International Journal of Modern Engineering Research (IJMER)
www.ijmer.com Vol.2, Issue.3, May-June 2012 pp-629-634 ISSN: 2249-6645

A Empirical Framework for Facial Expression Analysis

[1]Geeranjali Sharma,[2] H K Sawant
[1, 2] *Department Of Information Technology, Bharati Vidyapeeth Deemed University College Of Engineering, Pune-46*

ABSTRACT

For facial expression analysis different-different approaches have been implemented. First thing preprocessing step is done and after that face detection algorithm applied for face detection. After face detection facial feature tracking or feature is extracted on the behalf of obtained feature value. Different-different classifier likes ANN, HMM, SVM is used to train the feature value obtained and then after testing is done to classify the classes to whom this feature value belongs. If high dimension feature value is obtained than PCA does work to reduce the dimension. Our approach is that apply Gabor filter on face to extract the feature. After feature extraction SVM is used to training and testing the feature value and finally similarity measure is evaluated to classify the classes to whom it belongs means it be from happy, sad, disgust, surprise, angry class. As many algorithm has been used for showing the Facial Expression of a Human. They have used different-different feature extraction method. Consideration of some similarity function has taken. Sometimes PCA, LDA was considered for dimensions reduction and for better accuracy.

Keywords* - *Face detection, Facial Expression analysis segmented facial region Gabor filter, Support Vector Machine.

I. INTRODUCTION

Most research in facial expression recognition is limited to six basic expression and several combination.The expression are classified into emotion categories rather than another technique..It is difficult task to show all facial expressions because in everyday life six basic expression occur so frequently.Emotion is often communicated by small changes in one or two facial features,on the other hand the same facial expression may occured in more than one emotions.The presence or absence of one or more facial actions value may change its prediction.One man can show their facial expression in different manner than expressing the same facial expression by other person.The facial features value changes person to person for the same facial expression.The main problem in facial expression analysis is that no exact dataset is available that show 100% accuracy.The problem comes when the facial feature extraction is processed,sometimes frontal image is present and sometimes tilt in different orientation.The light condition also plays a different role so all these factors affect the preprocessing steps and the desired results are not obtained .If the main facial feature like tip of nose,eye,head,cheek is not properly extracted than no proper result comes out.During the extraction of these features value is very large so dimenson reduction technique like PCA,LDA is adopted to minimize this Problem.Different methodology is implemented to resolve the problem of emotions detection but yet not such good result has obtained because if we compare the same facial expression of two person the similarity measure threshold value does not match exactly.They show different value for the same emotions.So measure problem is in feature extraction and accurate face database.

Initially there were two methods to finding the facial expression activity.

a) Human observer based coding system.
b) EMG (electromyography)based system

Human observer based are basically based on manually done so it is time consuming and not fulfill the practical standards and EMG based requires the deployment of sensors on the human face which really restricts the natural observation of the facial expression analysis.
So for overcome this problem computer vision based recognition system is adopted. There are two different approach used in computer vision…

a) Facial expression recognition using 2D still image.
b) Facial expression recognition using image sequence.

Facial expression recognition using image sequences often use optical flow analysis to the image sequence and then apply pattern recognition tools on the observing value to show particular facial expression analysis. This method has need lots of information regarding multiple frames of images to recognize the expressions and thus has limitation in real time performance and showing more accuracy with more robustness. Still images based recognition basically based on feature extraction that can play an important parameter to define some thresholding and on the basis of that the recognition rate and accuracy is evaluated and that is acceptable in real time application.

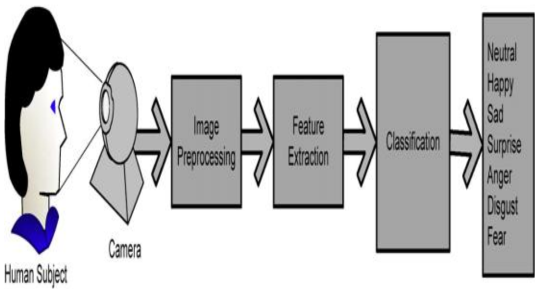

Fig1: Facial expression recognition using image

II. LITERATURE REVIEW

In this work the facial expression is based on the attributes of facial muscle that is hidden state of a HMM for individual image. Probability of the state is changed on the behalf of the feature vector obtained from image processing.[2]

Optical flow algorithm is used for the evaluation of velocity vector of two successive frames. After that FFT is applied to a velocity vector around the region of mouth and eye. The selection of feature vector is lower frequencies value and a mixture density is applied to scaled the output probability of HMM to detect the variation in human facial expression. Mixture density is very useful to enhance the accuracy as the mixture increases. Facial expression recognition is done by using the Support Vector Machine by modifying Kernels.[1] Facial expression is also recognized with the help of brain activity that is governed by the EEG signal.[3]

In this there is a proposed system that analyzes EEG system and classify them into 5 classes on the two emotional dimensions name like valance and arousal. However after using the 3-fold cross validation method the recognition rate for the valence dimension is 32% and for the 37% for the arousal and overall rate is approx 71%.

Facial expression recognition is also detected with the help Markov random Field.[4] In this the main concept is the essential factor that helpful for the expression detection is Eye and mouth expression. In this the first step is done image segmentation and skin detection for the Markov random field .second step is done for the eye and mouth feature extraction. The set of different color image is used as a training set.HLV color space that is responsible for the detection of the eyes and mouth region.3rd step module is in accordance with the detection of emotions in images with the help of edge detection and measurement of gradient of eye and mouth region.

There is also a facial expression detection is done by the using of coded form that consists of multiorientation ,multiresolution of the Gabor filters in which region of surface space and precise description of the place or region are mentions and that is aligned along the human face.[12] In this paper the similarity space is matched with the result obtained from the applying the Gabor filter and the result that is semantic value that obtained on the human observations. There is concepts that terminology is known as rank correlation which emphasize the semantic similarity and facial expression image similarity that is obtained after the applying the Gabor coding. The facial expression classification is also achieved of the frontal image by using of Eigen face [21].

In this paper whole face is not taken in consideration rather than classify the face in the regions that is beneficial for the facial expression classification and projecting that faces with the Eigen face and try to train with different types of facial expressions.Than taking the average of all different region faces showing different facial expression and after that making a mask. The important thing is that this masks fading the miss region and try to highlight the region that are changing during the different facial expression. For the recognition of the facial expression The new technique active appearance model (AAM) is used to trained the faces that is available in database that is used to represent the shape and texture variation that plays an important role in facial expression recognition. The features are those that are extracted from the parameter obtained from the AMM and is used to discriminate among the classification of different expression. The feature extraction with the help of AMM better than a simple classifier like Euclidean distance. The AMM makes a efficient method for the texture and shape to model such that It plays an important role and it is thoughtful like (SVM) support vector machine.

Facial expression has also detected with the accuracy of 85% with the help of facial feature vectors obtained with the help of Gabor filter and that feature value is convolved with the Log Gabor filter.[1]

In this whole face is taken into consideration and than its accuracy is tested over the classification with PCA principle component analysis and LDA liner discriminate analysis and the result are quite good .The result is achieved on low reso0lution image without specifying the fiducially points. Facial expression recognition is also achieved with the histogram sequence of the Local Gabor binary pattern.[17] Firstly the face image is convolved with multi orientation with the help of Gabor filter than after that Gabor coefficients map are extracted. Than after local binary pattern is applied on GCM to obtain the local Gabor binary pattern. Finally the SVM is used for the classification and the result of recognition rate is quite appreciable.

III. FACIAL EMOTION RECOGNITION SYSTEM

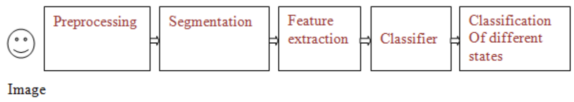

Fig 2 :block diagram

3.1 Preprocessing:--In the preprocessing step the environmental and other variations that are present in different images are minimized. The operation that is performed is contrast adjustment, image scaling, image brightness and other image enhancement method has done. Sometimes noise is associated with images with the variation in signaling and pixel variation so removal of this factor has become essential to achieving the better result.

3.2 Segmentation:-As we know that in many images processing the input is image and output is image. But for the facial expression analysis there is need of feature that is extracted from the input image. In other words we can say that input should be an image but output should be the feature value that has obtained from that input image.Segmentaion plays a major role in that direction.

Segmentation basically divides an image into its regions or objects. The subdivision depends on the level of problem solving. That is the Segmentation will stop when the the objects or the regions of interest in an application has been detected. The successes or failure of the any image processing or computer vision process depends on the Segmentations accuracy. So there is proper concentration is taken place in the finding the probability of accurate Segmentation. Most of the Segmentation algorithm basically depends on the two main properties of intensities values.

a) Discontinuity

b) Similarity.

The first category of Segmentation is taking place when the abrupt changes in the intensity such as Edge is found in an image. so In this case image is partitioned into regions.

In the second case of Segmentation an Image is partitioned into the regions that are similar for particular predefined criteria.

In my approach a facial image is segmented into 4 region means 2×2 grid of 128×128 size.

256×256

Fig 3 Before Segmentation

128×128

Fig 4 After Segmentation.

3.3 Feature extraction is the important method to define the any recognition System. Without extracting the key feature point it is not feasible to define the facial expression recognition System. In feature extraction method I have applied the Gabor filter on the four segmented region of the face. There is 68×68 features are coming outside from the one face region but due to application of absolute and scaling method we evaluate the major four prominent features like Gabor real part, imaginary part, magnitude and angle are evaluated. These extracted feature value is in the range of 0 and 1.so from the one face we calculates the 16 features that is defined by the Gabor filter outcome. We use the frontal face for the facial expression analysis. There is need to define such an unique parameter that can be differentiate between different facial expression of expressers. The set of parameter that can be extracted is known as feature vector and the information achieving from the feature vector is defines an uniqueness as aspect with the extraction technique. If the feature value extracted from the one expression matches with the feature extracted with the expressions of other faces than it is not known as a good feature extraction technique and it is known as feature overlap. The feature extraction should be different in comparison with the other so there is made a correlation and on behalf of this next procedure be implemented. so there is several method adopted for feature extraction and among these the Gabor filter bank based method is good. So we can say that Gabor based feature extraction technique is excellent for facial expression analysis and avoid the feature overlap condition.

3.4 Gabor Filter

$$\psi(k,x) = \frac{k^2}{\sigma^2}\exp\left(-\frac{k^2x^2}{2\sigma^2}\right)\left[\exp(ik.x) - \exp\left(-\frac{\sigma^2}{2}\right)\right].$$

A 2-D Gabor function is a plane wave with the wave vector K and having a Gaussian envelope function that is restricted with the relative width σ.The value of σ is fixed π for the image of resolution 256×256 Like in Lyons etal (1998). In the spatial Domain 2D Gabor filter is known as the Gaussian Kernel that has produced by sinusoidal plane wave. The important property of Gabor filter is that it is self similar because all filters can be generated from the one parent wavelet with the help of rotation and multi orientation. Gabor filter impulse response is also known as the convolution of the FFT of the harmonic function and the FFT of the Gaussian.

A discrete set of Gabor kernel is used that consists of the 3 spatial frequency association with the wave number K=π/4,π/8,π/16 and with the six distinct orientation from 0 to 180 degree differing with the interval of the 30 degree that makes a filter bank of total in 18 Gabor filters.

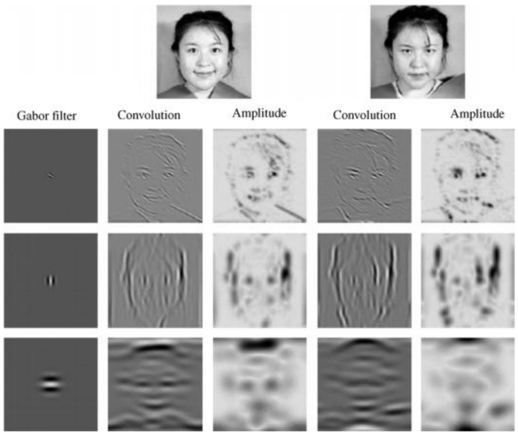

Fig 5:Gabor filter responses for the two sample images

The second term in bracket makes the Gabor wavelet kernel DC-free .as Gabor kernels is DC free than it is more robust against the variation in brightness in the image. Recognition using Gabor filter is done by the similarity function.

$$S_a(\mathcal{J},\mathcal{J}') = \frac{\sum_j a_j a'_j}{\sqrt{\sum_j a_j^2 \sum_j a'^2_j}}$$

In this equation jet J has been assumed from a fixed image position and jet j′=J′(x) taken a variable position

IV. OUR APPROACH

The image is frontal and the Gaussian function is centered at the origin. we consider the value of σ=π and taken the value of wave vector π/4,π/8,π/16.but in my approach we have seen there is not too much difference result comes out so mainly concentrate on the first one value. One applying the Gabor filter we see the outcomes of the result is classified into the four parts:--

a) Gabor filter output
b) Gabor kernel
c) Magnitude
d) Phase

As Gabor function is in complex form so its breaks into real and imaginary part and on the behalf of this it can be easily calculated the phase and magnitu7de that be responses of the Gabor filter on the particular facial image.

1) To get the real part and the imaginary part of filter output use real (gabout) and imag (gabout), respectively.

2) To get the magnitude and the phase of the complex filter output use abs (gabout) and angle (gabout), respectively.

As there is 68×68 real parts and imaginary parts value are coming outside on the Gabor filter output response that is also the same size of 68×68.we taken the absolute value of Gabor output response and finally consider one feature value. After evaluating the magnitude for one partition images similar we calculated the magnitude for the remaining 3 partition area. So for one face image we calculate the 4 features for magnitude only. Similarly we see the absolute value for imaginary part to calculate the Gabor kernel and evaluate one Gabor kernel among the 68×68 pixel value. the Gabor kernel value is basically in complex form but on taking absolute it just try to generate the real number. so for one facial image we calculates the 4 Gabor kernel value that value lies between the range 0 and 1.similarly the four absolute value of the real parts are evaluated and named their as Gabor filter. Similarly in the same fashion we have taken the maximum of maximum of the absolute value of real parts and maximum of maximum of absolute value of imaginary parts to calculate the magnitude. So there is 4 magnitude value is evaluated from one facial image. In the similar fashion we can evaluate the phase value and finally the 4 prominent value of phase is evaluated. So there is a 16 feature value is calculated from the one facial image.16 feature value is extracted because of the segmentation of the image into the four segmented part.

SVM is a useful method for the data classification. It is easier than using the Neutral Networks.SVM provides the accuracy and fast result for the data to be classified and belongs to the particular class. In SVM the data is partitioned into the two parts that one is called Training set and other is known as the testing set and each having the instances of the attributes. Each instances having one target means class labels and several attribute. The goal of the SVM is that to produce the model which predicts the target value of instances in the testing set which are given by only attributes value. It is based on the supervised learning methods. Support vector machine has a unique property that it creates the hyperplane or a set of hyperplane which has taken into the consideration for the classification. It defines the functional margin means a good separation is achieved by the hyperplane that has the largest distance to the nearest data points of any class. Greater the margin it is easy to find better accuracy and minimal errors has found.

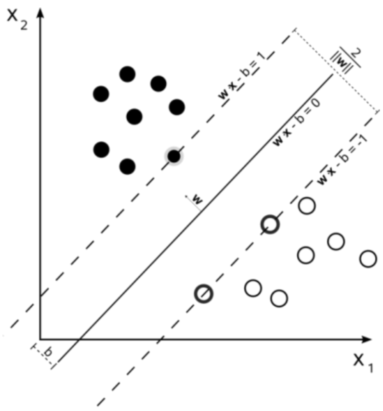

Fig 6 Hyperplane that discriminates between two classes.

Lets us assume that we have Data sets of having n features of the form

$$\mathcal{D} = \{(\mathbf{x}_i, c_i) \mid \mathbf{x}_i \in \mathbb{R}^p,\ c_i \in \{-1,1\}\}_{i=1}^{n}$$

Where C_i belongs to either -1 and +1 value indicates the which class the point X_i belongs.the equation of the hyper plane is given by in such a way

W·X-b=0

There is two cases to defining hyperplane

W·X-b≥1------

it indicates for class one

W·X-b≤-1-----

it indicates for the class others.

Here the training vectors X_i are projected into the higher dimensional space by the function ø. the main important characteristic property of SVM is that it works on the linear separable hyperplane with the minimal margin in the high dimensional space.

C>0 is the main term to define the error parameter. The kernel function is defined as $K(\mathbf{x}_i,\mathbf{x}_j) \equiv \phi(\mathbf{x}_i)^T\phi(\mathbf{x}_j)$.

There is four basic kernels that is defined as follows:-

linear: $K(\mathbf{x}_i,\mathbf{x}_j) = \mathbf{x}_i^T\mathbf{x}_j$.

polynomial: $K(\mathbf{x}_i,\mathbf{x}_j) = (\gamma\mathbf{x}_i^T\mathbf{x}_j + r)^d,\ \gamma > 0$.

radial basis function (RBF): $K(\mathbf{x}_i,\mathbf{x}_j) = \exp(-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2),\ \gamma > 0$.

sigmoid: $K(\mathbf{x}_i,\mathbf{x}_j) = \tanh(\gamma\mathbf{x}_i^T\mathbf{x}_j + r)$.

Where γ, r.and d are kernel parameters.

V. RESULTS

Facial Expression or Facial emotion detection through image database has been proposed: Basically Image Database is Japanese Female Faces and assumed to be frontal and preprocessed and some Indian database also has been used to evaluated the accuracy, After applying Gabor Filter on the four segmented facial region we extract the Features value that is fully defined by Gabor response,

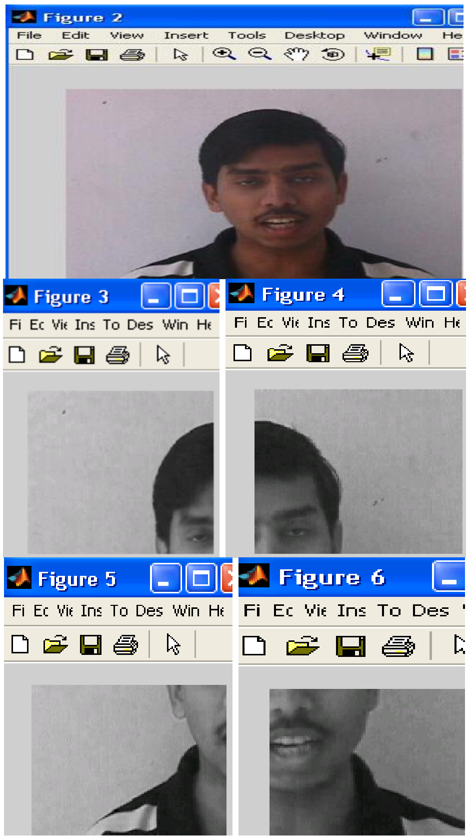

Fig 7: Input Image and Segmented Image of Surprised Class

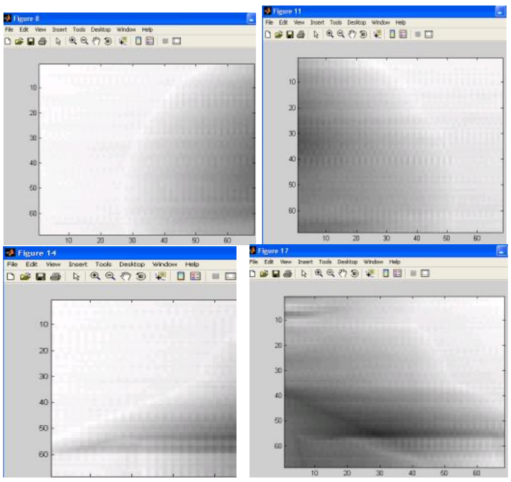

Fig 8: Gabor Filter Response

The main features that are extracted known as Gabor Real part, Imaginary part, Magnitude and lastly Phase .Basically all these features value are coming out from the segmented Face region under the response of Gabor Filter, After the feature extraction maximum of absolute of Gabor filter responses on the faces are evaluated, After the feature extraction rescaling is done so that all features value are lie between the range of 0 and 1, When the features of all the classes are extracted than finally classifier Support Vector Machine (SVM) applied for labels classification, Extracted Features are arranged into the SVM format means all the value should be written in a single row and showing the name of their respective classes, Than features value of different classes is divided into two domain one for the Training set and the second one for the Testing And in last the accuracy is calculated, accuracy is above 80%.

VI. CONCLUSION AND FUTURE PERSPECTIVE

System that can detects the Facial Expressions like Happy, Sad, Angry, Disgust, Surprise and Fear. We have used a still images that is frontal and preprocessed. The Facial Expression analysis plays a key role in the human Machine interaction. Our approach is different to previous work done and my accuracy is above 81%.Due to segmentation and rescaling the Gabor filter output response is remarkably good. So the extracted feature detection is lie in the range of [0,+1].

This range is essential when we are going to use SVM for the classification. After using the SVM the results obtained is satisfactory. In our future perspective we will try to enhance the accuracy up to 100% and try to make a robust system that can be applied to any database whatever the alignment of the Face. In future we will develop my work for the real time system that can be utilized in any sensitive area and surveillance system. Also the Facial expression analysis is useful for the security, Gaming, Intelligent tutoring system and human behavior recognition.

REFERENCES

[1] J. Cohn, T. Kanade Cohn-Kanade AU-Coded Facial Expression Database Carnegie Mellon University

[2] C.K. Chow, C.N. Liu. Approximating discrete probability distributions with dependence trees. IEEE Trans. Information Theory, 14:462–467, 1968.

[3] I. Cohen, N. Sebe, A. Garg, L. Chen, and T.S. Huang. Facial expression recognition from video sequences: Temporal and static modeling. Computer Vision and Image Understanding, 91(1-2):160–187, 2003.

[4] P. Ekman Strong evidence for universals in facial expressions. Psychol. Bull., 115(2): 268–287, 1994.

[5] J.H. Friedman On bias, variance 0/1-loss, and the curse-of-dimensionality. Data Mining Knowledge Discovery, 1 (1): 55–77, 1997.

[6] N. Friedman, D. Geiger, M. Goldszmidt. Bayesian network classifiers. Machine Learning, 29(2):131–163, 1997. [7] A. Garg, D. Roth. Understanding probabilistic classifiers. Proc. Eur. Conf. on Machine Learning, 179–191, 2001.

[8] D. Goleman. Emotional Intelligence. Bantam Books, New York, 1995.

[9] Intel Research Laboratories. OpenCV: Open computer vision library. http://sf.net/projects/opencvlibrary/.

[10] C.E. Izard. Innate and universal facial expressions: evidence from developmental and crosscultural research. Psychol. Bull., 115(2): 288– 299, 1994.

[11] R. Lienhart, J. Maydt. An extended set of haarlike features for rapid object detection. Proceedings of the IEEE International Conference on Image Processing, Rochester, New York, vol. 1, pp. 900-903, 2002.

[12] R. Lienhart, A. Kuranov, V. Pisarevsky. Empirical Analysis of Detection cascade of Boosted Classifiers for Rapid Object Detection. Intel Corporation, Technical report, 297–304, 2002.

[13] M. Pantic, L.J.M. Rothkrantz. Automatic analysis of facial expressions: the state of the art. IEEE Trans. PAMI, 22(12): 1424–1445, 2000.

[14] C. Papageorgiu, M. Oren, T. Poggio. A general framework for Object Detection. Proceedings of the International Conference on Computer Vision, Bombay, India, pp. 555-562, 1998.

[15] R. Schapire, Y. Freund. Experiments with a new boosting algorithm. Proceedings of the International Conference on Machine Learning, Bari, Italy, Morgan Kaufmann, pp. 148-156, 1996.

[16] R. Schapire. The strenght of weak learnability. Machine Learning, 5(1), 197-227, 1990.

[17] N. Sebe, I. Cohen, A. Garg, M.S. Lew, T.S. Huang. Emotion Recognition Using a Cauchy Naive Bayes Classifier. International Conference on Pattern Recognition (ICPR02), vol I, pp. 17– 20, Quebec, Canada, 2002.

[18] H. Tao, T.S. Huang. Connected vibrations: a modal analysis approach to non-rigid motion tracking. Proc. IEEE Conf. on CVPR, 735–740, 1998.

[19] P. Viola, M. Jones. Rapid Object Detection Using a Boosted Cascade of Simple Features. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii, vol. 1, pp. 511-518, 2001.