

Layered Recognition Scheme for Robust Human Facial Expression Recognition using modified Hidden Markov Model

Mayur Rahul, Rashi Agrawal
UIET, CSJMU, Kanpur
India
mayurrahul209@gmail.com
dr.rashiagrwal@gmail.com

Narendra Kohli
HBTU, Kanpur
India
kohli.hbti@gmail.com



ABSTRACT: Facial Expression Recognition is an application used for biometric software that can be used to recognize special expressions in a digital image by analysing and comparing different patterns. These software are popularly used for the purpose of security and are commonly used in other applications such as credit card verification, surveillance systems, medicines, home security, human-computer interface, etc.. Recognizing faces becomes very difficult when there is a frequent change occurs in facial expressions. In this paper two layer extension of HMM is used to identify continuous effective facial expressions. Partition technique is used for feature extraction. Two layered extension of HMM consists of bottom layer which represents the atomic expression made by eyes, nose and lips. Further upper layer represents the combination of these atomic expressions such as smile, fear etc. In HMM, Baum-Welch, Viterbi and Forward methods are used for parameter estimation for calculating the optimal state sequence and probability of the observed sequence respectively. This proposed system consists of three level of classification. Output of the first level is used for the training purposes for the second level and further this level is used for the third level for testing. Six basic facial expressions are recognised i.e. anger, disgust, fear, joy, sadness and surprise. Experimental result shows that Proposed System performs better than normal HMM and has the overall accuracy of 85% using JAFFE database.

Keywords: JAFFE, Gabor Wavelets Transform, PCA, Local Binary Patterns, Svm, Faps, Cohn-kanade Database, Markov Process, Static Modelling

Received: 9 September 2018, Revised 1 November 2018, Accepted 6 November 2018

DOI: 10.6025/jmpt/2019/10/1/18-26

© 2019 DLINE. All Rights Reserved

1. Introduction

The Facial Expression is the most useful medium for non-verbal communication in human daily life. Facial Expression Recognition can be applied in many interesting areas such as in Human-computer interface, Telecommunications, Behavioural Science, Video Games, Animations, Psychiatry, Automobile Safety, Affect sensitive music juke boxes and televisions, Educational Software, etc. Better classifier and better feature extraction method is very important because it significantly improves the accuracy and speed for identification and classification.

The Hidden Markov Modelling is useful in spectral and dynamic nature of speech pattern using acoustic state model [1]. The HMM becomes the most useful in speech recognition. Due to handling of time series data modelling and learning capability, HMM is very useful in classifying unknown feature vector sequence. Only small number of training samples required to train full HMM[2,3]. The discriminative strength of normal HMM is less suitable to handle more difficult task. Due to dynamic statistical modelling and time sequence pattern matching principle, HMM matches the most similar signal state as recognition results [4,5]. Hidden Markov Model is capable for processing continuous dynamic signal, can effectively use the timing signal moments before the state transition and after the state transition.

In this paper, partition based feature extraction method and two layer extension of HMM is introduced. New introduced HMM is consists of two layers. Bottom layer represents the atomic facial expressions such as expressions involving eyes, nose and lips separately. Upper layer represents the expressions which are basically the combination of atomic expressions such as smile, fear etc. Our assumption is that every facial expression is the combination of expressions of eyes, nose and lips. Our newly introduced system is capable of handling almost every situation like occlusion, brightness, background etc.

The rest of the paper are as follows: Basics of HMM is explained in section-II, Related Works are discussed in section-III, Proposed System in section-IV, Experiments and Results are discussed in section-V and finally concluded in section-VI.

2. Basics of HMM

Markov process is used for time series data modelling. This model is applicable for situations where there is a dependability of present state to previous state. For example, bowler bowls the ball in cricket game; ball first hit the ground then hit the bat or pad. This situation can be easily modelled by time series data modelling. In other example such as in a spoken sentence, the present pronounced word is depending on the previous pronounced word. Markov process effectively handles such situations.

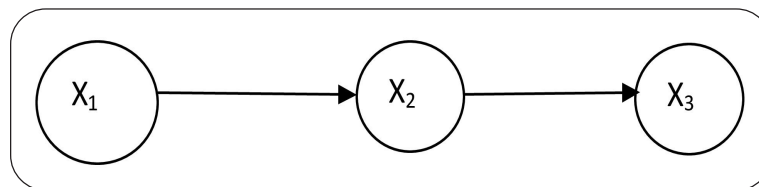


Figure 1. A Markov Model for 3 observations variables

When an observation variable is dependent only on previous observation variable in a markov model is called first order markov chain. The joint probability distribution is given by

$$p(X_1, X_2, \dots, X_N) = p(X_1) \prod p(X_N | X_{N-1}) \quad (1)$$

where $p(X_1, X_2, \dots, X_N)$ is the joint probability distribution of states X_1, X_2, \dots, X_N and $p(X_1)$ is the probability of state X_1 and $p(X_N | X_{N-1})$ is the probability of state X_N given state X_{N-1} .

A Hidden Markov Model is a statistical model which is used to model a markov process with some hidden states. This model is widely used in speech and gesture recognition. Hidden markov model is a set of observable states are measured and these variables are assumed to be depend on the states of a markov process which are hidden to the observer (Figure 2). Hidden states are states above the dashed lines.

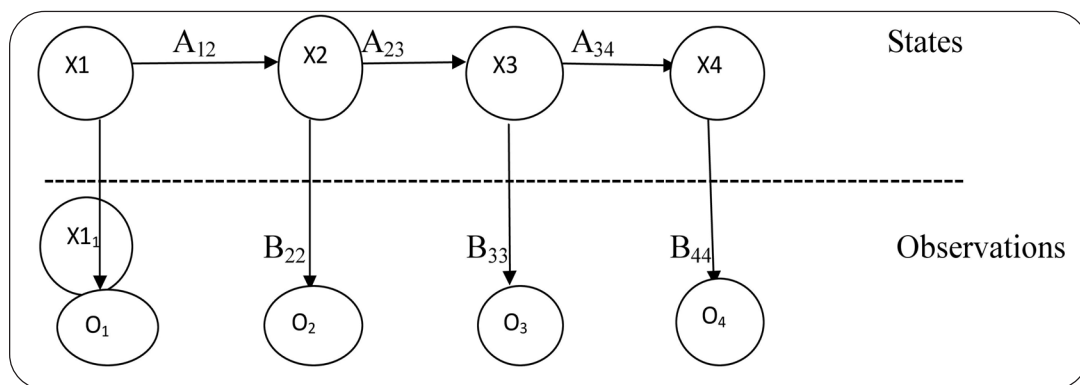


Figure 2. A hidden Markov Model for four variables

$$p(X_{t+1}=j | X_t = i) = a_{ij} \quad (2)$$

where $p(X_{t+1}=j | X_t = i) = a_{ij}$ is the conditional probability of the state X_{t+1} given the state X_t and a_{ij} is the i th row, j th column entry of the transition matrix $A = (a_{ij})$.

$$p(O_t = k | X_t = i) = b_{ik} \quad (3)$$

where $p(O_t = k | X_t = i) = b_{ik}$ is the conditional probability of the observation variable O_t given the state X_t and b_{ik} is the i th row, k th column entry of the emission matrix $B = (b_{ik})$.

An HMM is defined by A , B and Π . The HMM is denoted by $\partial = (A, B, \Pi)$

3. Related Works

The Facial Expression and Physiognomy was first started in the early fourth century. The Physiognomy is the study of the person's character from their outer appearance [6]. The interest on Physiognomy is minimized and the studies on Facial Expressions become active. The studies of facial expressions have been found in 17th century back. John Bulwer in 1649 was given the detail theory about the facial expressions and head movement in his book "Pathomyotomia". Le Brun in 1667 gave a detailed notes on Facial Expression, then it was later reproduced as a book in 1734 [7].

In the beginning of 19th century, Charles Darwin gives a theory which produced the greatest impact in today's automatic Facial Expression Recognition field. Darwin created a benchmark for various expressions and means of expressions in both animals and humans [8]. In 1970, Ekman and his colleagues done a great job in this field and further their work has great influence for today's new era of Facial Expression Recognition. Today this field is used in almost all area for example in clinical and social psychologists, medical practitioners, actors and artists. However in the end of the 20th century, with the advances in the fields of robotics, computer graphics and computer vision, animators and computer scientists started showing interest in the field of facial expressions. Suwa et. al. in 1978 creates a framework in automatic facial expression recognition by tracking 20 points to analyse facial expressions from the sequence of images.

Ekman and Freisan introduced a framework in Psychology to recognize facial expression [9]. In 1990, researcher used Ekman's framework to recognize facial expression in video and image[10]. Hidden Markov Model was also used to recognised facial expressions [11]. Further Cohen et. al. proposed a multilevel HMM to recognise emotions and further optimize it for better accuracy as compared to emotion specific HMM[10]. Pardas et. al.[2002] used the automatic extraction of MPEG-4 Facial Animation Parameters(FAP) and also proved that the FAPs provide the necessary information required to extract the emotions[12]. FAPs were extracted using an improved active contour algorithm and motion estimation. They used the HMM classifier. They also used the Cohn-kanade database for recognition. Their recognition rate was 98% for joy, surprise and anger and 95% for joy surprise and sad. The average recognition rate was 84%.

Aleksic et. al.[2006] showed their performance improvement using MS-HMM[13]. They also used PCA to reduce dimensionality before giving it to HMM. They used MPEG-4 FAPs, outer lip (group 8) and eyebrow (group 4) followed by PCA to reduce dimensionality. They used HMM and MS-HMM as classifiers. They used Cohn-kanade databases with 284 recordings of 90 subjects. Their recognition rate using HMM was 88.73% and using MS-HMM was 93.66%. Shang et. al. [2009] proposed an effective nonparametric output probability estimation method to increase the accuracy using HMM [14]. They worked on CMI database and get the accuracy of 95.83% as compared to non-parametric HMMs. Jiang [2011] proposed a method based on code-HMM and KNN further applied some discrimination rules [15]. Their proposed method achieves better accuracy to some extent. Suk et. al. proposed a real-time temporal video segmenting approach for automatic facial expression recognition[16]. They used SVM as a classifier and get the accuracy of 70.6%.

Wu et. al.[2015] incorporated a multi-instance learning problem using CK+ and UNBC-McMaster shoulder pain database[17]. They combine both multi-instance learning and HMM to recognize facial expression and outperforms state-of-the-arts. Sikka et. al.[2015]proposed a model based similarity framework and combine SVM and HMM[18]. They worked on CK+ and OULU-CASIA databases and get the accuracy of 93.89%. Ramkumar et. al. incorporated the Active Appearance Model(AAM) to identify the face and extracted its features[19]. They used both KNN and HMM to recognize facial expression and get the better results. Senthil et. al. proposed a method for recognizing facial expression using HMM[20].Xufen et. al. incorporated some modification in HMM to get better recognition rate[21].

Xiaorong et. al. proposed a framework for partially occluded face recognition using HMM and get the better result to some extent[22]. Punitha et. al. proposed a real-time facial expression recognition framework and get the better results as compared to existing framework[23]. Pagariya et. al. also proposed a system using Multilevel HMM for recognizing facial expression from the video[24].Islam et. al. incorporated both PCA and HMM for appearance and shape based facial expression recognition framework[25]. Singh et. al. introduced a 3 state HMM for recognising faces under various poses[26].

4. Proposed System

Different database requires different feature extraction method and different classifiers. It is not easy to conclude which feature extraction methods and classifiers best suited the situation. HMMs have been used in speech recognition so far. Due to handling of time series data modelling and learning capability, HMM is very useful in classifying unknown feature vector sequence. Only small number of training samples required to train full HMM [2,3]. The discriminative strength of normal HMM is less suitable to handle more difficult task. In this paper, we introduce a new framework which uses most powerful partition based feature extraction method and modified HMM as a classifier. New Multi-Stage HMM is the two layer extension of normal HMM. Bottom layer represents the atomic expressions made by eyes, lips and nose and Upper layer represents the combination of these atomic expressions such as smile, fear etc. The proposed framework uses the partition based feature extraction method, and then these extracted features are used to train the classifier. Some dataset are used to test the classifier. This framework is robust to occlusion, background and orientation. In this proposed System, Baum-Welch method is used for parameter estimation. Viterbi Method and Forward Procedure are used for calculating the optimal state sequence and probability of the observed sequence respectively. The proposed framework is shown in figure 3.

5. Experiments and Results

Following conditions have been applied during the experiments:

- (1) Faces are analysed from frontal view only.
- (2) No movement in the head.
- (3) No conversation during image taken.
- (4) No glasses during image capturing.

JAFFE database has 152 images from 13 subjects. They are 16 of anger, 27 of disgust, 19 of fear, 33 of joy, 30 of sadness and 27 of surprise. To train the HMM, we follow N fold cross validation rule. For example if there is k fold, then $k-1$ fold are used to train the HMM and remaining 1 fold is used to test HMM. For example in the set of nine images of anger. These 9 images are divided into nine folds. We are using 8 images to train and 1 for test. Further final result is the mean of all the 9 results.

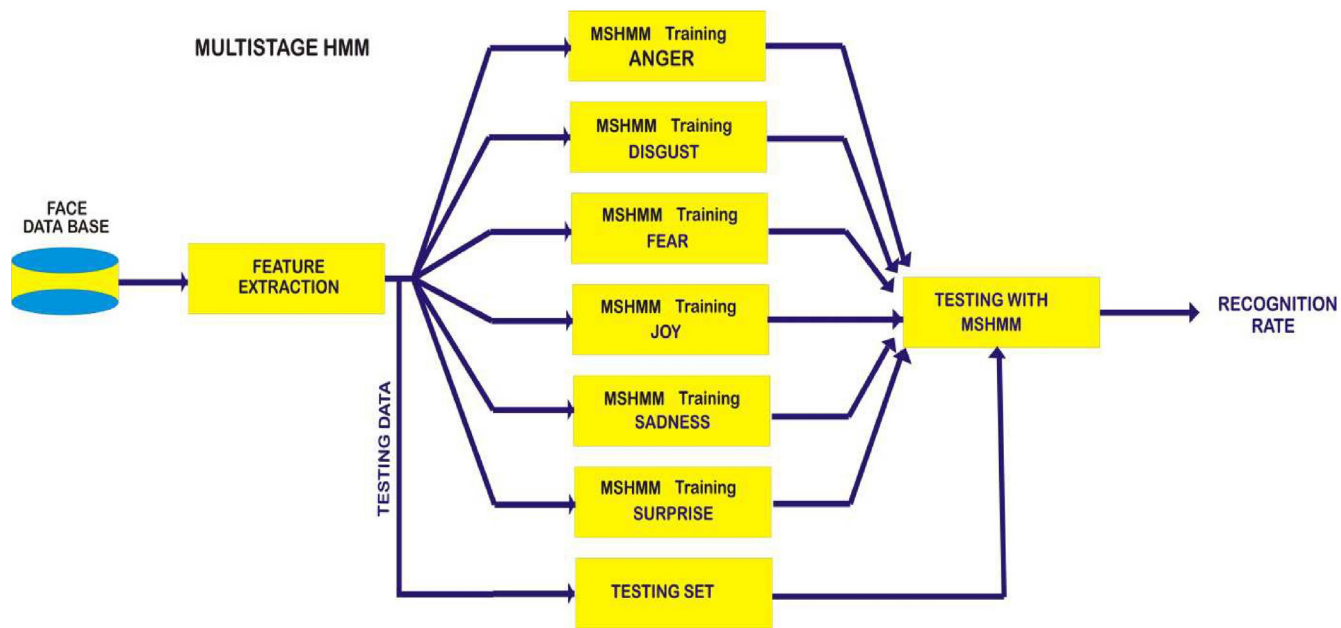


Figure 3. Block diagram for Facial Expression Recognition using Proposed System

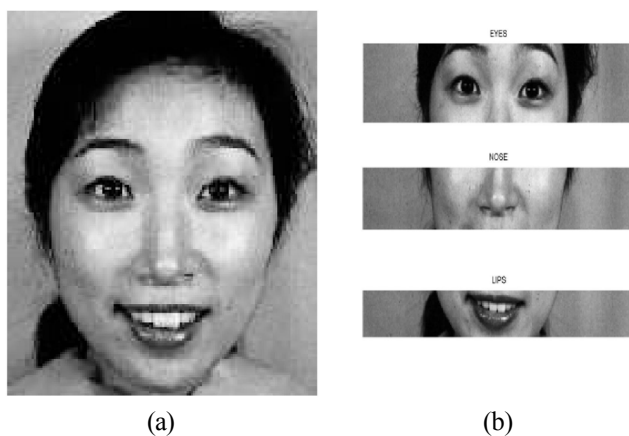


Figure 4. (a) Original Image (b) Partitioned Image

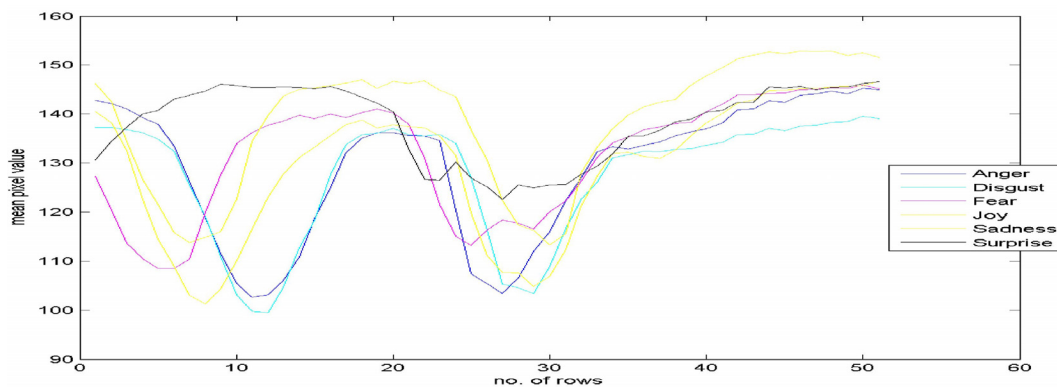


Figure 5(a). Plot of Eyes partition

Images are taken from the JAFFE database of facial expression. Each image is partitioned into three parts (see figure 4(a) and 4(b)). Partitioning of image is very effective method as compared to other state-of-the-art feature extraction methods. These partitions are easily distinguishable to achieve the higher accuracy (see figure 5). Plot shows the eyes, nose and lips of all six expressions of the same person.

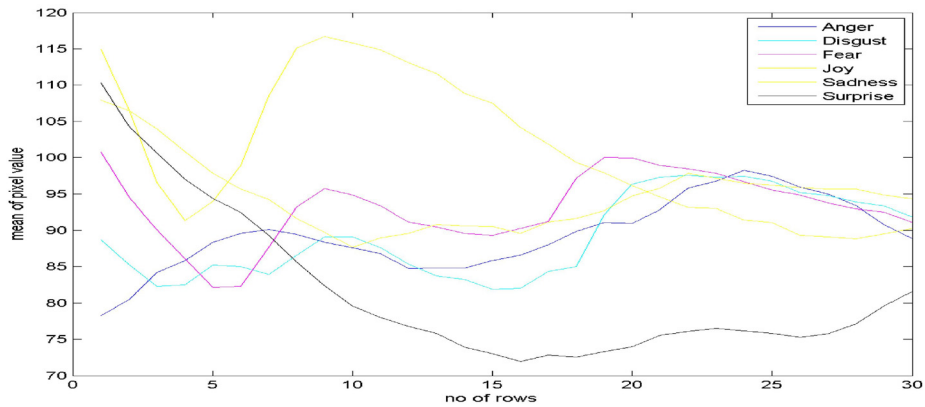


Figure 5(b). Plot of Lips Partition

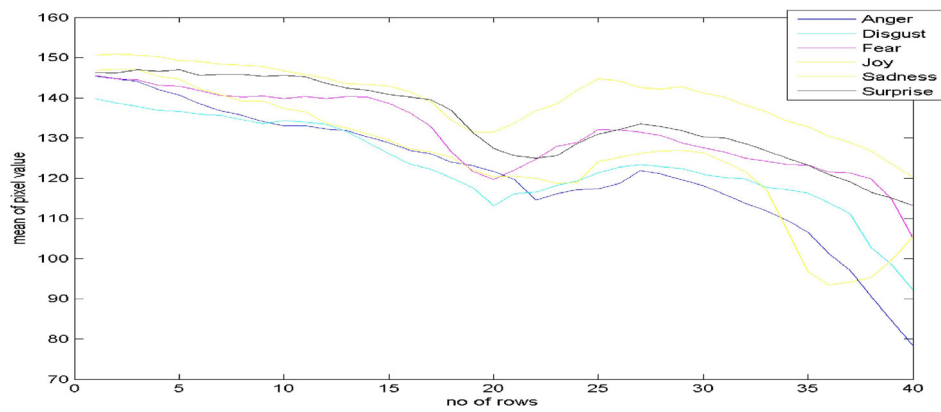


Figure 5(c). Plot of Nose partition

Again next plot shows the eyes, nose and lips partition of same “disgust” expression for different persons (see figure 6).

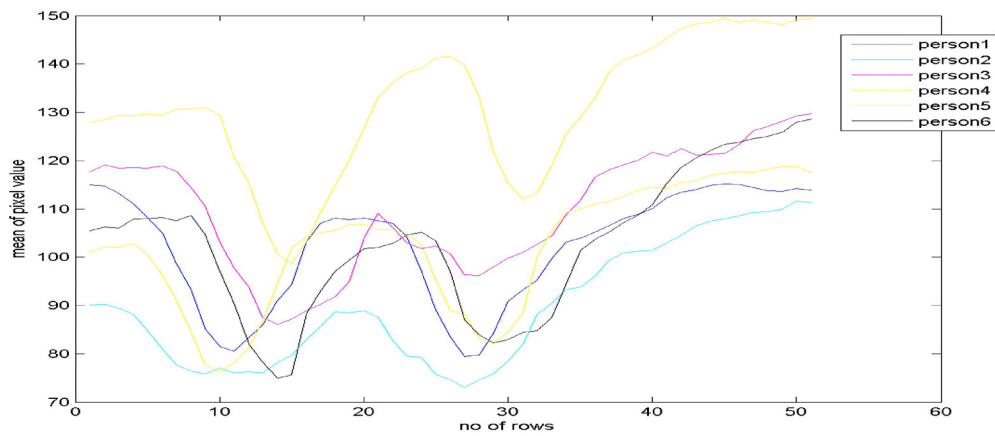


Figure 6(a). Plot of Eyes partition for “Disgust” expression of different persons

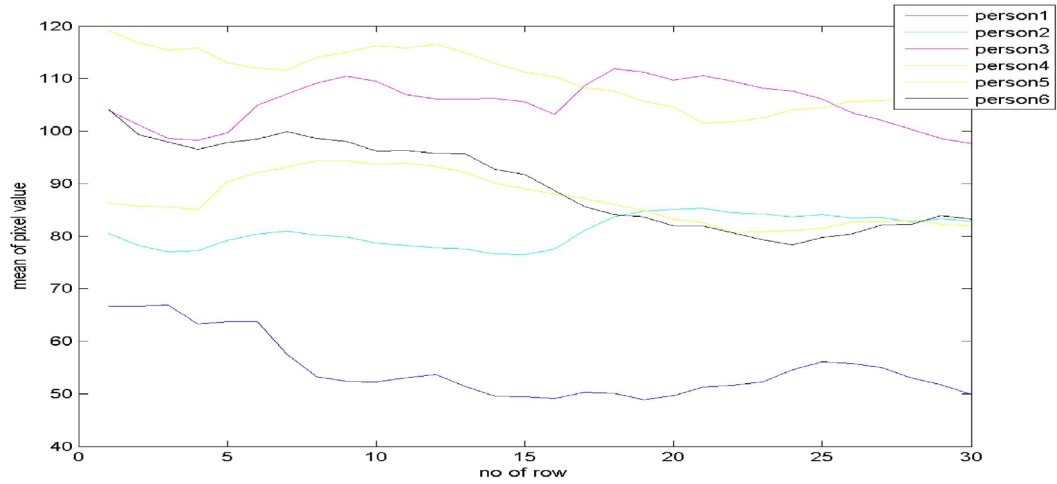


Figure 6(b). Plot of Lips partition for “Disgust” expression of different persons

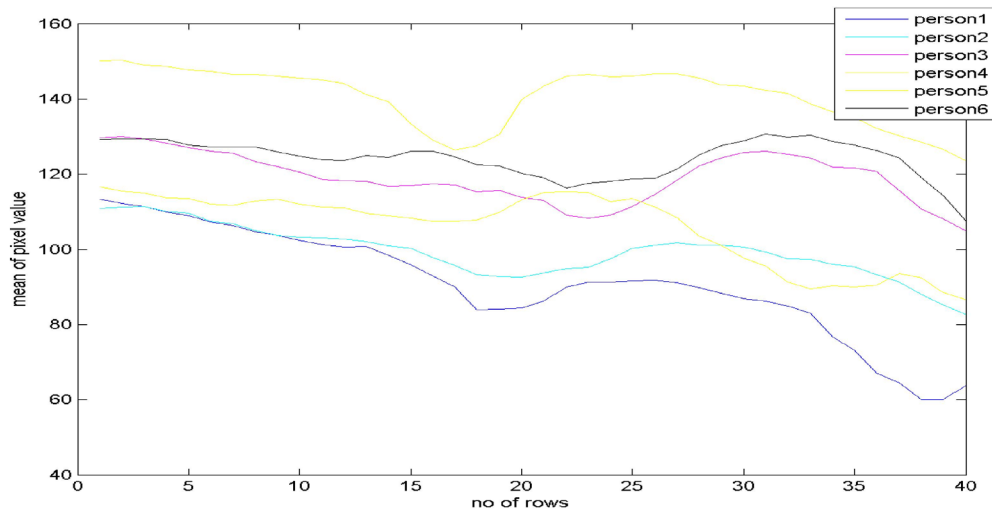


Figure 6(c). Plot of Nose partition for “Disgust” expression of different persons

	Anger	Disgust	Fear	Joy	Sadness	Surprise	Recognition Rate (%age)
Anger	12	0	2	0	0	2	75%
Disgust	0	22	0	0	2	3	81%
Fear	1	0	12	4	2	0	64%
Joy	0	0	0	33	0	0	100%
Sadness	6	0	0	0	24	0	80%
Surprise	0	0	0	5	0	22	81%

Table 1. Confusion Matrix for FER using Proposed System

These plots show the difference between different expressions of the same person and same expressions of the different persons. This partition based technique when incorporated with new introduced HMM gives much better results as compared to other state-of-the-art methods.

The recognition results using Partition Based feature extraction technique and modified HMM as a classifier is shown in table 1. From the results we can see that using the Proposed System will improve the recognition rate.

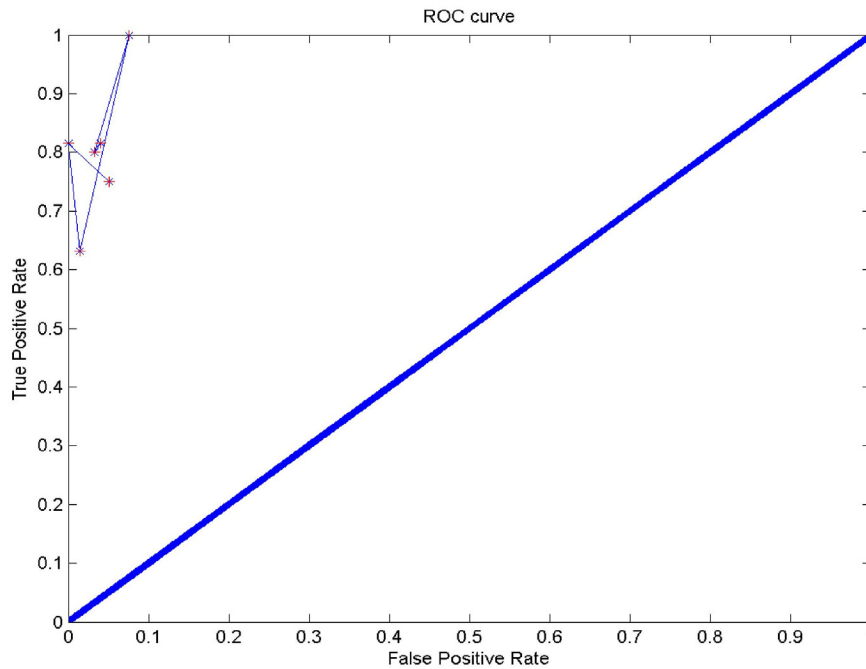


Figure 7. ROC curve for the given Confusion Matrix

The Receiver Operating Characteristics curve or ROC curve, is a graph used to represent the performance of the classifiers (see figure 7). It is the plot between True Positive Rate (TPR) versus False Positive Rate (FPR). TPR is the proportion of positive samples identified as positive samples and FPR is the proportion of negative samples identified as positive. The 45 degree line is called “line of no-discrimination”. The ROC curve above this line represents the goodness of the classifier and also indicates the classifying rate of identifying positive samples as positive is greater than its rate of misclassifying negative samples. This given ROC curve indicates the goodness of our classifier.

4. Conclusions

It is not easy to conclude which feature extraction method and classifier is best for this situation, but still modification of HMM significantly improves accuracy for the recognition of six basic facial expressions. This paper introduced a first framework for Partition Based Method as a feature extraction and modified HMM as a classifier for Facial Expression Recognition. HMM provides some modelling advantages over other feature based methods in Facial Expression Recognition. Due to generative nature of HMM, It is weak classifier as compared to other discriminative classifier such as SVM. Our proposed method makes HMM more powerful as a classifier with compared to other discriminative classifiers. We have also showed its strength using Confusion Matrix and Receiver Operating Characteristics (ROC) curve and found the overall accuracy of 82%.

References

[1] Katagiri, Shigeru., Lee, Chin-Hui . (1993). A new hybrid algorithm for speech recognition based on HMM segmentation and learning vector quantization. *Speech and Audio Processing, IEEE Transactions on* Volume 1(4) October. 1993 421 - 430

[2] Tsapatsoulis, N., Leonidou, M., Kollias, S. Facial expression recognition using HMM with observation dependent transition

matrix, *Multimedia Signal Processing*, 1998 IEEE Second Workshop on 7-9 Dec. 1998. p. 89-95

- [3] Lawrence R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition, *Proceedings of the IEEE*, 1989, 77 (2) 257—286.
- [4] Atlas, L., Ostendorf, M., Bernard, G. D. (2000). Hidden Markov Models for Monitoring Machining Tool-Wear, *In: IEEE International Conference on Acoustics, Speech, and Signal Processing*, 6: 3887—3890
- [5] Hatzipantelis, E., Murray, A., Penman, J. (1995). Comparing hidden Markov models with artificial neural network architectures for condition monitoring applications, *Fourth International Conference on Artificial Neural Networks*, 6. 369—374.
- [6] Highfield, R., Wiseman, R., Jenkins, R. (2009). How your looks betray your personality, *New Scientist*, February.
- [7] <http://www.library.northwestern.edu/spec/hogarth/physiognomics1.html>
- [8] Darwin, C., Darwin, F. (1904). *The Expression of the Emotions in Man and Animals*, 2nd edition, J. Murray, London, .
- [9] Ekman, P., Friesen, W. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto, 1978.
- [10] Cohen, I., Sebe, N., Chen, L., Garg, A., Huang, T. S. (2003). Facial expression recognition from video sequences: Temporal and static modelling. *In: Computer Vision and Image Understanding*, p 160–187.
- [11] Lien, J. J.-J., Kanade, T., Cohn, J. F., Li, C.-C. (200). Detection, tracking, and classification of action units in facial expression. *Robotics and Autonomous Systems*, 31(3) 131–146.
- [12] Parida, M., Bonafonte, A. (2002). Facial animation parameters extraction and expression recognition using Hidden Markov Models, *Signal Processing: Image Communication*, vol. 17, p.675–688.
- [13] Aleksic, P. S., Katsaggelos, A. K. (2006). Automatic Facial Expression Recognition Using Facial Animation Parameters and MultiStream HMMs, *IEEE Trans. Information Forensics and Security*, 1(1) 3-11.
- [14] Shang, L., Chan, K. P. (2009). Nonparametric discriminant HMM and application to facial expression recognition, *In: IEEE conference on CVPR*.
- [15] Jiang, (2011). A facial expression recognition model based on HMM, *In: International Conference on Electronic and Mechanical Engineering and Information Technology*.
- [16] Suk, M., Prabhakaran, B. (2015). Real-time facial expression recognition on smartphones, *In: IEEE Winter Conference on Applications of Computer Vision*.
- [17] Wu, C., Wang, S., Ji, Q. (2015). Multi-instance Hidden markov models for facial expression recognition.
- [18] Sikka, K., Dhall, A., Bartlett, M. (2015). Exemplar HMM for classification of facial expressions in video, *In: IEEE conference on CVPRW*.
- [19] Ramkumar, G., Lagashanmugam, E. (2016). An effectual facial expression recognition using HMM, *ICACCCT*.
- [20] Vijailakshmi, M., Senthil, T. (2014). Automatic human facial expression recognition using HMM, *IEEE 2014*
- [21] Jiang, Xuefang. (2011). A facial expression recognition model based on HMM, *IEEE*.
- [22] Xiorong, Pu., Zhiku, Zhou., Heng, Tan. (2014). Partially occluded face recognition using subspace HMM, *IEEE*.
- [23] Punith, A., Geetha, Kalaselvi. (2013). HMM based peak time facial expression recognition *IJETAE*, V 3 January.
- [24] Rani, R., Pagaria, Mahip, M., Bartere. Facial emotion recognition in video using HMM, *IJCER*, 3(4).
- [25] Islam, M. R., Toufiq, R., Rahman, M. F. (2014). Appearance and shape based facial system using PCA and HMM, *IEEE*.
- [26] Singh, K. R., Zaveri, M. A., Raghuwanshi, M. M. (2014). Recognizing face under varying poses with three states HMM, *IEEE*.