Available online at http://www.journalcra.com

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

International Journal of Current Research
Vol. 11, Issue, 04, pp.2799-2805, April, 2019

DOI: https://doi.org/10.24941/ijcr.34983.04.2019

ISSN: 0975-833X

RESEARCH ARTICLE

BEZIER CURVE AND DECISION TREE MODELLING FOR FACIAL EXPRESSION ANALYSIS

[1*]Humera Tariq, [2]Ayaz Ahmed, [3]Taimoor Raza, [4]Khalid Jamal, [5]Saad AKBAR and [6]Asia Samreen

[1,2,3,4]Department of Computer Science, University of Karachi, Karachi, Pakistan
[5]NED University of Engineering and Technology, Karachi, Pakistan
[6]Department of Computer Science, Bahria University Karachi, Karachi, Pakistan

ARTICLE INFO

Article History:
Received 20th January, 2019
Received in revised form
19th February, 2019
Accepted 27th March, 2019
Published online 29th April, 2019

Key Words:

Feature Extraction, emotion detection, Facial Expression Recognition (FER), Bernstein Polynomial, Bezier Curve, Pattern Recognition, Decision Tree, Curve Fitting.

Corresponding author:

ABSTRACT

It is crucial to handle missing or incomplete information obtained from various low level image processing tasks. Completion of such information requires foundation of typical mathematical, geometrical and linear algebra concepts. In this paper, emphasize is on completion of missing information in terms of detected pixels for expression recognition. Amongst seven standard expressions we chose to work with only Happy, Sad and Normal in this paper. Feature extraction, feature completion and trained decision tree model in XML format on novel and extended dataset has been demonstrated though classical edge detection and Bezier curve modeling. Experimental results show that the recognition rate of proposed system is 71% on CK data set, 76% on JAFEE dataset and 78% using PAKFE dataset for chosen principal emotions.

Citation: Humera Tariq, Ayaz Ahmed, Taimoor Raza, Khalid Jamal, Saad AKBAR and Asia Samreen, 2019. "Bezier Curve and Decision Tree Modelling for Facial Expression Analysis", *International Journal of Current Research*, 11, (02), xxxxxx-xxxxxx.

INTRODUCTION

Parameterization of primal facial regions: Eyes, Mouth and nose are essential to train decision tree for facial expression classification. The importance of our work is twofold: We establish an entirely new dataset Pakistani Facial Expression dataset (PAKFE) which is the very first datasets having images representing Asians. We then provide a comparison of PAKFE with two popular available datasets namely Cohn-Kanade AU-Coded Facial Expression Database(C-K) and Japanese Female Facial Expression (JAFFE). Secondly, we demonstrate the application of higher level mathematics to complete feature information to feed decision trees for facial expression recognition. There are many facial databases available for Facial Expression Recognition and for other research purposes. The most popular data sets are: (1) The Third Emotion Recognition used in The Wild Challenge and Workshop (Emoti W 2015) dataset (WCW) (2) Cohn-Kanade AU-Coded Facial Expression Database(C-K) (3) MMI Facial Expression Database (4) Japanese Female Facial Expression (JAFFE) Database (5) Affectiva-MIT Facial Expression Dataset (AM-FED). Expression Recognition System will segment detected face into three key parts left eye, right eye and mouth based on Voila Jones algorithm for real-time object detection.

Four important modules of Viola Jones method comprised of Haar Feature Selection, Integral Image Computation, Ada Boost and Cascade Classification. Voila Jones itself was trained by 5000 frontal images and 300 million non faces, that's why training takes time but detection is very fast (Viola, 2004; Viola, 2001). Detected facial regions are then pass through normalization, localization, smoothing and sharpening forimproved edge pixel detection (Hu, 2016; Zhang, 2019; Gonzalez, 2008).Once edge pixels has been detected, numerical analysis is applied to approximate shape and geometry of eyes and mouth regions (Alsmadi, 2016**;** Kucukoglu, 2019; Liu, 2005; Kenneth, 2001; Hill, 2001). Finally, Decision rules has been established to classify normal, sad and happy expressions based on approximate width and height of key facial features. Rest of the paper is organized as follows: Section II contains theory of polynomial approximation to model Bezier curve through detected edge pixels. Section III describes details of existing and extended dataset which follows description of methodology used to extract features. Section IV contain details of rules to build, train and test decision tree in XML format (Salmam, 2016; Ghimire, 2013; Revina, 2018). XML format is very useful technology for data transfer between backend and frontend. The overview of facial expression recognition methodology has been shown in Fig. 1. The proposed system will suppose to comprise these key building blocks: (1) Image Browser (2) Extraction of eyes, mouth and nose (3) Edge Detector (4) Curve Fitter (5) Feature Storage in XML File (6) Training Decision Tree and finally (7) Testing Decision Tree. Section V shows experimental results which follows conclusion and future work.

Polynomi AL Approximation: Problem of designing curves for eyes, mouth and other complexshapeshas been addressed by applying inspiring concept of partitioning unity for Tweening and Bezier curve modeling as given in Equation (1)-(4): Expansion of above basic formulasproduce*n+1*piecesthat can be used for linear interpolation or famous lerp operation [9] between given control points $P = (P_0, P_1, P_2 \ldots\ldots P_n)$ as parameter t varies from 0 to 1. Eq (5) expresses$P$ as affine combination of its Bernstein Polynomial components as follows:

$$1 = (1-t) + t\,; \quad (1)$$

$$1 = \left((1-t) + t\ \right)^2 \quad (2)$$

Similarly $1 = \left((1-t) + t\ \right)^3$ (3)

Generalizing$1 = \left((1-t) + t\ \right)^n$ (4)

$$P(t) = B_{i,n}(t)\,.P \quad (5)$$

$$B_{i,n}(t) = \binom{n}{i} t^i (1 - t)^{n-i} \quad (6)$$

$$for\ i = 0,1,2, \ldots\ldots n\ where \binom{n}{i} = \frac{n!}{i!\,(n-i)!}$$

Usually set $B_{i,n}$= 0, if $i < n\ or\ i > n$. The exponents on the "t " term increase by one as "i" increases and the exponents on the $(1 - t)$ term decrease by one as "i" increases. The Bernstein polynomial of degree 1 is written as:

$$B_{0,1}(t) = t;\ B_{1,1}(t) = (1-t) \quad (7)$$

Where t and $(1- t)$ are weights of two given Points A and B and demonstrates that one can start traveling from Point A and reaches to Point B on a parabolic path. The Bernstein polynomials of degree 2 are written as:

$$B_{0,2}(t) = (1-t)^2; B_{1,2}(t) = 2t\,(1-t);\ B_{2,2}(t) = t^2$$

Similarly, the Bernstein polynomials of degree 3 are:

$$B_{0,3}\,(t) = (1-t)^3$$
$$B_{1,3}\,(t) = 3t(1-t)^2$$
$$B_{2,3}\,(t) = 3t^2(1 - t)$$
$$B_{3,3}\,(t) = t^3$$

Let we obtained six pixel coordinates form eye image as (1, 13), (18, 1), (18, 30), (36, 2), (36, 28) and (54, 15) and laid them on a graph as shown in Fig 2. It can be observed in Fig 2 that direct joining of edge pixels will not yield smooth eye curve. Bernstein functions in Eq(5) and Eq (6) will generate smooth curve using in between points with the help of given control points as shown in Fig 3.After curve approximation, width and height of curve is computed using extreme points. Width is simply the distance between two extreme points while height is calculated by first finding midpoints and then using distance formulae as before.

$$mid - point = (x_{min} + x_{max})/2 \quad (8)$$

$$|d| = \sqrt{(x_{max} - x_{min})^2 + (y_{maz} - y_{min})^2} \quad (9)$$

Data ACQUISITION: We mainly used 3 datasets for training; C-K, JAFEE and PAKFEE. Readers can observe emotion samples in Fig The description of selected images for experimentation is as follows: There are about 2000 images of different subjects available in C-K dataset with several expressions. This database provides balanced saturation images neither dark nor bright. All these images are in Gray scale mode. The most unique part of this dataset is that it has both male and female of different ages. The dimensions of all these images is $\mathbf{640} \times \mathbf{490}$ with bit depth '**8**'. The size of all images of this dataset ranges from 119 KB to 140 KB, but there are some images which are bigger than 200 KB. These images take 1.4 to 1.9 seconds to be processed. From C-K dataset we have chosen 319 images of different persons having Happy, Sad and Neutral expression. Among 318 images, 147 happy, 94 neutrals and 78 sad images has been used to train the decision tree. JAFEE dataset is organized by a Japanese university for research purpose. There are around 216 frontal face images of several expressions. All these are in Gray scale mode. The dimensions of these images is $\mathbf{256} \times \mathbf{256}$ with '**8**' bit depth. Size of these images ranges between 64 to 65 KB. Processing time for these images ranges 0.3 to 0.4 seconds. From JAFEE, we have selected 73 frontal faces from which 216 images are used for training purpose. Among 73, 29 faces are happy while there are 22 images for sad and neutral expression each. We also build our own dataset PAKFE containing 300+ images of seven Pakistani subjects, again with three major expressions Happy, Sad and Neutral. 116 images are selected for training which include all three expressions. The dimensions, Size and Processing time of images should be considered for enhancing system performance. The description of mentioned attributes along with required processing time for all three datasets is summarized in Table I. Table I shows that images of CK dataset take more time as compared to others because of its greater dimensions (pixels). Benefit of this dataset is that it contains images of both genders. On the other hand, JAFEE dataset is quite fast in processing but its drawback is that it neither has images of male nor females of other regions except Japan. PAKFE dataset is quite average as compared to other two but the only issue can be that it does not have female faces. Bar chart in Fig includes comparison of dataset specifications used in all three face expression datasets.

Decision tree Modeling

The core algorithm for building decision trees called ID3 by J. R. Quinlan which employs a top-down, greedy search through the space of possible branches with no backtracking. ID3 uses entropy and information gain to construct a decision tree.A branch with entropy of 0 is leaf node. A branch with entropy more than 0 needs further splitting. The decision tree algorithm runs recursively on the non-leaf branches, until all data is classified. In decision tree training data has been organized into subsets which form tree branches and finally go down to leaf nodes.

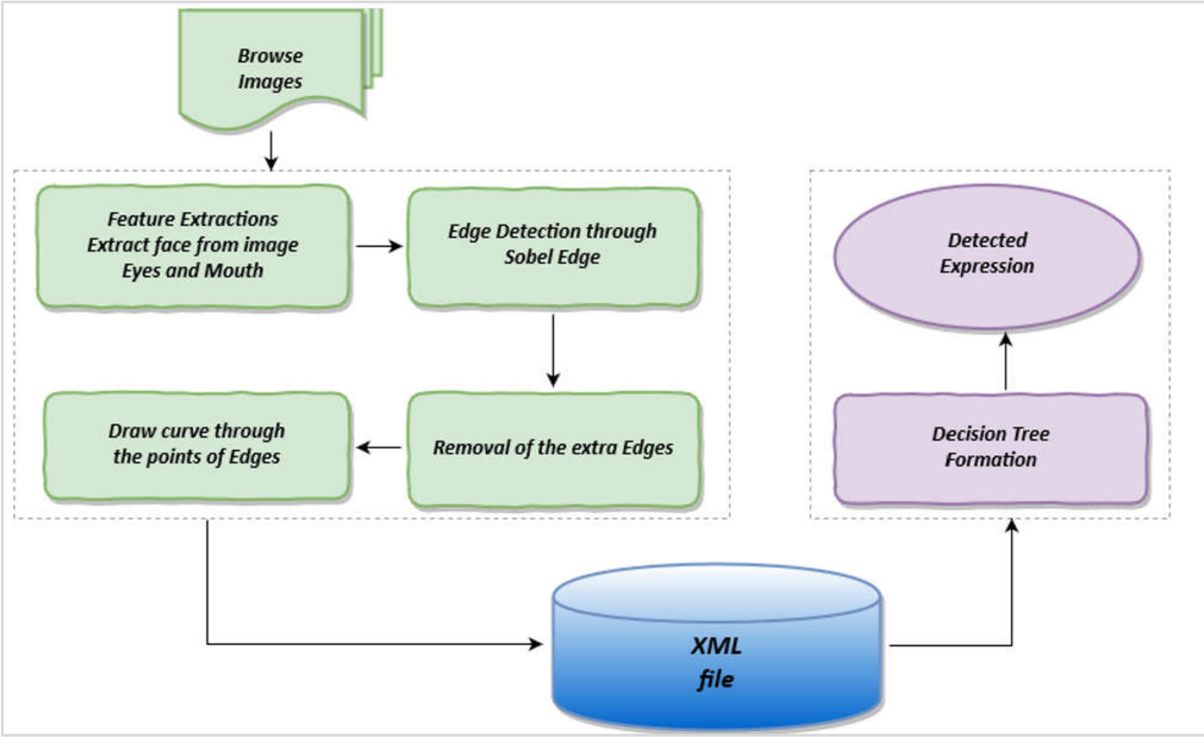

Fig 1. Overview of Decision Tree Based Facial Expression Analysis System

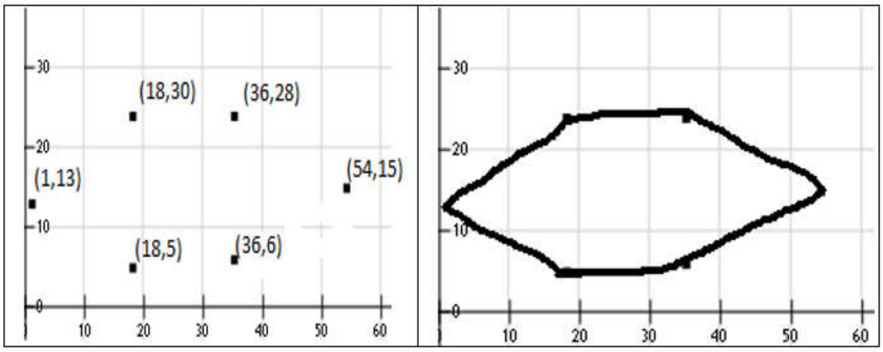

Fig 2. Limitation of joining edge pixels directly

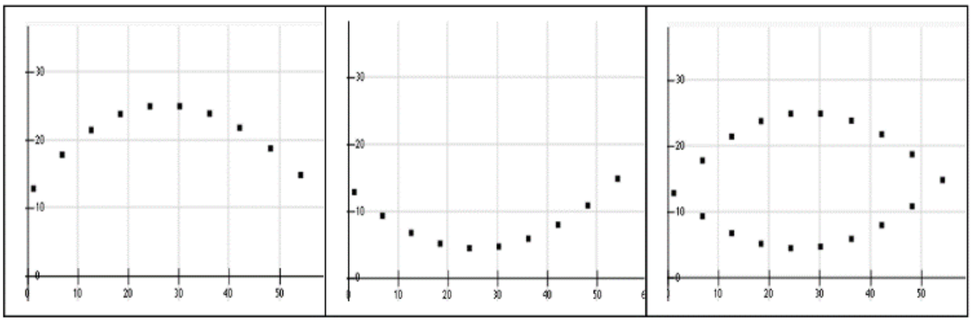

Fig. 3.Curve Modeling using Bernstein polynomial

Table 1.

	CK	JAFEE	PAKFE
Dimensions	640 x 490	256 x 256	300 x 300
Size (KB)	125 – 135	64 – 65	125 – 130
Mode	Gray scale	Gray scale	Digital
Gender	Both	Female	Male
Processing Time	1.4 – 1.9	0.3 – 0.4	0.7 – 1.0

Fig.4. Expression Dataset a) CK (b) JAFEE (c) PAKFE

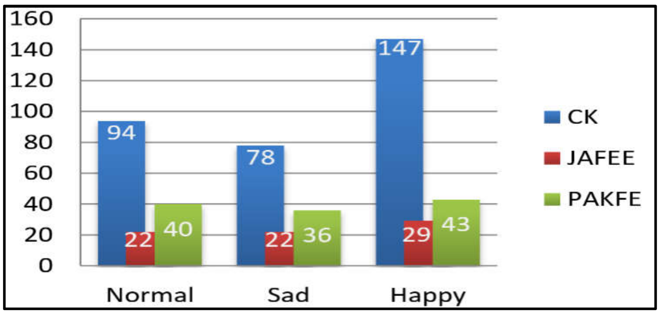

Fig. 5. Statistics of Expression Datasets

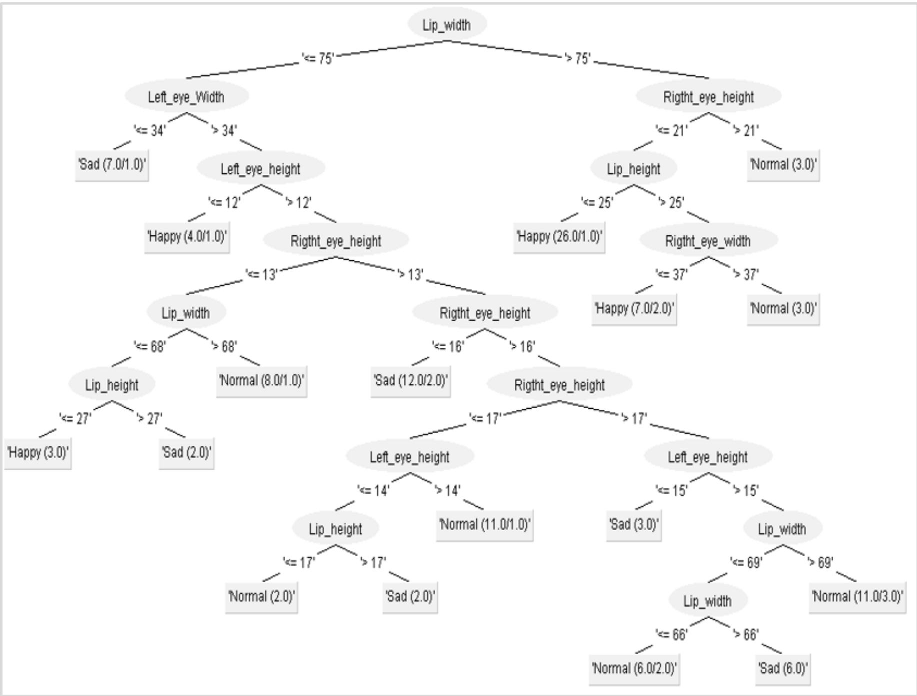

Fig. 6. Decision Tree Model

```
Lip_width <= 75
|   Left_eye_Width <= 34: Sad (7.0/1.0)
|   Left_eye_Width > 34
|   |   Left_eye_height <= 12: Happy (4.0/1.0)
|   |   Left_eye_height > 12
|   |   |   Rigtht_eye_height <= 13
|   |   |   |   Lip_width <= 68
|   |   |   |   |   Lip_height <= 27: Happy (3.0)
|   |   |   |   |   Lip_height > 27: Sad (2.0)
|   |   |   |   Lip_width > 68: Normal (8.0/1.0)
|   |   |   Rigtht_eye_height > 13
|   |   |   |   Rigtht_eye_height <= 16: Sad (12.0/2.0)
|   |   |   |   Rigtht_eye_height > 16
|   |   |   |   |   Rigtht_eye_height <= 17
|   |   |   |   |   |   Left_eye_height <= 14
|   |   |   |   |   |   |   Lip_height <= 17: Normal (2.0)
|   |   |   |   |   |   |   Lip_height > 17: Sad (2.0)
|   |   |   |   |   |   Left_eye_height > 14: Normal (11.0/1.0)
|   |   |   |   |   Rigtht_eye_height > 17
|   |   |   |   |   |   Left_eye_height <= 15: Sad (3.0)
|   |   |   |   |   |   Left_eye_height > 15
|   |   |   |   |   |   |   Lip_width <= 69
|   |   |   |   |   |   |   |   Lip_width <= 66: Normal (6.0/2.0)
|   |   |   |   |   |   |   |   Lip_width > 66: Sad (6.0)
|   |   |   |   |   |   |   Lip_width > 69: Normal (11.0/3.0)
Lip_width > 75
|   Rigtht_eye_height <= 21
|   |   Lip_height <= 25: Happy (26.0/1.0)
|   |   Lip_height > 25
|   |   |   Rigtht_eye_width <= 37: Happy (7.0/2.0)
|   |   |   Rigtht_eye_width > 37: Normal (3.0)
|   Rigtht_eye_height > 21: Normal (3.0)
```

Fig.7. Decision Rules

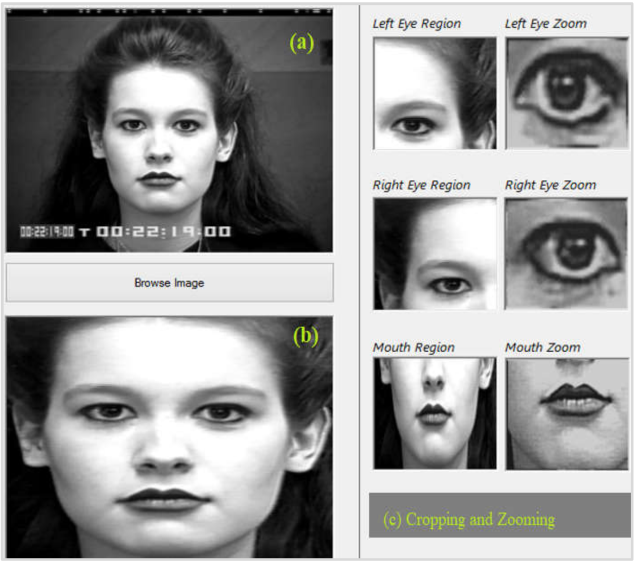

Fig. 8 a) Original Image (b) Extracted Face (c) Viola Jones Feature Detection, Cropping & Zooming

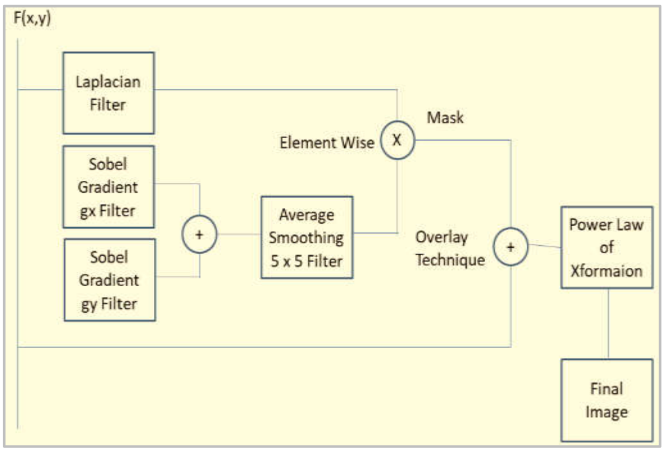

Fig. 9. Combining Filters for Enhanced Edge Pixels Detection

Table 2. Quantified Results of Emotion Recognition

ID	Left Eye Height	Left Eye Width	Right Eye Height	Right Eye Width	Lip Height	Lip Width	Expression
E_N_11	14	40	17	34	16	71	Normal
E_S_4	14	41	18	34	24	74	Sad
E_N_5	23	48	20	43	21	78	Normal
Z_H_13	25	48	19	43	22	83	Happy
Z_N_5	21	41	17	40	18	64	Normal
Z_N_14	22	45	20	44	16	73	Normal
Z_H_9	18	45	16	37	16	77	Happy
Z_N_11	22	43	20	42	20	71	Normal
Z_N_8	18	41	20	43	17	66	Normal
Z_H_1	23	46	18	40	17	73	Happy
E_S_21	13	38	14	32	19	70	Sad
SU_N_8	20	38	17	38	21	61	Normal
E_S_24	13	37	15	33	18	71	Sad
SU_N_15	19	39	18	39	33	66	Normal
E_S_11	14	41	17	34	18	71	Sad

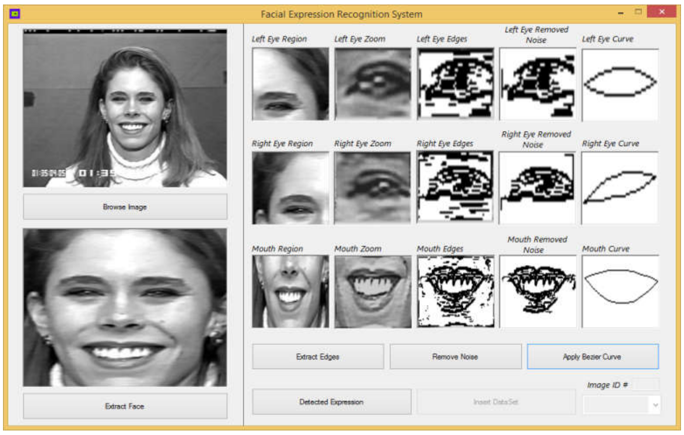

Fig. 10. Bezier curve for Normal Expression

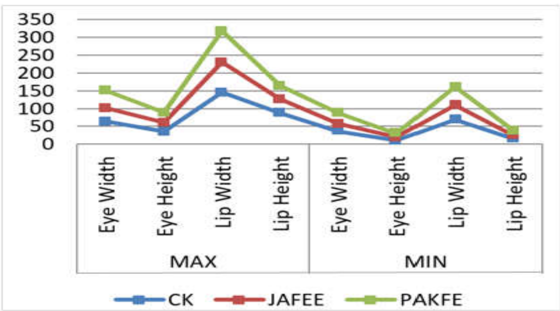

Fig. 11. Max and Min Value of Facial Features

In Fig 6. Lip width from facial expression set is kept at root node because it seems to be more informative then eyes or other facial features based on information gain. Entropy function to calculate gain value of a feature A over available data collection D is given in Eq. (10):

$$G(D, A) = entropy(D) \; - Average\ entropy(D, Values_A) \quad (10)$$

Where $entropy\ (D) = \ E(D) = \sum_{i=1}^{c} -a\ log_2 a$ (11)

'a' is count of positive examples for ith classthus entropy over entire data set is calculated by taking summation over all classes. While average entropy of an attribute or feature over is calculated by taking expected entropy overall possible values of that feature from Eq. (12) as follows:

$$Average\ entropy(S, Values_A) = \sum_i \frac{|S_i|}{|S|} E(S_i) \ (12)$$

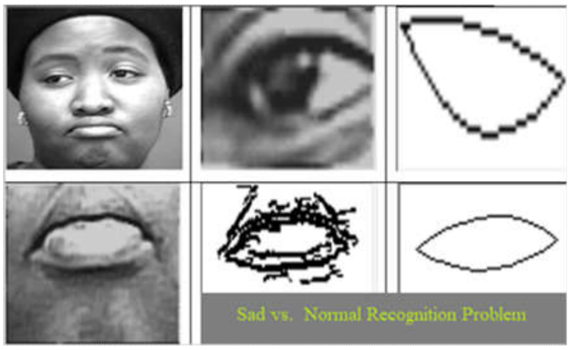

Fig. 12(a). Difficult Example I: Sad vs. Normal

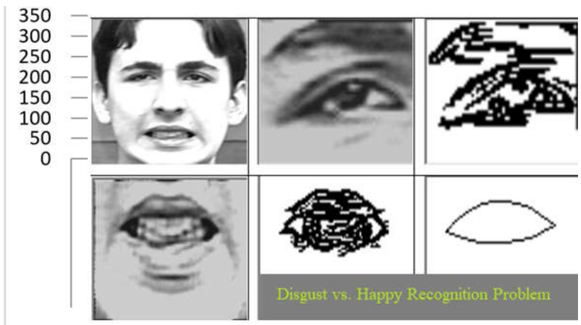

Fig. 12(b) Difficult Example II: Disgust vs. Happy

Table 3. Confusion Matrix over all three datasets

C-K Dataset	Smile	Normal	Sad	Total
Smile	116	18	12	146
Normal	15	66	13	94
Sad	9	24	45	78
Total	--	--	--	318
JAFEE Dataset	**Smile**	**Normal**	**Sad**	**Total**
Smile	25	1	3	29
Normal	1	17	4	22
Sad	4	4	14	22
Total	--	--	--	73
PAKFE Dataset	**Smile**	**Normal**	**Sad**	**Total**
Smile	29	11	3	43
Normal	7	30	3	40
Sad	5	12	16	33
Total	--	--	--	116

$\frac{|S_i|}{|S|}$ is the proportion of subset in which attribute 'A' attain **c** possible values. When a data set contains equal number of positive and negative examples, the entropy takes its maximum values as one showing that entropy always lied between zero and one; $0 < entropy < 1$. A decision tree can easily be transformed to a set of rules by mapping root node to the leaf nodes one by one as shown in Fig. 7.

MATERIALS AND METHODS

Experimentation started with Face Extraction and Viola jones Facial Feature detection as shown in Fig 8 (a) and Fig 8 (b). Outcome of face detection has been kept in the form of face rectangle which is then segmented into left eye, right eye and mouth patch as shown in Fig 8(c) through cropping and zooming. All the desired segments take form of rectangle and have been handled and manipulated by their bottom left corner, width and height with following specifications:

Face (face. rect.X, face. rect.Y, face. rect. Width, face. rect. Height);

Right Eye (0, 0, faceCrop. Width /2, face Crop. Height /2);

Left Eye: (faceCrop. Width/2,0, face Crop. Width / 2, faceCrop. Height / 2);

Mouth: (0, faceCrop. Height / 2, face Crop. Width, face Crop. Height / 2);

Explicit cropping removes eyebrows from eye image and similarly nose from the mouth image otherwise they will appear as noisy edges in edge detection step. The objective of edge detection is to identify facial feature pixels which will be needed during interpolation to complete curve of eyes and mouth. Edge detection is tricky and need combining various image processing filters for e.g. laplacian filter ∇^2 has been applied on original image $F(x, y)$ to highlight fine intensity transitions but at the same time it eliminates background features which will be recovered by combining scaled grayish laplacian image with orginal one if required. Soble gradient filters has been applied to enhance prominent edges but need to combine with smoothing to suppress noise due to derivative operation.The edge detection process combine the best details obtained by laplacian and sobel gradient in $G(x, y)$as shown in Fig 9. Mathematical forumlation of said filtering process can be stated as follows:

$$G(x,y) = c[\nabla^2 F(x,y)] .* \ (M(x,y) \otimes \overline{F(x,y))}$$

Where $M(x,y) = mag(\nabla F) = \sqrt{g_x^2 + g_y^2}$

Where c and γ are positive constants to enhance contrast of image for further processing. As we described earlier in methodology section that we detect edge pixels and use Bernstein polynomial to approximate Bezier curve for eyes and mouth features. Fig. 10 shows obtained Bezier curve sample while Table II show the sample quantified data of facial features obtained from training images. We have noticed max and min values of different features for all datasets and the result is shown in Fig 11. and Fig. 12 (b) shows typical cases where it is difficult to discriminate between sad and normal expression due to shorter eyes and mouth height. Similarly, conflict arises between disgust and happy expression. Results shows that normal and happy expressions are comparatively easy to detect as compared to disgust and sad expression. Confusion matrix for all three data sets based on testing data has been summarized in Table III.

CONCLUSION AND FUTURE WORK

It has been observed that highest lip width image must have happy expression. Minimum eye height has been considered as sad expression. The Expression which are difficult to recognize are sad and disgust and their numerical features will need regularization to properly discriminate from normal and happy expressions. Training of the decision tree based system for subset of expressions is satisfactory and reaches 80% but average testing recognition rate is 75% for all principal emotions namely Happy, Sad and Normal showing need of improvement in terms of including more facial landmarks, real time expression analysis and comparison with other data mining technique.

ACKNOWLEDGEMENT

We would like to thanks all participants who recorded and provide their data for facial expression analysis.

REFERENCES

Alsmadi. M. 2016. “Facial recognition under expression variation”. Int. Arab J. Info. Tech., vol. 13, no. 1A.

Ghimire D. and Lee. J. 2013. “Geometric feature-based facial expression recognition in image sequences using multi-class adaboost and support vector machines”. Sensors, vol 13, no 6, pp. 7714–7734.

Gonzalez, R. C., Woods. R. E. 2008. Digital image processing (3rd ed.). Upper Saddle River, N.J. Prentice Hall, 2008.

Hill. F. S. 2001. Computer Graphics Using OpenGL. EnglewoodCliffs, NJ: Prentice-Hall International, 2nd edition.

Hu, J., Tong, X., Xie, Q., Li. L. 2016. “An Improved, Feature-Centric LoG Approach or Edge Detection”. In: Gervasi O. et al. (eds) Computational Science and Its Applications . Lecture Notes in Computer Science, vol 9787. Springer,ICCSA.

Kenneth. I. J. 2001. On-Line Geometric Modeling Notes. Bernstein polynomials.

Kucukoglu, I., Simsek, B., Simsek. Y. 2019. “Multidimensional Bernstein polynomials and Bezier curves: Analysis of machine learning algorithm for facial expression recognition based on curvature”. App. Math. and Comp.

Liu, Y. S., Seng, W.and Liu. C.W. 2005. “A Comparison of Different Face Recognition Algorithms”. National Taiwan University.

Revina, I.M., Emmanuel. W.R. S. 2018. “A Survey on Human Face Expression Recognition Techniques”. J. King Saud University, Comp. and Info. Sci.

Salmam, F. Z., Medani, A., Kisi. M. 2016. “Facial Expression Recognition Using Decision Trees”.In 13th International Conference Computer Graphics, Imaging and Visualization. At: Beni Mellal City, Morocco.

The Third Emotion Recognition in The Wild (EmotiW) 2015 Grand Challenge. http://cs.anu.edu.au/few/emotionw2015.html.

Viola, P., Jones. M. J. 2004. “Robust Real-Time Face Detection”.Int. J. Comp. Vis., vol 57, no.2, pp.137–154.

Viola, P.A., Jones. M.J. 2001. “Rapid object detection using a boosted cascade of simple features,”, In: CVPR, issue 1, pp. 511–518.

Zhang, H., Zhou, X., Zhan, W., Yu. F. 2019. “Image Denoising Method Based onWeighted Total Variational Model with Edge Operator”. In: Proceedings of the Fifth Euro-China Conference on Intelligent Data Analysis and Applications. ECC 2018. Advances in Intelligent Systems and Computing, vol 891. Springer.

****