Indian Journal of Science and Technology, *Vol 8(35), DOI:10.17485/ijst/2015/v8i35/82202, December 2015*

ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Robust Face Recognition from Video based on Extensive Feature Set and Fuzzy_Bat Algorithm

Jana Selvaganesan[1]* and Kannan Natarajan[2]

[1]Mookambigai College of Engineering, Tiruchirappalli – 622502, Tamil Nadu, India; sjana95@yahoo.co.in
[2]Jayaram College of Engineering and Technology, Pagalavadi – 621014, Tamil Nadu, India; n.kannan.k@gmail.com

Abstract

Background/Objectives: This paper proposes a novel method of enhancing the face recognition process from video sequence with various pose and occlusion using an extensive feature set called Pose and Occlusion Invariant Feature set (POIF) and unsupervised learning technique. **Methods/Statistical Analysis**: Here an extensive feature set, POIF is created using local invariant feature namely Speeded Up Robust Feature (SURF), appearance features and weighted holo-entropy to find out the uniqueness of the face image. The Active Appearance Model (AAM) has been used to find the appearance based features in the face image. The proposed feature set, POIF is used to select the key frames in the video sequence and the key frame selection is optimised using unsupervised learning method namely, Fuzzy Clustering using Bat algorithm (FC-Bat). A dictionary of keyframes is then created, using which the faces from the test video is recognized. **Findings:** Experimental evaluation is done in MATLAB using McGill Real-World Unconstrained Face Video Database and Honda UCSD Dataset 1. The proposed system using FC_Bat algorithm is compared with Fuzzy optimized POIF feature set and Fuzzy c-means optimized POIF feature set and it is found that POIF with FC_Bat algorithm performs better with an accuracy of 97.5%. **Applications/Improvements:** The computational complexity of the proposed face recognition system is less as it uses unsupervised learning of features and best suits applications involving unlabelled data.

Keywords: Face Recognition, Fuzzy_Bat Algorithm, Occlusion, Pose, Unsupervised Learning

1. Introduction

Existing face recognition algorithms demonstrate satisfactory and reliable performance under constrained environments e.g., uniform illumination and fixed frontal poses[1]. Also existing face recognition algorithms are developed mainly for still images[2]. However, changes in illumination, pose, facial expression and partial occlusion of the face are the prime reasons for the decrease in recognition rate in a face recognition system. Face recognition from video sequences contribute significantly in areas of application where motion is used as an important attribute for face segmentation and tracking, and the presence of more data increases recognition performance. However, these systems have their own challenges. In addition to the need for tracking the video sequence, these systems require recognition algorithms that can integrate information obtained temporally for the complete video[3]. In traditional face image acquisition settings, such as passport agencies or police stations, attributes of consideration ranging from head pose to facial expression are controlled. In contrast, video surveillance systems exhibit a scenario where the activities of the recorded individuals and the effects of the environment vary significantly. Numerous performance evaluation efforts have demonstrated that face recognition algorithms that operate well in controlled environments tend to suffer in surveillance contexts. These issues have motivated the development of face recognition algorithms that make use of the abundant information provided by videos, to compensate for the poor viewing conditions encountered in uncontrolled viewing scenarios[4].

However, video based face recognition poses challenges such as low quality facial images, illumination changes, pose variations, occlusions and so on[5]. Head pose problem has been one of the bottlenecks for most

Author for correspondence

current face recognition techniques, because it changes significantly a person's appearance. With advancement in technology, manual operations are being replaced with robots in the area of human-machine interaction and this requires more robust face recognition algorithms[6]. Meanwhile, in surveillance environments and unconstrained scenarios, partial occlusion such as spectacles, scarf, beard etc., makes face recognition difficult and is a serious problem to be considered. The variation caused by factors such as occlusion, illumination, expression, pose, accessories and aging produces a performance drop in all the face recognition algorithms[7].

The poor performance of existing face recognition methods for video sequences under unconstrained environments particularly with varying pose and partial occlusions have motivated us to do this research in face recognition.

2. Proposed Methodology

Automatic recognition of faces by machines in a real world environment poses a serious challenge in the area of computer vision and video analysis. However, changes in illumination, pose, facial expression, partial occlusion of the face make face recognition in uncontrolled environments a challenging research area and are the prime reasons for the decrease in recognition rate in a face recgnition system. In this proposed method we present a robust Pose and Occlusion Invariant Feature (POIF) for face recognition from video sequences with various pose and occlusion. Here we use the concept of anomaly detection in an unsupervised manner wherein an anomaly is detected in an unlabeled dataset assuming that most of the objects are normal in that dataset. As this approach does not need labeled training dataset, it is most suitable for high dimensional dataset. The sequence of steps involved in the proposed method of face recognition has been illistrated in Figure 1.

Initially, the input video sequence is converted into frames. Then, face part is detected from each frame for further processing. Here we use Viola-Jones Face detection algorithm[8] for detecting the faces from the frames. After detecting the face exactly, the facial features are extracted. Here SURF features[9], appearance features and weighted holo-entropy[11] is used to find out the uniqueness of the face image. The Active Appearance Model (AAM)[10] is used to find out the appearance based features in the face image. From the combination of these three features a new feature set POIF is created using which key frames are selected from the video sequence with an unsupervised learning technique, namely Fuzzy Bat Clustering algorithm (FC-Bat). Here bat algorithm[12] is used for optimizing the Fuzzy rules. A library of keyframes are then created and the faces in the test video sequence is recognized using the library of keyframes. The implementation is done in Matlab and tested using various video from McGill Unconstrained Real World Face Database and Honda UCSD face dataset 1. The proposed algorithm for the training phase and testing phase is given in Figure 2 and Figure 3 respectively.

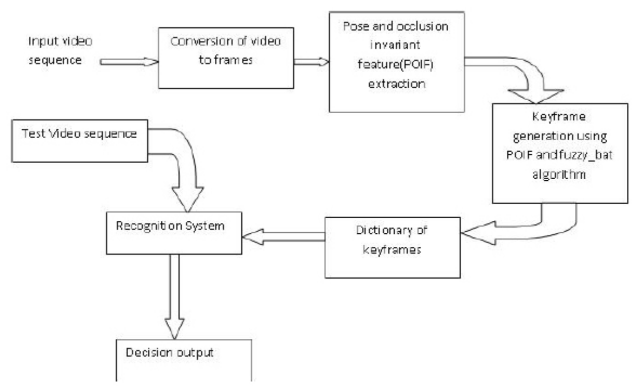


Figure 1. Steps involved in the proposed system.

Algorithm Training Phase

1. Get input video
2. Convert into frames
3. Detect faces in each frame using Viola Jones Face detector algorithm
4. Save each detected face as image file
5. For each image do the following
   - Get the invariant feature points
   - Select the strongest points
   - Get the location of these points
   - Determine the mean of these locations to form the SURF Feature set
   - Determine the modified entropy of each image to form Holoentropy feature
   - Implement Active Appearance Model on each image to form AAM feature
   - Concatenate the three features to form the feature set POIF
6. Learn the features from the sequence of images
7. Measure the similarity between Trained features and target features using squared euclidean distance
8. Cluster the features based on similarity measure
9. Optimize the features using Fuzzy _Bat algorithm
10. Choose the frames with more number of optimised features as keyframes.

Figure 2. Algorithm for the proposed face recognition system (training phase).

Algorithm: Testing Phase

1. Get input video
2. Convert into frames
3. Detect faces in each frame using Viola Jones Face detector algorithm
4. Save each detected face as image file
5. For each image do the following
   - Get the invariant feature points
   - Select the strongest points
   - Get the location of these points
   - Determine the mean of these locations to form the SURF Feature set
   - Determine the modified entropy of each image to form Holoentropy feature
   - Implement Active Appearance Model on each image to form AAM feature
   - Concatenate the three features to form the feature set POIF
6. Learn the features from the sequence of images from the test video
7. Measure the similarity between Test image feature and target features using squared euclidean distance
8. Compare this similarity measure with that obtained with the keyframe
9. If the similarity measure is less than the optimised threshold then a match is said to occur
10. Else the image from the test video is said to mis match with the images in train video.

Figure 3. Algorithm for the proposed face recognition system (testing phase).

3. Extensive Feature Set - POIF

Most of the existing video based face recognition algorithms work on local features obtained from known facial landmarks under controlled illumination, pose etc[13]. Global features lack the ability to capture invariant features leading to poor performance of face recognition algorithms. This drawback is overcome by local features from interest points which form an excellent descriptor. Some of the well known blob feature detectors that are insensitive to scale, rotation and illumination are Laplacian of Guassian, Difference of Gaussian and Determinant of Hessian. While SURF detects interest points using Determinant of Hessian, SIFT detects using Difference of Gaussian[14]. When compared to SIFT, SURF performs better in terms of speed, robustness, distinctiveness[9] and this has motivated us to use SURF in our proposed algorithm POIF. The Active Appearance Model (AAM) suits well for objects which vary in both shape and appearance. Though it involves complex computation, variations in face are not much in videos and hence best suits for applications involving robust pose and expression. Holoentropy is a novel concept which uses entropy and correlation information to detect outliers in a dataset. This motivated us to use the concept of holoentropy in face recognition and propose a modified entropy by assigning weights to the attributes, for the purpose of keyframe extraction. A brief overview of SURF detectors and descriptors, AAM Features and Holoentropy is given in this section.

3.1 SURF Features

Speeded-Up Robust Feature (SURF) is a robust interest point detector and descriptor which is insensitive to noise, displacement, scale, rotation and illumination. SURF has been widely used in areas such as human computer interaction, camera calibration, image registration, object recognition, image retrieval etc., where point correspondences needs to be established between two images or scenes. Basic steps for the implementation of SURF include selection of key-points or interest points at corners, blobs, T-junctions and determination of feature vector corresponding to each interest point.

3.1.1 Selection of Key Points using Integral Images

Selection of key point or interest point is done by calculating the Hessian–matrix using integral images and then finding the maximum of the determinant of the Hessian-matrix to which integer approximation is applied[9]. For a rectangular region formed by origin and any point $x = (x, y)^T$ the integral image $I_{int}(x)$ is given by the sum of all pixels within that region as given in Equation (1) and the Hessian matrix for the point $x(x,y)$ is given by Equation (2).

$$I_{int}(x) = \sum_{i=0}^{i \le x} \sum_{j=0}^{i \le y} I(i,j) \tag{1}$$

$$H(x) = \begin{bmatrix} Lxx(x) & Lxy(x) \\ Lxy(x) & Lyy(x) \end{bmatrix} \tag{2}$$

where $Lxx(x)$ is a second partial derivative in the direction x and $Lxy(x)$ is the mixed partial second derivative in the x and y directions. The derivatives of an image smoothed by gaussian function is given by $L(x) = g(\sigma_I) \otimes I(x)$, σ being the scale factor. The symbol '⊗' denotes convolution operation. The Determinant (D) and Trace (TR) of Hessian matrix is given by Equation (3) and Equation (4) respectively.

$$D = \sigma_I^2 \left(L_{xx}L_{yy}(\mathbf{x}) - L^2_{xy}(\mathbf{x})\right) \tag{3}$$

$$TR = \sigma_I \left(L_{xx} + L_{yy}\right) \tag{4}$$

The blob responses at point x is obtained using a 9×9 Gaussian approximation filter with a scale factor of 1.2[15].

Figure 4 shows the detected interest points in the sample image in a frame of the video input.

3.1.2 Interest Point Description

The interest point descriptor is determined by considering the intensity distribution in the rectangular region surrounding the interest point. Using Haar wavelet filters shown in Figure 5, the responses along x and y direction are computed. The extracted features comprising the interest points and their location, vary from one image to another and hence are unique to an image.

The spatial information is preserved by considering subregions of size 4×4 as shown in Figure 6. The response after the application of Haar wavelet filter, is represented as *dx* along horizontal direction and as *dy* along vertical direction. The sum of the gaussian weighted responses in each subregion is represented by dx and dy. Considering the absolute values of responses, the 4 dimensional vector v for each subregion is given by $(\Sigma dx, \Sigma dy, \Sigma|dx|, \Sigma|dy|)$ leading to a 64 length descriptor for the entire region. This descriptor contributes to the extensive feature set POIF.

3.2 Active Appearance Model

The Active Appearance Model (AAM) suits well for objects which vary in both shape and appearance. Though estimating the parameter is difficult using this model, it finds wide application in face recognition and tracking, as faces do not vary much in shape and appearance. The algorithm

Figure 4. Sample frame from 'fuji.avi' with 178 interest points.

Figure 5. Haar wavelet filter along x and y direction.

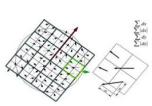


Figure 6. Subregion for Haar wavelet response computation[9].

optimizes by considering the difference between the estimate of target image appearance and the estimate of current appearance. The least squares technique is used to increase the computation speed during matching process.

Given a set of training images with keypoints that form the shape descriptors of the image, it is possible to generate a model for the variations in shape using statistics[16]. Each point is represented by a vector to which Principal Component Analysis (PCA) is applied. Let $\hat{x}$ represent the mean shape, P_s represent a set of orthogonal modes of shape variation and b_s represent the set of parameters that describe the shape. After applying the PCA to each data, the shape model can be described by Equation (5).

$$x = \hat{x} + P_s b_s \tag{5}$$

The grey-level appearance of the image is modeled by using a triangulation algorithm where the sample image is warped in such a way the control points coincide with the mean shape. A linear model thus obtained after application of PCA to the shape normalized data[17] is given by Equation (6).

$$g = \hat{g} + P_g b_g \tag{6}$$

Here $\hat{g}$ presents the mean normalised grey-level vector, P_g represents orthogonal modes having variations in grey level and b_g represents the group of model parameters of grey level. In summary, the vectors b_s and b_g describe the shape and appearance respectively. As there exists a relationship between shape and variations in grey-level, PCA is applied to the data thereby giving a vector as shown in Equation (7).

$$b = \begin{pmatrix} W_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} W_s P_s^T (x - \hat{x}) \\ P_g^T (g - \hat{g}) \end{pmatrix} \tag{7}$$

where W_s represents the weights for each parameter that describes shape. Similarly PCA is applied to this vector b in order to obtain a model for the grey-level appearance as given by Equation (8).

$$b = Qc \tag{8}$$

Here Q denotes the eigenvectors of b, and c denotes a vector of parameters with zero mean that describe appearance.

3.3 Holoentropy

In order to determine the keyframe we use an information theoretic approach which takes into consideration both entropy and total correlation[11]. We formulate a weighted modified entropy called weighted holoentropy.

Let χ be a dataset consisting of n objects $\{x_1, x_2, x_3, \ldots x_n\}$ where each object x_i, $1 \leq i \leq n$, denotes a m x 1 categorical attribute vector $Y = [y_1, y_2, y_3, \ldots y_m^T]$. Here y_i is assumed to be a random variable and subsequently Y as a random vector. Let $H_\chi(.)$, $I_\chi(.)$, and $C_\chi(.)$ denote the entropy, mutual information and correlation calculated on the dataset χ. Using the chain rule, entropy of Y is given by Equation (9).

$$H_\chi(Y) = H_\chi(y_1) + H_\chi(y_2|y_1) + \ldots + H_\chi(y_m|y_{m-1}, \ldots y_1) \quad (9)$$

The entropy gives a measure of uncertainity with respect to a random variable. The holo-entropy $HL_\chi(Y)$ given in Equation (10) is defined as "the sum of the entropy and the total correlation of the random vector Y, and can be expressed by the sum of the entropies on all attributes".

$$HL_x(Y) = H_x(Y) + C_x(Y) = \sum_{i=1}^{m} H_x(y_i) \quad (10)$$

3.3.1 *The Weighted Holoentropy*

The contribution of different attributes vary based on the application and hence we propose a modified form of holoentropy which assigns a weight to each attribute giving more importance to the attributes with less entropy. The weights are obtained directly from the data using inverse sigmoid function and is given by Equation (11).

$$w_\chi(y_i) = 2\left(1 - \frac{1}{1 + \exp(-H_\chi(y_i))}\right) \quad (11)$$

^

The weighted holoentropy of the random vector Y is obtained by adding up the weighted entropies of the individual attributes of Y and is given by Equation (12).

$$W_\chi(Y) = \sum_{i=1}^{m} w_\chi(y_i) H_\chi(y_i) \quad (12)$$

Here the entropy of each frame H(.) is calculated using Equation (9) and hence the vector Y has only one component. Thus holoentropy HL(.) is equivalent to the entropy H(.). Subsequently each frame is assigned a weight function w(.) which is obtained using Equation (11). The weighted entropy W(.) is given by Equation (12). This weighted entropy contributes to the extensive feature set POIF.

3.4 POIF Descriptor using Unsupervised Learning

The pose and occlusion invariant feature set is obtained by concatenating the SURF feature component, AAM feature component and Holoentropy feature on each frame. The squared Euclidean distance between the feature set of training frames and the feature set of target frame is calculated using Equation (13). If the distance measure between these two features is less than a threshold, a match is said to occur. The feature set so obtained is optimized using FC_Bat algorithm. All the frames with the best optimized feature set is assigned as keyframe. Considering two point $U(x_1, y_1)$ and $V(x_2, y_2)$, the squared Euclidean distance (d(U,V)) between them is given by

$$d(U,V) = (x_2 - x_1)^2 + (y_2 - y_1)^2 \quad (13)$$

4. Keyframe Generation using FC_Bat Algorithm

Learning of features can be carried out using supervised, semi-supervised or unsupervised algorithms. Unsupervised learning can be implemented on a dataset without human intervention. We propose an unsupervised learning technique to learn the POIF features of each frame and optimize the feature parameters using fuzzy bat algorithm, so that the feature descriptor of the invariant feature is unique. This domain specific feature descriptor provides the keyframes using which recognition of faces in test video is carried out. This increases the performance during classification process. Fuzzy Bat algorithm is used for the optimization of the POIF feature. A brief description of Bat algorithm and Fuzzy Bat algorithm is given in the following section.

4.1 Bat Algorithm

Xin-She Yang, inspired by the echolocation behaviour of bats, developed an optimization algorithm called Bat Algorithm (BA)[12]. The principle of echolocation is that "Each virtual bat flies randomly with a velocity v_i at position (solution) x_i with a varying frequency or wavelength and loudness A_i. As it searches and finds its prey, it changes frequency, loudness and pulse emission rate r. Search is intensified by a local random walk. Selection of the best continues until certain stop criteria are met". Microbats echolocation capability enables them to detect an object and determine its size, shape, position by sending a loud sound signal in the range of 25kHz to 150 kHz and listening to the echo received[12].

BA together with fuzzy optimization has been successfully used for unsupervised learning[18]. An approach combining these two concepts has proved to provide better matching performance between the actual data and the predicted data in the case of exergy models. Khan et al.[19] in their work used fuzzy bat algorithm to cluster office workplace. Khan and Sahari[20] in their work made a comparison of bat algorithm with other algorithms such as Particle Swarm Optimization and Genetic Algorithms and showed that BA performs better. This has motivated us to implement bat algorithm with fuzzy clustering for the purpose of enhanced face recognition under uncontrolled scenario.

The dynamic behaviour of bats could be controlled by tuning the frequency. Also to maintain a balance between exploration and exploitation, the algorithm dependent parameters could be varied[21]. The update equations for the bat's frequencies A_i, velocity v_i and position x_i in a d-dimensional search space are given in Equations (14), (15) and (16) respectively.

$$A_i = A_{min} + (A_{max} - A_{min}) \times \delta \tag{14}$$

$$v_i^t = v_i^{t-1} + (x_i^t - x_{best}^t) \tag{15}$$

$$x_i^t = x_i^{t-1} + v_i^t \tag{16}$$

Here δ denotes the random vector obtained from a uniform distribution whose value ranges between 0 and 1. The value x_{best}^t denotes the current best location at time t after comparison with all locations of all n bats. Since velocity is given by $v = A\lambda$ where λ is the wavelength at which the waves are emitted, the increment in velocity is based on varying either A or λ keeping the other value fixed. The minimum and maximum value of the frequency is determined by the size of the domain of the problem in consideration. The initial value of frequency for each bat is assigned any value in between A_{min} and A_{max}.

4.2 Fuzzy Clustering using Bat Algorithm

By the process of clustering, data objects are assigned to a particular group called cluster among a group of clusters, based on a similarity measure. In recent past, Fuzzy schemes have been popularly used for clustering where a particular sample may be simultaneously grouped under many categories but with different degrees of membership ranging from 0 to 1. Of various fuzzy clustering schemes, Fuzzy c-Means (FCM) clustering is popular because of its simple and easy implementation. However, the limitation of FCM is that it is sensitive to initial values and gets trapped in a particular optimum. Algorithms such as genetic algorithms and Particle Swarm Optimization (PSO) can be very useful, but they still have some drawbacks in dealing with multimodal optimization problems. However, Bat algorithm has been proved to be an efficient global metaheuristic optimization tool. This motivates us to use Bat algorithm combined with fuzzy clustering to find the keyframe with maximum number of unique features.

Fuzzy Inference System (FIS) has been extensively used in a number of fields such as data classification[22], decision analysis, automatic control and computer vision. It is based on a collection of membership functions and rules to map an input to a corresponding output. The rules in a FIS are usually of the form "If p then q" where p and q are fuzzy statements.

The structure of a FIS involves the following functional operations: Fuzzification, fuzzy inferencing, aggregation of all outputs and defuzzification. Two common types of fuzzy inference system based on the way in which outputs are determined, are Mamdani method and Sugeno method. As Sugeno type is computationally efficient, works well with linear techniques, optimization and adaptive techniques and suited for mathematical analysis in comparison to Mamdani method[23], we adopt Sugeno type inference system for classification of the test input frames.

In Sugeno type FIS, the output membership functions are either linear or constant. A typical sugeno fuzzy model can be described with the following rule.

"If Input 1 = x and Input 2 = y, then output z = ax + by + c"

A zero-order Sugeno model has a = b = 0 leading to a constant output z = c. Each output z_i is multiplied by the firing strength w_i by the fuzzy rule. For an AND rule, w_i is given by Equation (17).

$$W_i = \text{AndMethod}(F_1(x), F_2(y)) \tag{17}$$

where $F_{1,2}(.)$ are the input and output membership functions. A typical Sugeno fuzzy model is shown in Figure 7. The final output of the model with N rules is given by Equation (18). Table 1 summarizes the FIS model for the proposed system. The output of the Fuzzy inference system is used in determining the keyframe. Frames with optimised feature set is assigned as keyframe.

$$\text{Final output} = \frac{\sum_{i=1}^{N} w_i z_i}{\sum_{i=1}^{N} w_i} \tag{18}$$

5. Simulation Results

The proposed method was tested using 60 video sequences consisting of 31 females and 29 males, from McGill Real-World Unconstrained Face Video Database. The video sequences in this database are obtained in unconstrained environments such as changing illumination, different facial expressions, arbitrary face scales, changing head pose and partial occlusions[24]. The experiments were conducted with 70%, 75%, 80% and 85% of the frames as training input and the remaining frames as test input.

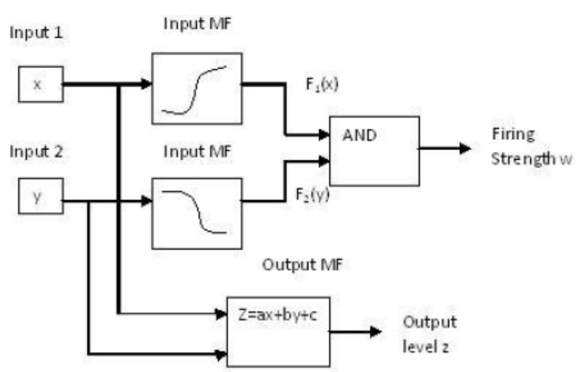


Figure 7. A typical sugeno type fuzzy model structure.

Table 1. Summary of FIS model for the proposed system

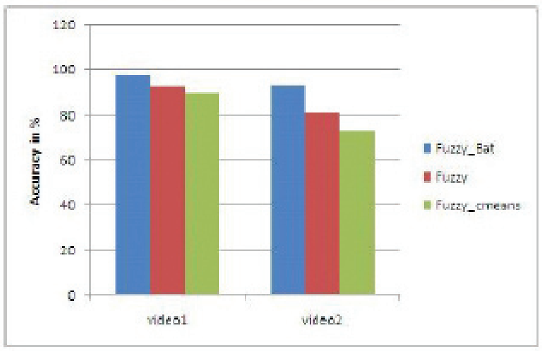
Parameter	Value
Type	Sugeno
And method	Prod
Or method	Max
Defuzzification method	Wtaver
Implementation method	Prod
Aggregation method	Max

Table 2. Comparison of accuracy for three algorithms using POIF feature for sample videos from McGill real-world unconstrained face video database with euclidean distance measure

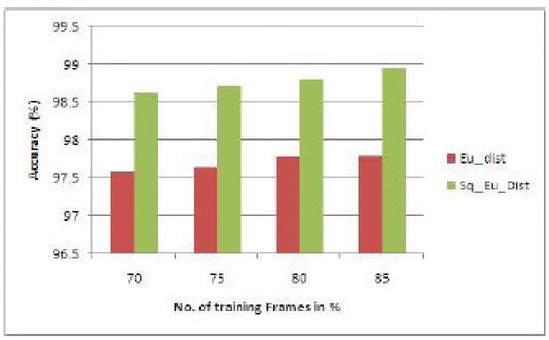
Accuracy / Data	FC_Bat (%)	Fuzzy (%)	FCM (%)
Video 1	97.58	92.79	89.55
Video 2	93.57	80.94	72.82

The sample data were tested with the POIF features and FC_Bat algorithm taking into consideration both the Euclidean distance and Squared Euclidean distance measures. The same data were also tested using POIF features with Fuzzy and Fuzzy c-means algorithm for the two cases of Euclidean distance and squared Euclidean distance.

Table 2 and Figure 8(a) gives the accuracy of POIF feature set implemented using FC_Bat, Fuzzy and FCM algorithm for Euclidean distance measure. The parameters for FC_Bat algorithm are summarized in Table 3. The accuracy is higher for FC_Bat algorithm compared to Fuzzy and FCM algorithm. The simulation results indicate that accuracy is higher. In Figure 8(b) the accuracy obtained when the POIF feature set was implemented with Euclidean distance and squared Euclidean distance measure for a sample data is shown graphically for various percentages of training frames. The accuracy is higher with squared Euclidean distance when compared to Euclidean distance measure. The accuracy obtained for three algorithms using POIF feature for sample videos



(a)



(b)

Figure 8. **(a)** Graphical representation of Table 1. **(b)** Comparison of accuracy for sample video 1 with POIF feature set using Euclidean distance and Squared Euclidean distance measures.

from McGill Real-World Unconstrained Face Video Database with Squared Euclidean distance measure is given in Table 4. Table 5 gives the accuracy for the three algorithms namely FC_Bat, Fuzzy and FCM when AAM and holoentropy features are used, and the proposed POIF feature set is used. The results show that POIF feature set enhances recognition rate when compared to an algorithm using only AAM and holoentropy features. This is due to the presence of SURF features in the POIF feature set. Presence of SURF feature enhances the face recognition under uncontrolled environments. Figure 9(a) and 9(b) illustrates the superior performance of POIF feature with FC_Bat algorithm in comparison to Fuzzy and FCM for the two distance measures namely squared Euclidean distance and Euclidean distance respectively.

Table 3. Summary of parameters for FC_Bat algorithm

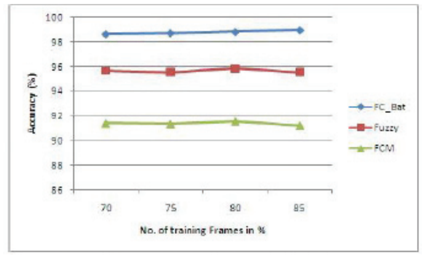
Parameter	Value
Population size	10
Number of generatiions	100
Loudness	0.5
Pulse rate	0.5
Minimum Frequency	0
Maximum Frequency	2
Dimension of search variable	629

Table 4. Comparison of accuracy for three algorithms using POIF feature for sample videos from McGill real-world unconstrained face video database with squared euclidean distance measure

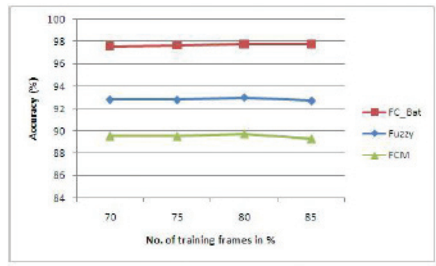
Accuracy / Data	FC_Bat (%)	Fuzzy (%)	FCM (%)
Video 1	98.62	95.65	91.36
Video 2	98.23	88.59	76.59

Table 5. Comparison of accuracy for three algorithms using AAM and Holoentropy for sample data from Mcgill real-world unconstrained face video database

Algorithm	AAM + Holoentropy	POIF
FC_Bat	90.82	98.62
Fuzzy	88.52	95.65
FCM	65.31	91.36



(a)



(b)

Figure 9. Comparison of accuracy for the POIF feature set with three algorithms, FC_Bat, Fuzzy, and FCM for sample data1 with. **(a)** Squared euclidean distance measure. **(b)** Euclidean distance measure.

6. Conclusion and Future Work

POIF is created by combining the robust invariant local features, the appearance features and the principle of outlier detection which is based on the concept of holoentropy. Entropy, a measure of uncertainty information, and the correlation information are combined to form holoentropy. Appropriate weights are assigned to the holoentropy using reverse sigmoid function to increase the effectiveness in the selection of keyframes which are then used to recognize the faces from the test videos. FC_Bat algorithm was used for clustering the frames and determining the keyframe from each cluster. A comparison of the proposed POIF feature implementation and optimization was made with that of fuzzy and Fuzzy c_ Means clustering (FCM) and was found that POIF with FC_Bat clustering performed better. An extensive experimental evaluation by simulation showed that the feature extraction using POIF along with unsupervised learning of frames using FC_Bat algorithm gives better results of accuracy under varying pose and partial occlusions when compared with other feature extraction and clustering algorithms. In addition to pose and occlusion invariance characteristic, robustness to resolution and depth could be taken as a future work. Also recognizing individuals from a crowd under uncontrolled scenario could be carried out in future.

7. References

1. Chng SI, Seng KP, Ang LM. Multiview-multiband face recognition system to solve illumination and pose variation. 3rd IEEE International Conference on Computer Science and Information Technology; Chengdu. 2010. p. 384–8.
2. Ayoob MR, Mathusoothana R, Kumar S. Face recognition using symmetric local graph structure. Indian Journal of Science and Technology. 2015 Sep; 8(24).
3. Xu Y, Chowdhury AR, Patel K. Integrating illumination, motion, and shape models for robust face recognition in video. EURASIP Journal on Advances in Signal Processing. 2008.
4. Barr JR, Bowyer KW, Flynn PJ, Biswas S. Face recognition from video: A review. International Journal of Pattern Recognition and Artificial Intelligence. 2012; 26(05):53.
5. Zhang Z, Wang C, Wang Y. Video-based face recognition: State of the art. Biometric Recognition in Computer Science. Berlin Heidelberg: Springer-Verlag; 2011. p. 1–9.
6. Raducanu B, Dornaika F. Pose-invariant face recognition in videos for human-machine interaction. Computer Vision – ECCV. Berlin Heidelberg: Springer-Verlag; 2012. p. 566–75.
7. Shermina J, Vasudevan V. Face recognition system with various expression and occlusion based on a novel block matching algorithm and PCA. International Journal of Computer Applications. 2012; 38(11):27–34.
8. Viola P, Jones MJ. Robust real-time face detection. International Journal of Computer Vision. 2004; 57(2):137–54.
9. Bay H, Ess A, Tuytelaars T, Gool L. Speeded-Up Robust Features (SURF). Computer Vision and Image Understanding. 2008; 110(3):346–59.
10. Edwards GJ, Cootes TF, Taylor CJ. Face recognition using active appearance models. Computer Vision-ECCV'98. Berlin Heidelberg: Springer-Verlag; 1998. p. 581–95.
11. Wu S, Wang S. Information-theoretic outlier detection for large-scale categorical data. IEEE Transactions on Knowledge and Data Engineering. 2013; 25(3):589–602.
12. Yang XS, Gandomi AH. Bat algorithm: A novel approach for global engineering Optimization. Engineering Computations. 2012; 29(5):464–83.
13. Mian A. Unsupervised learning from local features for video based face recognition. IEEE International Conference on Automatic Face and Gesture Recognition. Amsterdam. 2008. p. 1–6.
14. Lowe DG. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision. 2004; 60(2):91–110.
15. Du G, Su F, Cai A. Face recognition using SURF features. Proceedings of SPIE, 7496; 2009.
16. Cootes TF, Taylor CJ, Cooper DH, Graham J. Active shape models – Their training and application. Computer Vision, Graphics and Image Understanding. 1995; 61(1):38–59.
17. Cootes TF, Edwards GJ, Taylor CJ. Active appearance models. IEEE Transactions on Pattern Analysis and Machine Intelligence; Freiberg, Germany. 1998. p. 681–5.
18. Yang XS, He X. Bat algorithm: Literature review and applications. International Journal of Bio-Inspired Computation. 2013; 5(3):141–9.
19. Khan K, Nikov A, Sahai A. A fuzzy bat clustering method for ergonomic screening of office workplaces. S3T 2011, Advances in Intelligent and Soft Computing. 2011; 101:59–66.
20. Khan K, Sahai A. A comparison of BA, GA, PSO, BP and LM for training feed forward neural networks in e-learning context. IJISA. 2012; 4(7):23–9.
21. Khan K, Sahai A. A fuzzy c-means bi-sonar-based metaheuristic optimization algorithm. International Journal of Artificial Intelligence and Interactive Multimedia. 2012; 1(7):26–32.
22. Neshat M, Sepidname G, Eizi A, Amani A. A new skin color detection approach based on fuzzy expert system. Indian Journal of Science and Technology. 2015 Sep; 8(21).
23. Kaur A, Kaur A. Comparison of mamdani-type and sugeno-type fuzzy inference systems for air conditioning system. IJSCE. 2012; 2(2):323–5.
24. Demirkus M, Clark J, Arbel T. Robust semi-automatic head pose labeling for real-world face video sequences. Multimedia Tools and Applications. 2013; 70(1):495–523.