International Journal of Research in Advent Technology, Vol.7, No.5S, May 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

A Comprehensive Survey on HMM-Based Face Recognition Methods

C. Anand,
Assistant Professor, Dept. of Computer Science (SF),
Ayya Nadar Janaki Ammal College, Sivakasi.
anandc.chandrasekaran@gmail.com

Abstract--- Face Recognition is an active and significant research area since it has plenty of application domains in pattern recognition, image processing, biometrics etc. Researchers contributed lot of algorithms and techniques to uncover the mask of face recognition arena. In this paper, Hidden Markov Models (HMM) based various face recognition methods have been reviewed and its significant features, recognition rate and performance are analyzed.

Keywords--- Face Recognition, Hidden Markov Models (HMM), KLT Coefficients, MC-HMM, Pattern Recognition, SVD Coefficients.

1. INTRODUCTION

A. Pattern Recognition

Humans have developed highly sophisticated skills for sensing their environment and taking actions according to what they observe such as recognizing a face, understanding spoken words, reading handwriting, distinguishing fresh food from its smell etc. The similar capabilities that human would perceive are applied to machines result in pattern recognition [1].

A pattern is an entity that could be given a name such as finger print image, handwritten word, human face, speech signal or a DNA sequence. Pattern recognition is the study of how machines can observe the environment, learn to distinguish patterns of interest, make sound and reasonable decisions about the categories of patterns [1].

TABLE I
APPLICATIONS OF PATTERN RECOGNITION

Problem Domain	Application	Input Pattern	Pattern Classes
Speech Recognition	Telephone Directory Assistance	Speech Waveform	Spoken Words
Biometric Recognition	Personal Identification	Face, Iris, Fingerprint	Authorized Users
Medical	Computer Aided Diagnosis	Microscopic image	Cancerous /healthy Cell
Military	Automatic target recognition	Infrared images	Target type

B. Face Recognition

Humans often use faces to recognize individuals. Early face recognition algorithms used simple geometric models and the recognition process now matured into a science of sophisticated mathematical representations and matching processes. There are two predominant approaches that used in the face recognition problem are: feature-based and view-based [2].

Developed in the 1960s, the first semi-automated system for face recognition required the administrator to locate features on the photographs before calculating distances and ratios to a common reference point, which were compared to reference data. In the 1970s, Lesk used 21 specific subjective markers such as hair color and lip thickness to automate the recognition. The problem with both of these early solutions was that the measurements and locations were manually computed [3].

In 1988, Kirby and Sirovich applied principal component analysis, a standard linear algebra technique to the face recognition problem [4]. In 1991, Turk and Pentland discovered the eigen faces technique [5]. The technology first captured the public's attention from the media reaction to a trial implementation at the January 2001 Super Bowl, which captured surveillance images and compared them to a database of digital mug shots. Today, it is very much useful to combat passport fraud, support law enforcement, identify missing children, minimize fraud etc [6].

2. BACKGROUND

A. Hidden Markov Models

Hidden Markov Models (HMM) is a statistical Markov model which is modeled and assumed to be a Markov process with unobserved (hidden) states and it resembles a simple dynamic Bayesian network.HMM is closely related to optimal non-linear filtering problem such as a stochastic process. In HMM, the state is not directly visible but the output dependent on the state is visible and each state has a probability distribution over the possible output tokens.HMM can be considered as a generalization of a mixture model where the hidden variables which control the mixture component are to be selected for each observation are related to a Markov process [7].

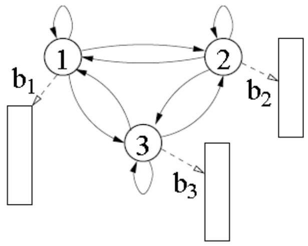

Fig.1 A simple HMM Model

B. Types of HMM

HMM models are of three types and Fig.2 represent the three HMM models. The Ergodic or fully connected HMM means that every state of the model could be reached in a single step from every other state of the model, i.e., an ergodic model has a property that every state can be reached from every other state in a finite number of steps. Left-Right or a Bakis model is underlying on the fact that as time increases, the state index increases or stays the same, i.e., the states proceed from left to right. The fundamental property of all left-right HMMs is that no transition is allowed to states whose indices are lower than the current state. A cross coupled connection of two parallel left-right HMM is a left-right model that has the flexibility over Bakis model due to its re-estimation procedure [1].

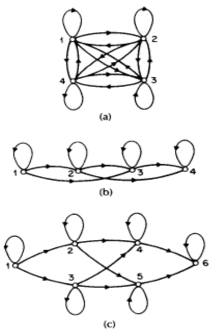

Fig.2 Illustration of 3-distinct types of HMMs
a) A 4-state Ergodic Model
b) A 4-state L-R Model c) A 6-state Parallel Path L-R Model

C. Applications of HMM

HMM is applied in a variety of fields such as speech recognition, language modeling, information retrieval, motion video analysis/tracking, protein sequence and genetic sequence alignment and analysis, financial time series prediction, etc. It is quite useful in applying probabilistic models for analyzing discrete sequence data in molecular and computational biology. A lot of interesting applications of HMM in Computational Biology includes finding whole genome shotgun sequence fragment assembly, multiple alignments of conserved sequences, splice site detection and implying phylogenetic trees [8]. HMM can be applied to other application domains such as E-Commerce to extract product/pricing information from many sites, convert information into structured format and to provide interface to look up product information. Users can consult a single site rather than navigating to and searching many sites [9].

3. SURVEY ON HMM BASED FACE RECOGNITION METHODS

A. Face Recognition using Hidden Markov Models

In 1994, Samaria F.S. introduced a face recognition system based on HMM, i.e., considered foremost research done using HMM [10]. The HMM method is based on matching image templates to a chain of states of a doubly-embedded stochastic model. In his work, parameterized 1D HMM in the shape of top-bottom models is used. In addition to this, top-bottom models are extended to pseudo 2D model which provides enhanced flexibility is also useful in research. The performance of HMM is evaluated in a fully automated system by cropping the database images automatically.

1) Face recognition using 1D HMM: In order to locate the facial regions, continuous density HMMs are used. The boundaries between the regions are represented by probabilistic transition between states and the actual images within a region are to be modeled by a multivariate Gaussian distribution. The system is trained and tested on ORL face database consist of 400 images of 40 subjects. The model is trained with 5 face images of a subject and each image generates an observation sequence. Experiments were carried out on Ergodic and Top-Bottom HMM's. Fig.3 illustrates the sampling technique for an Ergodic HMM on a given face image.

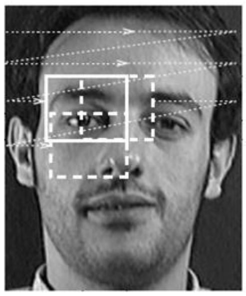

Fig.3 Sampling Technique for an Ergodic HMM

The sampling window moves from left to right and top to bottom and the observation sequence is created. When the right edge on the current line is reached, the sampling window moves back to the beginning of the line and an overlap between successive lines are created. An overlap of 25% occurs in each direction of sliding and the mean and standard deviation of each of the distributions was computed. Samaria introduced a five state top-bottom HMM for five facial regions of a given face image as depicted in Fig.4.

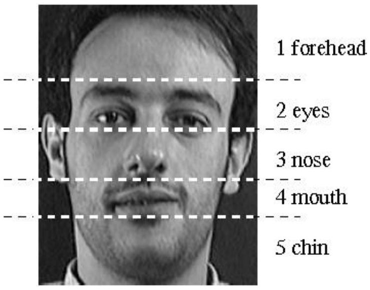

Fig.4 Top-Bottom 5-State HMM

Each training image is sampled by a block of 8 lines moving down in steps of 1 line. A high percent of overlap significantly increases the system performance. The 1D-HMM based on Ergodic and Top-Down model are tested against five untested faces on ORL database and achieves 85% accuracy [10].

2) Face recognition using Pseudo 2D HMM: In a 1D-HMM, each image is sampled using a block of lines and a horizontal alignment of face images is missing. Pseudo 2D HMM model allows both horizontal and vertical alignment. A typical P2D HMM and its equivalent 1D-HMM is shown in Fig. 5.

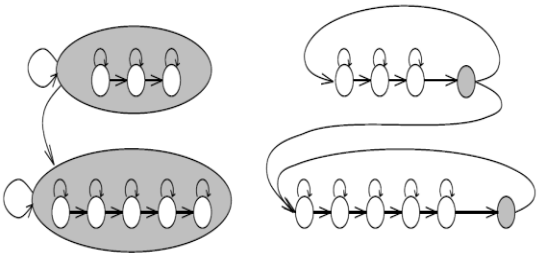

Fig. 5 Structure of a P2D HMM

In a P2D HMM, images are scanned from left-right and top-bottom and the scanned samples are arranged in a 2D lattice. P2D HMM structures are obtained by linking 1D left-right HMMs to form vertical super states and the network is not fully connected in two dimensions is called pseudo-2D HMM. Transitions are allowed only among the states of a super-state in horizontal direction and transitions occur among different super states in vertical direction. The sampling technique is slightly modified by adding a white frame at the end of each line of sampling. When an end-of-line state is reached, the model can either stay in the same row of states or jump to the next row of states. When the end-of-state line of the last row of states is reached, the model can either repeat the last row of states or terminate an observation. Experimental results show 95% accuracy on ORL database [10].

The bottleneck of P2D-HMM is that it requires greater number of states and super-states for a face recognition problem. Based on the topology, number of states is formed for every super-state. The Error rate also varies based on the topology which requires extensive experiments. The training time per face image is 25 seconds which in turn increasing the computational complexity and decreasing the system performance.

B. Hidden Markov Models for Face Recognition

In order to reduce the computational complexity of previous face recognition methods based on HMM, Nefian and Hayes introduced 2D-DCT feature vectors in 1998 [11]. Five State HMM is used to represent the five facial regions such as hair, forehead, eyes, nose and mouth in natural order that is shown in Fig.6. For improving the recognition rate, a sequence of overlapping blocks are generated each of width 92 pixels and height 10 pixels.

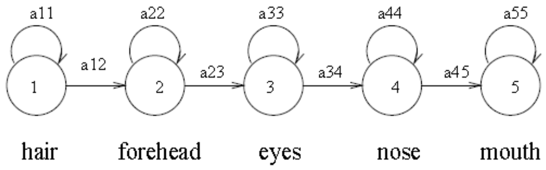

Fig.6 Five State HMM for Five Facial regions

Nefian and Hayes extracted the 2D DCT coefficients of each block which represent the compression properties of an image. A rectangular window of size 13×3 is chosen to represent the most significant coefficients. A set of 39 2D-DCT coefficients obtained from each block were used to form the observation vectors. The training data is uniformly segmented from top to bottom representing a five-state HMM and the observation vectors associated with each state are used as the initial estimates. Using the EM procedure, the model parameters are re-estimated. The system is trained and tested on ORL database and achieves a recognition rate of 84%. Nefian and Hayes decreased the processing time from 25 seconds in previous studies to 2.5 seconds per face image. The system developed by Nefian and Hayes reduce the computational complexity to a greater level [11].

Though the HMM system based on 2D-DCT coefficients is computationally effective, the recognition rates are compromising. For generating an observation vector, 39 2D-DCT coefficients are used. This would result greater number of DCT coefficients for generating a large number of observation vectors.

C. Face Detection and Recognition using Hidden Markov Models

Nefian and Hayes characterized the states of observation vectors using the Karhunen-Loeve Transform coefficients in 1998 [12]. The detection and evaluation performance of one-dimensional HMM for gray-scale images is investigated. Observation vectors are generated by dividing each face image of width W and height H into overlapping blocks. In order to overcome the difficulty of image noise, 2D-DCT coefficients are used as robust features. The size of the observation vectors are reduced significantly due to the choice of including DCT coefficients.

In addition to this, the observation vectors consist of KLT coefficients which represent the compression and de-correlation properties is properly used. Eigen vectors corresponding to the largest Eigen values of the co-variance matrix are used to form the KLT coefficients. The training data is uniformly segmented from top to bottom and the KLT coefficients associated with each state are used to obtain the initial estimates of the probability matrix. In the next iteration, uniform segmentation is replaced by the Viterbi segmentation and the final parameters of the HMM are obtained by using Baum-Welch recursive procedure.

The face detection system is trained and tested on MIT database consists of 48 images of 16 people with cluttered background under different illumination conditions. Manually segmented faces from 9 images were used in the training set and the other 39 images were used for testing. The system achieves a detection rate of 90%. During face recognition, the probability of the observation sequence given for each HMM model is computed using a Viterbi recognizer. The system is trained and tested on ORL database and achieves a recognition rate of 86% with a computational speed of 250 ms/face [12].

Though the face recognition system based on KLT coefficients shows improved computational cost than the previous HMM methods, it lacks in recognition rate. The method requires manual segmentation procedure that is simple for small set of testing images while it is difficult for a large data set.

D. Identity Confirmation System based on CTAG and HMM

Gyorodi *et al.* presented an embedded HMM approach to the face recognition problem for 2D face images in 2002 for identity confirmation systems [13]. In the previous studies, observation vectors are formed either by the pixel intensities or the DCT coefficients. Gyorodi introduced a flexible framework consist of Color Texture Adjacent Graf (CTAG) coefficients to form the observation vectors. The system consists of a set of super states in addition to a set of embedded states. The super states are used to model two-dimensional data in one direction while the embedded HMM is used for modeling the data in another direction that is represented in Fig. 7.

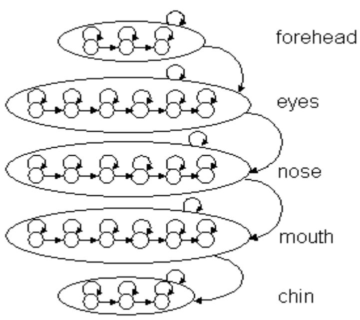

Fig.7 Five State Embedded HMM

Each and every state in the overall top-to-bottom HMM is assigned to a left-to-right HMM and the observation sequences are generated by scanning the image left to right and top to bottom. The overlap between adjacent windows in both vertical direction and the columns in the horizontal direction are obtained that is shown in the Fig. 8.

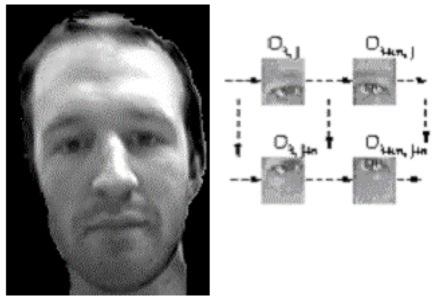

Fig. 8 Observation Sequence Generation

After extracting the observation vectors corresponding to the test images, the probability of an observation sequence for a given embedded HMM face model is computed. Highest likelihood value of the model is selected for identifying the unknown face. The system is trained and tested on ORL database consists of 400 images in .pgm format and half of the different images were used for training and testing. The system has greater efficiency due to the flexible nature of observation vectors [13]. Recognition rate is 83% on two observation vectors of length 28 pixels and 14 pixels. Though the face recognition method developed by Gyorodi shows improved computational costs, it suffers seriously from accuracy rate and complex HMM architecture.

E. HMM-Based Face Recognition using Selective Attention

Sequential methods used for face recognition task rely on the analysis of local facial features in a sequential manner. However, the distribution of discriminative information is not uniform over the facial surface. Eyes and mouth possess more information than the cheek. Salah *et al.* introduced 'Selective Attention' approach to the face recognition problem in 2007 by considering the above factors [14]. The selective attention approach guides the human eye by detecting salient locations and directing more resources at informative parts. Gabor wavelet filters are used for computationally effective saliency. The location selection is based on the Winner-Take-All network. The architecture of the saliency model is represented in Fig. 9.

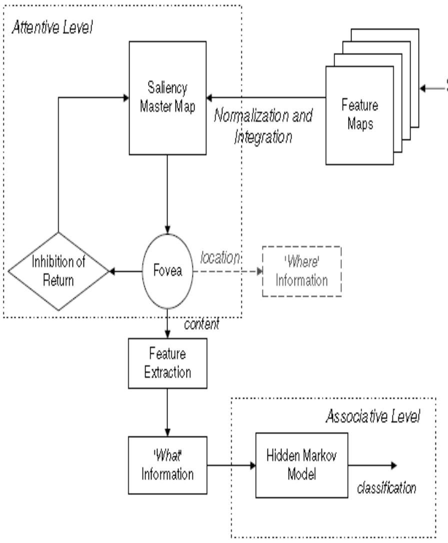

Fig.9 Saliency Model

Sequences of features are extracted from each face image. Scanning and feature extraction are the two essential steps needed to obtain a data sequence. A sequence of sub-images of fixed dimensionality is the initial step that is obtained by inspecting the face image while raster scan is used to acquire a sequence. Due to saliency-based scheme, the salient parts of an image are gaining robustness since the patches are extracted in decreasing order of importance. HMM classifier eliminates the negative effect of misleading background frames. From the two types of information extracted by the saliency scheme, only the "what" information is used, leaving out the "where" features.

Once the sequences are extracted, computing the features from each gathered sub-image is performed. The system is tested on BANCA database consist of 52 subjects and each subject has 12 images recorded under different conditions. All the images are preprocessed by a geometric interpretation followed by histogram equalization. Salah *et al.* used "Leave One Out" methodology for face recognition and the system is tested against different feature extraction methods such as Gray levels, DCT and HAAR wavelets. The system achieves an overall performance of 14.1%, 90.51% and 79.62% respectively on three different window sizes [14].

Some of the demerits of 'selective attention' approach are: system's performance is reasonable only on DCT and it shows very low accuracy when the gray level features are applied. When the face recognition problem is applied to a set of twins, the informative areas may vary, result in misclassification.

F. One Sample Face Recognition using HMM Model of Fiducial Areas

Ojo and Adeniran introduced an effective algorithm for recognition and verification with one sample image per class in 2011 [15]. The system uses 2D discrete wavelet transform to extract facial features from images.

A 1D discrete top-to-bottom HMM is used to segment every face image into five states for five facial regions that are illustrated in Fig. 10. A 2D DWT is used for feature extraction and for the generation of observation vectors. It also decomposes the image into its approximation coefficients, horizontal details, vertical details and the diagonal details. The approximation coefficients are coded using 256 gray levels producing a coded and reduced form of an input image. The coded image is then divided into sub-images and the overlap between successive sub-images subsequently increases the system performance [15].

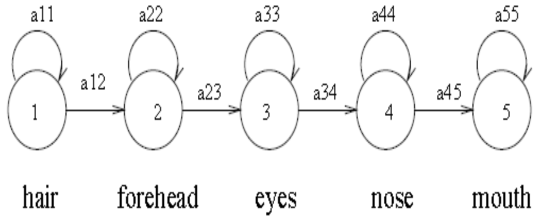

Fig. 10 Five State Top-Bottom HMM

2D sub-images were converted into a vector by extracting the coefficients' column-wise. The generation of observation vectors is illustrated in Fig.11.

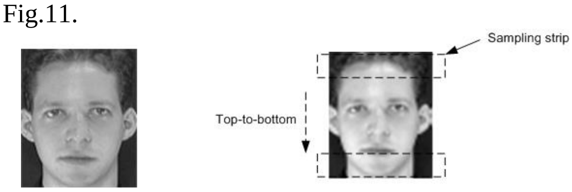

Fig.11 (a) Approximation Coefficients of an image (b) Segmentation into states

Baum-Welch re-estimation procedure is used to get the parameters which are used to optimize the likelihood of a training set observation vectors for each face. The maximum number of iterations for re-estimation is set to 5 and the model parameters are stored with the appropriate class names. Model likelihoods for all the models in the training set are calculated and the model with the highest log-likelihood is identified as a model representing the face. Euclidean distance measure is used to test if a face is in the training set or in a database. The system is tested on ORL database and a facial image per person is used for training and the remaining five images were used for testing. An accuracy rate of 80% was achieved and a misclassification of 20% is reached [15].

In the recent times, a face recognition system with an accuracy rate of 80% is only compromising and failed to apply for real-time situations where different challenges such as illumination conditions, lighting conditions and pose changes exist. Moreover, only one face image is used for training and the other images are used for testing. It is not fair for the face recognition system to focus on only one training image and at least half of the images for training would improve its accuracy.

G. Face Recognition using Maximum Confidence Hidden Markov Model

Raut and Patil introduced a new approach to the face recognition problem to increase the robustness against illumination changes in 2012 [16]. The performance of face recognition by MC-HMM heavily depends on the choice of model parameters and the usage of discriminative feature extraction method.

The system has two key elements: pre-processing based on pixel averaging for dimension reduction and energy normalization to reduce the effect of image brightness. Discriminative training algorithm based on approximation of maximum mutual information (MMI) and a discriminative feature extraction is used in the system. Five state HMM is used and each state has an embedded HMM associated with it and a total of 21 embedded HMMs are used in the system. There are five super-states for the five facial regions of a face that is depicted in Fig. 12.

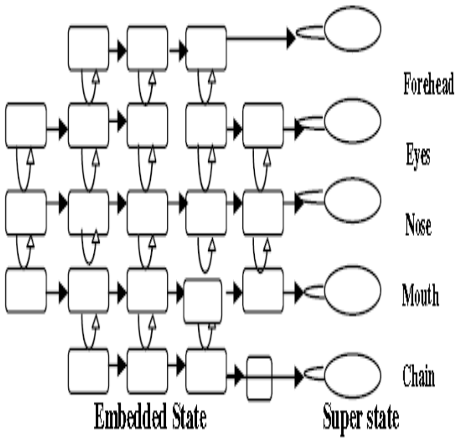

Fig. 12 Embedded HMM and five Super-states

In the initial stage, the data are uniformly segmented to obtain the initial estimates for model parameter. At the next iteration, uniform segmentation is performed with the help of initial features of a training image. HMM parameters are estimated using Viterbi algorithm and an input image is translated into super states and a set of embedded states is done at the third iteration. The system is tested on face database and a total of 25 face images were selected and manually cropped. The system is tested under three different set of approaches ML-HMM, MCE-HMM and MC-HMM. For a number of classes C=100, the system yields the recognition rate of 93%,89% and 90% while the training time per face image is 400,250 and 289 seconds respectively[16].

The system shows improved recognition rate than the previous studies. However, the parameter model follows a complex process since the system requires greater number of classes to yield better recognition rates. Moreover, the time required to train a face image takes at least five minutes for the three set of HMM approaches. This would result in increasing the computational complexity and the processing time of the CPU.

H. Algorithm for Face Recognition using HMM and SVD Coefficients

Anand and Lawrance included two new facial regions such as eyebrows and chin to the face recognition problem along with SVD coefficients for having better recognition rates in 2013[1]. Fig. 13 portrays the seven facial regions of a person and Fig. 14 shows the equivalent one-dimensional HMM representing every HMM state from left to right.

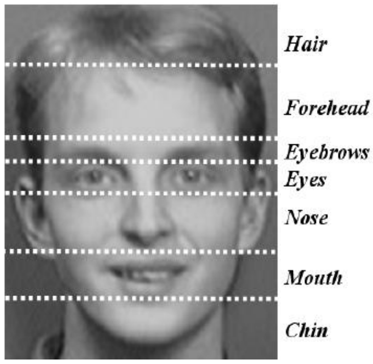

Fig. 13 Seven Facial regions of a person

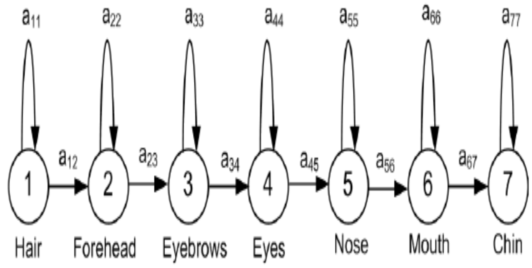

Fig.14 Seven State HMM for Seven Facial Regions

Order Statistic Filtering is used as a preprocessing operation for efficient computation. Observation vectors are generated by dividing each face image into overlapping blocks and SVD coefficients acts as a base for constructing the observation sequence. Based on SVD theory, the energy and information of a given signal is mainly conveyed by a few big singular values. The first three singular values are of more significance by the theory of SVD coefficients. Due to the discrete nature of HMM, quantization process is introduced to model the continuous observation vectors. Using the Baum-Welch algorithm, a HMM model is trained for every person in the database and tested on ORL face database consist of 400 images of 40 persons in Portable Gray Map (.pgm) format. After training process, each face (class) model is associated with a HMM. Face recognition problem is moved from single class to K-class classification problem and K distinct HMM models are used for representing K-classes. Five face images of a person are considered for training the system and tested against 200 unseen faces. The system achieves a recognition rate of 96.5% with a computational speed of 0.22 seconds per image. Choosing SVD coefficients as features increases the efficiency, in turn reducing the complexity [1].

4. COMPARATIVE ANALYSIS

Table II discusses the performance analysis of various HMM based face recognition methods. The algorithms such as [17], [16] and [1] shows much better recognition rates of having more than 90%.

TABLE III
COMPARISON OF RECOGNITION RATE AND COMPUTATIONAL COST OF HMM METHODS

Method	Recognition Rate	Computational Speed	Reference	Year
Top Down HMM+ Gray	87%	25sec	[10]	1994
Pseudo 2D HMM+ Gray	95%	25sec	[17]	1994
2D DCT	84%	2.5 sec	[11]	1998
2D DCT+ KLT Coeff.	86%	2.5 sec	[12]	1998
MC-HMM	90%	289 sec	[16]	2012
HMM+ SVD	96.5%	0.22sec	[1]	2013

In terms of computational speed, HMM+SVD algorithm proposed by Anand and Lawrance in 2013 [1], stood out greater with a whooping 0.22 sec/image. However, the recognition rates of the HMM based face recognition methods discussed, didn't provide cent percent accuracy. Our ultimate aim is to establish a more accurate face recognition system based on HMM classifier to provide nearly 100% accuracy in the future.

5. CONCLUSION

Face Recognition is an active research area that has immense use in various application domains in data mining, image processing, biometrics etc. Hidden Markov Model (HMM) is a statistical Markov process usually designed to model complex, non-linear and stochastic processes. HMM based face recognition systems are analyzed thoroughly and their efficiency in terms of recognition rate and computational speed are discussed. Still, the above discussed face recognition methods failed to provide 0% false rate during recognition. The authors' aims to establish a more accurate face recognition system based on HMM classifier in the future.

REFERENCES

[1] C. Anand and R. Lawrance, "Algorithm for Face Recognition using HMM and SVD Coefficients", *CiiT International Journal of Artificial Intelligent Systems and Machine Learning*, pp: 125- 130, vol. 5 (3), 2013.

[2] http://*www.ostp.gov/nstc*,National Science and Technology Council, USA.

[3] A .J. Goldstein, L.D.Harmon and A.B.Lesk," Identification of Human Faces", Proceedings *of IEEE*, 1971.

[4] L. Sirovich and M. Kirby, "A Low-Dimensional Procedure for the Characterization of Human Faces", *J. Optical Society of America*, 1987.

[5] M. A. Turk and A. P. Pentland, "Face Recognition Using Eigen faces", *Proceedings of IEEE*, 1991.

[6] Selim Aksoy, *Introduction to Pattern Recognition*, Departmentof Computer Engineering, Bilkent University, saksoy@cs.bilkent.edu.tr.

[7] http://www.wikipedia.com

[8] S. Roweis, *Tutorial: Hidden Markov Models*, University of Toronto, 2003.

[9] J. Geng and J. Yang," Automatic Extraction of BibliographicInformation on the web", *Proceedings of the 8th International Database Engineering and Applications Symposium (IDEAS)*, pp: 193-204.

[10] F.S. Samaria, "Face Recognition using Hidden Markov Models", PhD thesis, Engineering Department, Cambridge University, 1994.

[11] A.V. Nefian and M.H. Hayes, "Hidden Markov Models for Face Recognition", *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp: 2721-2724, 1998.

[12] A.V. Nefian and M.H. Hayes, "Face Detection and Recognition Using Hidden Markov Models", *Proceedings of IEEE International Conference on Image Processing (ICIP)*, vol: 1, pp: 141-145, 1998.

[13] R.S. Gyorodi, S. Holban, T. Coffey and C. A. Gyorodi,, "Identity Confirmation System based on CTAG and HMM",*Transactions of Periodica Polytechnicia on Automatic Control and Computer Science*, pp: 35-40, Vol.47 (61), 2002.

[14] A. A. Salah, M. Bicego, L. Akarun, M. Grosso, and M. Tistarelli, "Hidden Markov Model-based face recognition using selective attention", *Proceedings of SPIE 6492, Human Vision and Electronic Imaging XII*, 2007.

[15] J. A. Ojo and S. A. Adeniran, "One-sample Face RecognitionUsing HMM Model of Fiducial Areas", *International Journalof Image Processing (IJIP)*, volume 5: Issue-1, pp: 58-68, 2011.

[16] S. Raut and S. H. Patil, "Face Recognition using MaximumConfidence Hidden Markov Model", *Proceedings of International Journal of Advances in Engineering and Technology"*, pp: 574-581, Vol. 2(3), 2012.

[17] F. Samaria and A. Harter, "Parameterization of a StochasticModel for Human Face Identification," *Proceedings of IEEE Workshop on Applications of Computer Vision*, Florida, 1994.