

Recognizing Expression Variant and Occluded Face Images Based on Nested HMM and Fuzzy Rule Based Approach

Parvathi Ramalingam^{1*}, Shanthi Dhanushkodi²

¹Department of Information Technology, PSNA College of Engineering and Technology, Dindigul, India

²Department of Computer Science and Engineering, PSNA College of Engineering and Technology, Dindigul, India

Email: *parvathi@psnacet.edu.in, dshan71@gmail.com

Received 11 March 2016; accepted 20 May 2016; published 23 May 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The face recognition with expression and occlusion variation becomes the greatest challenge in biometric applications to recognize people. The proposed work concentrates on recognizing occlusion and seven kinds of expression variations such as neutral, surprise, happy, sad, fear, disgust and angry. During enrollment process, principle component analysis (PCA) detects facial regions on the input image. The detected facial region is converted into fuzzy domain data to make decision during recognition process. The Haar wavelet transform extracts features from the detected facial regions. The Nested Hidden markov model is employed to train these features and each feature of face image is considered as states in a Markov chain to perform learning among the features. The maximum likelihood for the input image was estimated by using Baum Welch algorithm and these features were kept on database. During recognition process, the expression and occlusion varied face image is taken as the test image and maximum likelihood for test image is found by following same procedure done in enrollment process. The matching score between maximum likelihood of input image and test image is computed and it is utilized by fuzzy rule based method to decide whether the test image belongs to authorized or unauthorized. The proposed work was tested among several expression varied and occluded face images of JAFFE and AR datasets respectively.

Keywords

Face Recognition, Fuzzy Rule Based Method, Expression and Occlusion Variation, Baum Welch Algorithm, Nested Hidden Markov Model

*Corresponding author.

1. Introduction

The face recognition has rapidly emerged as an active research area in the biometric field to provide various secure real world applications like security monitoring, law enforcement and surveillance systems [1]. Because of using a single image per person to identify the face, the database can avoid storing huge amount of face images which needs restricted storage capability [2]. Most of real world applications using face recognition technology would require identifying a person under some controlled conditions like variations in illumination, pose, and expression [3]-[5] and occlusion [6]. The objective is to identify a person from the database in any of these unpredictable limitations. Therefore, face recognition [7] is recently considered as most challenging environments rather than fingerprint, iris and speech recognition in Biometric field. In proposed work, the face is recognized under expression and occlusion variation with the highest recognition rate. All the face recognition approaches are basically classified as four major categories like geometric based methods, template matching methods, appearance based methods and statistical approaches. Geometric based methods [8] recognize the face, according to geometric relationship or spatial distances between facial features by locating feature points on the image [9] such as elastic bunch graph matching, landmark localization and feature tracking [10]. The template based methods [1] compare the given image with a set of stored templates that is generated by using statistical tools such as support vector machine (SVM), Linear discriminant analysis (LDA) and principle component analysis (PCA). The appearance based methods [11] builds 2D/3D a morphable model like human faces in which parameters of the model are used to recognize faces. Some of reconstruction methods are used to frame 2D/3D models [10] [12] [13]. In statistical approaches [14], some of the facial features are only taken and the relationship between these features is estimated to identify a person. Hidden Markov model (HMM) is one of statistical approach and it forms an observation vector sequence by considering every facial feature as a state in Markov chain. It calculates the similarity index with the training set to recognize faces. HMM technique is only the approach to produce the highest recognition rate when comparing with alternative techniques [15]. The proposed work uses Nested Hidden markov model (NHMM) with Baum Welch algorithm [16] to find the relevant characteristics of the image.

2. Related Work

Many of face recognition method do not considered face recognition with expression and occlusion variations. Bronstein [17] introduced a face recognition method to handle missing data and produces high recognition rates on a limited database. In his study, the missing data were synthetically derived from frontal scans of the image. Dibeklioglu [18] presented curvature based segmentation method to recognize person with significant poses variations. But, it could not apply to facial scans with yaw rotations greater than 45 degrees and it requires storing several samples per person. Jingo [19] reconstructed 3D generic elastic model (GEM) for each subject with pose invariant recognition by involving 2D image. The distance between synthesized image and test image is computed by using a normalized correlation matcher. It produces less accuracy for expression invariant or extreme pose invariant face recognition. Josef [20] presented an image matching method which is formulated on Markov random fields (MRF). The label pruning and error pre whitening measures were introduced to increase the accuracy that addresses the computational burden. Regressor based cross pose face representation [12] was improved by finding bias and variance in regression of different pose variation. The ridge regression and lasso regression were also explored to overcome the problem of subspace based face representation and Gabor features were used to improve recognition rates. Perakis [21] introduced the 3D face recognition method to handle pose variation. The face model is reconstructed from partial face model and it produces 83.7% recognition rate. Automatic landmark detector estimates pose and detect missed data within the partial face model. Rakesh [22] utilize directional and texture information from face images for face recognition. The scale adaptive digital filters and local descriptors were used to capture directionality and to extract features respectively. Rangan [23] improves the performance of face recognition system with pose variation by using truncated transform domain feature extraction (TTDFE) for feature extraction. The binary particle swarm optimization based feature selection algorithm was utilized to search feature space for optimal feature subset. Selva raj [24] investigated with feature extraction from facial electromyography (FEMG) signals for classifying six emotional characteristics like happy, sad, afraid, surprise, disgust and neutral. The k-nearest neighbor (knn) classifier was used to map the extracted features with respective emotions. Principle component analysis was employed for retrieving the emotional information from FEMG signals by analyzing efficiency of the features over conventional statistical fea-

tures. Vetter [13] proposed a 3D face reconstruction method by fitting 3D morphable model on 3D facial scans for producing 3D synthetic faces from scanned data. These frontal facial scans were tested among FERET database and the work does not consider when the yaw rotations of face image could not exceed 40 degrees. Zhisong Pan [25] developed two models, named as, empirical kernel sparsity preserving projection and empirical kernel sparsity score were employed for feature extraction and feature selection respectively. The nonlinear separable data were mapped into kernel space in which the nonlinear similarity can be captured, and then the data in kernel space was reconstructed by sparse representation to preserve the sparse structure.

3. Expression and Occlusion In-Variant Face Recognition

The expression and occlusion invariant face recognition involves five steps to recognize the face such as:

- Face detection.
- Fuzzy data conversion.
- Feature extraction.
- NHMM training.
- Face Recognition.

3.1. Face Detection

The proposed work is mainly focused on training and face recognition steps. The process of extracting facial portion from input image is known as face detection. It eliminates background, hair, ear and unwanted portions of the facial image. It identifies the face location and it extracts certain relevant facial regions by using principle component analysis (PCA) [26]. The image is normalized to crop only the face without having a background. The facial features are detected according to the skin color and the circle shape of the iris.

3.2. Fuzzy Data Conversion

The detected face image is converted into the fuzzy domain data by using the following algorithm which starts with the initialization of the image parameters; size, minimum, mid and maximum gray level. The fuzzy rule-based approach is an efficient method for many tasks in the image processing.

The algorithm to convert image into fuzzy data is given below:

- 1) Read the image and convert it into grayscale image if it is RGB image.
- 2) Find the size of the image ($M \times N$).
- 3) Find the minimum, maximum gray level of the image also find the average gray level of the image.
- 4) For $x = 0:M$, For $y = 0:N$
- 5) If gray_value between zero and min
Then $a = 0$;
- 6) Else if gray_value between min and mid
Then $a = (1/(mid-min)) * min + (1/(mid-min)) * data$;
- 7) If gray_value between mid and max
Then $a = (1/(max-mid)) * mid + (1/(max-mid)) * data$;
- 8) if gray_level between max and 255
Then $a = 1$;

This algorithm converts detected face images of both authorized and unauthorized persons into fuzzy domain data. Every fuzzy data (a) is assigned as fuzzy member of the fuzzy sets A and U , as represented in Equations (1) and (2) that are fuzzy sets for authorized and unauthorized persons respectively.

$$A = \{a_1, a_2, a_3, \dots, a_n\} \quad (1)$$

$$U = \{u_1, u_2, u_3, \dots, u_n\} \quad (2)$$

where, n is the number of face images, $a_1, a_2, a_3, \dots, a_n$ are fuzzy data which is obtained by using algorithm for authorized persons and they are collected together as members of fuzzy set A . Similarly, $u_1, u_2, u_3, \dots, u_n$ are members of fuzzy set U for unauthorized persons.

The general representations of these fuzzy sets [27] are denoted with the Equations (3) to (6) as,

$$A = \{a, \mu_A(a) | a \in X\} \tag{3}$$

$$U = \{u, \mu_U(u) | u \in X\} \tag{4}$$

$$\mu_A(a) = \begin{cases} 1, & \text{if } a \in A \\ 0, & \text{if } a \notin A \end{cases} \tag{5}$$

$$\mu_U(u) = \begin{cases} 1, & \text{if } u \in U \\ 0, & \text{if } u \notin U \end{cases} \tag{6}$$

where, a, u denotes common representation of fuzzy member for the fuzzy sets A and U respectively. X denotes the universe of disclosure. $\mu_A(a), \mu_U(u)$ indicates the membership functions [27]. These fuzzy sets are used in face recognition process to make decision by using fuzzy rule based method.

3.3. Feature Extraction

Haar wavelet transform [28] is employed for extracting features of individual images. The Haar transform is a mathematical analysis tool for the image decomposition and feature extraction in the wavelet transform using decomposition and reconstruction matrices. The image segmentation [29] is carried out by comparing co-occurrence matrix features of size $N \times N$ derived from wavelet transform by decomposing horizontally and vertically.

In this work, the Haar wavelet technique is used for extracting facial features of the face image. Each feature block of the detected face image is extracted based on Haar wavelet by considering following steps:

- 1) The face image is decomposed into sub images of size 4×4 or 8×8 in vertical and horizontal directions from left corner of the image [9] as shown in **Figure 1**.
- 2) Approximation coefficient matrix (CA) and detailed coefficient matrices such as vertical, horizontal and diagonal (Cv, Ch and Cd respectively) are computed from the sub images [29].
- 3) Step i and ii are repeated for each CA with the specified level of decomposition.
- 4) New feature blocks are generated as given in Equation (3) by finding pixel differences at the desired resolution level.
- 5) The new feature block is normalized by calculating the mean (μ) and variance (σ) of feature block by Equation (7) as,

$$f = \frac{F - \mu}{\sigma} \tag{7}$$

where, f and F specifies normalized feature block and unnormalized feature block respectively.

3.4. NHMM Training

A model is a statistical model which consists of a set of states to form a Markov chain [15]. A NHMM includes a set of observation data sequences which has two forms of stochastic finite process. The **Figure 2** depicts the transfer from one state to another and the probabilities between the states and the observed data.

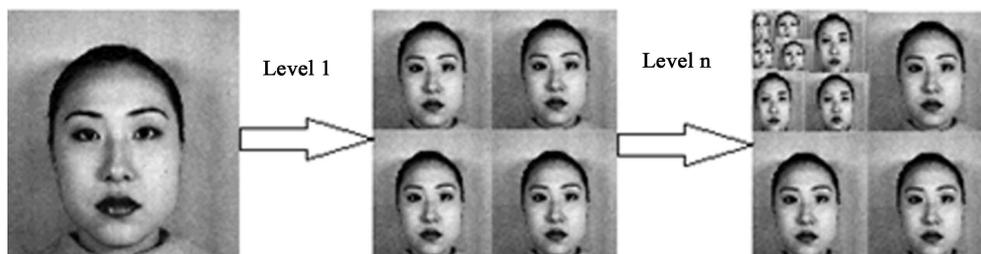


Figure 1. Image segmentation using Haar wavelet.

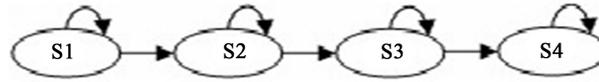


Figure 2. The one dimensional (1-D) HMM structure.

A NHMM [30] can be described with state transition probability matrix T , an initial state probability distribution π , and the emission probability associated with the observations for each state. A NHMM is defined as $\lambda = [\pi, T, E]$. In general, the transition between the states depends upon transition probability and emission probability. One dimensional NHMM structure (Figure 2) is suitable for analyzing 1-D random signals like speech signals. The NHMM structure (as shown in Figure 3) is developed from one dimensional HMM in which s1, s2, s3 and s4 specifies sub states of each feature like eyes, nose, mouth and chin respectively. Each super state indicates the every subject and the number of super states raises if number of subject is increased.

The NHMM involves three stages such as evaluation; decoding and learning to train the system with a different set of face images which are kept in a database.

3.4.1. Evaluation

A face image is considered as a super state which consists of four sub states like eyes, nose, mouth and chin (as shown in Figure 3) to construct NHMM. The number of super states is equivalent to a number of training images in the database.

Each normalized feature block is taken as a sub state which denotes facial features in NHMM. The normalized feature blocks are arranged in column wise to form a sub state sequence (Markov chain) $s = \{e, n, m, c\}$. The state transition matrix T and emission matrix E are calculated to find the probability of possible state transitions and probability distribution between states respectively. The probability of transition matrix $P(T|\lambda)$ and the probability of emission matrix $P(E|\lambda)$ are represented in Equations (8) and (9).

$$P(T | \lambda) = \frac{P(T \text{ and } \lambda)}{P(\lambda)} \tag{8}$$

$$P(E | \lambda) = \frac{P(E \text{ and } \lambda)}{P(\lambda)} \tag{9}$$

NHMM parameters such as number of states, number of symbols, state sequence, pseudo counts, pseudo transitions, transition tolerance and emission tolerance are calculated to model NHMM [15].

3.4.2. Decoding

The decoding stage of NHMM uses Baum Welch algorithm [16] to estimate the maximum likelihood for each sequence as follows:

- 1) Posterior state probability p , logarithm of sequence probability $\log P_{seq}$, forward probability f_s , backward probability b_s , according to scale s are calculated [15].
- 2) The maximum likelihood value is initialized for first sequence or updated for proceeding sequences according to the changes of NHMM parameters.
- 3) The overall transition summation and emission summation are found for each 2D state and up to the entire sequence length, where N and n are number of super states and sub states respectively, whereas, S and s represent each super and sub states respectively.
- 4) The overall Transition summation and emission summation value is updated according to the changes of parameters of NHMM.
- 5) The number of iterations is adjusted for the likelihood estimation of each sub state sequence and the overall probability for transition and emission matrices are estimated. The overall probability of transition matrices and the overall probability of emission matrices are denoted in Equations (10) and (11) as,

$$Q_T (T_s | \lambda) = \sum_{S=1}^N \sum_{s=1}^n P(T_s | \lambda) \tag{10}$$

$$Q_E (E_s | \lambda) = \sum_{S=1}^N \sum_{s=1}^n P(E_s | \lambda) \tag{11}$$

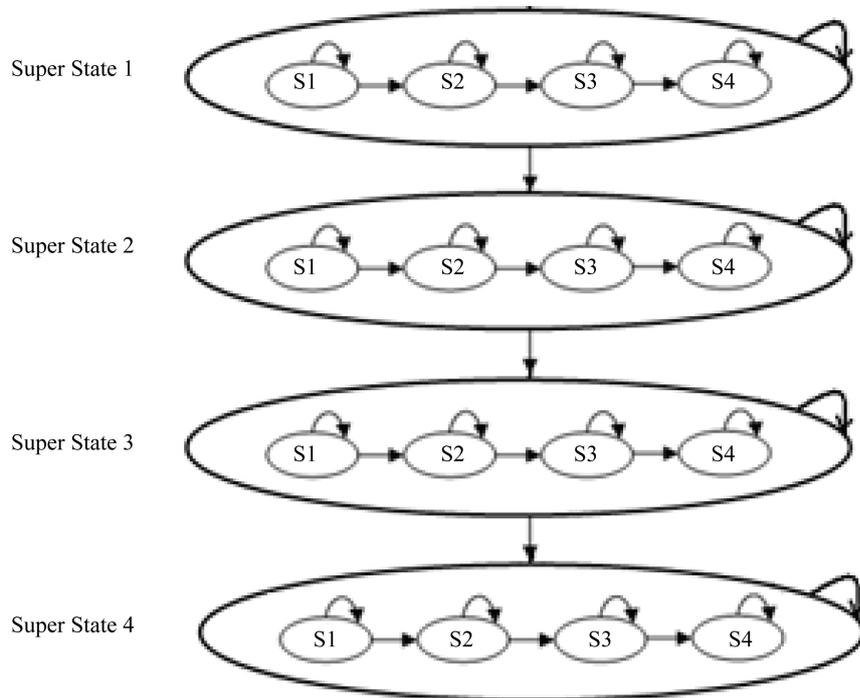


Figure 3. The NHMM structure.

3.4.3. Learning

In the training phase, each face image which is kept in the database is trained individually by following the Baum Welch algorithm to find maximum likelihood (as mentioned above). The number of iterations is increased according to the number of training images kept on the database. The accuracy of the training process depends on the number of iterations with the NHMM. The likelihood probability of training images L_s is represented by the Equation (12) as,

$$L_s = \sum_{S=1}^N \sum_{s=1}^n Q_T (T_s | \lambda) + \sum_{S=1}^N \sum_{s=1}^n Q_E (E_s | \lambda) \quad (12)$$

3.5. Face Recognition

The face recognition involves two procedures such as matching score estimation and fuzzy decision making.

1) Matching score estimation:

During recognition process, the expression and occlusion invariant face image is taken as test image which is to be compared with still images on the database. In some cases, an unauthorized face image may be taken as test image that can be recognized as unknown person.

After the facial region is selected by face detection process, the feature blocks are generated and normalized by using Haar wavelet transform. These feature blocks are arranged in column wise to form a Markov chain. The maximum likelihood is found for the test image by following NHMM training procedure (as shown in **Figure 4**).

The maximum likelihood of test image L_{TS} and maximum likelihood of input image, L_s likelihood probability of training images is taken to calculate the matching score, R , in Equation (13) as follows:

$$R = \arg \max_{0 < s < N} \left[\sum_{S=1}^N \sum_{s=1}^n L_s + \sum_{S=1}^N \sum_{s=1}^n L_{TS} \right] \quad (13)$$

2) Fuzzy decision making:

As discussed in fuzzy data conversion, the fuzzy sets A and U are formed and they are utilized for decision making by using the fuzzy if then rule which is depicted in Equation (14) as follows:

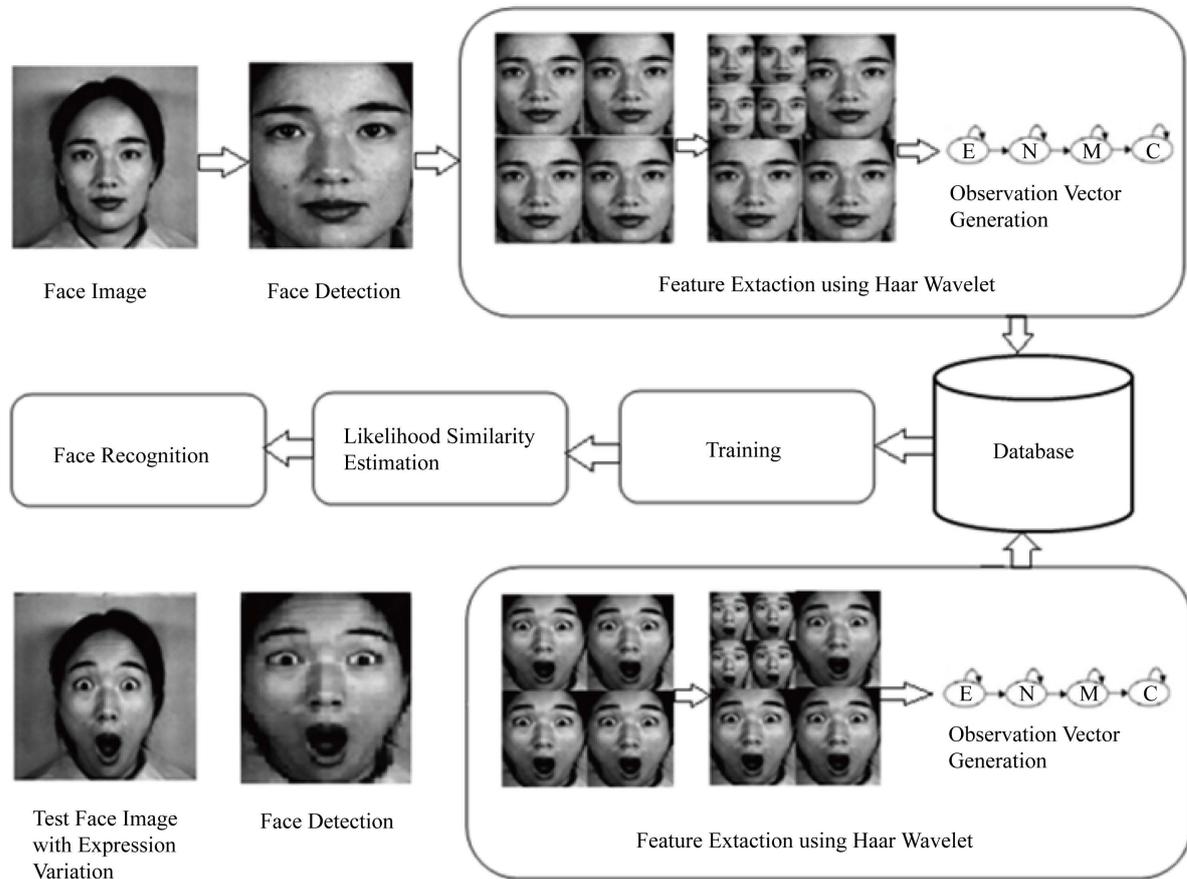


Figure 4. Architecture of expression and occlusion invariant face recognition.

$$\text{IF } R \text{ is high THEN } (a \in A \text{ ELSE } a \in U) \quad (14)$$

where, R is “high” is antecedent in which R is matching score and “high” is fuzzy predicate. “ A ELSE U ” is consequent in which A and U are fuzzy sets. If the R value is high, then the fuzzy data of test image moves towards the fuzzy set A , otherwise, it moves towards the fuzzy set U . The test image is considered as authorized if the fuzzy data of test image is the member of fuzzy set A , otherwise, it is considered as unauthorized. The matching score R becomes high when the test image is authorized person’s image.

4. Experimental Results and Performance Testing

The sample images (as shown in Figure 5) were collected from JAFFE (Japanese Female Facial Expression) database to implement the expression and occlusion invariant face recognition using MATLAB. The system was trained by one sample image per person whereas other images were used for recognition testing. An optimal size of all expression and occlusion varied face images is 256×256 pixels.

4.1. Experimental Result

The input images were cropped to remove the unnecessary background area by exploiting PCA [20]. The cropped face images were enhanced to easily predict the features like eyes, nose, mouth and chin (as shown in Figure 6). The facial region is detected based on the skin color and the simplicity of the background automatically. The eye positions were detected automatically based on color cue and the circle shape of the iris.

To decompose the detected facial region, Haar wavelet transform is used by computing coefficient matrices. The features were extracted and normalized by the Haar wavelet (as shown in Figure 7). These features were trained by NHMM and kept on the database. Observation vectors were generated from transformed image.



Figure 5. Sample images from JAFFE database.

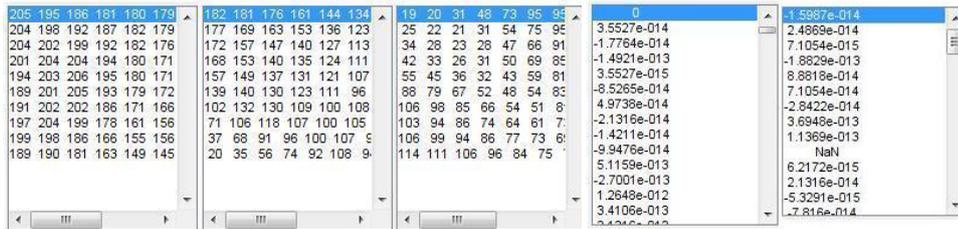


(a)



(b)

Figure 6. Detection Phase (a) Detect face; (b) Enhancement of face image to extract facial features.



(a)

(b)

Figure 7. Training Phase (a) Feature Extraction of face images (b) Normalization of the features.

During recognition process, the same procedure is to be followed for test image and trained feature set is compared with feature set which is kept in database. The maximum likelihood is found between input images and the test image with expression variations (as shown in Figure 8(a)). Based on the likelihood similarity, the test image can be recognized when it has a highest matching score. Otherwise, these test images were considered as an unrecognized person. The proposed work uses JAFFE dataset and AR dataset to recognize expression varied images (as shown in Figure 8(a)) and occlusion varied images (as shown in Figure 8(b)) respectively.

4.2. Performance Testing

The Expression invariant face recognition uses the JAFFE dataset which consists of different kind of expressions for thousands of subjects. This proposed work is tested among 212 face images of ten subjects with seven kinds of expressions like angry, happy, neutral, fear, disgust, sad and surprise (as mentioned in Table 1).

Table 1 shows number of recognized images (R) and number of unrecognized images (E) for each subject according to the seven kinds of expressions. Each subject has different number of expression varied images and all of them are tested. They are clearly represented in Table 1 with the result of whether they are recognized or unrecognized. For instance, the third subject has three angry face images in which two images are recognized

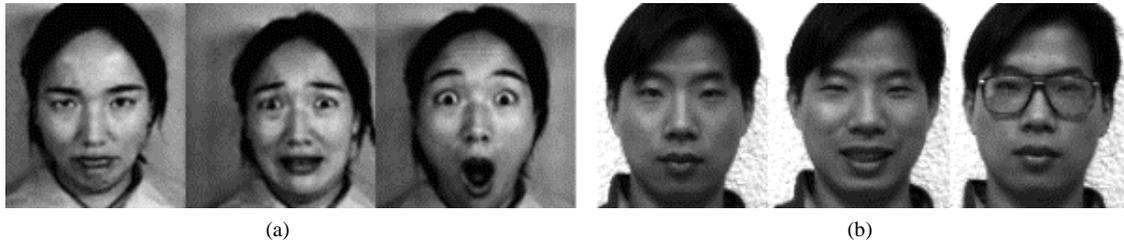


Figure 8. Recognition phase (a) Expression in variant face recognition using JAFFE dataset (b) Occlusion in variant face recognition using AR dataset.

Table 1. Experimental evaluation among JAFFE Dataset according to different kinds of Expressions. R denotes number of recognized images and E denotes number of unrecognized images (Error).

Subjects	Angry		Disgust		Fear		Happy		Neutral		Sad		Surprise		Total		
	R	E	R	E	R	E	R	E	R	E	R	E	R	E	R	E	T
1.	3	0	2	0	3	0	4	0	3	0	4	0	3	0	22	0	22
2.	3	0	4	0	3	0	3	0	3	0	3	0	3	0	22	0	22
3.	2	1	2	1	3	0	4	0	3	0	4	0	3	0	21	2	23
4.	3	0	3	0	3	0	2	0	2	0	3	0	3	0	19	0	19
5.	3	0	2	1	2	1	3	0	3	0	3	0	3	0	19	2	21
6.	3	0	3	0	3	0	2	1	3	0	3	0	3	0	20	1	21
7.	2	1	2	0	3	0	3	0	2	1	3	0	3	0	18	2	20
8.	3	0	3	0	3	0	3	0	3	0	3	0	2	1	20	1	21
9.	3	0	3	0	2	0	3	1	3	0	2	1	3	0	19	2	21
10.	3	0	2	0	3	0	4	0	3	0	4	0	3	0	22	0	22
Total	28	2	26	2	28	1	31	2	28	1	32	1	29	1	202	10	212

(R) and one image is not to be recognized (E). Similarly, third subject totally has 23 Expression varied face images in which 21 images are recognized (R) and remaining two images are not to be recognized (E).

Similarly, the proposed work was tested among 116 occluded face images and it recognizes the 110 images as shown in **Table 2**. The work was evaluated with different kinds of occluded images like person wearing glasses or scarf and they may have beard or hair fall on face.

The Proposed work recognizes 312 face images successfully among 328 expression and occlusion varied face images. The overall recognition rate is calculated as follows:

$$\text{Recognition rate} = \text{number of images recognized} / \text{number of test images} = 312 / 328 = 95.12\%.$$

$$\text{Error Rate} = \text{number of images unrecognized} / \text{number of test images} = 16 / 328 = 4.87\%.$$

Therefore, the system produces 95.12% recognition rate with seven types of expressions. The recognition rate and error rate were found and shown in **Table 3**.

The comparison of proposed work recognition rate with existing methods according to the expressions is shown in **Table 4**. The recognition rate is depicted in a graph (**Figure 9(b)**) according to the expressions. The proposed work produced the highest recognition rate when compared to existing methods like Active Appearance Model [31], Scaled Gaussian Process Regression (SGPR) [8], Coupled SGPR (CSGPR) [4] and the combination of SVM [32] and HM [16].

The proposed work produced the higher recognition rate such as 96.6, 93.3, 92.86, 96.55, 93.9, 97, 96.7 for expressions like neutral, angry, disgust, fear, happy, sad and surprise respectively and it produces 95.28 as average recognition rate. The performance analysis is also depicted with a graph (**Figure 9(a)**) that shows the increased recognition rate of proposed work among the existing works.

Table 2. Experimental evaluation among AR Dataset according to different kinds of Occlusion. S denotes Subject, T denotes total number of recognized images, R denotes number of recognized images and E denotes number of unrecognized images (Error).

S	Glass		Beard		Scarf		Hair		Overall		
	R	E	R	E	R	E	R	E	R	E	T
1	3	0	2	0	3	0	3	0	11	0	11
2	3	0	4	0	3	0	3	0	13	0	13
3	2	1	2	1	3	0	3	0	10	2	12
4	3	0	3	0	3	0	2	0	11	0	11
5	3	0	2	1	2	1	3	0	10	2	12
6	3	0	3	0	3	0	3	0	12	0	12
7	2	1	2	0	3	0	2	1	9	2	11
8	3	0	3	0	3	0	3	0	12	0	12
9	3	0	3	0	2	0	3	0	11	0	11
10	3	0	2	0	3	0	3	0	11	0	11
T	28	2	26	2	28	1	28	1	110	6	116

Table 3. Recognition rate Vs Error rate.

Expression	Total	Recognized	Error	RR	Error Rate
Angry	30	28	2	93.33	6.67
Disgust	28	26	2	92.86	7.14
Fear	29	28	1	96.55	3.45
Happy	33	31	2	93.94	6.06
Neutral	29	28	1	96.55	3.45
Sad	33	32	1	97	3.0
Surprise	30	29	1	96.67	3.33

Table 4. Comparison of proposed work recognition rate with existing methods.

Expression	Recognition Rate			
	AAM	SGPR	CSGPR	Proposed Work
Neutral	72.2	83.4	85.0	96.55
Angry	54.1	72.5	73.2	93.33
Disgust	58.7	73.0	74.8	92.86
Fear	60.4	68.3	69.9	96.55
Happy	72.9	87.2	89.1	93.94
Sad	57.2	68.9	70.2	97
Surprise	78.1	88.5	91.5	96.7
Average	64.8	77.4	79.1	95.28

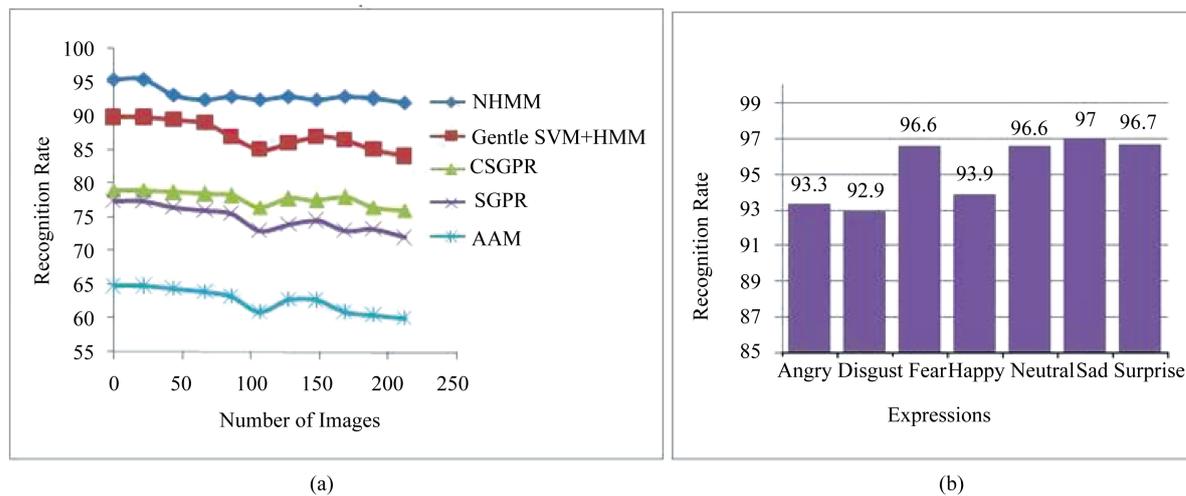


Figure 9. (a) Comparison of recognition rate with existing methods (b) comparison of recognition rate with different expression variations of proposed work.

5. Conclusion

This proposed work addresses the face recognition problem among seven kinds of varying expressions. The face recognition uses a new model to recognize the user by improving the recognition rate and it can operate under varying expressions and occlusion. When compared to existing work, the proposed work improves performance features like recognition rate, accuracy and recognition time. This work produced overall recognition rate of 95.12% for expression and occlusion varied face images and the 97% recognition rate for expression wise evaluation. The future work may improve the overall recognition rate by recognizing the face images which have combination of all constraints such as expression, occlusion, pose and illumination.

References

- [1] Yang, M.H., Kriegman, D. and Ahuja, N. (2002) Detecting Faces in Images: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**, 34-58. <http://dx.doi.org/10.1109/34.982883>
- [2] Pantic, M. and Patras, I. (2006) Dynamics of Facial Expressions: Recognition of Facial Actions and Their Temporal Segments from Face Profile Image Sequences. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **36**, 433-449. <http://dx.doi.org/10.1109/TSMCB.2005.859075>
- [3] Bartlett, M.S. (2005) Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behavior. *Proceedings of IEEE Conference in Computer Vision and Pattern Recognition*, **2**, 568-573.
- [4] Chang, Y. (2006) Manifold-Based Analysis of Facial Expression. *Image and Vision Computing*, **24**, 605-614. <http://dx.doi.org/10.1016/j.imavis.2005.08.006>
- [5] Pantic, M. and Rothkrantz, L.J.M. (2000) Automatic Analysis of Facial Expressions: The State of the Art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**, 1424-1445. <http://dx.doi.org/10.1109/34.895976>
- [6] Daugman, J. (1997) Face and Gesture Recognition: Overview. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**, 675-676. <http://dx.doi.org/10.1109/34.598225>
- [7] Valstar, M. and Pantic, M. (2012) Fully Automatic Recognition of the Temporal Phases of Facial Actions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **42**, 28-43. <http://dx.doi.org/10.1109/TSMCB.2011.2163710>
- [8] Tong, Y., Liao, W. and Ji, Q. (2007) Facial Action Unit Recognition by Exploiting Their Dynamic and Semantic Relationships. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**, 1683-1699. <http://dx.doi.org/10.1109/TPAMI.2007.1094>
- [9] Yan, Y., Wang, H.Z. and Suter, D. (2014) Multi-Subregion Based Correlation Filter Bank for Robust Face Recognition. *Pattern Recognition*, **47**, 3487-3501. <http://dx.doi.org/10.1016/j.patcog.2014.05.004>
- [10] Li, Y.Q. (2013) Simultaneous Facial Feature Tracking and Facial Expression Recognition. *IEEE Transactions on Image Processing*, **22**, 2559-2573. <http://dx.doi.org/10.1109/TIP.2013.2253477>

- [11] Koelstra, S., Pantic, M. and Patras, I. (2010) A Dynamic Texture-Based Approach to Recognition of Facial Actions and Their Temporal Models. *The IEEE Transactions on Pattern Analysis and Machine Intelligence*, **32**, 1940-1954. <http://dx.doi.org/10.1109/TPAMI.2010.50>
- [12] Lu, X. and Jain, A. (2006) Automatic Feature Extraction for Multiview 3D Face Recognition. *7th IEEE International Conference on Automatic Face and Gesture Recognition (FG2006)*, Southampton, 2-6 April 2006, 585-590.
- [13] Vetter, T. and Blanz, V. (2003) Face Recognition Based on Fitting a 3D Morphable Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **25**, 1063-1074. <http://dx.doi.org/10.1109/TPAMI.2003.1227983>
- [14] Valentin, D., Abdi, H., O'Toole, A.J. and Cottrell, G.W. (1994) Connectionist Models of Face Processing: A Survey. *Pattern Recognition*, **27**, 1209-1230. [http://dx.doi.org/10.1016/0031-3203\(94\)90006-X](http://dx.doi.org/10.1016/0031-3203(94)90006-X)
- [15] Jackson, P. HMM Tutorial. Centre for Vision Speech & Signal Processing, University of Surrey, Guildford.
- [16] Frazzoli, E. (2010) Introduction to Hidden Markov Models with Baum-Welch Algorithm, Lecturer Notes on Principles of Autonomy and Decision Making. Aeronautics and Astronautics Massachusetts Institute of Technology.
- [17] Bronstein, A., Bronstein, M. and Kimmel, R. (2006) Robust Expression-Invariant Face Recognition from Partially Missed Data. *9th European Conference on Computer Vision*, Graz, 7-13 May 2006, 396-408.
- [18] Dibeklioglu, H. (2008) Part-Based 3D Face Recognition under Pose and Expression Variations. Master's Thesis, Boğaziçi University, Istanbul.
- [19] Jingu, H., Prabhum, U. and Savvides, M. (2011) Unconstrained Pose-Invariant Face Recognition Using 3D Generic Elastic Models. *IEEE Transactions on Patterns Analysis and Machine Intelligence*, **33**, 1952-1961.
- [20] Kittler, J., Kim, T.K. and Cipolla, R. (2010) On Line Learning of Mutually Orthogonal Subspaces for Face Recognition by Image Sets. *IEEE Transactions on Image Processing*, **19**, 1067-1074.
- [21] Passalis, G., Perakis, P., Theoharis, T. and Kakadiaris, I.A. (2011) Using Facial Symmetry to Handle Pose Variations in Real-World 3D Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **33**, 1938-1951.
- [22] Mehta, R., Yuan, J. and Egiazarian, K. (2014) Face Recognition Using Scale-Adaptive Directional and Textural Features. *Pattern Recognition*, **47**, 1846-1858. <http://dx.doi.org/10.1016/j.patcog.2013.11.013>
- [23] Kodandaram, R., Mallikarjun, S., Krishnamuthan, M. and Sivan, R. (2015) Face Recognition Using Truncated Transform Domain Feature Extraction. *The International Arab Journal of Information Technology*, **12**, 211-219.
- [24] Jerritta, S., Murugappan, M., Khairunizam, W. and Ahmad, W. (2014) Emotion Recognition from Facial EMG Signals Using Higher Order Statistics and Principle Component Analysis. *Journal of the Chinese Institute of Engineers*, **37**, 385-394. <http://dx.doi.org/10.1080/02533839.2013.799946>
- [25] Pan, Z.S., Deng, Z.T., Wang, Y.B. and Zhang, Y.Y. (2014) Dimensionality Reduction via Kernel Sparse Representation. *Frontiers of Computer Science*, **8**, 807-815. <http://dx.doi.org/10.1007/s11704-014-3317-1>
- [26] Paul, L.C. and Sumam, A.A. (2012) Face Recognition Using Principle Component Analysis Method. *International Journal of Advanced Research in Computer Engineering and Technology*, **1**, 135-139.
- [27] Roger Jang, J.-S., Sun, C.-T. and Mizutani, E. (1997) Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Prentice Hall Publication, Upple Sddle River.
- [28] Whitehill, J. and Omlin, C.W. (2006) Haar Features for FACS AU Recognition. *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, Southampton, 2-6 April 2006, 5-101. <http://dx.doi.org/10.1109/FGR.2006.61>
- [29] Arivazhagan, S. and Ganesan, L. (2003) Texture Segmentation Using Wavelet Transform. *Pattern Recognition Letters*, **24**, 3197-3203.
- [30] Touj, S., Amara, N.B. and Amiri, H. (2005) Arabic Handwritten Words Recognition Based on a Planar Hidden Markov Model. *The International Arab Journal of Information Technology*, **2**, 318-325.
- [31] Lucey, S., Ashraf, A. and Cohn, J. (2007) Investigating Spontaneous Facial Action Recognition through AAM Representations of the Face. In: Kurihara, K., Ed., *Face Recognition Book*, Pro Literature Verlag, Germany, 395-406.
- [32] Samal, A. and Iyengar, P.A. (1992) Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey. *Pattern Recognition*, **25**, 65-77. [http://dx.doi.org/10.1016/0031-3203\(92\)90007-6](http://dx.doi.org/10.1016/0031-3203(92)90007-6)