

Detection of Mouth for Facial Expression classification using Genetic Algorithm

M.Mahadevi
MSU, Research Scholar,
Department of Computer Science,
S.D.N.B. Vaishnav college for women,
Chromepet, Chennai-44, Tamilnadu, India.
maxsaran@gmail.com

Dr.C.P.Sumathi,
Associate Professor & Head
Department of Computer science,
S.D.N.B. Vaishnav College for women,
Chromepet, Chennai-44, Tamilnadu, India.
drcpsumathi@gmail.com

Abstract— Day to day interactions of people in a communication involves expressions. Those expressions are mind reader of people emotions. The expressions are basically classified as Neutral, Happy, Sad, Surprise and Fear. The facial images are captured as digital images and are processed to classify the expressions. This paper involves in detecting mouth from the face to classify such facial expressions using genetic algorithm. Genetic algorithms are applied for image enhancement and classification using template matching. This paper applies an optimal filtering to the facial images using genetic algorithm and features of the face are detected by applying horizontal and vertical projections. Template matching is performed between a neutral expression and other expressions to retrieve the correlation value. Statistical parameter correlation which gives the relationship between two data sets is used to classify the expressions. The overall classification rate of the algorithm is 94%.

Keywords—Expressions; Genetic algorithm; Correlation; Template matching; Projections.

I. INTRODUCTION

In recent year's biometrics is the targeted research work. Mainly automating human and computer interaction had been a herculean task for the researchers even then much more new research work happens to be most promising. Still automating face and face based digital image processing are at the top in the research domain. Facial action coding system was developed by Paul Ekman [1] which was specially designed to identify facial expression based on action units. These action units were being researched by many researchers. Michael.F.Valstar et.al had made an analysis of the first facial expression challenge of facial action units[2]. Arman Savran et.al have researched on intensity estimation of facial action units[3]. Apart from facial action units there are many researchers who have developed several models to identify different facial expressions. Nidhi Tiwari and Mohd. Ahmed has detected facial features based on skin color model [4]. Facial expressions are classified using singular value decomposition and neural network [5]. Seyed Mehdi Lajvardi and Hong Ren Wu have applied color models with gabor filters and multiclass linear discriminant analysis was used for classification[6]. Preeti Augustine have used wavelets and Euclidean distance for facial classification [7]. Le Hoang Thai et.al have applied canny edge operator with principal component analysis and classification using neural network was implemented [8]. Deepak Ghimire and Joonwhoan Lee have used Histogram orientation gradient features to identify facial expression [9]. Shaohua Wan , J.K.Aggarwal have developed a robust model for spontaneous facial expression

recognition [10]. Boulbaba Ben Amor et.al have developed algorithm for 4D facial expression recognition[11]. Even though many researches have been made for different problems with different techniques, identifying facial expressions have been demanding research in the current research world. This work's main objective is to identify the facial expression of gray scale images with an optimal image preprocessing. Many soft computing techniques exist in facial expression classification, this work has implemented Genetic algorithm. Here image preprocessing is done using genetic algorithm and facial features are extracted and classified. Finally the extracted face region runs through template matching and correlation based classification is implemented.

II. PROPOSED METHOD

The proposed method presents mouth detection using projection analysis after preprocessing of the image. Ajit P. Gosavi, S. R. Khot [12] and Ramchand Hablani[13] have applied principal components and local binary patterns to identify the facial expression recognition. But this work have applied optimal pre-processing, as the image from the database is bound to some noises or disturbances which may make way to wrong identification of the features and thereby classification accuracy may get disturbed. Moreover the classification is done using mouth detection. The algorithm of the proposed method is given

A. Algorithm

- a. Pre-Processing.
 - Image enhancement using genetic algorithm
 - Normalization.
 - Binary image.
- b. Mouth Detection and Feature extraction using Horizontal and Vertical Projection.
- c. Template Matching and classification.

III. .PREPROCESSING

Digital images acquired are bound to noises and needs to be pre-processed to retrieve suitable information for the problem domain. The first step in pre-processing is to remove noise in the image and to improve the quality of the image. The algorithm has designed an optimal filter to remove noise from the image. Genetic algorithm is used to develop the optimal filter for the database images.

Median filter is a kind of smoothing filters which helps in preserving edges which is important for facial images to

identify the boundaries. In this paper optimal filter is developed using median filters running on genetic algorithm with varying window size. Peak Signal Noise Ratio (PSNR) is the metric used as fitness function in genetic algorithm for the chromosome to survive and enter into population.

The steps involved in developing the optimal filter.

1. Load an image as input from the database.
2. Salt and pepper noise is induced to the image and saved as noise image.
3. Calculate PSNR value for the noisy image.
4. Median filter is applied to the noisy image.
5. Calculate PSNR value for the filtered image.

The proposed steps are made to run in Genetic algorithm.

A. Genetic Algorithm

Genetic algorithms are adaptive heuristic search algorithms based on the evolutionary ideas of natural selection and genetics. M.Karthigayan et.al [14] have applied genetic algorithm by considering the eyes and mouth as ellipse and classification was accomplished by neural network for color images. Kai Hu et.al have also combined genetic algorithm with back propagation network for image enhancement for all images [15]. This work have concentrated mainly on the deciding the window size of the filter using genetic algorithm for facial images with PSNR value as the fitness value for the chromosome to generate new population. The important elements in Genetic algorithm [16] are individuals and population.

Fitness function

The fitness of an individual is the value of the objective function for its phenotype. For calculating the fitness, the chromosome has to be first decoded and the objective function has to be evaluated. This function quantifies the optimality of a solution, so that the chromosome is ranked with other chromosomes. The fitness function is the metric Peak Signal Noise Ratio (PSNR) value of the image. PSNR [17] is defined as the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. The PSNR (in dB) is defined in Fig.1.

$$\begin{aligned}
 PSNR &= 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \\
 &= 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \\
 &= 20 \cdot \log_{10} (MAX_I) - 10 \cdot \log_{10} (MSE)
 \end{aligned}$$

The window size is given as the inputs to GA to find the optimum value of PSNR of the image. Table 1 shows the parameters of the genetic algorithm that is used for the algorithm. The experiment results show the value of the PSNR of certain images which helps to decide the optimal window filter for the image enhancement. This work has taken JAFFE facial expression database images [18] which are in Tagged Image File Format (TIFF). The database contains 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) posed by 10 Japanese female models.

The noisy image when run under median filter resulted with unbalanced raise and decrease of the PSNR values as stated in the Table 2. The same noisy image is run through the optimal filter using genetic algorithm resulting in identifying the optimal window size for the median filter which makes it an optimal filter. The window size is decided as 4*4 since the PSNR value gets decreased after this point. Moreover the final point in genetic algorithm after the maximum generation was found to be 3.985 for window size 3*3 and 4.484 for window size 4*4 with approximate values of the PSNR values to be in the range of 370 to 380 which proves that the median filter when applied with window size of 4 *4 is an optimal filter.

Table 1 Parameters of Genetic algorithm

Parameters	Value
Generation	100
Population size	20
Fitness scaling	Rank
Selection Function	Roulette
Mutation	Uniform
Cross over	Scattered

Table 2 PSNR value for Median filter and optimal filter

Window size	Median Filter PSNR	Optimal Filter PSNR
1*1	161.44	161.73
2*2	458.79	449.27
3*3	1449.8	379.56
4*4	401.25	377.92
5*5	889.76	311.31

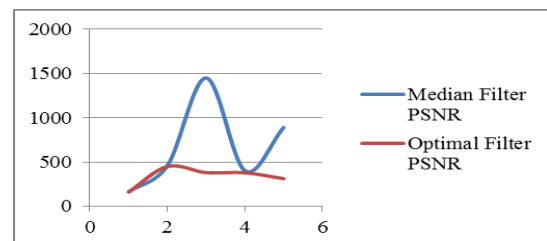


Fig.2. Graph showing comparative plots of median and optimal filters

In the Fig.2,the blue curve shows the median filter PSNR values wherein the red curve shows the optimal filter PSNR values. Since there is no much disturbance in the optimal filter curve each of the PSNR value generated for each window size

are optimum out of which the final point at window size 4*4 is taken as the window size for optimal filter. Fig.3.a shows the noisy image after salt and pepper noise and Fig.3.b shows the filtered image after using optimal filter developed using genetic algorithm. The filtered images all are resized to 128 * 128 and converted to binary image using Otsu 's threshold [19]. Thus derived binary image is given in the Fig.4.

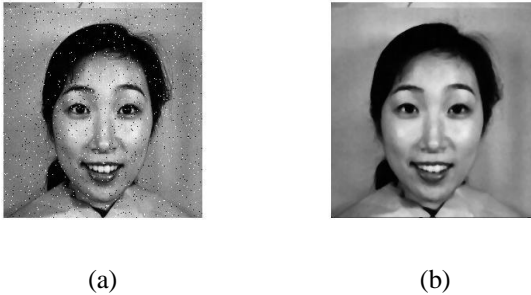


Fig.3. show the noisy image and filtered image after using optimal filter.

- (a) Noisy image
- (b) Filtered image



Fig.4. Binary image

IV. FEATURE EXTRACTION

A feature is defined as the function of one or more quantities that characterizes the objects in the image. Features can be classified into pixel features, Local features and Global features. Pixel features deals with pixel color and location. Local features consider subset of the image and Global features considers the whole image such as statistical and spatial features. Since facial expressions are identified easily with Mouth variations, the facial image is segmented for detecting Mouth. This work segments the Mouth from the binary image using the horizontal and vertical projection.

Horizontal and vertical projection [20] is defined as the metric that denotes a binary image as function of row coordinates and column coordinates. Horizontal projection is the row sum of the binary image and its pattern helps to identify the boundary of the face.

$$P_{hi} = \sum_{j=1} S(i,j), i=1,2,3...n \quad (1)$$

$$P_{vj} = \sum_{i=1} S(i,j), j=1,2,3...n \quad (2)$$

With the peak of the horizontal projection the face is divided horizontally into upper and lower face. Sobel edge operator is applied on to the divided face images resulting in edge detected upper and lower face as in Fig.5.a and Fig.5.b respectively.



Fig.5. Feature detected images

- (a) Upper face with eyes and eyebrows
- (b) Lower face with Mouth

Fig.5.a. clearly detects the features like eyes, eyebrows of the image and Fig.5.b detects the presence of mouth region with some connected components. The detected mouth is taken for feature extraction and for further classification phase. From Fig.5.b the lip region is matched with the template to classify the expressions.

V. TEMPLATE MATCHING AND CLASSIFICATION

Template matching is a technique to match small parts of an image to a template. The match between two images is evaluated using a statistical parameter - correlation. Correlation is the statistical relationship between two data sets involving dependence between the data sets. This work uses correlation to find the dependence between each of the expressions with the neutral expression. Since the database contains seven expressions including neutral expression a template of Mouth which exhibits a neutral state is developed manually to match with rest of the detected mouth images to classify the expressions.

The average image of the neutral expression passed through the phase of pre-processing and feature extraction to output the template image as shown in the Fig.6.



Fig.6. Average image of the neutral expression

Correlation between the neutral and any of the expression images is evaluated as there may be slight, moderate and extreme changes in the mouth region when a person gives expression from a neutral state. Such kind of correlation values may be in negative or positive correlated relationship. Correlation is computed between the template image and each of the database images after pre-processing and feature extraction. The standard correlation value ranges between -1 and +1 which states that there is relationship between the images. This research work has evaluated five expressions like happy, sad, disgust, anger and fear with the neutral expressions based on the correlation values. For all images correlation between the resultant feature image Fig.5.b and template image Fig.6. is calculated.

VI. RESULTS AND DISCUSSION

Jaffe database is considered for the classification of the expressions. The database contains 213 images with seven expressions. To test the algorithm a sample of 50 images were considered and the respective correlation values are calculated. Values of some images are displayed in table 3.

The correlation values which ranges from -0.02 to 0.01 is classified as expression 'Sad', 0.01 to 0.025 as 'Disgust', 0.026 to 0.042 as 'Fear', 0.043 to 0.055 as 'Anger', 0.056 to 0.09 as 'Happy'. By thresholding with these values for 50 images the Confusion Matrix derived for the expressions are given as in the table 4.

Confusion Matrix [21] is a table layout that gives the insight of the performance of the algorithm. The rows are the actual classes and the columns are the predicted classes of the algorithm. Table 4 gives the confusion matrix of the classification. The Diagonal elements are the true positives of the classes, column values predicted as some other class are false positives, row values that are wrongly labeled as actual class are false negatives [22] and the remaining that are predicted as actual are true negatives.

The true positives, true negatives, false positives and false negatives of each expression are given in the table 5

The Accuracy rate and Misclassification rate is given as

$$\text{Accuracy} = (\text{TP} + \text{TN}) / \text{Total} \quad (3)$$

$$\text{Misclassification rate} = (\text{FP} + \text{FN}) / \text{Total} \quad (4)$$

Accuracy and Misclassification rate for all expressions is given in the table 6.

The overall accuracy of the algorithm is found to be 94% with a misclassification rate of 6%.

Table 3 Correlation values of images

Image Name	KA.DI1 .42.tiff	KA.FE1. 45.tiff	KA.HA3.31 .tiff	KL.AN1. 167.tiff
Correlation	0.0126	0.0377	0.0579	0.0502
Image Name	KL.DI3 172.tiff	KL.SA2. 162.tiff	KL.SA3.16 3.tiff	KM.FE1. 23.tiff
Correlation	0.016	0.0049	0.0091	0.0419

Table 4 Confusion Matrix

	Anger	Disgust	Fear	Happy	Sad
Anger	6		3		
Disgust		13			1
Fear	1		12		
Happy				7	
Sad		2			5

Table 5 Performance of the algorithm

	True Positive	False Positive	False Negative	True Negative
Anger	6	1	3	40
Disgust	13	2	1	34
Fear	12	1	2	35
Happy	7	2 (not classified under any label)		41
Sad	5	1	2	42

Table 6 Accuracy and Misclassification

	Accuracy	Misclassification rate
Anger	92%	8%
Disgust	94%	6%
Fear	94%	6%
Happy	96%	4%(non-classified)
Sad	94%	6%

VII. CONCLUSION AND FUTURE WORK

This work has implemented genetic algorithm for image pre-processing and features were detected based on the projection profile technique and template matching technique was applied to extract features of the image which was used for classification of the expressions. The misclassification among the expressions has occurred because of the overlapping of the emotions between the anger and fear, disgust and sad. The Happy expression has higher rate of accuracy because that

specific emotion features are prominent and can be identified easily but the emotions like disgust and sad are overlapping emotions, anger and fear are slightly intersecting emotions. So the accuracy rate also varies. But the overall recognition rate of the algorithm is 94% and can be improved in future using an optimization algorithm or by including some more features of the expressions which will reduce the misclassification rate.

VIII. REFERENCES

- [1] P.Ekman and W.V.Friesen."Facial Action Coding System".Consulting Pshychologists Press,Inc.,577 College Avenue,Palo Alto,California 94306,1978.
- [2]] Michel F. Valstar, *Member, IEEE*, Marc Mehu, Bihan Jiang, Maja Pantic, *Fellow, IEEE*, and Klaus Scherer, 2012 "Meta-Analysis of the First Facial Expression Recognition Challenge", *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, vol. 42, no. 4, pp 966-979.
- [3] Arman Savran , Bulent Sankur, M. Taha Bilge 2012," Regression-based intensity estimation of facial action units", *Image and Vision Computing* 30 ,pp 774–784.
- [4]Nidhi Tiwari, Mohd. Ahmed., 2013 "Detection of facial features of frontal face color images" *International Journal of Advanced Computer Research* (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume-3 Number-3 Issue-11, pp 54-58.
- [5] Devi Arumugam, Dr. S. Purushothaman., 2011 " Emotion Classification Using Facial Expression", (*IJACSA*) *International Journal of Advanced Computer Science and Applications*,Vol. 2, No. 7,pp 92-98.
- [6] Seyed Mehdi Lajevardi, *Member, IEEE*, and Hong Ren Wu., 2012"Facial expression recognition in perceptual color space", *IEEE Transactions on Image Processing*, vol. 21, no. 8,pp 3721-3733.
- [7] Preethi Augustine,2013 "Uncontrolled Illumination Rectification Using Wavelet Normalisation for face recognition" *International Journal of Emerging Technology and Advanced Engineering*,ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 8, pp 182-186.
- [8] Le Hoang Thai, Nguyen Do Thai Nguyen and Tran Son Hai, *Member, IACSIT*,2011 "A Facial Expression Classification System Integrating Canny, Principal Component Analysis and Artificial Neural Network", *International Journal of Machine Learning and Computing*, Vol. 1, No. 4, pp 388-393.
- [9] Deepak Ghimire,Joonwhoan Lee, 2014" Extreme Learning Machine Ensemble Using Bagging for Facial Expression Recognition", *J Inf Process Syst*, Vol.10, No.3, pp.443-458.
- [10]Shaohua Wan, J.K.Aggarwal ,2014"Spontaneous facial expression recognition: A robust metric learning approach", *PatternRecognition*47,pp1859–1868.
- [11] Boulbaba Ben Amor, Hassen Drira, Stefano Berretti, Mohamed Daoudi, Anuj Srivastava,2014"4D Facial Expression Recognition by Learning Geometric Deformations" *IEEE Transactions on Cybernetics*, *IEEE*, 44 (12), pp.2443-2457. <hal-00949002>
- [12] Ajit P. Gosavi, S. R. Khot,2013" Facial Expression Recognition Using Principal Component Analysis", *International Journal of Soft Computing and Engineering (IJSC)* ISSN: 2231-2307, Volume-3, Issue-4, pp 258-262.
- [13] Ramchand Hablani, Narendra Chaudhari, Sanjay Tanwani, 2013" Recognition of Facial Expressions using Local Binary Patterns of Important Facial Parts", *International Journal of Image Processing (IJIP)*, Volume (7) : Issue (2) ,pp 163-170.
- [14]M. Karthigayan, M. Rizon, R. Nagarajan and Sazali Yaacob .,2008 " Genetic Algorithm and Neural Network for Face Emotion Recognition", *Affective Computing*, Jimmy Or (Ed.), ISBN: 978-3-902613-23-3, InTech,pp 57-68.
- [15] Kai Hu, Aiguo Song, Min Xia, XiaoLing Ye and YanYan Dou, 2013" An Adaptive Filtering Algorithm Based on Genetic Algorithm-Backpropagation Network", *Hindawi Publishing Corporation Mathematical Problems in Engineering* Volume , Article ID 573941, pp 1-8.
- [16] *Principles of Soft Computing*, Second Edition, S.N.Sivanandam, S.N.Deepa, Wiley Publications.
- [17] http://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio
- [18] www.kasrl.org/jaffe_download.html
- [19] https://en.wikipedia.org/wiki/Otsu%27s_method
- [20] *Digital Image Processing*, S.Sridhar, Oxford Higher education.
- [21] http://en.wikipedia.org/wiki/Confusion_matrix
- [22] en.wikipedia.org/wiki/False_positives_and_false_negatives.