

A Review on Emotion Recognition using facial expression and speech

¹Shinde A.R., ²Raut S.D. and ³Khanale P.B.

^{1,2}School of computational sciences, Solapur University, Solapur (India)

³Dnyanopasak College, Parbhani (India)

ARTICLE DETAILS

Article History

Published Online: 22 Dec 2018

Keywords

FFT, Discrete Cosine Transform, Discrete Wavelet Transformation.

ABSTRACT

This paper presents the literature related to the databases, features, pattern classifiers for emotion recognition from facial expression and speech. The interaction between human and computer will be more natural if computer are able to understand emotions of human being.

Discrete Cosine Transform (DCT), Discrete Wavelet Transformation (DWT, Fast Fourier Transform (FFT), Singular Value Decomposition (SVD), SUSAN edge detection operator, facial geometry, edge projection analysis, are used to extract the important features for recognizing emotion from facial expression. Hidden Markov Model (HMM) and Artificial Neural Network (ANN) has been proposed to classify emotions. Most of the researcher has be used facial and speech approach for finding human emotions. As per the literature review facial expression gives accurate recognition rate as compare to speech features.

1. Introduction

Recognition and extracting various emotions and then validating those emotions from the facial expressions have become important for improving the overall human computer interaction. Emotions play an important role in social interactions and facility rational decision making and perception.

If we go back then we will find facial analysis started in the ninetieth century. Darwin [1] presented already in 1872 the universality of facial expressions and their continuity in man and animals and claimed among other things, that there are specific inborn emotions, which originated in serviceable associated habits. Later on, in 1971, Ekman and Friesen[2] postulated six primary emotions that possess each a distinctive content together with a unique facial expression these are also referred to as basic emotions such as happiness, sadness, fear, disgust, surprise and anger.

To achieve specified objective of the research work, I have referred various research articles and survey papers regarding with emotion recognition system. I found there are broadly two approaches are used for recognizing human emotions. One is by using Facial Expressions and second is by using speech samples of different peoples in different languages. To carry out proposed research work i.e. Development of Emotion Recognition system I have referred research article in which both approaches are considered. In the following section I have studied and observed some observations of each paper.

Elham Bagherian, Rahmita. Et. Al. 2009[3], in this paper author writes that facial feature extraction has some problems which must be considered and solved in future work. Some problems of facial feature extraction are given as follows:

Small variation of face size and orientation can be affected the result. As the input image comes from the webcam, the room condition has different brightness; shadows and cleanness which can failed the process.

Human faces have a verity of emotions by many different expressions, but this system can detect the corner of the features in the case of sad, happy and surprise. The first stage is cropping the face area, as soon as the picture is taken from the webcam the second part of preprocessing is prepared by resizing that cropped image. To adjust the contrast and brightness of the image in order to remove noises built in MATLAB functions are used then it is converted to gray scale image, because the corner detector can only be applied on gray level.

The facial feature extraction method was implemented in MATLAB and examines in four different expressions are sad, happy, surprise and neutral.

author	technique	no. of feature	video/still-frontal/rotated
T. Kanade, 1997	geometry-based	eyes, the mouth and the nose	still-frontal
A. Yuille, D. Cohen, and P. Hallinan, 1989	template-based	eyes, the mouth, the nose and eyebrow	still-frontal
C. Chang, T.S. Huang, and C. Novak, 1994	color-based	eyes and/or mouth	still and video-frontal initially in a near frontal position and therefore both eyes are visible
Y. Tian, T. Kanade, and J.F. Cohn, 2002	appearance-based approaches	eyes and mouth	still-frontal and near frontal with different expression

In this paper researcher has been tried to review the works done in facial features extraction, and proposed the geometry based technique by the corner detection for extracting the facial feature points. Also their experimental results shows the system can works in the different orientation and expression.

Oh-Wook Kwon. Et. Al. 2003[4], author has discussed the motions recognition by speech in one of the broad research area for Human Computer Interaction(HCI) or affective computing. The major problem in emotion recognition is that there is no common database for it. So that researchers cannot

take the previous results for future work. Performance of emotion recognition depends on how we can extract the relevant features invariant to speaker, language and contents.

Author extracted the base features for each frame and formal feature streams by adding velocity acceleration components. For emotion recognition log energy, formant coefficients used as base features. The pitch and energy plays an important role in recognition of emotions. Using statistics computed from the base features, they analyzed the effects of the features in emotion recognition.

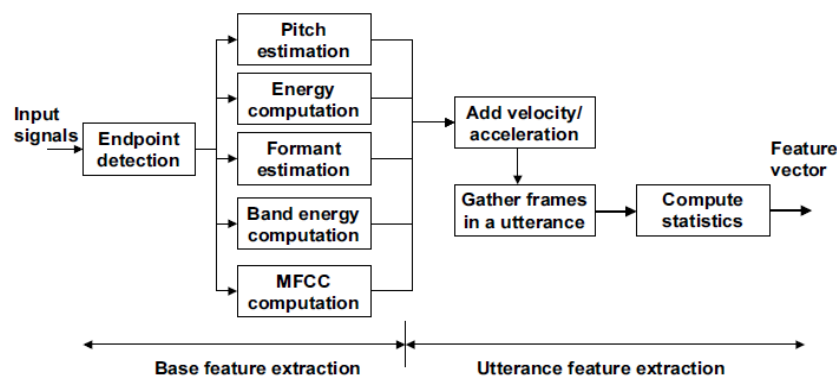


Figure 1: Block diagram of the feature extraction module.

Dom Morrison, et. Al. 2007[5], Machine based emotional intelligence is recruitment for more natural interaction between human and computers interfaces. Emotional intelligence categorized into four branches. Author process two new fields i.e. stacked seminisation and variation on majority voting. Stacked seminisation includes the several Meta learning algorithms. A variation on majority voting is collection of unique classifies. Author compares two emotional data sources natural, spontaneous emotional speech. This comparison gives advantages and disadvantages of both acquisition methods and how these methods affect on vocal emotion recognition.

J.K. Kani Mozhi et. al. 2009[6], in the proposed work author develops a system which describe and evaluates a face expressional recognition system using geometric structure representation of various expression like sad, angry, disgust, happy etc.

Charles Darwin defines the emotional expression as behaviors that evolved as a mechanism of communication. Author uses a facial feature extraction method which evolves geometry and behavior features. Author develops a facial expression on behavioral system which works on basic principle set of Facial Action Coding System (FACS). The FACS is used because it have comprehensiveness which could measure all possible, visible discernible facial action. The author uses another method i.e.geometry on expression for effectiveness, flexibility of facial recognition system.

Xia Mao et. al. 2007[7], author has discussed the Speech emotion recognition, as a vital part of affective human computer interaction has become a new challenge to speech processing. In this paper, a hybrid of hidden Markov

Model(HMM) and Artificial Neural Network(ANN) has been proposed to classify emotions, combining advantage on capability to dynamic time warping of HMM and pattern recognition of ANN; Optimal state sequences, exported from HMMs, are normalized to be one of the input of ANN; hence different methods of state normalization are compared. Adopting Beihang University Database of emotional speech (BHUDIES), comparison between isolated HMMs and hybrid of HMMs/ANN proves that the classifier introduced in this paper is more effective, and the average recognition rate of five emotion states has reached 83.9%

S.P. Khandait et. al. 2011[8], In this paper author highlights on problems of automatic facial feature extraction from a still frontal posed image and classification and recognition of facial expression and hence emotion and mood of a person is presented. Author used Feed forward back propagation neural network as a classifier for classifying the expressions. In this paper morphological image processing operations are used for face portion segmentation and localization. The basic facial features like eyebrows, eyes, mouth and nose are extracted using SUSAN edge detection operator, facial geometry, edge projection analysis. This proposed work carried out by using JAFFE facial expression database, and gives better result in terms of 100% accuracy for training images and 95.26% accuracy for test data set.

Fig. 1



Fig. 1 Few samples of facial expressions of person YM

Authors	No. of subjects Used	Images Tested	% accuracy
Kobayashi and Hara[28]	15	90	85
Zhang[12]	10	213	90.1
Lyons et. al.[28]	10	193	92
Sebe et. al.[28]	-	-	85-95
Kulkarni SS et. al.[28]	62	282	90.4
Chang JY,Chen JL [29]	08	38	92.1(for 3 expressions)
Our approach	10	30	96.42

Mandeep Kaur et. al. 2011[9], this paper explores and compares different techniques for automatic recognition of facial expression. The different techniques such as Principal Component analysis (PCA), PCA with SVD (Singular Value Decomposition) is used. In this paper author shows PCA with SVD is superior to former techniques in case of recognition rate. In this paper author used JAFFE standard database and achieves recognition rate for all the facial expression.

Jyoti Rani et. al. 2014[10], In this paper the author overviewed the automatic facial expression recognition systems and various research challenges. These systems involve face recognition, feature extraction and categorization. As per the survey, different kinds of techniques can be used for achieving better recognition rate. Emotion detection using facial expression is a universal issue and causes difficulties due to uncertain physical and psychological characteristics of emotions that are linked to the traits of each person individually. Therefore, research in this field will remain under continuous study for many years to come because many problems have to be solved in order to create an ideal user interface and improved recognition of complex emotional states is required.

Devi Arumugam et. al. 2011[11], In this paper researcher developed an algorithm which is used to identify exact emotion of the person by using facial expressions to identify researcher used a Radial Basis Function Network (RBFN) for classification and Fisher's Linear Discriminant (FLD), Singular Value Decomposition (SVD) for feature selection.

MU-CHUN SU et. al. 2007[12], The proposed system is able to automatically perform human face detection, feature point extraction and facial expression recognition from image sequences. The system executes in two steps, in the first step three multi-layer perceptrons (MLPs) are separately trained to recognize action units involving in the eyebrows, the eyes and the mouth regions. Then individual expression network was trained to recognize five basic facial expressions.

Shivani Gangwar et. al. 2013[13], In this paper feature extraction is done by using 2D Discrete Cosine Transform (2D DCT) which changes the image data from spatial to the frequency domain. 2D DCT is calculated for matrices of eye and mouth which is again resulted into two 64 x 64 matrices. In this matrix low frequency components are located at the top left corner of the matrix whereas high frequency components located at the bottom right corner of the matrix. Since high frequency components are more variant across images, low frequency components are selected. 66 coefficients for mouth portion and 66 coefficients for eye portion have been extracted. Thus there are 132 coefficients which form a feature vector and it is used to train the pattern recognition network i.e. Feed Forward Neural Network which uses back propagation algorithm to train the classifier for input data against given target data i.e. all six images. The trained classifier is then simulated to test new real world data and identify all six basic emotions i.e. happy, sad, anger, fear, surprise and neutral for simulation.

G.U. Kharat et. al. 2008[14], In this paper total seven emotions were considered for their recognition namely angry, disgust, fear, happy, sad, surprise and neutral. Various feature

extraction methods such as Discrete Cosine Transform (DCT), Fast Fourier Transform (FFT), Singular Value Decomposition (SVD) are used to extract the important features for recognizing emotion from facial expression. Various feature extraction techniques are compared to check which one is the best.

As per the results shown in this paper emotion recognition result on training and testing data set for optimally designed PCA when DCT is used. The accuracy of recognition is 95.47% on train dataset and 100% on test dataset for all the emotions. On the other hand emotion recognition results on training and testing data set for optimally designed MLP when FFT is used. The accuracy of recognition is 100% on the train data set and 88.10% on test dataset for all the emotions.

Pushpaja V. Saudagare et. al. 2012[15], In this research article the author proposed four steps: fact detection, pre-processing, principle component analysis (PCA) and classification. The proposed technique is depend on coding and decoding method. First of all extract the information and

encoded it then matched with the database of model. In this paper the automatic facial expression recognition system are overviewed. The neural network approach is based on face recognition, feature extraction and categorization. The approach of facial expression recognition method involves the optical flow method, active shape model technique, principle component analysis algorithm (PCA) and neural network technique.

2. Conclusion

As per the review of various papers including facial feature extraction and speech features emotion recognition can be achieved by using different methods. By considering two approaches i.e. Facial and speech most of the researcher tried to find human emotions. As compare to speech features, facial expression features plays a vital role in emotion recognition of human being. The facial features gives most accurate results than speech features.

References

1. Charles Darwin, "Expression of the Emotions Man and Animals", London: Jhon Murray, Albemarle street. 1872
2. Paul Ekman, "Facial Expressions", Handbook of Cognition and Emotion, John Wiley & Sons Ltd, 1999.
3. Elham Bagherian, Rahmita.Wirza.Rahmat and Nur Izura Udzir, "Extract of Facial Feature Point", "International Journal of Computer Science and Network Security", Vol.9, No.1, January 2009.
4. Oh-Wook Kwon, Kwokleung Chan, Jiucang Hao, Te-Won Lee, "Emotion Recognition by Speech Signals", "EUROSPEECH 2003 - INTERSPEECH 2003 8th European Conference on Speech Communication and Technology Geneva, Switzerland, September 1-4, 2003.
5. Domm Morrison, RuiLi Wang, Liyanage C. De Silva "Ensemble methods for spoken emotion recognition in call-centers" Speech Communication 49(2007) 98-112
6. J.K. Kani Mozhi and Dr. S.D. Wahida Banu "Face Expressional recognition using geometry & Behavioral Traits" International Journal of Computer Science and Network Security, Vol.9 No.8, August 2009
7. Xia Mao, Lijiang Chen Bing Zhang, International Journal of Computers, Vol.1, Issue 4., 2007.
8. S.P. Khandait, Dr. R.C.Thool, P.D. Khandait "Automatic Facial Feature Extraction and Expression Recognition based on Neural Network", International Journal of Advanced Computer Science and Applications, Vol.2, No.1., 2011
9. Mandeep Kaur, Rajeev Vashisht, "Comparative Study of Facial Expression Recognition Techniques", International Journal of Computer Applications, Vol. 13, No.1, January 2011
10. Jyoti Rani, Kanwal Garg, " Emotion Detection Using Facial Expressions -A Review", International Journal of Advanced Research in Computer Science and Software Engineering. Volume 4, Issue 4, April 2014
11. Devi Arumugam, Dr. S. Purushothaman, "Emotion Classification using Facial Expression" International Journal of Advanced Computer Science and Application Vol. 2, No. 7, 2011.
12. MU-CHUN SU, YIJWU HSIEH, AND DE-YUAN HUANG, "A Simple Approach to Facial Expression Recognition", International Conference on Computer Engineering and Application, Gold Coast, Australia, January 2007
13. Shivani Gangwar, Shshwat Shukla, Dr. Deepak Arora, "Human Emotion Recognition by using Pattern Recognition Network", International Journal of Engineering Research and Applications, Vol. 3, Issue 5, Sep-oct 2013.
14. G. U. Kharat, Dr. S. V. Dudul, "Performance Comparison of Facial Feature Extraction Techniques in Designing Human Emotion Recognition System Using Optimal PCA Neural Network", International Journal of Computer Science and Applications Vol.1, No.2, August 2008, ISSN 0974-1003.
15. Pushpaja V. Saudagare, D.S. Chaudhari, "Facial Expression Recognition using Neural Network- An Overview", International Journal of Soft Computing and Engineering, Vol. 2, Issue. 1, March 2012