Study of Facial Emotion Recognition Techniques

Minal Gajkumar Soundatte[1], Prof.J.A.Shaikh[2]

[1]Minal Gajkumar Soundatte, Padmabhooshan Vasantraodada Patil Institute of Technology, Sangli(Budhgaon)

[2]Prof.J.A.Shaikh, Padmabhooshan Vasantraodada Patil Institute of Technology, Sangli (Budhgaon)

Abstract—**This paper discusses various methods used for emotion recognition from human face images. The stages involved in facial emotion recognition involve various image acquisition, preprocessing, facial region extraction, feature extraction, feature selection and classification stages. Each stage has various aspects with respect to researchers view point. The techniques and their respective key points are collected for discussion in this paper. This paper contributes for forming a platform for understanding various aspects and methods involved in facial emotion recognition system and provide direction for researchers in the field.**

Keywords: Face Region, Facial Emotion, image processing, Feature selection, Classification.

I. INTRODUCTION

As of now a-days, Emotion affirmation is a basic locale in PC vision. It is electronic programming that serves to recognize human opinions, for instance, Fulfillment, Anger, Sadness, Fear, and Disgust, besides, Surprise. Likewise, feeling Affirmation has its foothold on open places. Cases of human sentiments on public spot like genuine dispute between two individuals, driving vehicle with shock and so on. By understanding sensation of the individual, prepared to prevent any inauspicious act or chance. In this paper, unique methodologies and data bases used in later papers are analyzed for look affirmation. Various experts are interested to do explore in facial explanation. Since by the Quote "One individual could be chosen by their reactions not by exercises". Like that, person's perspective can be easily recognized by their sentiments rather than their words.

The human face accepts a critical part in correspondence. The face can impart their opinions through sentiments. Face Expression approach [2] can be isolated into three huge advances so the face in an image is known for extra dealing with, facial part extraction which is the strategy used to ultimately address the looks portrayal which is the movement that bunches the features eliminated in the appropriate verbalizations. The facial enunciation is for recognizing the crucial human tendency like disappointment, fear, euphoria, pity, and shock. Figure1 under showing the general development of feeling affirmation. From the figure, there are three critical times of for the most part structure. First is face revelation task in which first observable features are removed and subsequently face is perceived. The resulting advance is where the facial part extraction and affirmation of focuses on the overall features are taken out at the completion of the last development in portrayal. The general structure of facial emotion recognition system is shown in figure 1.
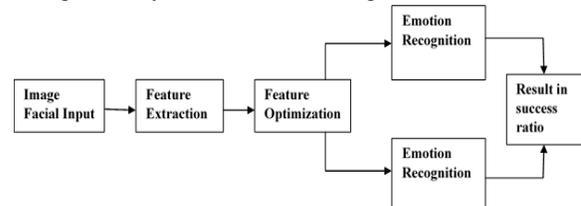

Figure 1: General Facial Emotion Recognition System

II. LITERATURE REVIEW

a.   Techniques for facial region extraction

Face revelation is the one of the crucial fear in PC vision. Viola-Jones face locater strategy utilized in clear face unmistakable proof methods by means of the creator Stefanos Zafeiriou et al in [1]. Ankit Srivastava et al [2] reviewed about Viola Jones assessment. Calculation accommodates four phases. Anyway, haar highlights are utilized to see properties, 2nd Intergal photo utilized for consolidate estimation, 1/3 one is AdaBoost classifier utilized for social affair the picture. The imperative advantage of utilising this adaBoost classifier is pointless statistics on the face are   disposed   of.   At   last,   streaming   classifier

calculations are utilized and the blessings of making use of this evaluation is assists with dealing with regular pictures, dismissal of beguiling superb ID and unbelievably low evaluation time.

In paper [3], producer dissected future face disclosure dataset and introduced WIDER FACE dataset which is extra outstanding and includes number sorts of appearances which include obstacle, positions, scale, etc. Past dataset like ALFW, FDDB, and PASCAL FACE do not have sufficient arranging information. They alluded to as, unique face vicinity strategies, produce poor models. They motive that WIDER dataset from a range of foundation, chips away at the graph with each poor and effective models. Here producer proposed a multi-scale two-stage flood shape to put together dataset. Different method attracted with the direction of section extraction and specific datasets like Cohn Kanade, AR enlightening assortment, JAFFE data base, and so forth was once proposed in [4].

In [5] the producer dissected about one-of-a-kind calculation, for example, Viola-Jones face affirmation and mind-set expulsion vector. While seeing face in the picture, troubles like brightening, aggravation emerges. To vanquish aggravations, pre-managing ought to accomplish face region. Four kinds of techniques have been examined for face affirmation. They are Template matching primarily based approach, Knowledge-based approach, Appearance-based approach, Feature invariant framework in [6].

1   Haar Classifier

Analyzing Haar factors involves increasing or decreasing the size of the pixel pack. It uses Haar-like elements to recognise an image. Using this method, items of various sizes can be diagnosed. In the getting ready stage itself, the Haar classifier will identify a number of factors that are mostly responsible for the face distinguishing proof problem. Because of the low estimation multifarious character, it is appropriate for face proximity in the planning stage as it would signal to high acknowledgement accuracy [27].

2   Adaptive Skin Colour

Adaptable skin-assortment mannequin functions face acknowledgment approach thinking about skin-assortment mannequin to apprehend the face area. This computation suggests a excessive precision as complexion is used for division. In this manner it will in standard be honestly to isolate the face place and non-face district. Anyway, this estimation would not work with distinctive stages of illumination. To keep away from this issue, adaptable gamma therapeutic system is smart to use (yet no longer appropriate regularly surroundings as a end result of excessive computational complexity) [27].

3   Adaboost Contour Points

Adaboost is reasonable to recognize face in continuous ecological components due to low computational complexity and extravagant precision. In this technique, different classifiers can be streamed. It traines the faces and fabricated areas of strength for a which prompts exorbitant acknowledgment accuracy. Then the new face will be strangely, with the existence measured model created through the classifier. It likewise used structure components to recognize face. The structure factors may besides supply a positive accuracy and by and large execution due to the truth the parts removed are considerably less which prompts low computational multifaceted design [27].

b.   Feature Extraction

To acquire excessive precision in face depiction Nianyin Zenga et al in [7] proposed Deep Sparse AutoEncoders (DSAE). The DSAE gain is, it take out characteristic questioning about free learning. It shops essentially indispensable data. [6] To dispense with specific phase in the human face is a sizable occupation in characteristic extraction. Various structures have been evaluated like Principal Component Analysis (PCA), Gabor wavelet, Linear Discriminant Analysis (LDA), Discrete Wavelet Transform (DWT), Bezier distort. A excessive layered solidify, via becoming a member of look and numerical parts. The maker in [9] used direction of motion of face datasets like MMI, CK+,FERA, DISFA, SFEW, MultiPIE. [9] In this paper, the maker proposed a response with the aid of becoming a member of photo preprocessing steps and Convolutional Neural Network.

[5] To get face depiction, two varieties of buildings are assessed in this paper. They are numerical section based totally method and substantially exhaustive layout based totally systems. LBP, Thermal Image Processing and Neural association, Fuzzy reasoning based totally feeling demand are in like manner audited to attain look. Different get-together opinions like Support Vector Machine (SVM), Euclidean Distance, chi rectangular distance are bankrupt down.

[1] Stefanos Zafeiriou et al. maker examined about include extraction methods. Pixel-based aspects make use of pixel fits and manage factor of set as parts. Nonexclusive direct components are made which is a getting sorted out limit. Experiences primarily based include used to be made for floor social match and nonexclusive openness of thing.

Ali Mollahosseini et al in [8], an arranging known as fundamental idea relationship to manipulate the seem problem over many face datasets. The proposed community has two layers known as 4 Inception layer and max pooling.

[10] Author used PCA, LBP and HOG for contain extraction techniques are used for depiction of face. The shut through elements are taken out in every aspect the use of Gabor wavelets with picked scales and path have been proposed by using the maker Ithaya et al [11]. These components are exceeded to a get-together classifier for seeing the region of face locale. The made method is carried out 98% exactness the use of CK dataset.

Rule Components Analysis (PCA) is a method for overseeing seeing plans in data, and supplying the information to consolidate their related features and differences. The appear affirmation the usage of Eigen faces in which PCA is used to avoid components from enter picture. Explicitly they make planning dataset to look at result. Once inputted face image is pre-overseen and distinction and getting equipped dataset which are at this factor figured proper presently thinking about the idea, they separated the fame set into six essential training as displayed through boundless expression(Happy, Surprise, Disgust, miserable, Angry, Fear)[12].

Neighborhood matched Pattern (LBP) primarily based contain extraction framework is used inferable from its incredible mild invariance property and low computational complexity [13]. The close by traits is restricting via the middle well worth and the end result is handled as a twofold number. Expecting that the run pixels regard is greater vital than the neighbor's well worth development 1, anyway zero consequently, it encodes the close by facts beneficially.

Dynamic Appearance Model (AAM) is a certified method for form and floor look and phase extraction. It has been typically used in PC imaginative and prescient applications. AAM produces quantifiable look fashions via becoming a member of a mannequin of structure assortment with a floor gathering. So the AAM makes the shape, floor combo mannequin of orchestrating facial image improvement "Surfaces" are pixel powers of the aim photo [14].

Facial Action Coding System (FACS) used to be made through Paul Ekman and Wallace Friesen in 1976 is a graph for assessing look. FACS depends on the examination of the members of the family between muscle withdrawal and modifications in the face appearance. The Face can be disengaged into Upper Face and Lower Face Action devices [15].Action Units are adjustments in the face completed by using one muscle or a combo of muscles. There are forty six AUs that tackle modifications in seem to be and 12 AUs associated with eye gaze route and head bearing. Haar classifier based totally technique is picked for face openness inferable from its excessive affirmation exactness and relentless execution [16]. Consolidates in truth setting apart associated rectangular shapes in which the cost of the section is the capability of extent of pixel.

c. Features reduction

1 Linear Discriminant Analysis (LDA)

LDA is a method for learning coordinated subspaces. As it might be organised soon, paper [16] used a straight LDA-based request plot for FER. While this was going on, paper [13] utilised LDA to view enunciations by utilising LBP characteristics. In paper [13], LDA is employed to look for projections of hatchets where the records of activity for various lessons are distant from one another while calling for the records of activity for comparable lessons to be close to one another.

2 Fisher Face Method

The Fisher Face approach calls for exact normalisation, registration, and mapping of internal face features [16]. The Fisherface approach was proposed by Peter N. Belhumeur, Joao P. Hespanha, and David J. Kriegman [17] and has been tested on the Harvard and Yale face databases with results that demonstrate lower error rates than Eigenface. Fisher's Linear Discriminant (FLD) approach served as the foundation for Fisherface. The ratio of between-class scatter to within-class dispersion will be maximised [17].

3 Principle Component Analysis (PCA)

The most frequently used strategy is PCA. It is a face affirmation technique considering straight appearance.

The information speculation approach behind PCA isolates input face photos into a foreordained number of specific component pictures. The "Eigen Face" feature picture is one of the fundamental pieces of the basic planning set of face pictures. PCA is in like manner utilized in clinical picture assessment, handwriting examination, lip scrutinizing, talk affirmation, and hand movement assessment [29].

d.   Classification

Support Vector Machine classifier use Eigen values to parent out Eigen faces in [5]. In Euclidean Distance, the rectangular root summation view is regarded as the viable end result of the exceptional picture. In chi rectangular distance, least distance yields the most over the pinnacle well worth amongst the two precis of limits.

The instructive assortments, for example, Cohn-Kanade statistics base and JAFFE instructional listing are utilized to accomplish feeling insistence. Ghulam Ali et al in [10] proposed classifier helped NNE (mind affiliation outfit) assortments relies upon upon three undertakings. A first errand is to set up the blueprint of baselevel classifier. Second mission is to bring gathering set of sport plans. Third undertakings it to deal with the trouble of becoming a member of choices. Creator introduced NB classifier (Naïve Bayes classifier) to settle becoming a member of preferences issue. Convolution frontal cortex community flood is utilized in [2]. It perceives massive edifying collection. It accomplishes extraordinary velocity and greater accuracy. [4] Classifier is utilized to bundle disposed of elements. SVM, Nearest Neighbor and SVM classifiers are utilized. Planning of Convolutional Neutral Network proposed with the aid of the creator in [9]. It consists of 5 layers, for example, convolutional layers, sub-surveying layers and definitely associated layer. Here, first layer remove visible parts. First layer from an ordinary standpoint spin round corners, shapes, etc. Spatial purpose is diminished in the following layer.

Face components can be perceived in the 0.33 and fourth layer. The final layer receives the components in time-honored and produce eventual effect of stage of sureness of the alluded to verbalizations. [8] Author using waiting for pass records base. Creator

introduced essential frontal cortex format and accomplishes stage of reflection and portrayal to tune down tangled plan. [7] In this paper, producer proposed a necessary forlorn autoencoders, a large frontal cortex community which is made via stacked little autoencoders and subtle maxclassifier. While organizing arranging, establishing development is to coordinate the stack forlorn autoencoders and their consequence is feed to the greater layer of DSAE. Next mess up engendering method is utilized to make overhauls of cutoff factors which are utilized in the necessary frontal cortex affiliation.

Well functions a bunch of factual mannequin to painting the measurable way of behaving of a signal [21]. In paper [21], they lookup 5 selections of HMM fashions which are; (I) Left-Right HMM, (ii) Ergodic HMM, (iii) Emotion-Specific HMMs, (iv) Multilevel HMM and (v) Mixture of HMM and intelligence organization. End is, the HMM accomplishes higher grouping both three or 5 kingdom mannequin utilized free articulation or combo of articulation making use of staggered nation models.

NN executes a nonlinear reduce of the information dimensionality. It plans a factual desire about the classification of the articulation that has been noticed. Each end result unit will gauges on the probability of the analyzed articulation has a area with the associated classification [22].

Support Vector Machine (SVM) is one of the nicely recognized measurable tactics utilized in AI to damage down records utilized for order and relapse examination. SVM utilized exclusive piece functionality to format facts in enter area into high-layered spotlight areas [21].

Bayesian Network (BN) is a graphical mannequin which equipped to exhibit a life like and herbal connections by using making use of probabilistic methodology amongst a bunch of attribute [21].

Scanty Representation Method (SRM) makes use of a direct combine of all the education exams to "meagerly" tackle and signify the face picture. "Meagerly" is alluding to sure coefficients of the direct combine are equal or close to zero [18].

Table 1 offers define of the facial district extraction techniques.

Algorithm	Accuracy	Performance in real-time
Haar classifier	Accuracy is high for face detection due to suitable Haar features	Computational complexity is very less due to a set of features which contribute the maximum, for the face detection problem in a training phase
Adaptive skin color	Accuracy is good as skin color in identified easily but fails in different levels of illumination	Approaches such as adaptive gamma corrective method is used to get rid of illumination problem which leads to high computational complexity and is not suitable in real-time environment
Adaboost contour points	High detection accuracy due to strong classifier single face is detected using contour points due to which the accuracy is good	Computational cost is less due to trained model low computational complexity due to less number of features

Table 1: Comparative study of different facial region extraction methods

Table 2: Comparative different existing methods for facial emotion recognition

Paper	Expressions	Face Detection	Feature Extraction	Expression Classification	Data set	Accuracy
[3]	Happy, Anger, Sad, Surprise, Normal	Luxand Face Recognition	Angle And Distance Method	SVM	Real-time	85.6%
[5]	Anger, Contempt, Disgust, Fear, Happy, Sadness, Surprise, Neutral	AAM, Hog, PCA	DSAE	Softmax Classifier	CK+ JAFFE	95.79 %
[7]	Happy, Sadness, Surprise, Anger, Disgust, Fear, Neutral	Define Distance between Two Centers of Both Eyes	Wavelet Entropy	Jaya Algorithm	[15]	96.8%
[8]	Happy, Sadness, Surprise, Anger, Disgust, Fear, Neutral	Viola-Jones's Haar-like feature cascade detector	PCA, Fisher Face + HOG	SVM	CK+	81%
[9]	Neutral, Fatigue	Hog	Facial Landmarks Points (Cascade of Regression Tree)	SVM	CK+ PICS	82.79 %
[11]	Joy, sad, surprise, fear, anger, disgust, neutral	Hog	Face Alignment with Regression Tree (Landmark Detection Algorithm)	Multilayer Perceptron	JAFFE	88.03 %

III. CONCLUSION

This essay adds to the body of research on techniques for facial emotion recognition. Different face region extraction techniques from the input image are used by the facial recognition system. Then, features are taken from the extracted face region. Following that, the retrieved characteristics are optimised to choose the signature features. An artificial neural network classifier is trained using signature features. Thus, the trained classifier is prepared to recognise facial emotions. Typically, multiclass support vector machines outperform traditional classifiers. For better feature selection and optimization for quicker classifier convergence, the particle swarm optimization and micro genetic algorithm might be taken into consideration.

In the near future, the research of facial emotion detection may also improve social feedback and interactions between Human Robot Interfaces (HRI).

The geometric portion of the face is mostly used for emotion detection (eg; eyes, eyebrow, and mouth). The review takes controlled experimentation, real-time data, and uncontrolled imagery into account. The primary concerns are on the validity of the research methods used.

REFERENCE

[1] Stefanos Zafeiriou, Cha Zhang and Zhengyou Zhang, "A Survey on Face Detection in the wild: past, present and future" Computer Vision and Image Understanding Volume 138, September 2015, Pages 1-24.

[2] A survey of face detection algorithms Ankit Srivastava,Suraj Mane, Aaditya Shah, Nirmit Shrivastava, Prof. Bhushan Thakare,2017 IEEE International Conference on Inventive Systems and Control.

[3] "WIDER FACE: A Face Detection Benchmark", Shuo Yang1 Ping Luo, Chen Change Loy, Xiaoou

Tang, IEEE Conference on Computer Vision and Pattern Recognition

[4] "Literature survey on Face and Face Expression Recognition", Anil J, Dr. L. Padma Suresh, 2016 International Conference on Circuit, Power and Computing Technologies.

[5] "Facial Expression Recognition: A Literature Survey", Assoc Prof. Mrs. Jyothi S Nayak, Preeti G, ManishaVatsa, Manisha Reddy Kadiri, Samiksha S International Journal of Computer Trends and Technology (IJCTT) – Volume 48 Number 1 June 2017.

[6] "A Review: Facial Expression Detection with its Techniques and Application", Neha Bhardwaj and Manish Dixit, International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.9, No.6 (2016), pp.149- 158

[7] "WIDER FACE: A Face Detection Benchmark", Shuo Yang1 Ping Luo, Chen Change Loy, Xiaoou Tang, IEEE Conference on Computer Vision and Pattern Recognition

[8] "Literature survey on Face and Face Expression Recognition", Anil J, Dr. L. Padma Suresh, 2016 International Conference on Circuit, Power and Computing Technologies.

[9] "Facial Expression Recognition: A Literature Survey", Assoc Prof. Mrs. Jyothi S Nayak, Preeti G, ManishaVatsa, Manisha Reddy Kadiri, Samiksha S International Journal of Computer Trends and Technology (IJCTT) – Volume 48 Number 1 June 2017.

[10] "A Review: Facial Expression Detection with its Techniques and Application", Neha Bhardwaj and Manish Dixit, International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.9, No.6 (2016), pp.149- 158

[11] "Facial expression recognition via learning deep sparse autoencoders", Nianyin Zenga ,Hong Zhanga , Baoye Song, Weibo Liuc, Yurong Li d,e, Abdullah M. Dobaief, Neurocomputing· September 2017.

[12] "Going Deeper in Facial Expression Recognition using Deep Neural Networks", Ali Mollahosseini, David Chan, and Mohammad H. Mahoor, 2016 IEEE Winter Conference on Applications of Computer Vision.

[13] "Facial Expression Recognition with Convolutional Neural Networks: Coping with Few Data and the Training Sample Order", Andre Teixeira Lopes, Edilson de Aguiar , Alberto F. De Souza , Thiago Oliveira-Santos, Pattern Recognition · July 2016.

[14] Andrew Ryan;Jeffery F. Cohn, Simon Lucey;Adam Rossi, "Automated Facial Expression Recognition System," 2009 IEEE

[15] C.P. Sumathi, T. Santhanam and M.Mahadevi, "A Automatic Facial Expression Analysis A survey," International Journal of Computer Science & Engineering Survey (IJCSES) Vol.3, No.6, December 2012.

[16] Ms.Aswathy.R "A Literature review on Facial Expression Recognition Techniques," IOSR Journal of Computer Engineering (IOSR-JCE) 2013.

[17] P. Hespanha, D. J. Kriegman, and P. N. Belhumeur, "Eigenfaces vs . Fisherfaces : Recognition Using Class Specific Linear Projection," vol. 19, no. 7, pp. 711–720, 1997.

[18] Y. Xu, Q. Zhu, Z. Fan, D. Zhang, J. Mi, and Z. Lai, "Using the idea of the sparse representation to perform coarse-to-fine face recognition," Inf. Sci. (Ny)., vol. 238, pp. 138–148, 2013.

[19] D. A. Pitaloka, A. Wulandari, T. Basaruddin, and D. Y. Liliana, "Enhancing CNN with Preprocessing Stage in Automatic Emotion Recognition," Procedia Comput. Sci., vol. 116, pp. 523–529, 2017.

[20] S. L. Happy, A. Dasgupta, P. Patnaik, and A. Routray, "Automated alertness and emotion detection for empathic feedback during Elearning," Proc. - 2013 IEEE 5th Int. Conf. Technol. Educ. T4E 2013, pp. 47–50, 2013.

[21] M. H. Siddiqi et al., "A Brief Review of Facial Emotion Recognition Based on Visual Information," 2018 IEEMA Eng. Infin. Conf. eTechNxT 2018, vol. 5, no. 1, pp. 196–201, 2018.

[22] M. Pantic, M. Pantic, and M. Rothkrantz, "Automatic Analysis of Facial Expressions," IEEE Trans. Pattern Recognit. Mach. Intell., vol. 22, no. 12, pp. 1424 – 1444, 2000.

[23] Z. J. Wang et al., "Article in press," vol. 0, pp. 1–11, 2017.

[24] S. Tolba, A. H. El-baz, and A. A. El-Harby, "Face Recognition : A Literature Review," Int. J. Signal Process., vol. 2, no. 2, pp. 88–103, 2006.

[25] Y. Gao and K.H. Leung, "Face recognition using line edge map," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 6, June 2002.

[26] S. Du, Y. Tao, and A. M. Martinez, "Compound Facial Expressions of Emotion," Proc. Natl. Acad. Sci., vol. 111, no. 15, pp. E1454– E1462, 2014.

[27] S. Deshmukh, M. Patwardhan, and A. Mahajan, "Survey on RealTime Facial Expression Recognition Techniques," IET Biom., pp. 19, 2015.

[28] Y. Tian, "Evaluation of Face Resolution for Expression Analysis", CVPR Workshop on Face Processing in Video, 2004.

[29] M. Rahman, "A comparative study on face recognition techniques and neural network," arXiv Prepr. arXiv1210.1916, vol. 3, no. 5, pp. 155– 160, 2012.