the right singular vectors for corresponding singular values. Therefore, the SVD of X can be expressed as follows [30]:

$$X = [U_1, U_2, \ldots., U_k] \begin{bmatrix} \Sigma_{11} & & \\ & \ddots & \\ & & \Sigma_{kk} \end{bmatrix} [V_1, V_2, \ldots., V_k] \qquad (3)$$

3.3 Hidden Markov Model

HMM is a 1D model in which the transitions between the states are invisible (hidden) and only the outputs of the states are observed. The parameters of HMM are [36]:

1) N; Number of states in the model. The interconnections of the states depend on the topology of the model. The individual states are denoted as $S = \{S_1, S_2, \ldots, S_N\}$, and the state at time t as q_t .

2) M; Number of symbols in each state or the range of values employed in the observation vectors (codebook). The individual symbols are denoted as $V = \{v_1, v_2, \ldots., v_M\}$.

3) $A = [N\text{x}N]$; the transition matrix which comprises the probabilities of the state transitions, where

$$A = \{a_{ij}\} \qquad 1 \leq i, j \leq N \qquad (4)$$

$$\{a_{ij}\} = \mathrm{P}(q_{t+1} = S_j \mid q_t = S_i), \qquad 1 \leq i, j \leq N \qquad (5)$$

$$\sum_{j=1}^{N} a_{ij} = 1, \qquad 1 \leq i \leq N \qquad (6)$$

4) $B = [N\text{x}M]$; the emission matrix or the observation symbols probability distribution in the states, where

$$B = \{b_j(k)\}, \qquad 1 \leq j \leq N \qquad (7)$$

$$b_j(k) = P(O_t = v_k \mid q_t = S_j),\ 1 \leq j \leq N, 1 \leq k \leq M \qquad (8)$$

$$\sum_{j=1}^{N} b_j(k) = 1, \qquad 1 \leq k \leq M \qquad (9)$$

Each observation vector (O_t) comprises number of elements and $O = (O_1, O_2, \ldots, O_T)$ be the observation sequence at T different observation instances, whereas the corresponding state sequence be $S = (q_1, q_2, \ldots, q_T)$, where $q_t \in$ {1, 2,..,N}.

5) $\pi = [N\text{x}1]$; the initial state vector, $\boldsymbol{\pi} = \{\pi_i\}$, where

$$\pi_i = P(q_1 = S_i), \qquad 1 \leq i \leq N \qquad (10)$$

$$\sum_{i=1}^{N} \pi_i = 1 \qquad (11)$$

The model is indicated by the following notation:

$$\lambda = (A, B, \pi) \tag{12}$$

For more details about HMM and the associated problems with the suggested solutions, the reader may refer to the most popular article introduced in [36].

4. THE PROPOSED WORK

The proposed FR system was performed employing the three steps; feature extraction, training and testing processes. Fig. 1 illustrates the main steps of the FR system.

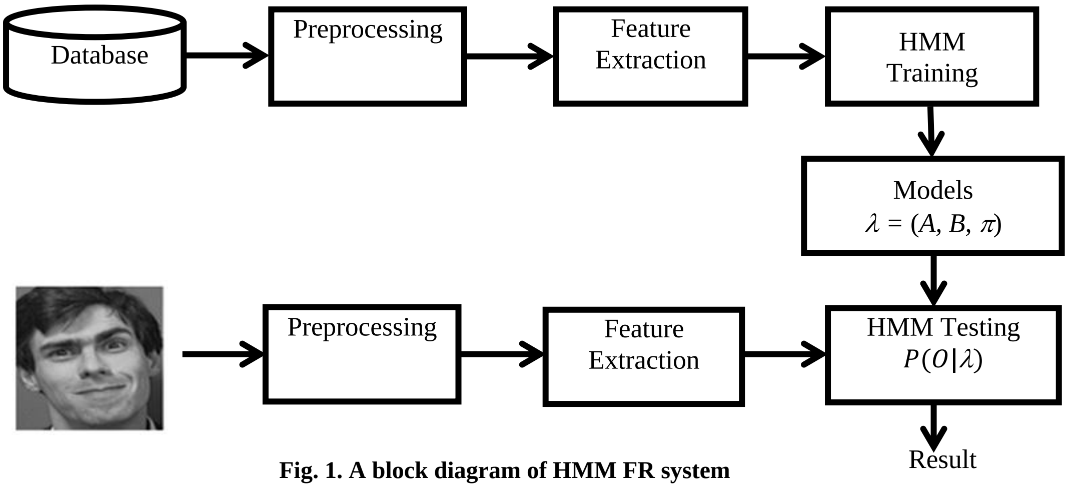


Fig. 1. A block diagram of HMM FR system

The first step of the procedure proceeds with the image being resized to 56x56 and one of the Order Statistic Filters (Minimum or Median) is used to improve the processing and eliminate the effect of noise and illumination caused by camera flash. The general Matlab expression for Order Statistic Filter is given by:

$$Y = ordfilt2(X, order, domain) \tag{13}$$

where Y and X are output and input images respectively, order is the order of the filter output value and domain is the filtering window. We use a filter window of size 3x3, $(order = 1)$ for minimum filter, and $(order = 5)$ which represents the midpoint of domain matrix for median filter. Another Matlab function for median filter $(medfilt2)$ can be used instead of $(ordfilt2)$ which gives the same result. Fig. 2 shows a sample image before and after filtering.

(a) **(b)**

Fig. 2. Image filtering, (a) A sample image before filtering (b) Filtered image.

The second step is to extract features from filtered images. A sequence of overlapping blocks is generated with block size 5x56 and an overlap of 4 pixels in height with the same width as shown in Fig. 3.

The number of blocks obtained can be calculated according to the following equation:

$$N = \frac{H - L}{L - P} + 1 \tag{14}$$

where, H is the height of the image, L is the height of the block and P is the amount of overlap. After substituting $H = 56$, $L = 5$, and $P = 4$, in (14), yielding $N = 52$ blocks.

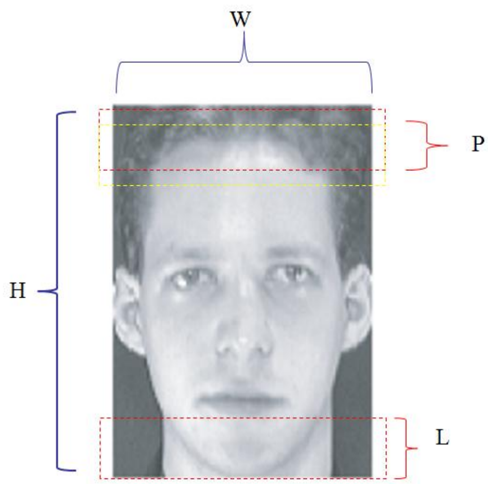


Fig. 3. Blocks extraction

The SVD of each block is calculated followed by using only the top left element of U matrix (U_{11}) and the first two non-zero elements of the singular value vector Σ $(\Sigma_{11}\ and\ \Sigma_{22})$ to represent block features. Thus, the block included 280 pixels is implemented by three numbers only. Therefore, this reduction in the size of observation vectors leads to high reduction in computational complexity of the system.

The next step is the quantization of the features to provide a precise separation of the blocks. The quantization process is done as follows; let $A = [U_{11}\ \ \Sigma_{11}\ \ \Sigma_{22}]$ be the vector of the three coefficients and $L = [L_1\ \ L_2\ \ L_3]$ is the vector of quantization levels, where each coefficient would be quantized to a distinct level L_i. The difference between two successive quantized vectors can be calculated according to the following equation:

$$D = \frac{A_{max} - A_{min}}{L} \tag{15}$$

where D is the difference vector, A_{max} and A_{min} are the maximum and minimum values of two successive A vectors respectively. The maximum and minimum values are consequently replaced with new values depending on the comparison between the two vectors, whereas L remains constant. The final value of D is obtained when all differences between the successive vectors are computed. So the quantized vector of A is obtained as follows;

$$Q = \frac{A - A_{min}}{D} \tag{16}$$

The vector A is replaced with its quantized vector $Q = [Q_1 \;\; Q_2 \;\; Q_3]$ for all 52 blocks and each Q vector is implemented by a distinct label according to the following formula

$$Label = Q_1 * W_1 + Q_2 * W_2 + Q_3 * W_3 \tag{17}$$

The vector $W = [W_1 \;\; W_2 \;\; W_3]$ is chosen experimentally with constant values for all calculations. The last step produces a sequence of integer values for each image which represents an input data for HMM.

In the training process, only three states of left-to-right HMM topology are used as shown in Fig. 4.

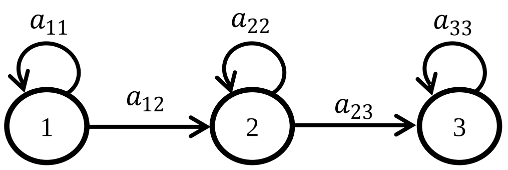


Fig. 4. State diagram of 3-state left-to-right HMM

The initial estimations of the three parameters of the model (π, A, B) are chosen as follows;

$$\pi = [1 \;\; 0 \;\; 0] \tag{18}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} \tag{19}$$

The possible values of (a_{11}, a_{12}) and (a_{22}, a_{23}) are (0.9, 0.1), (0.8, 0.2), (0.7, 0.3), (0.6, 0.4) and (0.5, 0.5), whereas $a_{33} = 1$ according to the constraint in (6). The matrix B is estimated using the following Matlab equation:

$$B = \frac{1}{M} * ones(3, M) \tag{20}$$

where M represents the number of observation symbols per state, and in this approach it ranged between 180 and 220. The HMM (one model for each person) is trained using Baum Welch algorithm shown in Fig. 5, where five images for each person are trained and the algorithm iterates two times.

In the testing process, each test image is treated similarly as training image in obtaining observation vectors. The probability of the observation vector of the test image with all models trained is obtained using Viterbi algorithm as illustrated in Fig. 6. The model which gives a maximum probability is assumed to be an index to the identified person.

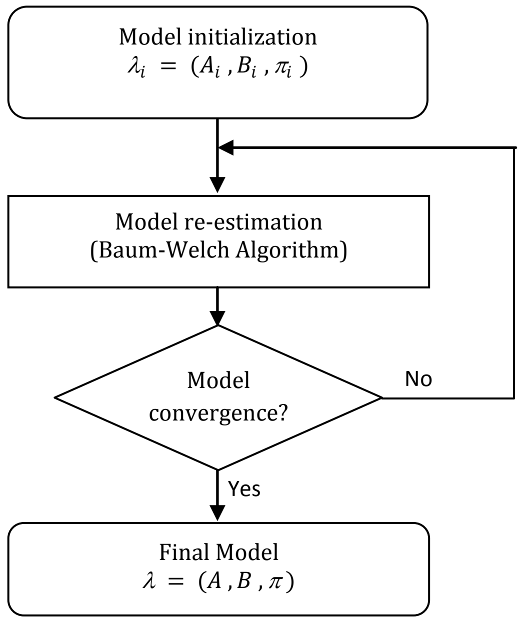


Fig. 5. Training process algorithm.

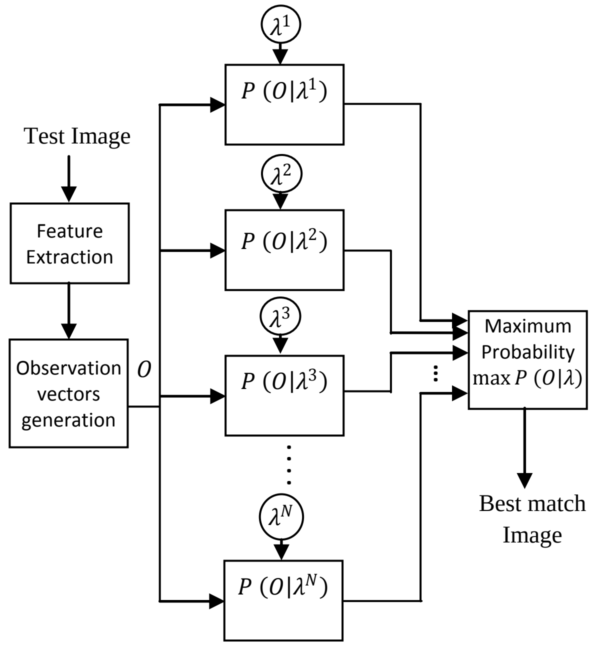


Fig. 6. Recognition process algorithm.

5. EXPERIMENTAL RESULTS

The experiments were conducted on ORL database [40], which consists of 400 gray-scale face images corresponding to 40 persons (10 images for each person). The size of each image is 112×92 pixels and each pixel is represented by 256 gray levels or eight bits in binary. For each individual, five images were chosen for training and the rest for testing.

Each image was resized to 56x56 and several tests were carried out using two types of order filter, as well as without filtering. The experiments exhibited the utility of using filters to improve the performance of the system. The results of using median filter are better than those of minimum filter.

Further test was done by replacing the first coefficient (U_{11}) with the maximum variance of the block and achieved 100% recognition rate as illustrated in Table 1.

Table 1. Experimental results

Filter Type	No. of Observation symbols (M)	Coefficients	Training Time for one image (sec.)	Testing Time for one image (sec.)	Recognition Rate (%)
Non	220	$[U_{11} \quad \Sigma_{11} \quad \Sigma_{22}]$	0.019	0.043	97.5
Minimum	220	$[U_{11} \quad \Sigma_{11} \quad \Sigma_{22}]$	0.020	0.044	99
Median	180	$[U_{11} \quad \Sigma_{11} \quad \Sigma_{22}]$	0.020	0.044	99.5
Non	200	$[Var. \quad \Sigma_{11} \quad \Sigma_{22}]$	0.023	0.047	95.5
Minimum	200	$[Var. \quad \Sigma_{11} \quad \Sigma_{22}]$	0.024	0.048	97.5
Median	200	$[Var. \quad \Sigma_{11} \quad \Sigma_{22}]$	0.024	0.048	100

The system accomplishes a high training speed due to using only three states of HMM and the number of iterations which is no more than two. The value of L and W vectors were chosen experimentally, where L ranges from 7 - 15, and W is ranged between 1 and 25. Table 2 shows a comparison between the proposed work and the others which have used HMM and ORL database.

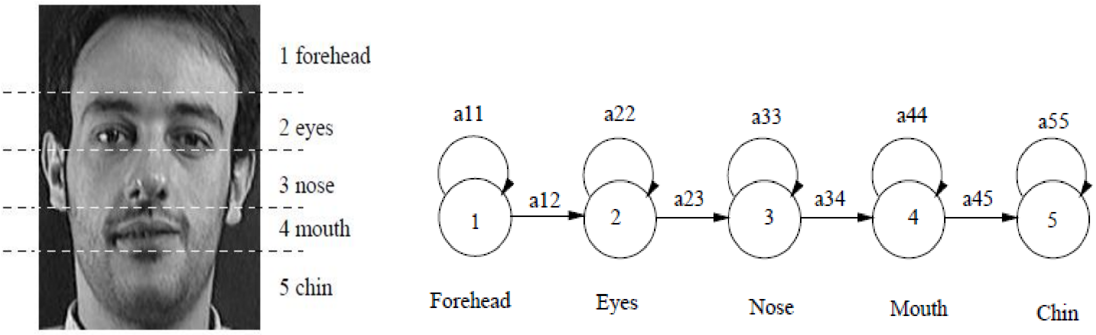
Table 2. Comparative results of different HMM FR approaches on ORL database

Year	Ref. No.	Feature Extraction Method	No. of States	Training Time for one image (second)	Testing Time for one image (second)	Recognition Rate (%)
1998	9	DCT	5	23.5	3.5	99.5
1999	10	DCT	5	NA	NA	98
2003	15	WT	NA	NA	NA	100
2003	16	DWT	5	1.13	0.3	100
2003	17	DWT	NA	NA	NA	97.5
2006	25	ANN	5	NA	NA	100
2008	28	SVD	7	0.63	0.28	99
2009	18	WT	5	NA	NA	97.1
2009	29	FFT+PLS	NA	NA	NA	83.4
2010	20	WT+PCA	5	NA	1.2	96.5
2010	30	CCA	5	8.63	33.2	95
2011	12	DCT	5	5.2	0.25	91
2011	21	DWT	5	1.3	0.8	100
2012	14	DCT	3	12.6*	71.7*	82.76
2013	22	DGWT	7	NA	NA	99
2014	24	GF+LDA	15	NA	NA	97.5
2015	This Paper	SVD	3	0.024	0.048	100

(*) The authors did not indicate the time mentioned is returned to a single image or set of images.

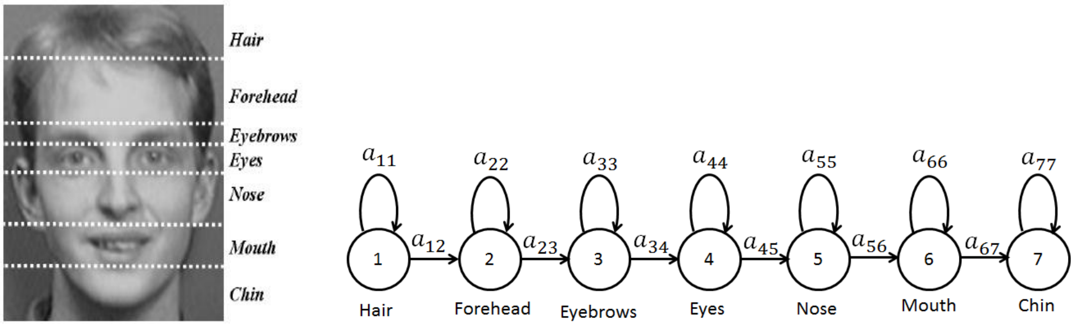
It is evident from Table 2 that the proposed system attains maximum recognition rate and lowest processing time due to the minimum number of states used.

Most of the researchers used 5-state HMM in FR system depending on their persuasion about the relationship between the number of states and the facial region. They divide the face image into five regions (forehead, eyes, nose, mouth and chin) or (head, forehead, eyes, nose and mouth) as shown in Fig. 7. Others added two more states to the system according to seven regions of the face image as illustrated in Fig. 8.



(a) (b)

Fig. 7. 5-state face modeling; (a) Facial regions, (b) Left-to-right HMM



(a) (b)

Fig. 8. 7-state face modeling; (a) Facial regions, (b) Left-to-right HMM

The following factors are deduced to reject the relationship between the number of states and the facial regions:

1. Facial regions are five or more.
2. Each facial region consists of overlapped blocks, where each block represents the output of one state. Therefore, no such facial region appears in separate state.
3. In the proposed system, the overall three states operate in a small part of the facial region at one cycle which obviously supports our trend.

Based on our knowledge in the literature, only the research introduced in [14] had been conducted using three states of HMM FR. The maximum recognition rate achieved is 82.76% using ORL database which is rather far from our result in the proposed work, as well as in the processing time.

The computational complexity of our system can be viewed from the transition and emission matrices. The size of the transition matrix (NxN) is 3x3 and the size of the emission matrix (NxM) is 3x200 for the best result, whereas in [28], the two matrices are of size 7x7 and 7x1260 respectively.

The experiments were carried out in Matlab (Matrix Laboratory) environment (R2014a) under Windows 7 (32-bit) operating system with 2.53 GHz processor speed.

6. CONCLUSIONS

In this work, a fast FR system is proposed using only three states of HMM, so the idea dominated for the last decades about the relationship between the number of states and the facial regions has been refuted. Some researchers had divided the face image into five regions and others believe that more details can be extracted if the image is split into seven regions. In this approach, a scientific concept has been rigorously proven. The concept boils down to the fact that there is not any relationship between the number of states in HMM and the facial regions. Using only three states in HMM with two training iterations reduces the computational complexity and the memory used. A high reduction in the time consumed is obtained while achieving a recognition rate of 100% that supports this trend significantly.

In future work, focus will be on using only two states of HMM and also try to obtain more reliable feature extraction methods.

7. REFERENCES

[1] R. Brunelli and T. Poggio, "*Face Recognition: Features versus Templates*"; IEEE Trans. on Pattern Recognition and Machine Intelligence; No. 10, PP. 1042-1052, Vol. 15, 1993, October.

[2] T. Karim, M. Lipu, M. Rahman, and F. Sultana, "*Face Recognition Using PCA-Based Method*", In Proc. of the International Conference on Advanced Management Science (ICAMS), Vol. 3, PP. 158-162, IEEE, 2010.

[3] F. Chelali, A. Djeradi, and R. Djeradi, "*Linear Discriminant Analysis for Face Recognition*", In Proc. of the International Conference on Multimedia Computing and Systems (MMCS. 2009), PP. 1-10, IEEE, 2009.

[4] M. Bartlett, J. Movellan, and T. Sejnowski, "*Face recognition by independent component analysis*", IEEE Trans. On Neural Networks, No. 6, PP. 1450-1464, Vol. 13, 2002, Nov.

[5] G. Guo, S. Li, and K. Chan, "*Face recognition by support vector machines*", In Proc. of IEEE International Conference on Automatic Face and Gesture Recognition (FG '00), Grenoble, France, PP.196-201, IEEE, 2000.

[6] S. Nazeer, N. Omar, and M. Khalid, "*Face Recognition System using Artificial Neural Networks Approach*," in International Conference on Signal Processing, Communications and Networking (ICSCN '07), Feb. 2007, Chennai, India, PP. 420-425, IEEE, 2007.

[7] L. Rabiner and B. Juang. "*An Introduction to Hidden Markov Models*", IEEE ASSP Mag., Vol. 3, No. 1, PP. 4-16, IEEE, Jan. 1986.

[8] F. Samaria and F. Fallside. "*Face identification and feature extraction using hidden markov models*", In G. Vernazza, editor, Image Processing: Theory and Applications. Elsevier, 1993.

[9] V. Kohir and U. Desai, "*Face recognition using a DCT-HMM approach*", In Proc. of Fourth IEEE Workshop on Applications of Computer Vision, WACV '98, Oct 1998, Princeton, NJ, PP.226-231, IEEE, 1998.

[10] A. Nefian and M. Hayes, "*An Embedded HMM-based Approach for Face Detection and Recognition*", In Proc. of IEEE International Conference On Acoustics, Speech and Signal Processing, Mar 1999, Vol. 6, PP. 3553-3556, IEEE, 1999.

[11] S. Eickeler, S. Muller, and G. Rigoll, "*Recognition of jpeg compressed face images based on statistical methods*", Image and Vision Computing, Vol. 18, No. 4, PP. 279–287, Elsevier, Mar. 2000.

[12] B. Vaseghi and S. Hashemi, "*Face Verification Using D-HMM and Adaptive K-Means Clustering*", In Proc. of 4th IEEE International Conference on Broadband Network and Multimedia Technology (IC-BNMT), Oct. 2011, PP. 270 – 275, IEEE, 2011.

[13] P. Xiaorong, Z. Zhihu, T. Heng, and L. Tai, "*Partially occluded face recognition using subface hidden Markov models*", In Proc. of 7th International Conference on Computing and Convergence Technology (ICCCT), Dec. 2012 , Seoul , PP.720-725, IEEE, 2012.

[14] K. Singh, M. Zaveri, and M. Raghuwanshi, "*Recognizing Faces under Varying Poses with Three States Hidden Markov Model*", In Proc. of IEEE International Conference on Computer Science and Automation Engineering (CSAE), May 2012, Zhangjiajie, China, , Vol. 2, PP. 359–363, IEEE, 2012.

[15] M. Bicego, U. Castellani, and V. Murino, "*Using Hidden Markov Models and Wavelets for face recognition*", In Proc. of 12th International Conference on Image Analysis and Processing (ICIAP '03), Mantova, Italy, September 2003, PP. 52–56, IEEE, 2003.

[16] H. Le and H. Li, "*Simple 1D Discrete Hidden Markov Models for Face Recognition*", In Proc. of 8th International workshop (Visual Content Processing and Representation), (VLBV 2003), Sep. 2003, Madrid, Spain, PP. 41–49, Springer-Verlag Berlin Heidelberg, 2003.

[17] L. Bai and L. Shen, "*Combining Wavelets with HMM for Face Recognition*", In Proc. of the 23rd International Conference on Innovative Techniques and Applications of Artificial Intelligence (SGAI '03), Dec. 2003, Cambridge, UK, PP. 227-233, Springer-Verlag London, 2003.

[18] I. Makaremi and M. Ahmadi "*A Robust Wavelet Based Feature Extraction Method for Face Recognition*", In Proc. of IEEE International Conference on Systems, Man, and Cybernetics (SMC 2009), Oct. 2009, San Antonio, TX, USA, PP. 2173-2176, IEEE, 2009.

[19] R. Shrivastava and A. Nigam, "*Analysis and performance of face recognition system using Gabor filter bank with HMM model*", In Proc. of 2nd International Conference on Trendz in Information Sciences & Computing (TISC 2010) , Dec. 2010, Chennai , PP. 239-244, IEEE, 2010.

[20] L. Peng and L. Jiao, "*Implement of face recognition system based on Hidden Markov Model*", In Proc. of sixth International Conference on Natural Computation (ICNC 2010), Aug. 2010, Yantai, Shandong, Vol. 7, PP. 3344-3348, IEEE, 2010.

[21] E. Abbas and H. Farhan, "*Face Recognition using DWT with HMM*", Engineering & Technology Journal, UOT, Baghdad, Iraq Vol. 30, No. 1, PP. 142-154, 2012.

[22] M. Srinivasan and N. Ravichandran, "*A new technique for Face Recognition using 2D-Gabor Wavelet Transform with 2D-Hidden Markov Model approach*" In 2013 International Conference on Signal Processing Image Processing & Pattern Recognition (ICSIPR), Feb. 2013, Coimbatore, PP. 151-156, IEEE, 2013.

[23] M. Srinivasan and N. Ravichandran, "*A 2D Discrete Wavelet Transform Based 7-State Hidden Markov Model for Efficient Face Recognition*", In 4th International Conference on Intelligent Systems Modelling & Simulation (ISMS 2013) , Jan. 2013, Bangkok, PP. 199-203, IEEE, 2013.

[24] Z. Elgarrai, O. Meslouhi, H. Allali, M. Kardouchi, and S. Selouani, "*Face Recognition System Using Gabor Features and HTK Toolkit*", In Tenth International Conference on Signal-Image Technology and Internet-Based Systems (SITIS), Nov. 2014, Marrakech, PP. 32-36, 23-27, IEEE, 2014.

[25] V. Bevilacqua, L. Cariello, G. Carro, D. Daleno, and G. Mastronardi, "*A face recognition system based on Pseudo 2D HMM applied to neural network coefficients*", Journal of Soft Computing-A Fusion of Foundations, Methodologies and Applications, Vol. 12, No. 7, PP. 615-621, Springer-Verlag, 2008.

[26] J. Cao and C. Tong, "*Facial Expression Recognition Based on LBP-EHMM*", In Congress on Image and Signal Processing (CISP '08), May 2008, Sanya, China, vol. 2, PP. 371-375, IEEE, 2008.

[27] N. Vu and A. Caplier, "*Patch-based similarity HMMs for face recognition with a single reference image*", In Proc. of the 20th International Conference on Pattern Recognition (ICPR 2010), Aug. 2010, Istanbul, Turkey, PP. 1204-1207, IEEE, 2010.

[28] H. Naimi and P. Davari, "*A New Fast and Efficient HMM-Based Face Recognition System Using a 7-State HMM Along With SVD Coefficients*", Iranian Journal of Electrical & Electronic Engineering, Vol. 4, Nos. 1&2, PP. 46-57, January 2008.

[29] H. Yegang and L. Benyong, "*Face Recognition Based on PLS and HMM*" in Chinese Conference on Pattern Recognition (CCPR 2009), Nov. 2009, Nanjing, China, PP. 1-4, IEEE, 2009.

[30] J. Cai, H. Ren, and Y. Yin, "*Non-overlapped Sampling Based Hidden Markov Model for Face Recognition*", 3rd International Congress on Image and Signal Processing (CISP), Oct. 2010, Yantai, China, Vol. 4, PP. 1901-1904, IEEE, 2010.

[31] M. Islam, R. Toufiq, and M. Rahman, "*Appearance and shape based facial recognition system using PCA and HMM*", In 7th International Conference on Electrical & Computer Engineering (ICECE), Dec. 2012, Dhaka, Bangladesh, PP. 1-4, IEEE, 2012.

[32] K. Sawada, A. Tamamori, K. Hashimoto, Y. Nankaku and K. Tokuda, "*Face recognition based on separable lattice 2-D HMMS using variational bayesian method*", In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), March 2012, Kyoto, Japan, PP. 2205-2208, IEEE, 2012.

[33] A. Tamamori, Y. Nankaku, and K. Tokuda, "*Image recognition based on separable lattice trajectory 2-D HMMS*", in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2013, Vancouver, BC, Canada, PP. 3467-3471, IEEE, 2013.

[34] R. Gonzalez, R. Woods and S. Eddins, "*Digital Image Processing using Matlab*", 2nd edition, Prentice Hall, 2009.

[35] I. Pitas and A. Venetsanopoulos, "*Order statistics in digital image processing*", Proc. IEEE, Vol. 80, No. 12, PP. 1893-1921, IEEE, Dec. 1992.

[36] L. Rabiner, "*A tutorial on hidden markov models and selected application in speech recognition*", In Proc. of IEEE, vol. 77, No. 2, IEEE, PP. 257-286, 1989.

[37] L. Welch, "*Hidden Markov Models and the Baum-Welch Algorithm*", IEEE Information Theory Society Newsletter, Vol. 53, No. 4, Dec. 2003.

[38] B. Juang and L. Rabiner, "*The segmental K-means algorithm for estimating parameters of hidden Markov models*", IEEE Trans. On Acoustics, Speech and Signal Processing, No. 9, PP. 1639 – 1641, Vol. 38, IEEE, 1990.

[39] A. Viterbi, "*A Personal History of the Viterbi Algorithm*", IEEE Signal Processing Magazine, Vol. 23, No. 4, PP. 120-142, IEEE, Jul. 2006.

[40] AT&T Laboratories, Cambridge, U.K., "*The ORL Face Database*", Available at:http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html