Received: November 29, 2016

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Geometrical Approach for Emotion Recognition from Facial Expressions Using 4D Videos and Analysis on Feature-Classifier Combination

Suja Palaniswamy[1]*   Shikha Tripathi[2]

[1]*Department of Computer Science and Engineering*
[2]*Department of Electronics and Communication Engineering, Amrita School of Engineering, Bengaluru, Amrita Vishwa Vidyapeetham, Amrita University, India - 560 035*
* Corresponding author's Email: p_suja@blr.amrita.edu

Abstract: Emotion recognition from facial expressions using videos is important in human computer communication where the continuous changes in face movements need to be recognized efficiently. In this paper, a method using the geometrical based approach for feature extraction and recognition of six basic emotions has been proposed which is named as GAFCI (Geometrical Approach for Feature Classifier Identification). Various classifiers, Support Vector Machine (SVM), Random Forest, Naïve Bayes and Neural Networks are used for classification, and the performances of all the chosen classifiers are compared. Out of the 83 feature points provided in the BU4DFE database, optimum feature points are identified by experimenting with several sets of feature points. Suitable "feature-classifier" combination has been obtained by varying the number of feature points, classifier parameters, and training and test samples. A detailed analysis on the feature points and classifiers has been performed to learn the relationship between distance parameters and classification of emotions. The results are compared with literature and found to be encouraging.

Keywords: Feature extraction, Feature points, Classifier, Emotions, SVM, Random forest, Naïve Bayes, Neural network;

1. Introduction

Emotion recognition has become an emerging area of research to provide support in applications like patient monitoring, driver fatigue detection, robotics, animation, forensics, medical aid, psychology, and surveillance, etc. Many researchers have presented methods for automatic recognition of emotions from videos. For recognizing emotions using BU4DFE, a method was proposed using Iterative Closest Point (ICP), Free Form Deformation (FFD), vector projections and Hidden Markov Model [1]. An emotional avatar image concept using FERA 2011 database has been developed in [2]. Wan using geometrical approach proposed a method that uses Euclidean distance, Principal Component Analysis and SVM [3]. A model for emotion recognition using facial Points Localization Model has been developed in [4]. Sandbach made a comprehensive survey on the developments of 3D and 4D facial expression recognition, and reviewed the tracking and alignment methods [5]. A novel phase congruency based descriptor for dynamic facial expression analysis which is robust to image scale and illumination variations was introduced in [6]. Appearance, geometric and curve based are the different approaches for feature extraction.

The recognition of six basic emotions namely, anger, happy, fear, disgust, sad and surprise in videos is challenging when pose and illumination vary, occlusion of objects, uncertainty in face motion, etc., Factors such as apex frame extraction, number and location of feature points, method to form feature vector and choice of classifier play vital roles in emotion recognition. A suitable "feature-classifier" combination improves the accuracy of emotion recognition. To address the feature-classifier combination and identification of optimum number of feature points, we proposed a dynamic method for apex frame extraction and geometric based approach for feature extraction in our earlier works [7][8]. Video sequences from BU4DFE database [9] had been used.

The apex frame and a suitable classifier are the key elements for emotion recognition. The accuracy obtained in [7] was 84.12%. To improve the accuracy, we determine optimum feature points starting from 39 [10] out of the 83 feature points provided in the BU4DFE database by experimenting with different sets of feature points 8, 11, 15, 17, 19, 20, 23, & 25. Feature points are added and removed on the face regions by identifying the significant feature points iteratively based on the distance measure and the accuracy obtained for each set of feature points. Support Vector Machine (SVM), Random Forest (RF), Naïve Bayes (NB) and Neural Network (NN) classifiers are used and the accuracy obtained by these classifiers for basic emotions are compared. For the given emotions and test samples we compare the accuracy of all the four classifiers and utilise the classifier that yields the best results. A user friendly Graphical User Interface (GUI) has been developed for this experimentation. The proposed method using geometrical approach for determining optimum feature points and "feature-classifier" identification is named as GAFCI (Geometrical Approach for Feature Classifier Identification).

GAFCI uses two frames for feature extraction and gives comparable results with a lucid methodology when compared to complex algorithms in literature that use multiple frames for feature extraction. The accuracy obtained using GAFCI is better than curve based approach [11] and mesh model method [12]. The major highlight of this work is in identifying a suitable "feature-classifier" combination. If a specific pool of feature points and set of classifiers are given, the GAFCI arrives at a suitable classifier and a set of optimum feature points respectively. Determination of optimum feature points, appropriate classifier, comparative analysis on the performance of the classifiers for different number of feature points, identification of suitable training and test samples and "feature-classifier" combination are the important highlights of this work. As per the knowledge of the authors such an analysis has not been carried out earlier.

The remainder of this paper is organized as: Section 2 gives a detailed view of the proposed system GAFCI, Section 3 discusses the results, Section 4 describes the identification, analysis of classifiers, Section 5 describes "feature-classifier" identification and Section 6 concludes our work.

2. Proposed system (GAFCI)

The procedure for emotion recognition is shown in Figure.1. The major steps in recognizing emotions in videos involve apex frame extraction (detecting the frame where the peak of an emotion is expressed) [7] followed by feature extraction and classification. The number of frames varies from 65 to 100 for a 4 sec video. A video expressing emotion will have frames containing neutral, onset, apex and offset of that emotion. For the chosen feature points in a single frame, distance parameters are determined between them. This is performed for apex and neutral frames. The difference between these distance parameters is calculated for apex and neutral frames. The distances form the feature vectors and the feature vectors for all the emotions and subjects are given to classifier for classification.

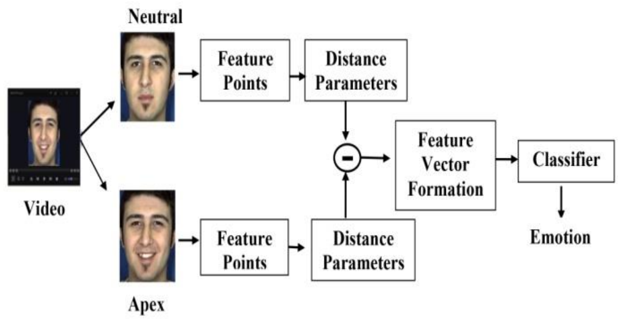

Figure.1 Procedure for emotion recognition

2.1 Feature extraction

The apex frame extraction is followed by feature extraction. In geometric based feature extraction method, feature points are marked on the face which describes the geometric information about features of the face like eyebrow, eye, nose and mouth. BU4DFE database has 83 feature points marked on face for every frame. 39 feature points are selected out of 83, for both neutral and apex frames and is shown in Figure.2. When facial expression changes, the vertical displacement is determined by the movement of eye, eyebrow and mouth; horizontal displacement is determined for corner points in mouth. The horizontal and vertical distances between the feature points l_i , l_j in one frame are represented as $(l_i,l_j)_X$ and $(l_i,l_j)_Y$ respectively. Horizontal and vertical distances are calculated as

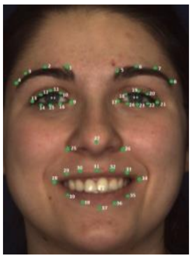

Figure.2 A frame representing 39 feature points

$k_z = \left|(l_i - l_j)\right|$, z =1 to 39, (i, j)∈ ⟦1,39⟧ (1)

To calculate the distance in GAFCI for feature extraction, either '*x*' or '*y*' coordinate of the feature points is considered based on horizontal or vertical distance respectively. For example, k_{26} refers to horizontal distance calculated by using the '*x*' coordinates of feature points 34 & 28, and the remaining vertical distances are calculated by taking the '*y*' coordinates of the feature points. Set of sample horizontal and vertical distances are shown in Table 1. After calculating the 39 distance parameters for neutral and apex frames individually, the difference between the corresponding distances in neutral and apex frame are calculated by subtracting the distance parameters. They are then concatenated to form a feature vector which is given as $[dk_1, dk_2, ..., dk_{39}]$ where *dk* is the difference between the distances for a feature point in neutral and apex frame of an emotion. Only two frames have been used for the formation of feature vector instead of the complete video sequence. This process is repeated for 60 subjects and 6 emotions. The feature vector W_e^s thus formed for 39 feature points is given in Equation (2). So, the dimension of feature vector is 606 x 39 for 39 feature points. The feature vector is fed to classifiers for classifying into emotions. Let 's' be the number of subjects and 'e' be the number of emotions.

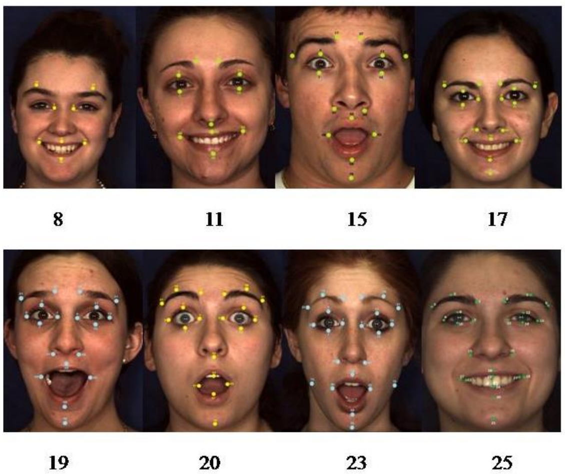

Figure.3 Different sets of feature points

Table 1. Sample distance parameters

k_i	Distance	Feature	k_i	Distance	Feature
k_1	$(l_1, l_2)_Y$	Left Eyebrow	k_{36}	$(l_{27}, l_{28})_Y$	Chin
k_4	$(l_5, l_6)_Y$	Right Eyebrow	k_{24}	$(l_{27}, l_{26})_Y$	Nose
k_7	$(l_9, l_{13})_Y$	Left Eye	k_{26}	$(l_{28}, l_{34})_X$	Mouth
k_{15}	$(l_{17}, l_{21})_Y$	Right Eye	k_{27}	$(l_{31}, l_{37})_Y$	Mouth

$W_e^s = [dk_1, dk_2, \ldots, dk_{39}]^T$ (2)

$e = 1, 2, \ldots, 6,\ \ s = 1, 2, \ldots, 60$

We now proceed to extract different sets of feature points 8, 11, 15, 17, 19, 20, 23 and 25 that are marked on a frame as shown in Figure.3. These set of feature points are iteratively extracted. The procedure includes feature extraction, feature vector formation, classification by 4 classifiers and analysis of results. To the 39 feature points, few feature points are added in mouth, chin and lower cheek region and few feature points are eliminated to arrive at 25 feature points as shown in Figure.4. For 25 feature points set, the horizontal and vertical distances for outer and inner corner points on mouth are represented as k_1, k_2, k_3 and k_4 respectively. k_5 and k_6 denote two vertical distances from upper and lower mouth to chin respectively. The vertical distances from eyes to eyebrows are denoted as k_7 and k_8. Vertical distances from nose corners to lower mouth centre point are represented as k_9 and k_{10}. Determination of distance parameters for the aforesaid feature points are shown in Figure.5. The distances between the feature points in neutral and apex frames are determined using Equation (1) by varying the indices in the variables depending on the number of feature points chosen. Similar to feature points 39, after calculating the distance parameters for neutral and apex frames individually, the difference between the corresponding distance parameters in neutral and apex frame are calculated by subtracting them. They are then concatenated to form a feature vector by using Equation (2) by varying the indices of *dk*. This procedure is repeated for 60 subjects and 6 emotions. The feature extraction is related to accuracy and is elaborated in

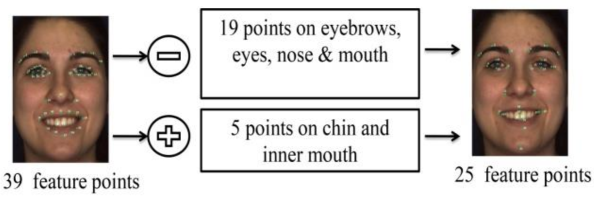

Figure.4 Significant feature points extraction

Received: November 29, 2016

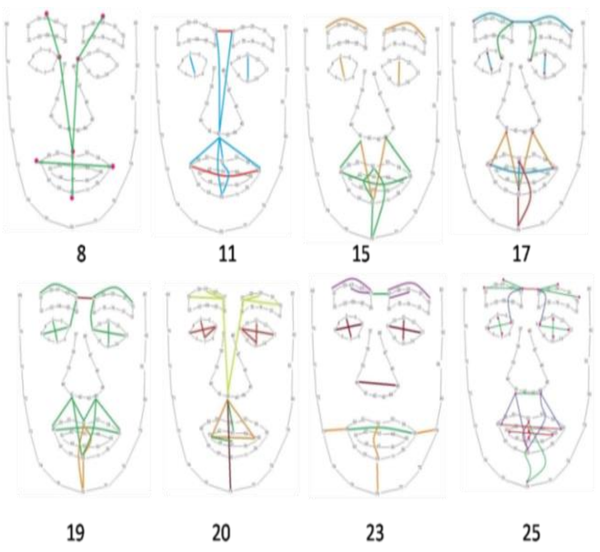

Figure.5 Diagrammatic representations for determining distance parameters

Section 4.1 for remaining sets of feature points. The feature vectors are given to classifier for classifying into 6 basic emotions which is discussed in the next subsection.

2.2 Classification

Four classifiers viz. SVM [13], NN [14], RF [15] and NB [16] are used in this work. The feature vectors formed are divided into training and test samples and they are fed to each of the classifiers separately for training and testing. The classifier initially gets trained to the feature vector and then classification is performed for the test samples. The emotion closest to the threshold distance of the target emotion is identified as the expected emotion

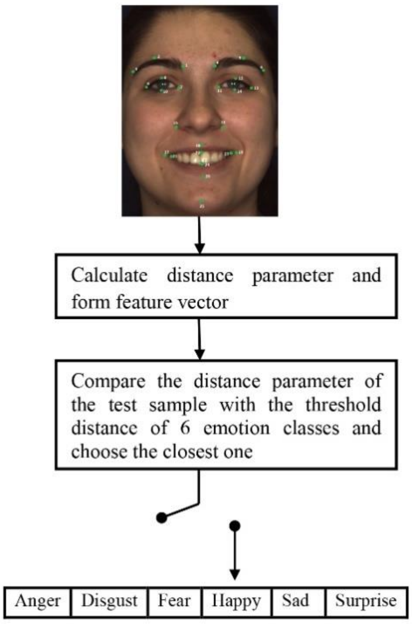

Figure.6 Procedure for classification

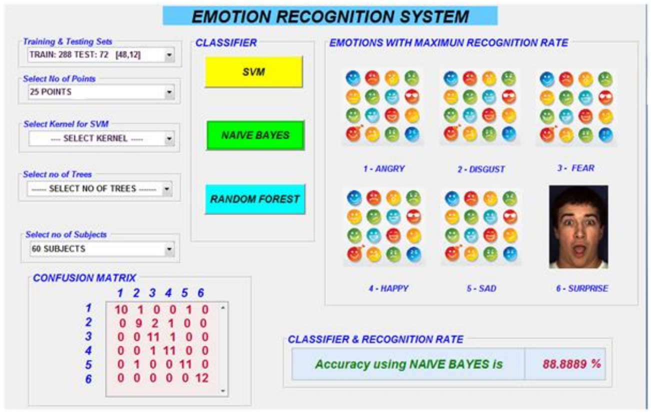

Figure.7(a) GUI displaying output for Naïve Bayes

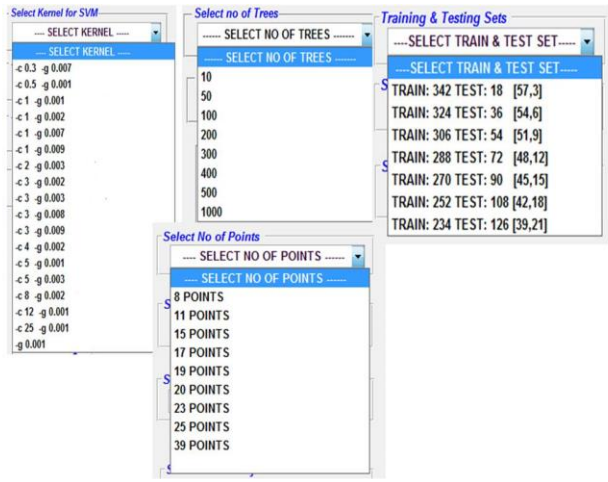

Figure.7 (b) Options for classifier parameters

which is pictorially shown in Figure. 6. A detailed analysis is carried out on the performance of classifiers by varying number of training and test samples. Optimum feature points are identified from the set of feature points chosen. A suitable kernel for SVM, and number of trees for RF are identified.

A Graphical User Interface (GUI) is developed to choose different sets of feature points, classifier, kernels, number of trees, number of training and testing set etc., and to determine the accuracy with user friendly options. The GUI screen and the options are shown in Figures 7(a), & 7(b). Classification using SVM, NB and RF are invoked using GUI. Neural network is executed directly using Neural Pattern Recognition (NPR) tool in Matlab. After performing training and classification, result for the given input is displayed in the GUI.

The average accuracy for all the samples and the confusion matrix are displayed in GUI. In addition, one image for an emotion that resulted in maximum accuracy is also displayed. The numbers 1-6 in confusion matrix in rows and columns as headers represents six basic emotions in order as shown in the GUI and it represents the target and output class.

Received: November 29, 2016

For SVM, RF & NB classifiers, selecting training and test data sets, number of subjects, and number of feature points are common. For SVM, kernel is chosen and in RF, number of trees is chosen. The results obtained for all the options of testing by choosing feature points, classifiers, and its parameters are discussed in detail in Section 3.

3. Results

This section gives a detailed insight into the performance of the feature points and classifiers chosen. For the four classifiers, different sets of training and test samples, and the sets of feature points chosen, the emotion-wise classification accuracy is determined. Standard Deviation (SD) is used to define the consistency of a classifier. A classifier is said to be consistent if its SD is the lowest among the four classifiers used.

The behaviour of each classifier with respect to their input features and the accuracy for different emotions are discussed in detail in the subsections. Matlab R2014b with Intel i7 and Windows 7 is used for experimentation. The experiment is repeated 10 times and accuracy and time taken are obtained by averaging over all trials. We require the classifier to recognize emotions within 5 sec for suitable real time applications.

3.1 Support Vector Machine

We tested with kernels like linear, polynomial, Sigmoid and Radial Basis Function and determined that Sigmoid kernel yields the best accuracy. The Sigmoid kernel is given as

$$k(\boldsymbol{x_i}, \boldsymbol{x_j}) = tanh(\gamma \boldsymbol{x_i}^T . \boldsymbol{x_j} + c) \quad (3)$$

The $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are different training vectors, 'c' is a cost parameter and 'γ' is the margin band. The performance for different sets of feature points and training and test data sets is shown in Figure.8. It is observed that for 17, 19, 20, 23, and 39 feature points the accuracy increases starting from (57,3) reaches a maximum for (48,12) and again reduces

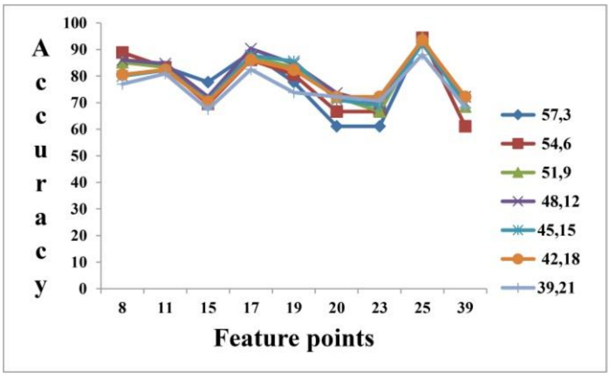

Figure.8 Performance of SVM

for (39,21). For remaining feature points, the maximum and minimum accuracy is obtained for (57,3) and (39,21) respectively. The parameters 'c' and 'γ' in Equation (3) are varied and their respective best values for different feature points are determined by trial and error. For 25 feature points, the average accuracy is maximum for 'c' = 1 and ' γ ' = 0.001. The average time taken for classification is 4 sec. The performance of SVM is further explained in detail in Section 4.

3.2 Random forest

The average accuracy obtained for various training and test samples, 500 trees and different feature points is plotted in Figure. 9. Training and test samples (48,12) yields the maximum average accuracy for all the feature points. 17 feature points result in the highest accuracy 92.36%.

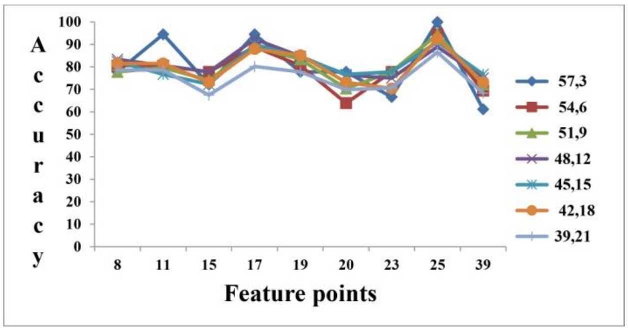

Figure.9 Performance of RF classifier

The influence of different number of trees (10, 50, 100, 200, 300, 400, 500 and 1000) on accuracy is analysed. The emotion-wise accuracies for 17 feature points for (48,12) for 4 values of trees are shown in Table 2. The emotion 'anger' is recognized at 100% accuracy for all the values of trees considered. Apart from 'happy', remaining emotions are well recognized.

Table 2. Emotion-wise accuracy & time for different trees of random forest classifier for 17 feature points

	50	200	500	1000
Anger	100.00	100.00	100.00	100.00
Disgust	90.00	90.00	90.83	96.66
Fear	93.33	96.66	97.50	98.33
Happy	69.16	68.33	68.33	68.33
Sad	88.33	93.33	97.50	94.16
Surprise	95.00	99.16	100.00	100.00
Average accuracy	89.30	91.24	92.36	92.91
Average time(sec)	2.70	3.96	4.80	7.64

The performance analysis of RF for emotion 'happy' is discussed in Section 4.2. As the number of trees increases, accuracy improves but time taken increases. A trade-off between accuracy and time taken needs to be considered when choosing the number of trees. The percentage of increase in accuracy from 500 to 1000 is just 0.55 but there is an increase of 1.5 times in classification time. Therefore, 500 is chosen as the best value of tree as it gives the highest accuracy in 5 sec.

3.3 Neural network

Figure.10 represents the performance of NN for various feature points for (48, 12) training and test samples. It shows that 25 feature points has maximum accuracy as 84.7%. Compared to other classifiers NN perform poorly. Increasing number of subjects and samples will improve the accuracy as the performance of NN depends on the volume of samples used for training. The classification time also increases which is not preferable for the applications where time is a major criterion.

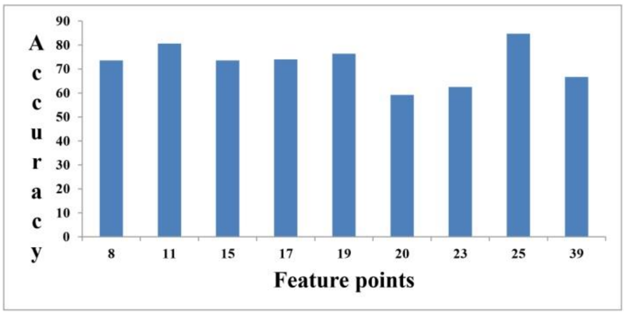

Figure.10 Performance of NN classifier

3.4 Naïve Bayes

For NB classifier, the accuracy obtained for different feature points and various training and test samples is shown in Figure.11. Naïve Bayes achieves maximum accuracy 88.89% for both 25 and 17 feature points for (48,12) samples. The average time taken for Naïve Bayes is 5 sec.

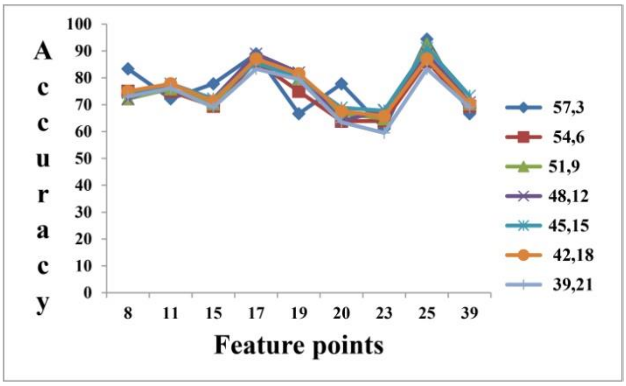

Figure.11 Performance of NB classifier

4. Identification and analysis of classifiers

This section describes the extraction of different sets of feature points, identifies optimum feature points, training and test set, and suitable classifier. The performance evaluation of the two best feature points set and classifiers are discussed in detail.

4.1 Identification of efficient training & test set, feature point & classifier

Here, we identify efficient training and test combination from a list of training and test samples experimented. In addition, a suitable classifier and feature point set are also determined in this section. Analysis is made on the number of samples used for training and testing for all the classifiers and for all the feature points which is shown in Figure.12. The average accuracy obtained is maximum for (48,12) that includes 80% training and 20% testing samples. In literature, this combination is widely used and our test results coincide with the literature. So, (48,12) is chosen as the suitable training and test data set.

The accuracy is related to the location of feature points and the procedure for feature points extraction is described in this section. Initially, experiment was performed for 39 feature points and later 20, 19, 11, & 8 feature points are extracted iteratively based on the references [3], [17], [13], & [4] respectively. It is observed that in each iteration the accuracy either increases or decreases for few of the classifiers instead of uniform increase. So, experiment is performed to increase the number of feature points. For choosing feature points 15, 17, 23 and 25, the face is divided into regions namely left eyebrow, right eyebrow, left eye, right eye, nose, and mouth region which includes chin and lower part of cheek. Iteratively feature points are added and removed on face regions. To make the features of "lower" part of the face more prominent, a feature point is added on chin [17].

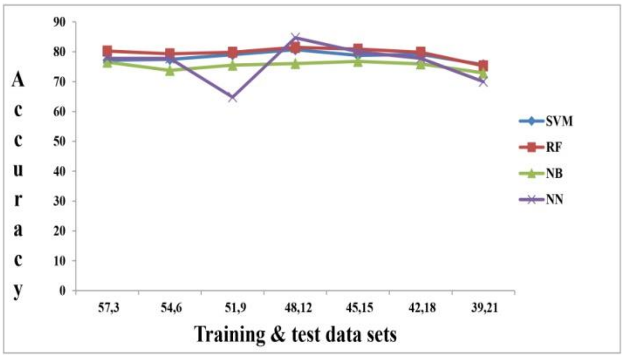

Figure.12 Performance of four classifiers

Received: November 29, 2016

Table 3. Average accuracy for the 4 classifiers and all feature points with (48,12) training and test samples, number of trees in RF is 500

Feature Points	SVM	RF	NB	NN
8	86.11	81.94	72.22	73.60
11	84.72	77.78	80.56	80.60
15	72.22	77.78	72.22	73.60
17	90.28	92.36	88.89	74.00
19	84.72	84.72	81.94	76.40
20	73.61	76.39	63.89	59.70
23	69.44	75.00	68.06	62.50
25	93.05	88.89	88.89	84.70
39	72.22	75.00	70.83	66.70

The "upper" part of the face comprises of eyes and eyebrows and when significant feature points are extracted on those regions, the accuracy for 15 feature points did not improve for NB and NN, instead it reduced for SVM and RF. Similarly, when two feature points are added on inner corner of eyes, the 17 feature points set improved accuracy remarkably for three of the four classifiers. Additionally, when two feature points are added to cheeks, the accuracy drastically reduced for 23 feature points set. So, to improve the accuracy two feature points are added in the inner region of mouth and the 25 feature points performed better than 23 for all the classifiers. In this way the feature extraction is performed. The average accuracies of 4 classifiers and different sets of feature points are compared using Table 3.

From Table 3 it can be observed that SVM, NB and NN achieve maximum accuracy for 25 feature points and RF for 17 feature points. Naïve Bayes gives the same maximum accuracy for 17 and 25 feature points. 25 feature points is identified as the optimum feature points as it gives the best average accuracy among three classifiers. Next, appropriate classifier needs to be identified.

Table 4. Average accuracy & SD for 25 feature points for (48,12) samples

	SVM	RF	NB	NN
Anger	83.33	81.66	83.33	91.70
Disgust	91.66	95.83	75.00	75.00
Fear	91.66	83.33	91.66	66.70
Happy	100.00	97.50	91.66	100.00
Sad	91.66	75.83	91.66	75.00
Surprise	100.00	100.00	100.00	100.00
Average accuracy	93.05	89.03	88.89	84.73
SD (σ)	6.28	9.99	8.60	14.35

For 25 feature points, (48,12) samples emotion-wise accuracy for all four classifiers is shown in Table 4. Number of trees for RF is 200, Sigmoid kernel for SVM and 20 hidden neurons for NN are considered in Table 4. In NB, apart from the low accuracy for emotion 'disgust' it is comparable to SVM. Neural network performs poorly for all emotions except 'anger' compared to other classifiers. SVM achieves highest accuracy and RF performs similar to SVM. 'Surprise' is recognized at 100% accuracy by all the classifiers and 'happy' is recognized at 100% by SVM and NN. It is observed that SVM is performing consistently well as its SD is less compared to other classifiers.

4.2 Analysis of performance of classifiers

Different sets of feature points are extracted from 39 as described in Section 4.1. The locations of feature points contribute to the performance of the classifiers. The question is whether GAFCI has improved the performance of classification or not. GAFCI through which the appropriate feature points are derived has an effect on accuracy. When a person expresses an emotion, the muscles stretch and the movements make the feature points change their locations in vertical and horizontal directions. For example, when eyebrows and mouth move, the feature points on nose, mouth and eyes will change their location. The distances calculated in the emotion image is subtracted from neutral and the difference is a measure of how much each feature point has moved from neutral i.e. actual movement of the feature point. When eyes and mouth are open, the distances k_1 to k_4, k_7, and k_8 contribute to the accurate recognition of emotion 'surprise'. Emotions, 'happy' and 'disgust' are well recognized by widening of the mouth. The movement in the lower part of the face is contributed by k_5 and k_6. For few distance measures the average distances for 10 samples are shown in Table 5. The distance measures for 'neutral' is also shown as it is used to estimate the movement of the feature points. The distance k_4 is maximum for 'surprise'.

Table 5. Significant distance measures for basic emotions

	k_3	k_4	k_5	k_6
Anger	34.94	2.02	49.39	26.55
Disgust	33.52	9.98	46.74	19.31
Fear	37.97	7.09	46.58	27.15
Happy	51.73	9.54	51.13	30.54
Sad	37.12	1.73	43.47	28.18
Surprise	35.60	14.50	53.76	29.21
Neutral	39.37	1.02	41.65	27.21

25 and 17 feature points result in the highest accuracy for all the classifiers and their accuracies for SVM & RF are shown in Table 6. It is evident from Table 6 for 25 feature points and SVM, emotions 'happy' and 'surprise' result in 100% which is contributed by the distance k_4.

Similarly, for 17 & 25 feature points for SVM, emotions 'sad' and 'fear' are recognized well. For 25 and 17 feature points, 'anger' and 'happy' results in the lowest accuracy respectively. For both 'anger' and 'happy' out of 12 test samples, one sample is misclassified into 'disgust' while another into 'sad'. Here, the misclassification happens as the vertical and horizontal distances for those samples fall into the bins of 'disgust' and 'sad'. Whenever a subject expresses blend of emotions, the classifier tries to classify the emotion into the bin which is very close to the threshold as shown in Figure. 6.

In SVM, out of 288 samples, 238 samples are support vectors and the maximum distances for each of them are determined. Few of those distances are k_1, k_2, k_4, k_5 and k_6. The distances k_1 and k_2 in mouth are included for all the feature points set, k_5 and k_6 are determined for 15, 17, 23, & 25. Though all these distances contribute to the accuracy, it has been identified that the vertical distance k_4 of inner mouth feature points which is determined only for 25 feature points contributes significantly to the 93.05% accuracy. Emotion 'happy' and 'surprise' achieve 100% accuracy because of this distance measure. It is concluded that the significant feature points on the face regions extracted by GAFCI method are sufficient to yield good accuracy. From the above discussion, we choose 25 feature points and SVM as the best combination.

Table 6. Comparison of emotion-wise accuracy for 17 & 25 feature points for SVM & RF (500 trees) classifiers

	17		25	
	SVM	RF	SVM	RF
Anger	100.00	100.00	83.33	81.66
Disgust	83.33	90.83	91.66	95.83
Fear	91.66	97.50	91.66	83.33
Happy	66.66	68.33	100.00	97.50
Sad	100.00	97.50	91.66	75.83
Surprise	100.00	100.00	100.00	100.00
Average	90.28	92.36	93.05	89.03

4.3 Comparison of results with literature

The results obtained for GAFCI are compared with literature and is shown in Table 7. The features considered for comparison are the database, algorithm, number of feature points, and accuracy. The number of feature points is a very important factor for comparison as the accuracy depends on the distance parameters in feature vector. The procedure for emotion recognition employs selection of feature points, distance measure and difference between them for neutral and apex frame. There is a quantum leap in accuracy when using 25 feature points in comparison to 26 feature points [12]. The results of GAFCI when compared with [18] shows that 25 feature points are sufficient instead of 83 to achieve a similar accuracy. This shows that the accuracy of GAFCI is higher than the results reported in literature.

Table 7. Comparison of results of GAFCI with literature

Ref	Database	Algorithm, classifier	No of points	Accuracy (%)
[4]	BU4DFE	PDM,SVM	8	83.89
[12]	RML & DVD	DM model & Isomap	26	88.20
[18]	BU4DFE	DVF, RF	83	93.00
GAFCI	**BU4DFE**	**Distance parameter, SVM**	**8**	**86.11**
			25	**93.05**

5. Feature-Classifier identification

We define optimum feature points and classifier as those that gives maximum accuracy. If only a specific pool of feature points and classifiers are given, arriving at suitable classifier and set of optimum features points respectively can be very useful. Based on the results presented in the previous section, we determine the optimum number of feature points as well as the optimum classifier in this work. In addition, optimum classifier-emotion combination is also identified.

5.1 Optimum feature point identification

Given a set of test samples and classifier, the GAFCI identifies the optimum feature points as shown in Figure.13.

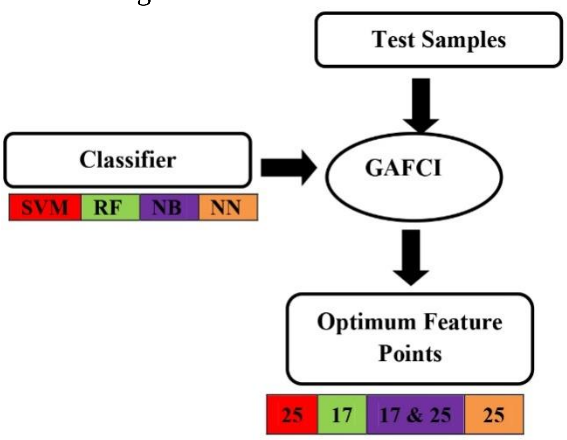

Figure.13 Optimum feature points identification

The optimum number of feature points is identified from Table 3 by selecting the corresponding feature points for the maximum accuracy. The "classifier-optimum feature points" combination is represented in matching colours. From Figure.13, it can be inferred that when test samples and RF classifier is given, 17 is the optimum number of feature points. When 17 and 25 feature points are compared, the former performs well for RF and NB while the later does so for SVM, NB, and NN. For majority of classifiers the choice of feature points is 25.

5.2 Optimum classifier identification

Given a set of test samples and feature points, GAFCI determines optimum classifier as shown in Figure.14. The optimum classifier is identified from Table 3 by choosing the classifier that yields the highest accuracy for each set of feature points (maximum accuracy in a row). When the performance of RF and SVM is compared, for six different sets of feature points RF is the optimum classifier. However, if accuracy is a major criterion then SVM is the best classifier. When test samples and 19 feature points are given as input, the GAFCI identifies SVM and RF as the optimum classifiers. Determination of the best classifier in this situation is discussed in Section 5.3.

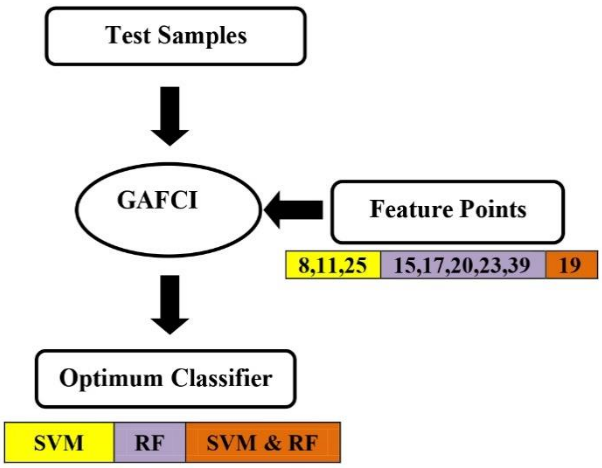

Figure.14 Optimum classifier identification

5.3 Optimum classifier-emotion identification

The next important contribution is appropriate classifier identification for an emotion. Suppose for a given feature point set, test samples are tested by all the four classifiers and if the classifiers' accuracies are different then which classifier's results should be considered? This is addressed using the output displayed in the GUI. When the classifier is tested for test samples and set of feature points, the GUI displays confusion matrix (Figure. 7a) and also the image(s) of the emotion which has the highest accuracy. For NN, the NPR tool displays a confusion matrix. The confusion matrices given by the classifiers are compared to find the maximum recognition for a particular emotion and the emotion identified by that classifier is considered as the emotion for the given test samples. SVM and RF are identified as the optimum classifier in Figure.14 and from Table 4 it can be derived that for emotion 'anger' NN gives the best accuracy. Naive Bayes achieves maximum accuracy for the emotions 'fear', 'sad' and 'surprise' on par with SVM.

6. Conclusion

We have performed a comparative analysis on the classifiers SVM, RF, NB, and NN to identify the classifier that yields the best results. The location and number of feature points are very essential and plays a significant role in formation of feature vector that contribute to the recognition of emotions. We have experimented on this aspect by choosing significant feature points out of 83 feature points given in the BU4DFE database and discussed the role of horizontal and vertical distances between the feature points. The set of 25 feature points perform well for three of the classifiers used and it is considered to be the optimum. SVM achieves the maximum accuracy of 93.05% for 25 feature points and chosen as the appropriate classifier. The suitable parameters for SVM and number of trees for RF are determined. We also identified RF as one of the appropriate classifiers for a few feature points.

An exclusive study has been made on the number of training and test samples for four different classifiers and it has been identified that (48,12) training and test samples perform well for 60 subjects. Analysis on misclassification of emotion is performed. A significant contribution is optimum feature-classifier identification, where a set of efficient feature points and classifier are identified. In addition, an optimum classifier for each emotion has been determined. Implementing emotion recognition using GAFCI with pose and illumination variations will be considered as future work.

Acknowledgement

The authors wish to express their sincere gratitude to Mr. Kalyan Kumar and Dr. Karthik R.M. for their support in this work.

References

[1] G. Sandbach, S. Zafeiriou, M. Pantic, D. Rueckert, "Recognition of 3D facial expression dynamics", *Image and Vision Computing*, Vol. 30, pp.762-773, October 2012.

[2] S. Yang, B. Bhanu, "Understanding Discrete Facial Expressions in Video Using an Emotion Avatar Image", *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, Vol.42, No.4, pp.980-992, August 2012.

[3] C. Wan, Y. Tian, S. Liu, "Facial Expression Recognition in Video Sequences", In: *Proc. of 10th World Congress on Intelligent Control and Automation*, Beijing, China, pp.4766-4770, 6-8 July 2012.

[4] A.Saeed, A.Al-Hamadi, R.Niese,"The Effectiveness of using Geometrical Features for Facial Expression Recognition", In: *Proc. of 2013 IEEE International Conf. on Cybernetics (CYBCONF)*, Lausanne, Switzerland, pp.122-127, 13-15 June 2013.

[5] G. Sandbach, S. Zafeiriou, M. Pantic, L. Yin, "Static and dynamic 3D facial expression recognition: A comprehensive survey", *Image and Vision Computing*, Vol. 30, No.10, pp.683-697, October 2012.

[6] S. Shojailangari, W.-Y. Yau, E.-K Teoh,., "A novel phase congruency based descriptor for dynamic facial expression analysis", *Pattern Recognition Letters*, Vol. 49, pp. 55-61, 2014.

[7] S. Palaniswamy, V.P. Kalyan Kumar, S. Tripathi, "Dynamic Facial Emotion Recognition from 4D Video Sequences", In: *Proc. of 8th Int. Conf. on Contemporary Computing*, IC3, Noida, India, pp 348-353, August 2015.

[8] V.P. Kalyan Kumar, S. Palaniswamy, S. Tripathi, "Emotion Recognition from Facial Expressions for 4D Videos Using Geometric Approach", In: *Proc. of Advances in Signal Processing and Intelligent Recognition Systems*", Springer, Tvm., India, pp 63-73, December 2015.

[9] L. Yin, X. Chen, Y. Sun, T. Worm, M. Reale, "A High-Resolution 3D Dynamic Facial Expression Database", In: *Proc. of the 8th International conference on Automatic Face and Gesture Recognition*, Amsterdam, Netherlands, 17-19 September 2008.

[10] O. Rudovic, M. Pantic, and I. Patras, "Coupled Gaussian Processes for Pose-Invariant Facial Expression Recognition", *IEEE Transactions on Pattern Analysis an Machine Intelligence*, Vol. 35, No.6, pp. 1357-1369, 2013.

[11] N. Bajaj, S.L. Happy, "Dynamic Model of Facial Expression Recognition based on Eigen-face Approach", In: *Proc of Green Energy and Systems Conference 2013*, Long Beach, CA, USA, November 2013.

[12] Y. Tie, L. Guan, "Human Emotion Recognition Using a Deformable 3D Facial Expression Model", In: *Proc. of IEEE International Symposium on Circuits and Systems (ISCAS)*, Seoul, Korea (South), pp 1115-1118, 20-23 May 2012.

[13] P. Sohail, Bhattacharya, "Classifying facial expressions using point-based analytic face model and Support Vector Machines", In: *Proc. of IEEE International Conference on Systems, Man and Cybernetics, ISIC*, Montreal, Que, Canada, pp.1008-1013, October 2007.

[14] G.P. Zhang, "Neural Networks for Classification : A Survey", *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and reviews*, Vol.30, No.4, pp.451-462, November 2000.

[15] L. Breiman, "Random forests", *Machine. Learning*, Vol. 45, No.1, pp. 5–32, 2001.

[16] N. Sebe, M.S. Lew, I. Cohen, A. Garg, "Emotion recognition using a Cauchy Naive Bayes classifier", In: *Proc. of the 6th International conf. on Pattern Recognition*, 1, Quebec, Canada, pp. 17-20, August 2002.

[17] M. Pantic and L.J.M. Rothkrantz, "Facial Action Recognition for Facial Expression Analysis From Static Face Images"**,** *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, Vol. 34, No.3, pp.1449-1461, 2004.

[18] H. Drira, B. B. Amor, M. Daoudi, A. Srivastava "3D Dynamic expression Recognition based on a Novel Deformation Vector Field and Random Forest", In: *Proc of 21st International Conf. on Pattern Recognition (ICPR)*, Tsukuba, Japan, pp. 1104 – 1107, 11-15 November 2012.