Emotion Recognition from Facial Expression and Voice for Single User

V. Geetha

Associate Professor, Department of IT, Pondicherry Engineering College

vgeetha@pec.edu

Abstract: Affective computing is one of the emerging trends in computer field. In the emotion recognition system machine act as human friend. Machines are mobile phones, laptop, tablets and desktop system. A person emotion can be analyzed by facial expression and voice (speech or audio). In this paper propose a machine interact to the user and analysis the user emotion. Input as taken from webcam (image or picture) of user. Images compared to training set and recognize the user emotion. Emotion recognition system used in different field E-Learning, Artificial Intelligence, Medical field, Games (Entertainment), and Psychology.

Keywords: Emotion recognition system, machine interact and facial expression.

1. INTRODUCTION

Human Computer Interaction are play efficient role in Emotion Recognition System. The Affective computing, Dr. Rosalind Picard of MIT Media Laboratory coined the term Affective Computing in 1994 and published the first book on Affective Computing in 1997. In this field related to Artificial Intelligence (AI), Virtual Reality (VR) and Human Computer interaction (HCI). Many researchers interest to detect the human emotion and mood. In past few decades many researches done in the field of emotion recognition system. A machine predict the user emotion by facial expression, voice, hand gesture, images, videos, movies, audio, body gesture etc., Then analysis the user emotion and giving the correct response for those emotion state.

Emotion detection is currently immense area in the research of medical and computer science field. Many researchers doing big amount research in this field for patient's psychological problem and speech- impaired children and lonely person. In computer oriented field many applications are used surveillance camera, database management system for image based, Tutoring system, Smart homes having robots and entertainment. In social emotion detection main aims is communication. Some speech impairment children and autistic person uses the emotion recognition system. Using non-verbal communication and face to face communication are recognized their feelings by emotion only. Mainly for people suffered from autism to understand the other person's mood. A portable emotion detectors are used autistic people to interact with the outside world by recognizing the facial expression.

Emotion recognition system detected by different methods are facial expression, speech(audio and voice), body gesture, input devices, brain signal, textual input, keystroke dynamics and mouse movements, physiological signal, multimodal features, Bio-signals(Brain Signal, skin temperature, blood pressure, heart rate, respiration rate) and body movements(for example limbic movement). Device used for recognize the emotion are mobile phones, camera, movies, videos, surveillance camera, FLIR thermal camera, desktop webcam, Electroencephalogram (EEG), Electrocardiography (ECG), Electrodermal activity (EDA), Electromyogram (EMG), speaker, microphone, keyboard and mouse movements.

Researchers classify basic emotion states are happy, sad, fear, surprise, disgust and anger. In facial expression classify the positive and negative state of motion as happy and anger respectively. The first step of emotion detection is face recognize system where face is present in image. Face detection is complex task because variable in face gesture, pose and lighting condition. Some of the challenges in face detection are variation in shape, color, size, structural component, imaging condition, lighting direction, camera position may change, direction of image taken, noisy in image acquisition, facial expression, occlusion in group of people.

In cognitive science and neuroscience, there have been two leading models describing how humans perceive and classify emotion. The emotion are classify into the continuous and the categorical model. The continuous model defines each facial expression of emotion as a feature vector in a face space. In contrast, the categorical model consists of C classifiers, each tuned to a specific

emotion state as happy or surprise.

Face detection algorithm are use to analyses the facial data to training set compare data set from the database. Different types of algorithm are used to emotion recognition system. Classifiers are used to detect the accurate emotion detection. So selecting the classifier is the important process because successful classifiers are used to detect accurate and quick emotion detection. Different classifiers are linear discriminant classifiers (LDC), k-nearest neighbour (k-NN), Gaussian mixture model (GMM), support vector machines (SVM), artificial neural networks (ANN), decision tree algorithms and hidden Markov models (HMMs).

In the many of the system database are data-dependent. It makes biggest challenges in voice database as may used in speech. Speech recognition recognise the emotion of user. In the database has the training set of classifier are present choosing the correct set from that. But it gives high success rate of emotion detection by offering the high quality audio and balanced classes. These databases are use in method of evaluating and comparing different affect recognition systems.

Human computer interactions are used in many application. In past few years HCI are improved rapidly. E-learning is one of the application to recognize the student emotion. He/she able understand teacher slide. Using the emotion detection system to analyses the every student emotion. For example in MOOC production system able to predict the student mood and doubt can be clarified. Tutoring system recognise the emotion and make positive attitude to maximize the learning process. To make more interactive and effective way of learning.

Then give feedback and guide towards a solution of student problem. Smart Home with intelligent robot is the recent trend application in this emerging field. By user mood robot act different process lighting setup, smart kitchen, smart interactive mirror, smart Air conditioner, Music system and give appropriate response to the user. Robot only recognise the user emotion by facial expression, body gesture and bio-signal predict the health and feelings of the user. Interactive mirror is a smart artifact, developed for smart home applications. It Supports Face recognition and emotion recognition. Measurement of physiological parameters like body weight, height and display of health progress chart (Body Mass Index, Body metabolic Rate)

In the medical field to reduce stress ,depression and anxiety for patients recognize emotion and giving the treatment by music therapy. Positive effects on Alzheimer are recognized. In the call center using voice to detect the emotion of customer. To recognize the voice of anger or happy customer can prioritize the angry calls. Then only

satisfy the customer needs and make more product sale. In marketing side emotion recognition process make more impact on advertisement. It makes great deal of attention and engagement of customer. To improve the sale make emotion vital in purchasing decision.

The paper is organized as follows: Section 2 discusses the related work. Section 3 presents a proposed system emotion recognition system. Section 4 presents the algorithm and process steps. Finally, Section 5 concludes with a general discussion.

2. RELATED WORK

In 1970 onwards research works have been doing in recognizing the human emotion by facial expression. Face detection by many features presence eye, nose, mouth, eyebrow, lip, etc., Real value parameters are Eyebrow raise distance, Upper eyelid to eyebrow distance, Inter-eyebrow distance, Upper eyelid − lower eyelid distance, Top lip thickness, Lower lip thickness, Mouth width, Mouth opening. The Binary parameters are upper teeth visible, lower teeth visible, forehead lines, eyebrow lines, nose lines, chin lines, nasolabial lines.

The challenges in the speech recognition are Frequency characteristics, Time related features, Voice quality parameters and energy descriptors. Frequency characteristics are Accent shape, Average pitch, Contour slope, Pitch range, Final lowering. Time related features are speech rate and stress frequency. Voice quality parameters and energy descriptors are breathiness, brilliance, loudness, pause discontinuity, pitch discontinuity. In Table 1.1 shows survey of emotion recognition system. Study refers the paper and the author name. Attributes refer to face, voice, gesture, EEG signal, ECG signal, Body movements and music. These are attributes are taken as input and early make as training set. Techniques are referred as RGB, thermal images, grayscale images and classifier used in system.

Algorithm referred as what are different types of algorithm specified in paper. It states main goal of the method. This may single algorithm or mixture of algorithm fuse to form new algorithm. Some are Active appearance model, particle swarm Optimization and genetic algorithm. Response time refers the static or dynamic. Static represent the still images or videos used to detect the emotion recognition system. Dynamic represent the spontaneous or immediate response to user from the real time implementation for example surveillance camera detects person emotion from crowd dynamic taking input from the user. Input format refers to audio, video, image, Mpeg, jpeg, bmp, movies and etc,. Limitation refers to the problems or dis advantage in this

method for example increase in computation al cost, device cost, low performance in device and accuracy may low. Best features refer to improvement in methods, best result, high accuracy and used in application oriented.

Chavan et al[11] presents "real-time implementation of a facial emotion recognition system" using the windows desktop devices. It code dynamically for basic emotions are happy, sad, fear, disgust, surprise, neutral and anger. It automatically detects the user face and codes the facial expression dynamically. It is user dependent because single face only detects by this paper. Accuracy may prominently reduce in multiple face detection so it needs dynamic way of predicting emotion in movies by facial expression and audio(music).Histogram of face area and video compressibility are used to recognize the face expression. First audio only and video only format are separately used for emotion detection. Later fusion method combining audio and video make mixture modal in multimodal system. In audio may failed to detect the emotion because laughing scene in movies unable to predict villain or hero. In the animated movie audio and scene are unable to predict.

Kamlesh Mistry et al[8] proposed the micro genetic algorithm embedded with particle swarm optimization. He used different classifiers to detect the seven dimension of emotion state and proposed hvnLB operator performs horizontal and vertical neighborhood pixel comparison to retrieve the initial discriminative facial features. It give better performance compared with other classical algorithm present in that paper.

Valentina Bono et al[13] present emotion detection method by EEG based feature classify emotional states are happy, fear and netural. EEG recordings are processed in bandpass filter to get EEG data for feature extraction and it classify face perception. this method extract large number of neurophysiological features and effectively reducing the dimensionality using machine learning techniques. This support for real time emotion classification.

Yong-Hwan Lee[2] presents mobile camera are used to detect the facial expression emotion recognition system. Weighted fuzzy k-NN Classification are used to compare upgrade to the multiple face detection.

Muneeba Raja et al[6] discuss the emotion detect via RF based on the gesture and body movements. Emotion can be sensed by the sensor. Attributes used to detect emotions are text, speech, facial expression, hand gesture, body gesture, physiological movements and multimodal. It is user independent to detect the single user from the multiple persons as like crowded area. This method is device free and incorporate with the RSSI values to recognize emotion.

Ankit Goyal et al[9] suggested the mixture of expert model for multimodal emotion predict in movies. It is a the result. Delaunay triangulation form triangle meshes to map the vectors in face image. Active Appearance model is used in this paper to extract feature set of input. It is dynamic process taking image in sequence of video. Pose of face restricted in camera view may change the position and direction of face. Only happy, sad and neutral can be recognized by this method.

Gloria Zen et al[14] propose the personalized model for facial recognition. Transductive Parameter Transfer (TPT), a framework for building personalized classification models and define some application. Using the visual data and the gesture movements detect the user emotion. It is user- independent also provide accuracy and reduce the computational cost. Gesture recognition and facial expression are assumed to "one size fit all". It is mainly designed for the average person applicable for action unit, pain detection and Smart watch based gesture recognition.

Peng Liu et al[4] for spontaneous facial expression analysis. In this paper, "present a novel infra-red thermal video descriptor in order to improve spontaneous emotion recognition". It represent as thermal video clips. Mainly based on temperature spontaneous movement of body and muscle movement produce heat. It captures by FLIR thermal camera. This method is used to identify the low back pain is real or fake. It is useful to detect spontaneous changes of emotion can be detected.

Table 1.1 Survey of Emotion recognition systems

Study	Attributes	Technique	Algorithm	Response Time	Input format	Limitations
Yong-Hwan Lee[2]	Yong-Hwan Lee[2]	Fuzzy k-NN	Active Appearance Model (AAM)	Dynamic	MPEG-4 & iPhone camera	three bas types ha neutral (average of 76%)
Kamlesh Mistry et al[8]	Face	Micro-GA	Particle Swarm Optimization (PSO)	Static	Images from Cohn Kanade and MMI	70% of each run
Ankit et	Face, Audio (music)	Video Compressibility and Histogram of Facial Area (HFA)	Mixture of Experts (MoE)-based Fusion Model	Dynamic	Video(movies)	laughing villains unable t
Sotiris Malassiotis et al[10]	Face	Active Shape Model (ASM)	Point Distribution Model (PDM)	Static	2D and 3D images	ASMis thanAA
Chavan et al[11]	Face	RGB	Bezier curves detection	Dynamic	Desktop (webcam)	unable compou emotions
Leo et Pauly al[12]	Face	The Haar-like features used for face detection	Fisher Faces Algorithm	Dynamic	webcam captures the video frames	Gender takes alg
Valentina Bono et al[13]	Face, Brain signal	EEG based ERP	artifact reduction algorithm	Dynamic	EEG face-evoked Dataset and grey scale images	Only cl emotion
Peng Liu et al[4]	Face	Thermal images	scale-invariant feature transform (SIFT)	Dynamic	Infrared thermal camera	It is positive
Gloria et	Face, gesture	Local Binary Pattern Histograms and TPT	novel transfer learning approach	Dynamic	Audio, images, video	User-ind process

3. PROPOSED SYSTEM

This paper proposes automatic emotion detection by facial expression for user dependent. In this work, the facial expressions have been recognized only by static image. The expression such as happy, fear, sad, neutral, surprise, disgust are considered for the experiment of face recognition system. Image taken by webcam and analyses the facial expression User login to the system (mobile, tablets). Webcam take snapshot of the user. Image taken by machine compared to the training set which already stored in database. Fuzzy K-NN classifier are used to feature extraction. This may result to get high accuracy result. After the template matching recognize the user mood and response to user.

Machine reason for emotion and test voice of the user and give appropriate result to the user. If user in sad emotional states. Users favorite music, video, images are displayed which is already stored in database.   This makes user make normal or feel good to seen that videos or images. Likewise all emotional states user's already

stored the images and videos are displayed. Every emotional states has different music, speech or images are shown to the user. This makes feel user not in alone and feel some better. To make machine act as companion to the user. Analyzes the user mood and respond the appropriate behavior .shows in figure 1. Machine acts as a human friend.
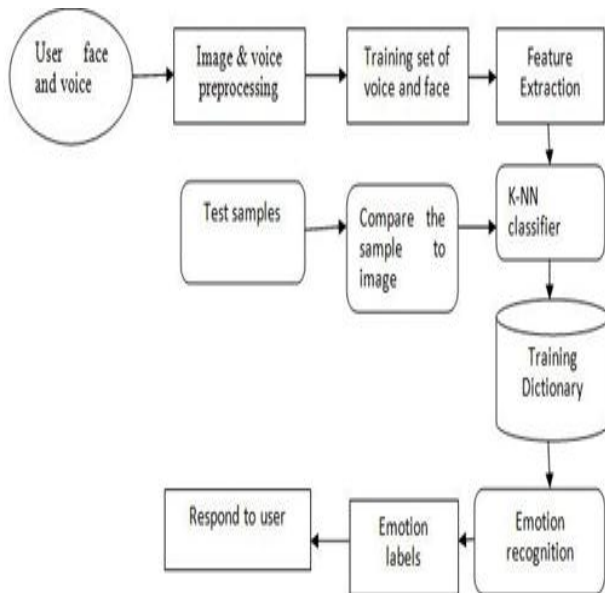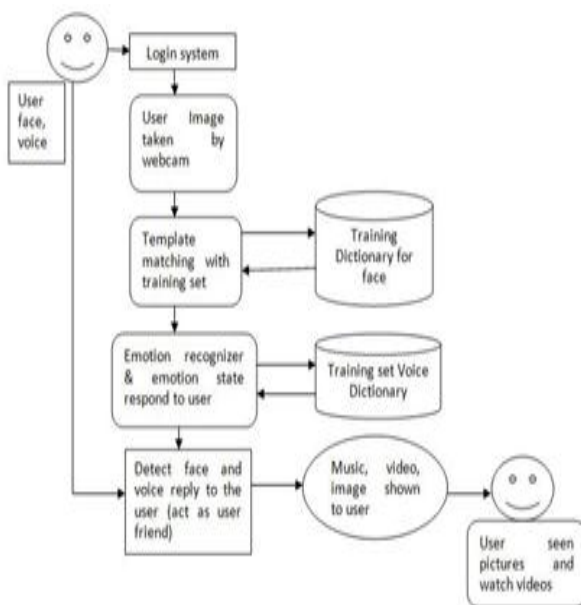

Figure 1. Architecture

Figure 2. Module diagram

4. ALGORITHM

1. Parameter user, input image, classifier c, dataset, Test sample, emotion e, e1, e2, e3, e4, e5,voice v1,v2
2. Initialize the image taken by webcam.
3. Select the user
4. If user is registered already in system do
5. System take image and sent to dataset comparison
6. Select the user dataset and need any update done
7. Goto classifier
8. End if
9. Else new user for using the system do
10. First taking training set of user mood
11. Basic types of emotion are taken data set
12. Next test samples are stored in database
13. End else
14. Goto classifier(c)
15. Classifier(c)
16. Select c user data by classifier
17. For each c do
18. Compare with the test samples
19. For each e do
20. For each v1 do
21. Recognize the emotion e
22. Assign the emotional state
23. End for
24. End for
25. End for
26. Goto to voice
27. Voice v1
28. For each v1 do
29. Goto classifer c
30. Voice v1 recognize the user
31. Respond to the user emotion
32. For each v1 do
33. Recognize emotion e by user voice v1
34. If user voice is different do
35. Voice v2 does not match with v1
36. Assign v2 is new voice
37. Request user to update voice in training set.
38. End if
39. Switch for emotion e
40. Case1:e1 consider as 'happy'
41. Do
42. Play fast music, favorite picture of user.
43. Break
44. Case 2:e2 consider as 'sad' do
45. Play the motivational songs, favorite movies or videos
46. Break
47. Case 3: e3 consider as 'angry' do
48. Play the favorite melody song and background music
49. Break
50. Case 4: e4 consider as 'surprise' do
51. Play the videos and music
52. Break
53. Case 5: e5 consider as 'netural' do

54. Play all old songs and slide show of memorable images

55. Break

56. End for

57. End switch

58. To make person feel better good.

5.   CONCLUSION

In this paper, presented and implemented a simple approach for recognition of the facial expression analysis. The algorithm is performed two major steps: one is a detection of emotional state of user and change the user mood if he/she in negative feelings. This acts as friend for lonely person. Many application are used medical field, neuroscience and machine intelligence.

REFERENCES

[1] Kevin Dullemond, Ben van Gameren, Arie van Deursen, Margaret-Anne Storey "Fixing the 'Out of Sight Out of Mind' Problem One Year of Mood-Based Microblogging in a Distributed Software Team ", IEEE, 2013.pg 267-276.

[2]   Yong-Hwan Lee,Woori Han,Youngseop Kim,Bonam Kim "Facial Feature Extraction Using an Active Appearance Model on the iPhone" in Eighth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing,IEEE,2014

[3]   Dolly Reney, Dr.Neeta Tripaathi "An Efficient Method to Face and Emotion Detection" in Fifth International Conference on Communication Systems and Network Technologies, IEEE, 2015.

[4] Peng Liu and Lijun Yin, "Spontaneous Facial Expression Analysis Based on Temperature Changes and Head Motions" IEEE Transactions on, 2015.

[5] Zankhana H. Shah, Dr.Vikram Kaushik "Performance Analysis of Canny Edge Detection for Illumination Invariant Facial Expression Recognition" in International Conference on Industrial Instrumentation and Control (ICIC), 2015.

[6] Muneeba Raja, Stephan Sigg "Applicability of RF-based Methods for Emotion Recognition: A Survey" in Contact-free Ambient Sensing, IEEE ,2016.

[7] Evangelos Sariyanidi, Hatice Gunes, and Andrea Cavallaro "Automatic Analysis of Facial Affect: A Survey of Registration, Representation, and Recognition" in Pattern Analysis and Machine Intelligence ,Vol 37, No.6,IEEE,2015

[8] Kamlesh Mistry, Li Zhang, Member,Siew Chin Neoh, Chee Peng Lim, and Ben Fielding "A Micro- GA Embedded PSO Feature Selection Approach to Intelligent Facial Emotion Recognition" IEEE, 2016.

[9] Ankit Goyal1, Naveen Kumar, Tanaya Guha, Shrikanth S. Narayanan "A Multimodal Mixture-Of- Experts Model For Dynamic Emotion Prediction In Movies" IEEE ,2016

[10] Sotiris   Malassiotis   and   Filareti   Tsalakanidou "Recognizing Facial Expressions from 3D Video: Current Results and Future Prospects"

[11] P. M. Chavan, Manan C .Jadhav, Jinal B. Mashruwala, Aditi K. Nehete, Pooja A. Panjari "Real Time Emotion Recognition through Facial Expressions for Desktop Devices" in IJESE,2013.

[12] Leo Pauly, Deepa Sankar "A Novel Online Product Recommendation System Based on Face Recognition and Emotion Detection" in lCCICCT , 2015.

[13] Valentina Bono, Dwaipayan Biswas, Saptarshi Das and Koushik Maharatna "Classifying Human Emotional States using Wireless EEG based ERP and Functional Connectivity Measures" in IEEE, 2016.

[14] Gloria Zen, Lorenzo Porzi, Enver Sangineto, Elisa Ricci, Member, IEEE, and Nicu Sebe "Learning Personalized Models for Facial Expression Analysis and Gesture Recognition" in IEEE, 2016.

[15] Kamlesh Mistry, Li Zhang,John Barnden"Intelligent Facial Expression Recognition with Adaptive Feature Extraction for a Humanoid Robot" in IEEE, 2015.

[16]   Christos   Georgakis,Yannis   Panagakis,Maja Pantic" Discriminant Incoherent Component Analysis" in IEEE Transactions on Image Processing, VOL. 25, NO. 5, 2016.

[17]   Wataru Sasaki,Tadashi Okoshi,Yuki Furukawa, Jin   Nakazawa,Yuuki   Nishiyama"Poster   Abstract: SmileWave: Sensing and Analysis of Smile-Based Emotional Contagion over Social network" in IEEE, 2016.

[18]   Zankhana H. Shahan and Dr. Vikram Kaushik "Performance Analysis of Canny Edge Detection for Illumination Invariant Facial Expression Recognition" in International Conference on Industrial Instrumentation and Control (ICIC), 2015.

[19]   Krystian Radlak and Bogdan Smolka "High Dimensional Local Binary Patterns for Facial Expression Recognition in the Wild" in Proceedings of the 18th Mediterranean Electrotechnical Conference, 2016.

[20]   Yee-Hui Oh,Anh Cat Le Ngo,Raphael C.-W. Phan, John See and Huo-Chong Ling "Intrinsic Two-Dimensional Local Structures For Micro-Expression Recognition" in IEEE, 2016.