Emotions recognition in human computer interaction: An audit

Shikha kamal
Department of Computer science, Himachal Pradesh University, Shimla, Himachal Pradesh

Jawahar Thakur
Department of Computer science, Himachal Pradesh University, Shimla, Himachal Pradesh

ABSTRACT

Human computer interaction, most widely known as its acronym HCI. It is concerned with understanding how people make use of devices and systems that assimilate or embed computation, and how such devices and systems can be more useful and more usable. Two channels have been distinguished in human interaction. One is implicit and another is explicit. Explicit channel defines about anything or nothing on the other hand implicit message conveys the message about speaker themselves. Both linguistic and technology takes lots of efforts to understand explicit message. Emotions are associated with implicit messages. Emotions can be analyzing through vocal signals, facial expressions and multimodal information. This paper aim is to deliver a brief guidebook of approaches, application area and comparative studies used to analyze the emotions over past decades researches. After reviewing the results of emotions recognition, it is found that happiness has the highest accuracy rate than other emotions.

Keywords: Emotion recognition, FACS, feature extraction, emotion classification.

1. INTRODUCTION

Human face to face communication is an ideal model for architect a multimodal human computer interface. Channel is a medium to communicate and there are different methods to sense the signals from outside world. Basically, one person perceives the others person emotional condition through their facial expressions. The results of human communication are through acoustic channel and visual channel. Acoustic channel consist of speech, vocal signals and visual channel carries facial expressions and body movements. The studies in computer-assisted recognition of facial expression started in 1990s. In 1991 Masse was the first researcher to use the image processing technique to analyze emotional expression. Emotion is the fundamental component of being human. It is the key to express your feelings while communicating. Emotional facets contain large impact on social intelligence such as communication understanding. Decision plays an important role to understand behavioral aspects of human. Human can easily understand mental condition of a person through its emotional behavior. Mostly few existing works defines to analyze emotions through facial expressions as well as vocal signals [2],[5]. Psychological research in analyzing emotions also interpret that verbal part effects 7% of the message, vocal part 35% and facial expression 55% to the effect of the spoken messages [12]. This defines that facial expression plays an important role in human communication. As the demands grows for system with human computer interaction. Researcher found that emotion can be recognized through electroencephalography (EEG) technique. This technique is mainly used to monitor the brain activity. When someone is in coma, the level of brain activity can be determined through EEG.

To recognize emotion through EEG, input signals are taken from the scalp of the hair and KDE method used to extract the feature and emotion is classified through

artificial neural network [21]. A lot of researches has been done in recognize emotion through EEG based on novel regression model [10, 11]. As the recent developments in facial expression recognition techniques, the major focus is on artificial neural networks (ANN) and Support Vector Machine (SVM) in emotion classification. The technique first analyses the information conveyed by the facial regions of the eye and mouth into a merged new image and using it as an input to a feed forward neural network trained by back propagation. SVM showcases the use of Oriented Fast and Rotated (ORB) on a single frame of imagery to extract texture information, and the classification is completed using SVM [24].

This paper is categorized as: section 2 survey the various methods used to categorize the emotion based on vocal signals and multi-modal information. Section 3 summarized the recognition based on facial expressions. Section 4 and 5 characterize the various application area and databases available publically. Last section summarizes the average accuracy rate of various emotions.

2. ACOUSTIC AND MULTIMODAL INFORMATION BASED ANALYSIS

Acoustic information transmits data about vocal signals, utterance of the voice and quality of speech. Speech information comes under verbal communication. Some classification that have been used are: Anger, sadness, happiness, neutral [5], Normal, happy, angry, sad, afraid [16], Happiness, sadness, boredom, disgust [13]. Speech information process data into two main parts: speech processing and emotion recognition. Input of speech signals is the first step of analyzing emotions

through speech. Input is taken as the utterances of the voice and calculated by the feature extraction process. Feature is being calculated by the tone of the voice and utterance of the voice. Utterance of the voice is divided into number of periods. Each no. of frames contains unique feature which is extracted by the various methods. Speech processing flow are defined as:

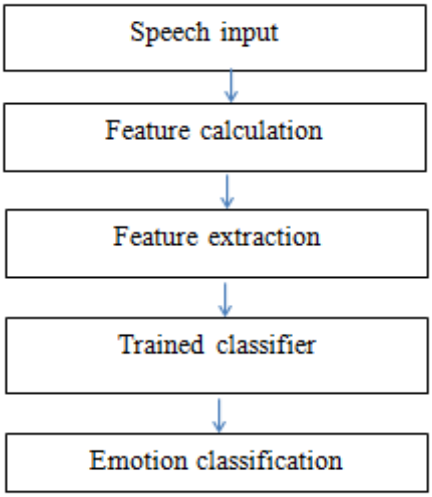

Figure 1: Processing flow

FEATURE EXTRACTION

Speech is consisting of pitch, tone, frequency and utterance. There are two mainly type of speech features: prosodic features and spectral features. Prosodic features deals with the pitch, intensity and formant frequency profit as well as voice quality measures. Spectral features deals with the actual power of voice signals. Spectral features use the various algorithms to extract the feature: Mel-frequency cepstral co-efficient (MFCC), linear prediction cepstral co-efficient (LPCC), log-frequency power co-efficient (LFCC), Perceptual linear prediction [2]. As pointed by Nicholson [13] the speech feature extract from utterance of the voice using linear prediction capestral co-efficient for training and testing. In LPCC first step determine the beginning and ending point of an utterance. LPCC is comes under spectral features. Speech power is compared with predefined threshold value. Utterance of the voice neglects when speech power drops from the threshold value. After knowing the starting and ending point of an utterance it will be divided into two equal length in time. For example 10 intervals are expressed as the vectors f1, f2, f3……. F10. Each of the 10 vectors in consists of feature parameter for intervals, (P, p, c1, c2 …c10, d) and the value of feature parameter is defined in matrix form. Vectors are defined as:

$f_v f_v = [f_1, f_2, f_3 \ldots\ldots\ldots\ldots f_{10} f_1, f_2, f_3 \ldots\ldots\ldots\ldots f_{10}]$ (1). The feature vector $f_v f_v$ is used as an input for the emotion classification purpose.

Another approach to extract feature by calculating statistics for the vocal information: frequency, energy, speaking rate and their bandwidths (BW1, BW2, and BW3). Statistics were calculated in the form of mean, standard deviation, minimum, maximum, and range. RELIEF-F algorithm is used for feature selection [16].

EMOTION CLASSIFICATION

After perceive the speech and its calculating features the next step is to classify the emotion on the basis of given input. There are different approaches were used to classify the emotions: K-nearest neighbors, neural network, ensemble of neural network classifiers, set of neural network approach, discrete hidden markov model [16], [5].

- **K-nearest neighbor**

K-means algorithm was used to compress the number of training feature vectors for each class and estimates the probability of each experimental class. In the experiment of petrushin [16],[25]. 70% of data set as a database of cases for comparisons and 30% as a test sets. The average accuracy of classification through k-NN is 55% using 8 features. Anger has a highest accuracy rate i.e. 65%. Accuracy rate of fear was very poor i.e. 13%.

- **Neural networks**

No. of features is taken as the input and the no. of output is determined as specific emotion. Basically there are two different neural network used as a classifier i.e. one class in one neural network (OCON), two layer backpropagation neural network architecture. Two layer neural networks define the 65% of average accuracy rate. Fear has the poorest recognition rate i.e. 25-50% [16].

- **Ensemble of neural network :**

This classifier consist of highest accuracy rate i.e. 70%**.** It consists of an odd no. of neural network classifier & the decision were making on the basis of voting principle. Accuracy of happiness is 65%, accuracy of fear is 53% and for sadness is 73% [16].

TABLE 1 : Method used to analyse speech signals multimodal information.

Authors	**Feature extraction Method**	**Classifier**	**Test cases**	**Accuracy**
Petrushin [16]	RELIEF- F algorithm	Ensemble of neural networks, 2 layer back- propagation neural network, K-nearest neighbor	23 subject s , 700 utterances	70%
Nicholson [13]	Linear predictive coding parameters	Single-layer-neural network based on LVQ (linear vector(quan	100 speakers	50%
Rieger jr. [25]	Linear predictive	Ensemble of K-NN	6 subjects	95%
Busso [5]	Sequential backward feature selection technique	Naïve – bayes classifier	-	89.1%
Emerich [2]	-	Naïve bayes classifier	2 subjects	90%

3. FACIAL EXPRESSION ANALYSIS

It is easy to recognize the emotional condition of a person through their facial expressions. A facial expression plays an important role in communication between two persons. As the rapid growth of technology facial expression also plays an important role to interaction between human- computer systems. Majorly there are two type of communication: verbal and non- verbal communication. Facial expression comes under non-verbal communication. In computer assisted recognition the expressions was analyzed on the basis of facial action coding system (FACS). First, before the facial expression analyzed, the face should be detected. Detection of face may be in 2D model or in 3D model system. Feature extraction is the next process to analyze the facial expression and it has been extract on the basis of action unit. Feature can be extract in real time image, image sequences or video. Action unit is taken as the input for extraction process after detecting the image. for example the conversation between teacher and student: The teacher asked: “all the best for the final examination.” Student kept smiling (lip corner raiser AU12) and asked thank you.

As given in the above conversation lip corner raiser describes the action unit used for the feature extraction. Apart from FACS there are also another methods [18], [22] used for feature extraction process. After extracting the feature last process is to classify the emotion on the basis of extracted feature. This section discusses various methods and approaches used to analyze the basic three stages of emotion recognition based on facial expressions.

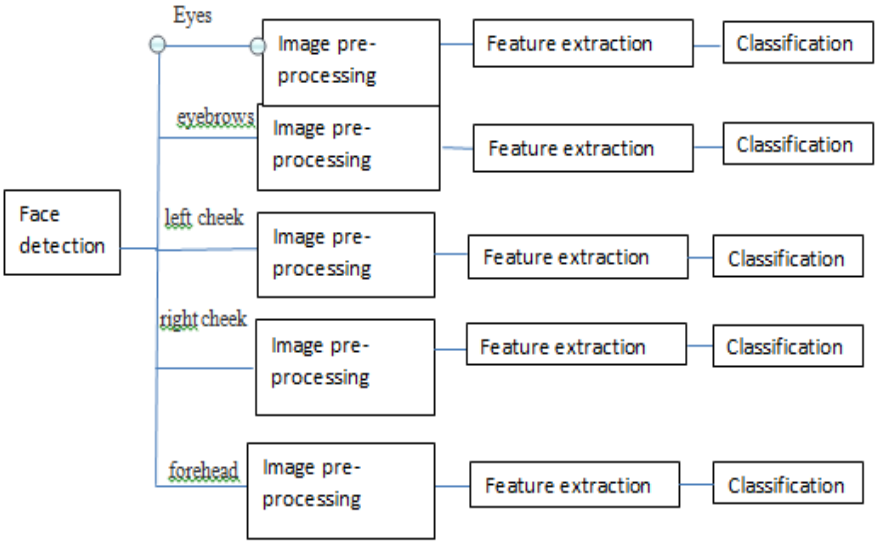

Figure 2. Facial expression analyzing process

FACE DETECTION

The first step in recognition process is detection of image and image is a face having different view i.e. frontal view, 3D view. Input is taken from any device like camera. In the experiment of tarnowski [19] input image is detected through kinect device. Kinect device is a scanning resolution having infrared emitter and two cameras but a relatively high rate of image registering (30 frames/s). One of the cameras record visible light, while the other operates in infrared and is used for measuring the depth. Infrared rays reflected from the user's body allow creating a 3D model of a face. OpenCV face detector is used to identify the large face within the image.

For most works in automatic facial expression analysis, the conditions under which a facial image or image sequence is obtained are controlled. Usually, the image has the face in frontal view. Hence, the presence of a face in the scene is ensured and some global location of the face in the scene is known a priori. However, determining the exact location of the face in a digitized facial image is a more complex problem. First, the scale and the orientation of the face can vary from image to image. If the pictures are taken with a fixed camera, faces can occur in images at various sizes and different angles due to the movements of the subject. It is little difficult to find the template in the image. Noises are presence in an image in the form of facial hair, mole, and eye glasses etc. noise is being removed by feature extraction process. Two-gray-scale image with 100 to 200 pixels with lower limit of detection of a face by human observer. Another characteristic of the human visual system is that a face is perceived as a whole, not as a collection of the facial features. It is difficult to detect an image by computer with occluded picture [12].

FEATURE EXTRACTION

After detecting the face from any device the next step is to extract the feature from the detected image. Extraction of image is based on facial muscle action and each action is defined as the action unit. Facial feature can be extracted through geometrical method. Rather than geometric points there are various method though which feature can extract efficiently. Using active shape method (ASM)[3], some new feature like distances between two particularly selected markers are generated as high-level nfacial shape feature. Eucledian distance is calculated by applying these formula [6].

Eucledian distance (u,v)=

$$\sqrt[2]{(x1-y1)^2}+(x2-y2)^2 \dots\dots\dots\dots (xn-yn)^2 \quad (2)$$

Usually, the main purpose of feature extraction is to fetch the properties of facial components. Adaboost algorithm[18] is the another method used to extract feature using rectangular feature classifier. Adaboost algorithm is based upon SVM classifier to detect and locate face. Adaboost algorithm constraining on weak classifiers to depend on only a single feature. Some factors make difficult to extract feature i.e. hair, glasses, mole etc. Rather than these problems image size input, and orientaion of the face would also obscure the facial features. In general, three types of face representation method is used in facial expression analysis: holistic, analytic, hybrid [12]. Weak classifier is expressed as:

$f_m f_m = p_w p_w(y=1|x) - p_w p_w(y=-1|x)$ (3)

Linear discriminant analysis (LDA) is a method to extract dominant feature. It is based upon domain specific information and increases the rationality between domain specific and within domain scatter matrix. LDA provide more efficient feature space can be providing that uniquely identified feature of each domain [7, 15]. In every phase of emotion recognition various approaches was used to sense, detect and classify the facial expression. Various approaches to detect emotions are: Traditional FER approaches, deep learning based FER approaches.

- **Traditional FER approach**

Traditional approaches illustrate the geometric based feature, appearance based feature and hybrid features. Geometric feature defines the skin color of a model and the statistical theory based face detection [18].

The first geometric feature based approach introduced in 1978. Geometric feature mainly focus on corner feature, edges, marks and salient points of the image. These features extracted from various location of facial elements such as eyes, eyebrows, nose, mouth and chin. These various elements landmarks points were tracked by the point tracker [22]. Calculation of angle and Euclidian distance has been calculated within the frame of image.

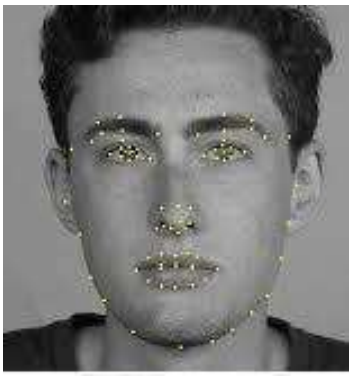

Figure 3 . Geometric points on facial elements

Appearance based feature extraction usually based upon global face region and this approach is mainly used in real time facial emotion analysis system. Extracted region-specific appearance features by dividing the enitire face region into domain-specific regions. This approach gives the better recognition accuracy [17].

- **Deep learning based FER approach**

Deep learning based algorithm used convolutional neural networks(CNN) and recurrent neural network (RNN) algorithms to feature extraction, classification and recognition tasks.

This approach mainly recognize seven basic emotions. CNN is basically used for feature extraction and RNN for the classification purpose. CNN contains three different types of layers: convolutional layer, max pooling layer, and fully connected layers. Convolutional layers take input in the form of images and represents a spatial arrangements of the facial image. Second max polling and also known as subsampling layers. Max pooling layer reduce the dimension and ignore the marks over facial image. The

last fully connected layer of a CNN structure evaluate the score on the original image. Mostly deep-learning-based methods have adapted a CNN directly for AU detection [17].

Table (2) : Template based method used for feature extraction

Authors	Method	Comment
Shuai [18]	Geometric feature method is used based on invariant feature, skin color, priori knowledge.	Frontal view Faces without facial hair growth
Myunghoon [3]	Active shape model applies to extract fudicial facial feature.	Extract facial feature as well as new fetaures like distance between two particular marker. 77 facial landmarks are located.
Jones [23]	Adaboost algorithm to measure the intensity between eyes and the region across the upper cheeks.	Frontal view Integral image Feature based system
Xia [7]	Linear discriminant anlayses (LDA) is performed to extract to dominant feature.	3D view 200 sequences

FACIAL EMOTION CLASSIFICATION

Classification of emotion is the last stage of FER. It is basically the output of feature extraction process. The system based on facial expression reveals that eyebrows gives poorest performance while recognizing emotions and cheeks area provide valuable information for classification of emotions [5].

Adaboost algorithm is used for both feature extraction as well as classifications of emotions. It boost the classification performance of a simple learning algorithm. This is done by combining the collections of weak classifier and form a strong classifier. Weak classifier gives 51% result while classification of emotions [8].

- **Emotion classification using image sequence**

Support vector machine (SVM) is used to analyze emotions. Feature extraction and detection of image is through geometrical feature. Geometrical feature is extracted through facets such as eyes, eyebrows, cheeks etc. similarly, emotions classification take input from feature extraction process and further classifies emotion through SVM classifier. SVM classify six basic emotions i.e. fear, anger, surprise, happy, sad and disgust. SVM classifier analyzes surprise facets much better than other emotions. It creates some confusion between emotions. Here are some results through SVM [8].

- **Support vector machine classifier (SVM)**

SVM are used in pattern recognition problem and highly solved structural minimization problem. In many cases, SVM gives better result than other competing methods. SVM generally solve linear as well as non-linear problems. Non- linear problems are solved by a method called the kernel in which transformation of the input feature vectors into generally higher dimensional feature space by a mapping function. Maximum discrimination is obtained by an optimal placement of the maximum separation between the borders of two classes. SVM can handle two-class problems but a variety of strategies exists for multiclass discrimination. To construct an optimal hyper plane, SVM minimize the error functions using iterative training algorithm. Various kernel is used in SVM i.e. linear, polynomial, radial basis function(RBF) and sigmoid. RBF and polynomial gives highly accurate result. In case of classification SVM combine the length of normalized emotion patterns. Length normalization and Sub sampling is the technique used to transform the dynamic feature length sequence into a static one [2, 4].

Table (3): Summarized methods for facial expressions based on emotion classification

Authors	Method	Facial actions	Test cases	Accuracy
Tarnowski [19]	MLP and KNN used as a classifier in two division form: ● Random division ● Natural division	Six action units	Each subject take 3 session trials 26 subjects	3-NN: 95% MLP: 73%
Xia [7]	Discrete Hidden Markov model	10 action unit RGB images	10 subjects	90.92%
Seneschal [4]	SVM RBF(Radial basis function) kernel	3 action units	200 samples	94%
Emerich [2]	SVM with POLY and RBF kernel	-	270 images	90.7%

- **Artificial neural network:** An artificial neural network consist of various neurons. Neural network transforms the input into meaningful output to match the target value. An ANN is defined by three types of parameters: The interconnection pattern between the different layers of neurons, learning process for updating the weights of the interconnections and activation function that converts a neuron's weighted input to its output activation. In ANN, the input layer contains the number of neurons and hidden layers vary according to the features of the input image, whereas the number of neurons in the output layer is nothing but the four emotional states [14]. ANN is used as a classifier to categorize emotions through electroencephalography EEG signals [21].

Table (4): Emotion classification using neural network based method

Authors	**Method**	**Test Cases**	**Accuracy**
Authors	**Method**	**Test Cases**	**Accuracy**
Lahane[21]	Artificial neural network	-	-
Li [11]	Novel regression model (GRSLR)	14 subjects	83.46%
Horlings [10]	Support vector machine	10 subjects	76%

4. APPLICATION AREA

Due to highly developed era, it is essential to build an intelligent machine which is human friendly. Emotion recognition is the active area in the field of computer science research zone. There are several application area of emotion recognition [1, 15] and some of the applications area are:

- Council the medical state of patient.
- During healthcare, determining patients feeling and comfort level about the treatment.
- Determining struggling counts of expressions in autistic person.
- In the case of e-learning, determines the emotion and adjust the learning technique and presentation according to the style of learner.
- Determining weakness in the case of driving and alerting in advance.
- In the case of call centre, determine anger and stress levels in the voice and compute angry calls.

5. DATABASE

In the field of FER there are numerous databases used for experiments and contained trained datasets. Some databases contain still images and few have 2D images and video sequences. Here are the some brief descriptions of few datasets available publically:

- IEMOCAP, SEMAINE, MELD, DailyDialog and Emocontext are the datasets mainly used for analyze emotion recognition in conversation. It contains multimodal data contains visual textual and acoustic information. Apart from SEMAINE datasets rest of the datasets contains categorical emotion labels [9].
- Cohon-Kanade (CK) database was come into existence in 2000 for the purpose of research in automatic analysing facial expression. It consist 593 pictures of FACS. CK is available in two versions and third is in preparation. Version 1 contains 486 images from 97 posers. Images starts from neutral expression and go forward to peak level. Peak level images are fully facial action coded and given an emotion label. Version 2 are extended version of CK and contains various expressions too. CK+ provide protocol and baseline result for facial feature tracking and action unit and emotion recognition. Version 2 contains validated emotion labels [17].
- Bao face database contains lots of pictures most of the images are Asians. There are 149 images of single person. Most often resolution are 149×207 to 192×273 pixels with DPI 96. This database contains two parts and second part contains 221 pictures. Their resolutions are from 348×432 to 800×600, DPI 72. Pictures are clicked in poor lightning condition [20].
- CMU/VASC image database: This database test set was originally assembled as part of work in neural network based on face detection. This database contains 4 subdirectory: newest, rotated, test, test-low. Usually all datasets are

trained under test condition. Apart from these databases some other databases are: Caltech faces 1999 databases, NIST Mug-shot images databases, Yale face databases [20].

6. CONCLUSION

Analysis of emotions is an general problem which is solved by the human very easily. Three main stages were used to analyze the emotional state of a person i.e. face detection, feature extraction and emotion classification. In human computer assisted system emotion can be analyze through speech, facial expression, and multimodal. Emotions are the basically depends upon the actions which is performed under mental ability condition of a person. This paper conclude the all extensive efforts made by researcher from past over years. There are various method to analyze emotions. Emotions can be analyzed through visual information (facial expression, image sequences etc.), speech information, multimodal information (speech, image, and video sequence), and textual data. Various methods and techniques used to analyze these all information. As this paper presented such approaches: conventional FER approach and deep learning based approach. Conventional FER approach is further divide recognition process into three steps: face detection, feature extraction and emotion classification and conventional FER approach used Adaboost algorithm, SVM to classify and feature extraction process. In case of deep learning based FER approach it uses conventional neural network approach (CNN) for analyze expressions. This paper is also gives brief introduction about datasets and databases used to detect FER.

REFERENCES

[1] M. Dubey and L. Singh, 'Automatic Emotion Recognition Using Facial Expression : A Review', Int. Res. J. Eng. Technol., vol. 3, no. 2, pp. 488–492, 2016.

[2] S. Emerich, E. Lupu, and A. Apatean, 'Emotions recognition by speech and facial expressions analysis', Eur. Signal Process. Conf., no. Eusipco, pp. 1617–1621, 2009.

[3] M.Suk, B.Prabhakaran. 'Real time facial expression recognition on smartphones'. IEEE Winter Conference on application of computer vision, 2015.

[4] T.Senechal, D.Mcduff, R.Kalou by " Facial action unit detection using active learning and effiecient non-linear kernel approximation." Waltham, MA, 02452, USA.

[5] C. Busso, Z. Deng, S. Yildirim, and M. Bulut, 'and multimodal infoAnalysis of emotion recognition using facial expressions, speeBusso, C., Deng, Z., Yildirim, S., & Bulut, M. (2004). Analysis of emotion recognition using facial expressions, speech and multimodal information. … on Multimodal …, 205–211', Icmi, pp. 205–211, 2004.

[6] N. Gupta and P. N. Kaur, 'Design and Implementation of Emotion Recognition System by Using Matlab', Int. J. Eng. Res. Appl., vol. 3, no. 4, pp. 2002–2006, 2013.

[7] Lu Xia, C. Chen, J. K Aggarwal "View invariant human action recognition using histogram of 3D joint." Computer and Vision Research center.

[8] Rajesh A Patil, Vineet Sahula, A. S. Mandal. Expression Recognition In Image Sequences Using Active Shape Model (2019).

[9] S. Poria, N. Majumder, R. Mihalcea, and E. Hovy, 'Emotion Recognition in Conversation: Research Challenges, Datasets, and Recent Advances', IEEE Access, vol. 7, pp. 100943– 100953, 2019.

[10] R.Horlings, D.Leon, J. M. Rothkrantz. Emotion Recognition using Brain Activity. International Conference on Computer Systems and Technologies.

[11] Y. Li, W. Zheng, Z. Cui, Y. Zong, and S. Ge, 'EEG Emotion Recognition Based on Graph Regularized Sparse Linear Regression', Neural Process. Lett., vol. 49, no. 2, pp. 555–571, 2019.

[12] M. Pantic and L. Ü. M. Rothkrantz, 'Automatic analysis of facial expressions: The state of the art', IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 12, pp. 1424–1445, 2000.

[13] J. Nicholson, K. Takahashi, and R. Nakatsu, 'Emotion recognition in speech using neural networks', ICONIP 1999, 6th Int. Conf. Neural Inf. Process. - Proc., vol. 2, pp. 495–501, 1999.

[14] K. Hulliyah, N. S. A. A. Bakar, and A. R. Ismail, 'Emotion recognition and brain mapping for sentiment analysis: A review', Proc. 2nd Int. Conf. Informatics Comput. ICIC 2017, vol. 2018-Janua, pp. 1–6, 2018.

[15] K. Ho, M. J. Chung "Congnitive Face analysis system for future interactive TV."

[16] V. A. Petrushin, 'Emotion recognition in speech signal: Experimental study, development, and application', 6th Int. Conf. Spok. Lang. Process. ICSLP 2000, no. Icslp, pp. 6–9, 2000.

[17] B. C. Ko, 'A brief review of facial emotion recognition based on visual information', Sensors (Switzerland), vol. 18, no. 2, 2018.

[18] S. Liu, W. Wang 'the application study of learner's face detection and location in the teaching network system based on emotin recognition' 2nd interanation conference on network security, wireless communication and trusted computing, 2010.

[19] P. Tarnowski, M. Kołodziej, A. Majkowski, and R. J. Rak, 'Emotion recognition using facial expressions', Procedia Comput. Sci., vol. 108, pp. 1175–1184, 2017.

[20] M. Magdin, F. Prikler 'Real Time facial expression recognition using webcam and SDk affectiva', 2017.

[21] P. Lahane and A. K. Sangaiah, 'An approach to eeg based emotion recognition and classification using kernel density estimation', Procedia Comput. Sci., vol. 48, no. C, pp. 574–581, 2015.

[22] S. Dhall, P. Sethi 'Geometric and appearance feature analysis for facial expression recognition' International Journal of Advanced Engineering Technology.

[23] P. Viola, M. j. Jones 'Robust real time face detection' International journal of computer vision, 2004.

[24] T. Kundu, C. Saravanan 'Advancements and recent trends in Emotion Recognition using facial image analysis and machine learning models' International Conference on Electrical & Electronics Communication, Computer and Optimization Techniques (ICEECCOT), 2017.

[25] S. A. Rieger, R. Muraleedharan, and R. P. Ramachandran, 'Speech based emotion recognition using spectral feature extraction and an ensemble of kNN classifiers', Proc. 9th Int. Symp. Chinese Spok. Lang. Process. ISCSLP 2014, no. 2, pp. 589–593, 2014.

.