

Shape and Texture based Facial Expression Recognition using Neural Network

J. Suneetha
Associate Professor
Dept. of CSE, SIET, Puttur.

Abstract: - Facial expressions are the changes in the face of a person according to his internal emotional states. This paper presents a new method of Facial Expression Recognition System that is a combination of shape and texture features from eyes, nose and mouth parts of the face. To develop this model, the Region of Interest (ROI) is applied for eye, nose and mouth parts. The ROI of eye is given as input to feature extraction module which adopts Active Appearance Model (AAM) to extract shape features and also ROI of eye is given as input to another feature extraction module which extract texture features using Gray-Level Co-Occurrence Matrix (GLCM). The obtained shape and texture features are combined to form a combined feature vector of eye. Similarly the ROI of nose and mouth are given as input to feature extraction modules which adopts Active Appearance Model (AAM) to extract shape features and Gray-Level Co-Occurrence Matrix (GLCM) is used to extract the texture features. The obtained shape and texture features of nose are stored in a combined feature vector and similarly the shape and texture features of mouth are stored in another combined feature vector. The combined feature vector of eyes, nose and mouth are merged to form a final feature vector and is given as input to the recognition module. In this paper, Neural Network model is developed and trained using well known Gradient Descent Feed Forward Back Propagation Algorithm in order to classify the facial expressions.

Keywords: Active Appearance Model, Gray-Level Co-Occurrence Matrix, Neural Network

1. INTRODUCTION

Emotion is a state of feeling like thoughts, psychological changes and expressions which positively affects intelligent functions like decision making, perception and empathetic understanding. Human beings express emotions in day today interactions and emotions are reflected in voice, hand, body gestures and mainly through facial expressions[1]. Facial Expressions are a form of nonverbal communication, which plays a vital role in interpersonal relations and social communications

Facial expressions are dynamic features which communicate the speaker's attitude, emotions, intentions, and so on. The face is the primary source of emotions. Facial features play an essential role in the human facial analysis and features are classified as permanent or transient. Eyes, lips, brows and cheeks are examples of permanent features whereas facial lines, brow wrinkles and deepened furrows are examples of transient features[2].

The combination of different features of a face like eyes, mouth and nose contribute to express different types of emotion. For example happiness is represented by a larger separation between the left and right corners of the mouth and eyes tend to be relaxed for a happy expression. On the other hand, surprise is generally characterized by the mouth being wide open, which means a smaller separation between the left and right corners of the mouth as compared to the upper and lower lips and also, eyes tend to be wide open and hence larger in size for a surprised person [3]. Angry is represented by eyes are wide opened, height of eye corners increased, raised nostrils and compressed mouth whereas for disgust expression is represented by the wrinkle nose and raised upper lip for a disgust person. Fear is represented by eyes are opened, eye corners width extended, mouth stretched backward and curved for a fearful person. For a sad person, the eyes are slightly closed and lip corners are pulled down. Finally, neutral expression does not represent any changes in the eye, mouth and nose part.

Shape feature of an object are very powerful for similarity search and retrieval process because shape of objects is strongly linked to functionality and identity of object. This property helps to distinguish shape features from other eliminating visual features such as a colour or texture. Texture is observed in the structural patterns related to the surface of objects like wood, grain, sand, gross and cloth. The texture term is used to represent repetition of basic texture elements called textures which contain several pixels whose placement could be periodic or random.

2. SYSTEM ARCHITECTURE

In the proposed model, a combination of both shape and texture features are eyes, nose and mouth parts are considered as features and are used to classify the facial expressions. In the first experiment, a facial expression recognition model is developed by utilizing shape and texture features of eye, nose and mouth parts. This recognition system consists of mainly three phases such as preprocessing, feature extraction and facial expression recognition. The various processing steps involved in the recognition system are shown in the following Figure 1.

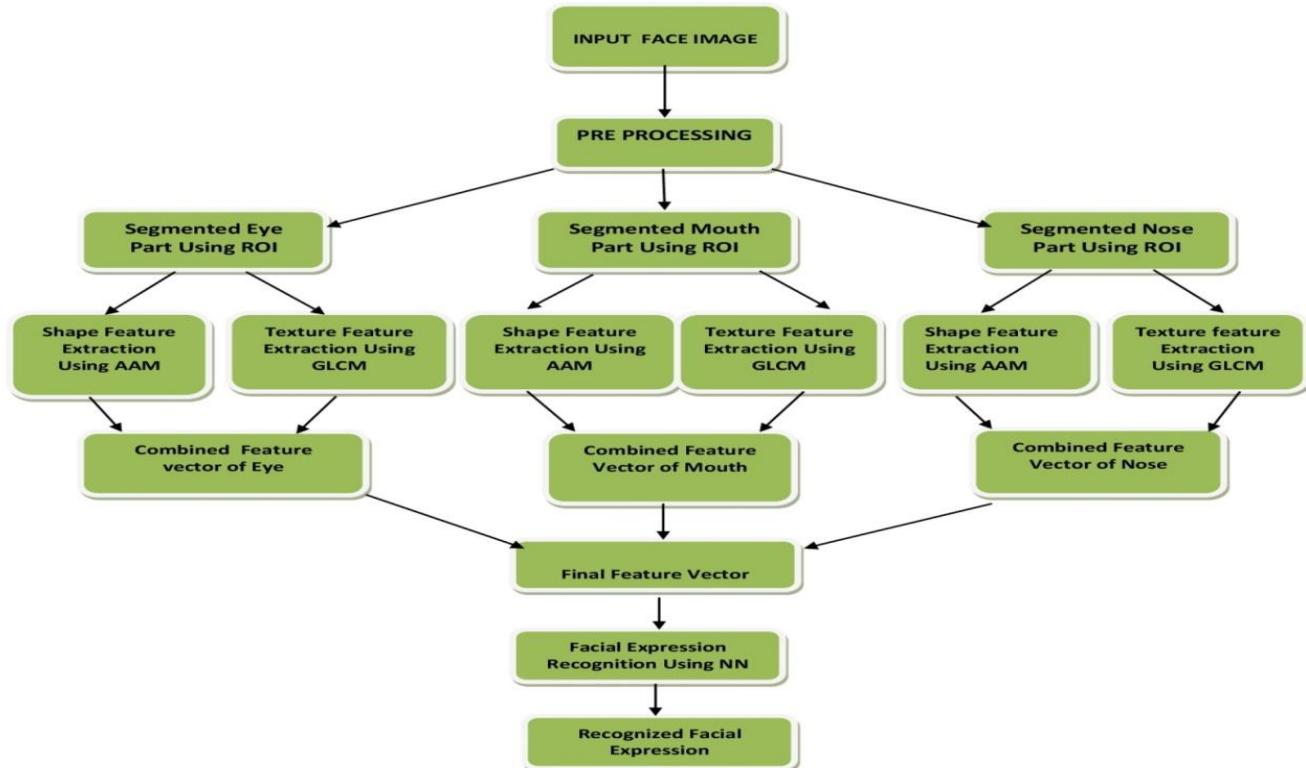


Figure 1: Phases in Facial Expression Recognition System

The first step in the proposed facial expression recognition is preprocessing which is mainly used to improve the quality of the image and then from this image by applying Region of interest (ROI) the eye, nose and mouth parts are segmented from the given input image. The segmented eye part is given as input to two feature extraction modules. The first feature extraction module is used to extract shape features of eye using Active Appearance Model (AAM). The second feature extraction module is used to extract texture feature from eye part using Gray-Level Co-Occurrence Matrix (GLCM) technique. The outputs from these feature extraction modules are stored in a combined feature vector. In the similar way shape and texture features by using AAM and GLCM respectively from the segmented mouth part and nose part are extracted to form a combined feature vector of mouth part and combined feature vector of nose part. The individual combined feature vectors of eye, mouth and nose parts are merged to form a final feature vector. This final feature vector is used to train the Neural Network using Gradient Descent Feed Forward Back Propagation Algorithm. This trained Neural Network can be used to classify the facial expression into distinct classes such as happy, angry, disgust, fear, surprise, sad and neutral. The detailed description of the techniques which are applied in each phase of the proposed model is given in the following section.

3. PROPOSED MODEL

The proposed model consists of three phases such as preprocessing, Feature extraction, Facial expression Recognition

3.1 Preprocessing

Preprocessing is the first phase of the proposed model and is used to suppress unwanted distortions that will be relevant for further processing and analysis task. The preprocessing process includes operations are convert an image into grayscale, noise removal, contrast enhancement and segmentation. The image information is represented in digital form as two dimensional array and various operations can be performed on two dimensional array such as noise removal, enhancement, feature extraction, classification etc. In the proposed model, preprocessing comprises operations like converting an RGB image to grayscale image, reduction of noise using an Adaptive Median Filter, contrast enhancement using Histogram Equalization and segmentation using Region of Interest (ROI)[4,5].

3.2 Feature Extraction

Feature extraction is a special form of dimensionality reduction. When the input data is too large to be processed and suspected to be redundant then the data is transformed into a reduced set of feature representations. In the proposed model, shape and texture features are considered for facial feature extraction. The feature extraction technique Active Appearance Model (AAM) is adopted in the development of shape based feature extraction. The feature extraction technique Gray Level Co-

Occurrence Matrix (GLCM) is also adopted in the development of texture based feature extraction. In the proposed work, the shape features are extracted from the eye part by using Active Appearance Model (AAM) and four optimized features are selected using inverse compositional algorithm then stored in a feature vector. The segmented eye texture features are extracted by applying Gray Level Co-Occurrence Matrix (GLCM) technique. When GLCM technique is applied on eye part, four important texture features such as Angular Second Moment (Energy), Inverse Difference Moment (Homogeneity), Contrast, and Correlation are obtained as the output of GLCM and values are stored in a another feature vector. The shape feature vector of eye and texture features vector of eye are stored in a combined feature vector. In the similar way, above eye feature extraction procedure is applied for nose and mouth then obtained individual combined feature vectors of nose and mouth. The combined feature vectors of eyes, nose and mouth are fused to form a final feature vector and is given as input to the recognition module.

3.3 Facial Expression Recognition

After feature extraction, the most important task is the selection of a proper classifier which is fast and robust. Recognition is a process of classifying the different input patterns into distinct defined expressions such as anger, disgust, fear, happy, sad, surprise and neutral. In this experiment, obtained shape and texture combined feature vector of eyes, nose and mouth are fused to form a final feature vector and is given as input to the recognition module. In this experiment, Neural Network model is developed and trained using well known Gradient Descent Feed Forward Back Propagation Algorithm. The Neural Network model is adopted in the development of proposed model is Gradient Descent Feed Forward Back Propagation algorithm. The performance of the trained model is tested with sample data after applying above discussed preprocessing and feature extraction techniques. The experimental results of the proposed system are given in the following section.

4. IMPLEMENTATION

Active Appearance Model (AAM): Active Appearance Model (AAM) is a statistic model incorporating both shape and intensity information and extracting the prominent geometric based shape features from the preprocessed image. In Active Appearance Model (AAM), the shape of a segmented input image is represented by a vector consisting of the positions of the landmark points[6]. Active Appearance Model (AAM) is a computer vision algorithm, template matching statistical model is used for building shape and appearance of facial features by automatically locating landmark points that define shape and appearance of objects in an image. AAM is particularly suited to the task of interpreting faces

in images. Faces are highly variable, deformable objects and manifest very different appearances in images depending on pose, lighting, expression and identity of the person. Interpretation of such images requires the ability to understand this variability in order to extract useful information

Gray Level Co-Occurrence Matrix (GLCM): Gray Level Co-Occurrence Matrix was proposed by Haralick and it is widely used for texture analysis[7,8]. GLCM is a method for extracting texture features from preprocessed image. The GLCM is created from a gray scale image and it calculates how often a pixel value i occurs either horizontally, vertically or diagonally to adjacent pixels with the value j . Haralick proposed two steps for texture feature extraction, the first step is computing co-occurrence Matrix and the second step is calculating texture based features from co-occurrence Matrix

Gradient Descent Feed Forward Back Propagation Algorithm

Initially all the edge weights are randomly assigned. For every input in the training dataset, the ANN is activated and its output is observed. This output is compared with the desired output and the error is propagated back to the previous layer. This error is noted and the weights are adjusted accordingly. This process is repeated until the output error is below a predetermined threshold [9,10]. The various steps in the Gradient Descent Feed Forward Back Propagation Algorithm are as follows

➤ Algorithm Steps for Gradient Descent Feed forward Back Propagation

Step 1: The loads for the neurons of hidden layer and the output layer are designed randomly choosing the weight. However, the input layer possesses the constant weight.

Step 2: The anticipated Bias function and the activation function are evaluated by means of

- The loads for each and every neuron are attached separately from the neurons in the input layer.

$$\hat{y} = ML(x, W), \text{ with } x = (x_1, x_2, \dots, x_n) \text{ and } \hat{y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_m)$$

Where W is the set of parameter $\{W_{ij}^L, W_{i0}^L\}, \forall i, j, L$

- For each unit i of layer L of the ML . The net input is evaluated as

$$S = \sum_i y_j^{L-1} w_{ij}^L + w_{i0}^L \quad \text{----- (1)}$$

- The activation function is

$$F(s) = \frac{1}{1 + e^{-s}} \quad \text{----- (2)}$$

- The learning error is represented as follows

$$E^{BP} = \sum (T_k - O_k)^2 \quad \dots (3)$$

Where T_k is desired output and O_k is actual output. The Multi layer uses the algorithm of Gradient Descent Feed Forward Back Propagation Neural Network for training to update the weights.

for the Feed Forward Back Propagation Neural Network.

Step 3: The Back Propagation Error is evaluated for each node and thereafter the weights are modernized as per the following Equation 4

$$w_{(n')} = w_{(n')} + \Delta w_{(n')} \quad \dots (4)$$

Step 4: The weight $\Delta w_{(n')}$ is updated as per Equation 5

shown below.

$$\Delta w_{(n')} = \delta \cdot X_{(n')} \cdot E^{(BP)} \quad \dots (5)$$

Where, δ - Learning Rate, which is habitually in the range of 0.2 to 0.5.

$E^{(BP)}$ - Back Propagation Error.

Step 5: The process is repeated till the Back propagation (BP) error is reduced to the minimum i.e. $E^{(BP)} < 0.1$. The experimental results of the proposed work are discussed in the following section

5. RESULTS

The proposed Facial Expression Recognition System is implemented in MATLAB platform. In this proposed work, the Japanese Female Facial Expression (JAFFE) database is used for facial expression recognition. From this database, the images are considered for training and testing purpose. The 80% of facial expressions samples from JAFFE database are considered in this thesis work in order to develop the Facial Expression Recognition System. During this development process, preprocessing techniques such as converting an RGB image to grayscale image, reduction of noise using an adaptive median filter, contrast enhancement using Histogram Equalization and segmentation using Region of Interest (ROI) are applied for each sample. Shape features from eye part of each sample are extracted by using Active Appearance Model (AAM) and texture features from eye part of each sample are also extracted by using Gray Level Co-Occurrence Matrix (GLCM) and these are stored in a combined feature vector. For each sample shape and textures related to all the features of nose are stored in a combined feature vector. Similarly all the shape and texture extracted features from mouth part are also stored in a combined vector. The individual combined

vector related to mouth, nose and eye are merged to get a final feature vector. For each sample image this type of final vector is formed and is stored in a database. The considered Neural Network is trained by using this database.

After completion of training next phase is testing. A few samples are considered for testing purpose. For these test samples the above discussed preprocessing techniques, feature extraction techniques and formation of final feature vector is applied to form a test database. The performance of trained network is analyzed by using this database.

The combined feature vector of eyes, nose and mouth are fused to form a final feature vector and is given as input to the recognition module. In this experiment, Neural Network model is developed and trained using well known Gradient Descent Feed Forward Back Propagation Algorithm. Back Propagation algorithm attempts to minimize the error at each iteration. The weights of the network are adjusted by the algorithm such that the error is decreased along a descent direction. Finally the developed Neural Network model is classified the expressions into seven distinct types such as angry, disgust, fear, sad, happy, surprise and neutral. The screen shot of the proposed model is shown in the following Figure 2.

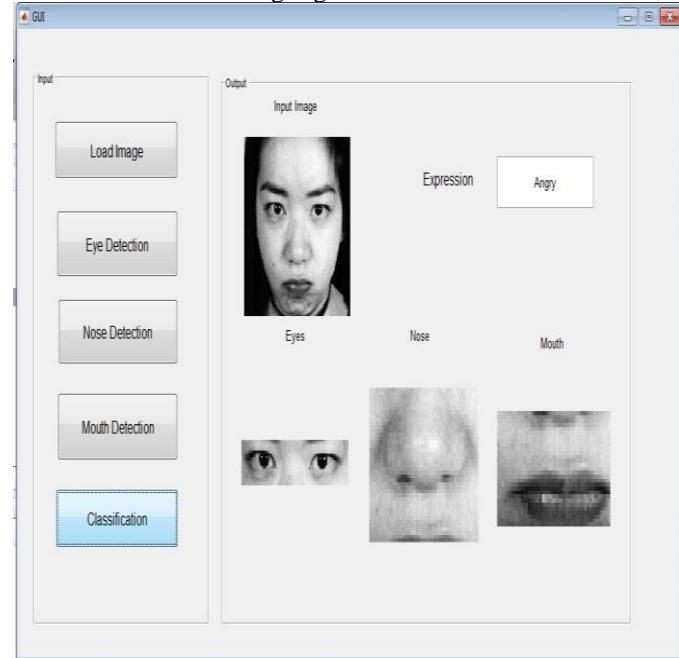


Figure 2 : Classified Expression for Given Image

6. PERFORMANCE ANALYSIS

The performance of the proposed system measures through sensitivity, specificity and accuracy

Specificity:

The specificity is the proportion of negatives which are accurately recognized. It is related to the capability of test to recognize negative results.

Sensitivity:

The sensitivity is the proportion of actual positives which are accurately recognized. It is related to the capability of test to recognize positive results.

Accuracy:

The following equation can be used to compute accuracy of the developed classifier.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100$$

Where TP- True positive, TN- True negative, FP- False positive, FN- False negatives.

The following Table 1 shows the performance measures of proposed work measuring specificity, sensitivity and accuracy.

Table 1: Performance Measurements for proposed work

Expression	Specificity (%)	Sensitivity (%)	Accuracy (%)
Angry	88	89	91
Disgust	86	87	93
Fear	89	89	95
Happy	94	95	97
Sad	93	98	98
Surprise	95	98	95
Neutral	94	96	96

The graphical representation of the above discussed metrics related to the proposed work as shown in following Figure 3.

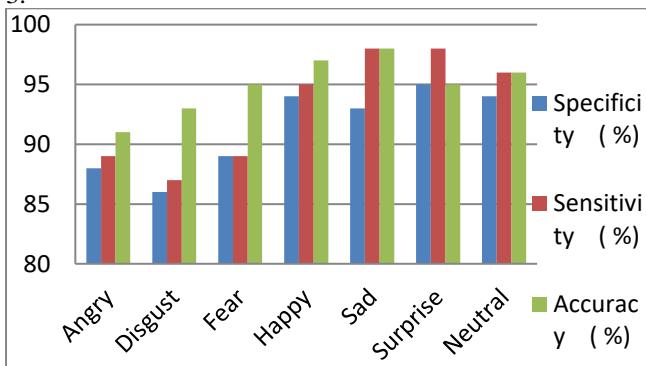


Figure 3: Graphical Representation of Evaluation Metrics

The proposed work provides good emotion classification results. The above Table 1.gives information about performance analysis for the image classification in specificity, sensitivity and accuracy. The developed ANN model with shape and texture features of eye, nose and mouth parts, obtained maximum specificity 95% for surprise expression and minimum specificity 86% for disgust expression. Maximum sensitivity 98% is obtained for sad and surprise expressions and minimum sensitivity is 87% for disgust expression. The developed ANN model, maximum accuracy obtained 98% for sad expression and minimum accuracy 91% for angry expression. The average value obtained for specificity, sensitivity and accuracy are 91%, 93% and 95% respectively by using the ANN model. The developed model achieved good recognition accuracy but it requires more time for training and also complexity in the

structure. Keeping in view of this, in this thesis work, developed another Facial Expression Recognition Hybrid model. The detailed explanation of Facial Expression Recognition Hybrid model is discussed in the following section.

7.CONCLUSION

In this proposed work, a Facial Expression Recognition model is proposed and developed with a combination of shape and texture features of eyes, nose and mouth parts of face image. The significant shape features are extracted by using Active Appearance Model (AAM) and texture features are extracted by using Gray-Level Co-Occurrence Matrix (GLCM) from the segmented eyes, nose and mouth parts of face image. The final vector is formed which is a combination of extracted shape and texture features of eyes, nose and mouth parts of face image. The Neural Network model is developed and trained using Gradient Descent Feed Forward Back Propagation Algorithm with the help of final feature vector to classify the seven facial expressions. The performances of all the proposed and developed models in this thesis work are analyzed in terms of Sensitivity, Specificity and Accuracy and achieved good results. The proposed work accuracy 95% obtained by using Neural Network.

8. REFERENCES

- [1] A. Mehrabian, "Communication Without Words Psychology Today", Vol. 2, No. 4, pp. 53-56, 1968.
- [2] P. Ekman, W.V. Friesen, "Constants Across Cultures in the Face and Emotion", J. Personality Social Psychol. Vol. 17, No.2, pp. 124-129, 1971.
- [3] M. Pantic, L. Rothkrantz, "Expert System for Automatic Analysis of Facial Expression", Image Vision Computing, Vol. 18 (11) , pp.881–905, 2000.
- [4] Gonzalez, R. C. and Richard E Woods. "Digital Image Processing", Pearson Edition, Second Edition.
- [5] Priyanka Kamboj and Versha Rani, "A Brief Study of Various Noise Model & Filter Techniques", Vol.4, No. 4, Journal of Global Research in Computer Science, 2013.
- [6] Thanh Nguyen Duc, Tan Nguyen, Huu, Luy Nguyen Tan, "Facial Expression Recognition Using AAM Algorithm", 2009.
- [7] A. Punitha, M. Kalaiselvi Geetha, "Texture Based Emotion Recognition From Facial Expressions Using Support Vector Machine", International Journal Of Computer Applications, Volume 80 - No. 5, October 2013.
- [8] P.Mohanaiah, P.Sathyarayana, L.Gurukumar, "Image Texture Feature Extraction Using GLCM Approach", International Journal of Scientific and Research Publications, Vol 3, Issue 5, May 2013.
- [9] Shubhangi Giripunje, Preeti Bajaj, "Recognition Of Facial Expressions For Images Using Neural Network", International Journal of Computer Applications, Vol.40, No.11, Feb 2012.
- [10] Ashish Kumar Dorga, Nikesh Bajaj, Harish Kumar Dorga, "Facial Expression Recognition Using Neural Network With Regularized Back-Propagation Algorithm", International Journal Of Computer Applications, Vol. 77, No.5, September 2013.