Automatic Analysis of Human Facial Expression Recognition in Image Sequence

Geethu G S¹ & Usha Kingsly Devi K²
¹,²Department of Electronics and Communication Engineering, Anna University Regional Campus – Tirunelveli, Tamilnadu, India.

Abstract: Facial expression conveys nonverbal cues among humans. Recognition of facial expression is an essential research area in the field of Human Machine Interfaces (HMI). Most of the existing automated system for facial expression analysis has an impact over accuracy. This paper proposes Multi-class Linear Discriminant Analysis (LDA) to obtain better recognition rate of facial expressions. The seven basic facial expressions are used in this work are happy, surprise, sad, anger, fear, disgust, and neutral. The face in an image sequence is detected using Viola Jones algorithm. Geometric and Appearance-based feature extraction approaches are used here. The geometrical features are extracted using Image Normalization and Thresholding techniques and optimal facial feature points are selected by an Entropy calculation. Local Directional Pattern (LDP) is utilized to extract the appearance-based facial feature which contains edge, spots and corner information of an image, and then the Local Directional Pattern Variance (LDPv) descriptor obtains contrast contribution during expression changes. Finally, the performance of the system is compared with machine learning techniques. The proposed method is trained and tested using images from Cohn-Kanade database.

Index Terms—Facial Expression Recognition (FER), Human Machine Interfaces (HMI), Viola Jones detection, Local Directional Pattern (LDP), Local Directional Pattern variance (LDPv), Linear Discriminant Analysis (LDA), K-means clustering, Cohn-Kanade (CK).

1. Introduction

Automated facial expression recognition has attracted much attention in scientific community. It has many applications to human-machine interaction, data-driven animation, video conferencing, psychological theory, human emotion analysis etc. The impact of facial expression on the above-mentioned application area is constantly growing due to many factors includes subtlety, complexity, variability found in the different set of facial images. To overcome this, several research efforts have been done regarding facial expression recognition system. Psychologically, there are six basic facial expressions namely happy, surprise, sad, fear, anger and disgust. Human can easily interpret the facial expressions, but for a machine this task is rather difficult. The system that would perform the following process, such as face detection, feature extraction, and classification in real time with high accuracy brings big achievements for human-machine interaction. Although an extreme large variety of lighting conditions, resolution, pose and orientation affect the recognition rate of facial expressions.

Generally, facial expression recognition is a process to analyze the shape and motion of facial features. Shape feature contains head height, mouth width, eyebrows vertical position, eyes width etc and motion feature contains lip stretch, nose wrinkle, inner brow raiser, outer brow raiser etc. Feature extraction is the keypoint for facial expression recognition. In this paper, two different feature extraction processes are used: Geometric and Appearance-based feature extraction. Shape features are deal with Geometric-based feature extraction and the edge features are deal with Appearance-based feature extraction. These features help to boost the accuracy of FER system.

In our system, Multi-class Linear Discriminant Analysis classifier is proposed. This classifier can handle arbitrary number of classes and it finds the linear combination of features that characterizes or separates two or more classes of objects or events. The performance of proposed classifier is compared with K-Means Clustering depends on recognition rate accuracy. The structure of this paper is organized as follows: Section II deals with the related work done for this system and Section III describes the proposed facial expression recognition system. The experimental results are presented in Section IV. Conclusion and future scope are discussed in Section V.

2. Related Work

Ekman and Friesen [1] proposed that there are a neutral and six basic facial expressions namely happiness, surprise, sad, fear and disgust. These seven facial expressions are used in this work. Facial expression conveys the emotional state of an individual to observers, which is in the form of non-
verbal communication. Facial Expression Recognition system is a computer application capable of identifying or verifying a person’s expression from a digital image or video frame. A survey for facial expression recognition from image sequence can be found in [2, 3]. Lajevardi and Margaret [4] proposed a method for face detection in image sequences which is fully automatic. They have used Viola Jones detection method [5] based on Haar-like features and AdaBoost learning algorithm. The statistical characterization of the motion patterns in specific regions of the face is proposed by Yacoob and Davis [6]. They developed a region tracker for rectangles enclosing the facial features.

The two most common approaches to the facial feature extraction are the geometric-based methods and the appearance-based methods [7]. Appearance-based features are affected by the facial appearance variations between subjects and geometric-based features are affected by shape deformations corresponding to each facial expression. Differ from the local texture; geometric feature is an important visual cue for facial expression recognition.

Detection and location of the face as well as extraction of facial features from images is an important stage for facial image interpretation tasks. Facial feature points normally include corners of eyes, eyebrow corners, lip corners, nostrils, and eyeballs. There are two types of methods for facial feature point detection: texture-based and shape-based methods. Local texture around a given feature point, is considered in texture-based methods [8, 9, 10]. While in shape based methods all facial feature points are treated as a shape [11, 12]. Neural network based feature detector proposed in [8]. This feature detector detect an eye from images by scanning an input window of i*j pixels across the image, where each gray value in the input window serves as an input for the neural network. A log Gabor wavelet for automatic facial feature point detection specifically identifies the facial feature points required for lip and head trackers [9]. The seven facial feature points were detected, namely the outer eye corners, one nostril, and the outer corners and top and bottom mid points of the lips. In [10], a two level hierarchical wavelet network for facial feature localization is presented. Hu et al. [11] proposed a method, called Active Wavelet Networks, in which a Gabor wavelet network representation is used to model the texture variation in the training set. The Gabor wavelet network approach represents a face image through a linear combination of 2D Gabor functions whose parameters position, scale, orientation, and weights are optimally determined to preserve the maximum image information for a chosen number of wavelets. Direct appearance model (DAM) [12] is applied for multi view face alignment. Haar feature based Adaboost classifier [13] combined with the statistical shape models to localize the relevant facial features. Boosting algorithm [14] determines the facial feature point candidates for each pixel in an input image and then uses a shape model as a filter to select the most possible position of feature points, like eyes, nostrils and lip corners. Vukadinovic et al. [15] propose a method using Gabor wavelets and AdaBoost classifier. They detect 20 facial feature points in images of expressionless face. The method consists of four steps: face detection, region of interest detection (ROI), feature extraction, and feature classification. For face detection they use fast and robust face detector based on a cascade scheme consisting of a set of Haar feature based Gentleboost classifier. The detected face region is then divided into 20 ROIs, each consisting of one facial feature using a combination of heuristic techniques based on the analysis of vertical and horizontal image histograms. In general, although some of these detectors seem to perform quite well when localizing a small number of facial feature points such as the corners of the eyes and the mouth, none of them detects all facial feature points and, more importantly, none performs the detection with high accuracy. Kimura and Yachida [16] use integral projection method to detect the facial features. Here the input image is normalized using the centre of the eyes and mouth. Rajesh et al. [17] proposed a technique to achieve high accuracy for feature point detection. This method detects the facial feature point in any image from image sequences with tilted faces under varying illumination conditions. It is comprised of five steps: Face detection, Region of interest detection, image normalization, thresholding, and corner detection.

Feature selection is a process of selecting a subset of relevant features (variables, predictors). The advantage of feature selection are as follows, simplification of models makes them easier to interpret by researcher/user, short training time enhance generalization by reducing the overfitting and improve the performance of the classifier.

Variance and entropy are the strong metrics for measuring the information content of data. The data obtained are informative around mean with known variance, such as Gaussian distribution [18]. Entropy provides reliable information about the variations occurred at particular moments in time and analyzes the evolutionary process over time. In general, entropy is a statistical measure of randomness that can be used to characterize the features. There are several applications of entropy in the field of industrial organization and innovation studies [19]. The concept of entropy originated from the studies of Ludwig Boltzmann in 1877 and then probabilistic interpretation in information theory by Claude Shannon in 1948. Later, Henri Theil developed several applications of information theory in economics and information theory (1967) and
statistical decomposition analysis (1972) based on entropy [19]. In [20] entropy measurement extracts the feature points which are highly affected by expression deformation. More entropy denotes that the feature point is highly affected and it carries more information about the expression. Low entropy means a feature point is stable which does not affected by an expression deformation.

Appearance-based extraction methods deal with the changes occurred in facial appearance and these deformations are extracted using image filters such as Gabor-wavelet, Curvelet and local binary pattern [21, 22 & 23]. Local Binary Pattern has many applications in the field of face detection, face recognition and facial expression recognition [23, 24]. Although LBP is efficient for better computation and robustness to the monotonic illumination changes, it is sensitive to the non-monotonic illumination variation and obtains poor performance in the presence of random noise. Tasked and others proposed Local Directional Pattern (LDP) method to overcome the weakness of Local Binary Pattern (LBP) [25]. This technique applied successfully in the field of face recognition, gender recognition, object description and facial expression recognition [26, 27]. LDP feature extraction method represents curves, edge and texture characteristic of face [28]. Most existing appearance-based method only considers the whole facial feature to classify expressions. Due to changes of facial expression, the region of eyes and mouth has powerful influence. Therefore, highlight the local region which has more contribution on expression changes. Jabid et al. introduced additional contrast information to Local Directional Pattern descriptor and this weighted LDP method called Local Directional Pattern Variance (LDPv) [29]. It adjusts the different contributions of LDP descriptor using the variance of local structure and account that texture with significant contrast should impact more such as eyes and mouth that are more sensitive to high contrast regions. In [30] first extract the global LDP features which can guarantee basic expression difference and then LDPv descriptor extracts the local regions of eyes and mouth to extrude the distinction between expressions.

Classification of facial expression depends on deformation of face corresponds to each expression. The facial expression classified into seven types: happy, surprise, sad, anger, fear, disgust, and neutral. A lot of classifier is used for facial expression recognition such as Naïve Bayesian classifier, Neural Networks, Bayesian classifier, K-Nearest Neighbour, Hidden Markov Model (HMM), Linear Discriminant Analysis, Multi-class SVM, K-Means Clustering and its performance varies according to the recognition rate accuracy obtained for each facial expressions. NB classifier is used for facial expression recognition in [4]. Wenning et al. introduces the Gabor wavelet transformation method which converts the 34 geometric points into a labelled graph vector. For facial expression recognition uses Kernel Canonical Correlation Analysis (KCCA) tested on Japanese female facial expression (JAFFE) [31]. The recognition rate for six expression using KCCA achieved as 77.05%. Petar and Aggelos [32] extracts the facial animation parameter using active contour algorithm and Multistream Hidden Markov Model (MHMM) based automatic facial expression recognition have used. This technique achieves high recognition rate for fear and sadness facial expressions apart from another four expressions. Kotsia and Pitas [33] have proposed a method which is based on mapping and tracking of facial features from the video frame. They have used Candide wire frame model for feature tracking. Recognition is semiautomatic, in the sense means that the user has to manually place the candied grid nodes on facial landmarks depicted at the first frame of image sequence.

Multi-class Support Vector Machine classifier classifies the six expression into its corresponding classes, neutral state is not considered here. It obtains high recognition rate accuracy. Hamid et al. proposed a fixed geometric model for geometric normalization of facial images. Multiclass Support Vector Machine with polynomial kernel representation for classifying the selected Extended Cohn-Kanade datasets. This eliminates geometric variability in emotion expressions and obtain high recognition rate for fear, surprise and happiness expressions with less computational cost [34]. Viola-Jones Algorithm detects the region of interest face and extracts the features using image normalization and thresholding technique. Active Appearance Model (AAA) [35] tracks the facial feature points and classifies the seven facial expressions using Multiclass SVM. This method is experimented on Cohn-Kanade and IMM database.

In [36, 37] presented a Linear Discriminant Analysis (LDA) for facial expression classification. Lyons et al. uses Gabor filter at different scales and orientations, and 34 fiducial points applied for each convolved image to construct the feature vector for representing each facial image. After, PCA is applied to reduce the dimensionality of feature vectors, and LDA is used to identify seven different facial expressions. The recognition rate achieved as 92% for JAFFE database [37]. Dubuisson et al. combined the sorted PCA as feature extractor with LDA as the classifier to recognize facial expressions [36]. Their recognition rate is 87.6%. Wang [38] proposed the K-Means Clustering with Cosine similarity distance measurement for facial expression recognition.

3. Proposed Framework

Figure 1 illustrates the proposed framework, which is composed of face detection, feature point
detection, feature extraction and classification. Face detection is performed by Viola Jones algorithm and then crop the facial image according to the rectangle enclosed on the face. The geometric and appearance based features are used here. Facial feature points from the image sequence, such as eyebrow corners, eye corners, nostrils and lip corners are detected using image normalization and thresholding techniques. After, Shannon Entropy selects the deformable feature points which carry more information about the human emotions. Local Directional Pattern feature descriptor is an appearance-based feature extraction method. A LDP feature is obtained by computing the edge response values in all eight directions at each pixel position and generating a code from the relative strength magnitude. Each face is represented as a collection of LDP codes for the recognition process. In addition to this edge features, the contrast feature is extracted here. Local Directional Pattern variance (LDPv) characterizes the contrast variance of local texture information for more accurate facial expression recognition performance. The extracted geometric and appearance-based features are given as input to the classifier. Multi-cals Linear Discriminant Analysis and K-Means Clustering classifier are trained for seven facial expressions using Cohn-Kanade database. Both the classifier classifies the expression into one of the seven classes such as happy, anger, sadness, surprise, disgust, fear, and neutral. Finally, the performance of this classifier is compared based on accuracy.

3.1. Face Detection

The facial expression recognition system is to be fully automatic in real time applications. The high variability present in different types of face is difficult to detect the face. In this work, Viola Jones detection technique is used. In general, Viola Jones algorithm is a object detection technique helps to detect the object in real time with good detection rate, at least two frames per second is processed. Viola Jones algorithm based on Haar features detects the face area includes eye, nose and mouth regions. These regions are selected by enclosing a rectangle box on the face. Crop the outer portion of the rectangular box; thereby remove the ear and hair portions. When compared to eye, nose and mouth region, the ear and hair portion doesn’t give any information about the emotion. Detected facial feature region is shown in Figure 2.

![Figure 2. Detected facial feature region](image)

3.2. Geometric-based Feature Extraction

The shape features play an important role in facial expression recognition system. These features define the shape of each facial region changed during expression deformation. Geometric based feature extraction consists of two steps: feature point detection and selection of optimized feature points. The feature points, such as eye corners, eyebrow corners, eyeballs, lip corners and nostrils are detected using [17].

For feature point detection, divide the face into different regions. First divide the cropped facial image into three equal halves horizontally as upper part, middle part, and lower part. The upper part contains eye region, middle part contains nostril region, and lower part contains mouth region. Then the upper part is divided vertically into two parts such that each part contains one eye. Each eye region is again divided horizontally into two parts, thereby separate the eye and eyebrow. The position of the eyeball is detected using horizontal and vertical histogram analysis. Vertical histogram shows the intensity difference between the row pixels wise. The peak of this histogram gives the y coordinate of the eyeball. Horizontal histogram shows the intensity difference between the column pixels wise. The peak of this histogram gives the x coordinate of the eyeball. In each separated facial region, perform image normalization and thresholding techniques. Image normalization is a linear process. For example, if the intensity range is 50 to 180 and the desired range is 0 to 255, subtract the minimum value of image from each of pixel intensity, making the range 0 to 130. Then the pixel intensity is multiplied by 255/130, making the desired range 0 to 255. Compute the average of normalized image using (1).

$$avg = \frac{1}{N_xN_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} R(x, y)$$ (1)
where, N1×N2 is the size of the image and R(x,y) is a pixel value at coordinate (x,y). Select the threshold value as \( \frac{\text{avg}}{3} \), which converts the gray scale image into binary image using equation (2).

\[
\text{if } R(x,y) > \frac{\text{avg}}{2} \Rightarrow R(x,y) = 1, \text{ else } R(x,y) = 0
\]

(2)

Then detect facial feature points from each region. Entropy based feature selection chooses the accurate feature points which are highly affected by expression deformations. More entropy means the feature point is more affected and therefore carries information about the expression. Low entropy means a stable feature point which does not change throughout an expression deformation.

Shannon entropy is used here to select the deformable features of seven facial expressions. The computation of entropy is shown in equation (3). Shannon entropy is defined as:

\[
H(x) = -\sum P(x_i) \log_b P(x_i)
\]

(3)

Here, \( b = 2 \) (entropy uses two bins of logical arrays). \( P(x_i) \) is the probability of occurrence for each symbol encountered in the image. Normalize the image by the total number of pixels to get the probability distribution function.

### 3.3. Appearance-based Feature Extraction

Local Directional Pattern (LDP) extracts the appearance changes of face during expression deformation. LDP descriptor contains detail information of face such as edges, spots, corners and other local textures. This feature is obtained by computing the edge response values in all eight directions at each pixel in an image and encoding the edge response value into an 8 bit binary number using the relative strength of the edge responses. LDP features are less sensitive to noise and illumination changes, which describes the local primitives in a more stable manner and also retains more information.

Eight directional edge response values \( \{m_i\} \), \( i = 0, 1 \ldots 7 \) of a particular pixel is computed using Kirsch mask in eight different orientations \( M_i \) centered on its own position. It is a 3×3 matrix shown in the Figure 3, and Figure 4 shows eight directional edge response positions. Here, \( m_0 \) is the k-th most significant directional response value. The code formation is shown in Figure 5.

\[
LDP_x = \sum_{i=0}^{8} b_i(m_i - m_k).2^i, b_i(a) = \begin{cases} 1 & a \geq 0 \\ 0 & a < 0 \end{cases}
\]

(4)

![Figure 3. Eight directional Kirsch edge response masks](image)

![Figure 4(a). Eight directional edge response positions. (b). LDP binary bit positions](image)

![Figure 5. LDP code formation with k=3](image)

LDP histogram for an input image with size M×N is represented in (5).

\[
H(i) = \sum_{r=1}^{M} \sum_{c=1}^{N} f(LDP_x(r,c),i), f(a,i) = \begin{cases} 1 & a = i \\ 0 & a \neq i \end{cases}
\]

(5)

Compute all the LDP code for each pixel \( (r,c) \), where \( i \) is the LDP code value. For a particular value \( k \), there are \( C_k^2 \) different bins for the histogram. Resulting histogram vector has the size of \( 1 \times C_k^2 \) is produced for the image.

Local Directional Pattern variance (LDPv) method is considered to obtain corner information to LDP operator. Therefore the variance \( \sigma \) is introduced as an adaptive weight to adjust the contribution of the LDP code in the histogram generation.
where, $\bar{m}$ is the average of all directional responses $\{m_i\}$ which is calculated for a position $(r,c)$.

### 3.4. Multi-class Linear Discriminant Analysis

Linear Discriminant Analysis is a method used in statistics, pattern recognition and machine learning to find a linear combination of features. The basic idea of LDA is to find a linear transformation that best discriminates among classes and the classification is then performed in the transformed space based on some metric such as K-Nearest Neighbor, Euclidean distance etc. Mathematically a typical LDA implementation is carried out via scatter matrix analysis.

Two-class LDA was first introduced by Fischer and its idea was to transform the multivariate observations $x$ to univariate observations $y$ such that the $y$’s derived from the two classes were separated as much as possible. If the number of classes is more than two, then a natural extension of Fischer Linear Discriminants exists using Multiple Discriminant Analysis. As in two class case, the transformation suggested still maximizes the ratio of intra-class scatter to the inter-class scatter. But unlike the two-class case, the maximization should be done among several competing classes.

Multi-class LDA is a generalization of standard two-class LDA that can handle arbitrary number of classes. It is based on the analysis of two scatter matrices: with-class scatter matrix and between-class scatter matrix.

Given a set of samples $x_1, \ldots, x_p$ and their class labels $y_1, \ldots, y_n$; The with-class scatter matrix is defined as:

$$S_\omega = \sum_{i=1}^{p}(x_i - \mu_{y_i})(x_i - \mu_{y_i})^T$$

where, $n$ is a number of classes and $\mu_{y_i}$ is a sample mean of particular class $y_i$.

The between-class scatter matrix is defined as:

$$S_B = \sum_{k=1}^{m} n_k (\mu_k - \mu)(\mu_k - \mu)^T$$

where, $m$ is the number of classes, $\mu_k$ is the sample mean of the $k$-th class, $n_k$ is the number of samples in the $k$-th class, $\mu$ is the overall sample mean given by $\mu = \frac{1}{m} \sum_{i=1}^{n} \frac{n_k}{m} \mu_k$.

After obtaining $S_\omega$ and $S_B$, find a set of linear combinations (with coefficients $w$) that maximizes the ratio of the between-class scattering to the within-class scattering as:

$$\hat{w} = \frac{|w^T S_B w|}{|w^T S_\omega w|}$$

If $S_\omega$ is non singular, Equation (11) can be solved as a conventional eigenvalue problem and $\hat{w}$ is given by the eigenvectors of matrix $S_\omega^{-1} S_B$. Thus, the transformation $\hat{w}$ can be obtained by solving the generalized eigenvalue problem.

$$S_B \hat{w} = \lambda S_\omega \hat{w}$$

Generally, at most $m-1$ generalized eigenvectors are useful to discriminate between $m$ classes. Multiple Discriminant Analysis provides an elegant way of classification using discriminant features. Once the transformation $\hat{w}$ is given, then the classification is performed in the transformed space based on K-Nearest Neighbour. The principle behind nearest neighbour is to find a predefined number of training samples closest in distance to the new point. Steps followed to compute K-nearest neighbours:

1. Determine parameter $k$ = number of nearest neighbour).
2. Calculate the distance between the query instance and all the training samples.
3. Sort the distance and determine nearest neighbour based on the $k$-th minimum distance.
4. Gather the category $Y$ of the nearest neighbour.
5. Use simple majority of the category of nearest neighbours as the prediction value of the query instance.

### 3.5. K-Means Clustering

Clustering is the task of assigning a set of objects into groups (called clusters) so that the objects in the same cluster are more similar (in some sense or another) to each other than to those in other clusters. K-Means Clustering algorithm is the most popular and simplest algorithm of clustering. It groups the input data points into the specified number of clusters based on the features. The clusters are defined based on the facial expressions. The algorithm is composed of the following steps:

1. Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the K centroids.
4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of
the objects into groups from which the metric to be minimized can be calculated.

The distance between the test sample and training sample is calculated by Cosine Similarity measures[38]. It measures the similarity between two vectors depends on the cosine of the angle between them.

4. Experimental Results

This section deals with the performance comparison of the proposed approach. The performance of K-means clustering and the proposed Multi-class Linear Discriminant Analysis classifier is compared here.

4.1. Cohn-Kanade Database

The proposed algorithm has been tested on Cohn-Kanade database [39]. This database includes 388 image sequences from 100 subjects. Each sequence contains 12-16 frames. The subject’s age ranges from 18 to 30 years. Sixty-five percent of subjects were female and thirty-five percent were male. Fifteen percent of subject comes from the Africa-American background and three percentage from the Latin-American and Asian based subjects. The image sequences represented 100 different subjects expressing different stages of expression development, starting from a low arousal stage, reaching a peak of arousal and then declining. Facial expressions of each subject represent seven basic expression include anger, disgust, fear, happy, surprise, sad. The seven facial expressions from this database are shown in Figure 6.

4.2. Recognition Performance

The performance of the facial expression recognition system is evaluated based on accuracy. In this experiment 40 images per expression are selected randomly from Cohn-Kanade database for training and testing. The size of each image in this database is (640×490). Viola Jones algorithm successfully detects the face in an image sequence with good detection rate. Two types of extraction process are done to obtain better features: appearance and geometric based extraction. Extracted features are trained using Multi-class LDA and K-means clustering. Then, the performance of both classifiers is compared with their recognition rate.

Table 1 shows the Confusion matrices and accuracy using Multi-class LDA classifier. The confusion matrix is a n×n matrix containing the information about the actual class label (in its columns) and the label obtained through classification (in its rows). The diagonal entries of the confusion matrix are rates of facial expression that are correctly classified while the off-diagonal entries correspond to misclassification rates. Recognition rate level is reduced due to the corresponding expression makes confusion with other expression. Happy makes confusion with surprise and disgust, while sad makes confusion with disgust and fear. Neutral obtains high recognition rate and disgust has low recognition rate. Here the geometrical and appearance features are combined to train the classifier. Therefore, the overall accuracy obtained for seven expressions using Multi-class LDA is 87.5%.

Table 2 shows the Confusion matrices and accuracy using K-means clustering. Neutral expression has high recognition rate, while disgust obtains low recognition rate. The geometrical and local texture features are combined to train the K-Means classifier. Therefore, the overall accuracy obtained for seven expressions using K-means clustering is 82.5%. The experimental results show that Multi-class LDA is better than K-means clustering for recognizing the seven facial expressions.

The performance measure for feature descriptor is shown in Table 3 and Table 4. Using geometric based feature extraction, the accuracy obtained from Multi-Class LDA is 81% and from K-Means Clustering is 73%. Using LDP+LDPv feature extraction, the accuracy obtained from Multi-Class LDA is 74% and from K-Means Clustering is 62%.

Table 1. Confusion matrices and accuracy using Multi-class LDA

<table>
<thead>
<tr>
<th>Expression</th>
<th>Happy (%)</th>
<th>Surprise (%)</th>
<th>Sad (%)</th>
<th>Anger (%)</th>
<th>Disgust (%)</th>
<th>Fear (%)</th>
<th>Neutral (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>99</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Surprise</td>
<td>8</td>
<td>92</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Sad</td>
<td>0</td>
<td>0</td>
<td>83</td>
<td>6</td>
<td>12</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Anger</td>
<td>2</td>
<td>0</td>
<td>9</td>
<td>87</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Disgust</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>7</td>
<td>79</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Fear</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>88</td>
<td>0</td>
</tr>
<tr>
<td>Neutral</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>98</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 2. Confusion matrices and accuracu using K-Means Clustering

<table>
<thead>
<tr>
<th>Expression (%)</th>
<th>Happy (%)</th>
<th>Surprise (%)</th>
<th>Sad (%)</th>
<th>Anger (%)</th>
<th>Disgust (%)</th>
<th>Fear (%)</th>
<th>Neutral (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>88</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Surprise</td>
<td>6</td>
<td>93</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sad</td>
<td>0</td>
<td>0</td>
<td>84</td>
<td>2</td>
<td>9</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Anger</td>
<td>7</td>
<td>0</td>
<td>12</td>
<td>81</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Disgust</td>
<td>0</td>
<td>21</td>
<td>0</td>
<td>73</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fear</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>79</td>
<td>0</td>
</tr>
<tr>
<td>Neutral</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>96</td>
</tr>
</tbody>
</table>

Table 3. Performance measure for geometric based feature descriptor

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-class LDA</td>
<td>81</td>
</tr>
<tr>
<td>K-Means Clustering</td>
<td>73</td>
</tr>
</tbody>
</table>

Table 4. Performance measure for LDP+LDPv feature descriptor

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-class LDA</td>
<td>74</td>
</tr>
<tr>
<td>K-Means Clustering</td>
<td>62</td>
</tr>
</tbody>
</table>

5. Conclusion

The proposed system recognizes the seven facial expressions with high accuracy. The geometrical and appearance features are combined to train both the classifier. Multi-class Linear Discriminant Analysis achieves 87.5% of recognition rate, while K-Means Clustering achieves 82.5% of recognition rate. Therefore, the performance of facial expression recognition system is improved using Multi-class LDA which is experimented on Cohn-Kanade Database. Our approach fulfills requirement for a real time system as it employs automated procedure. This system is applicable for frontal faces and face with seven expressions. The recognition of facial expressions in image sequence with tilted faces is a challenging problem in present days. The head movement is required by many applications such as computer graphics animation, communication etc. Further, efficient techniques can be used to recognize the tilted faces with high accuracy and also to include more samples for testing.

5. References


