

# Emotion Recognition using Speech and EEG Signal – A Review

Priyanka Abhang<sup>1</sup>  
Research Student

Shashibala Rao<sup>1</sup>  
Research Student

Bharti W. Gawali<sup>1</sup>  
Associate Professor

Pramod Rokade<sup>2</sup>  
Assistant Professor

<sup>1</sup> Department of CS and IT, Dr. B.A.M. University, Aurangabad, Maharashtra, India.

<sup>2</sup> Department of Zoology, R.B. Attal College, Georai, Beed.

## ABSTRACT

In recent years the research interest is improving in the field of human computer interaction. This paper focus on one of the aspect of human computer interaction in concern with, the recognition of emotion in a person with the help of Electroencephalogram (EEG) signals and speech. EEG uses an electrical activity of the neurons inside the brain. EEG machine is used for acquisition of the electrical potential generated by the neurons when they are active. The Brain cells communicate with each other by sending electrical impulses. Emotions allow people to express themselves beyond the verbal domain. Speech is the most natural form of communication. A much of work is been done in speech recognition in various languages. It is one of the components that closely related to emotions. Very less work has been carried out using combine aspects of speech, emotion and EEG. Thus this paper attempts to review the combine efforts of EEG brain signal and Speech to recognize the emotions in humans.

## General Terms

Speech Recognition, EEG.

## Keywords

Electroencephalogram, Speech recognition and Emotion.

## 1. INTRODUCTION

Human communication, nonverbal information, such as intention and emotions, plays an important role, especially by using information of emotion people can communicate with each other more smoothly. The emotion recognition systems in speech or facial expressions which have been used include several emotional states such as joy, fear, sadness, disgust, anger, surprise and neutral[1]. Emotion recognition in humans is an important research area. Recognizing emotions from speech has gained increased attention recently, because automatic emotion recognition can help people to develop and design many applications about human-machine communication [2]. The most widely used speech cues for audio emotion recognition are global-level prosodic features such as the statistics of the pitch and the intensity. Therefore, the statistical measures such as the means, standard deviations, ranges, maximum values, minimum values and the energy were computed using various speech processing software [3]. Speech recognition is the process of converting, an acoustic signal, captured by a microphone or a telephone, to a set of words. The recognized words can be the final results, as for applications such as commands & control, data entry, and document preparation. They can also serve as the

Input to further linguistic processing in order to achieve speech understanding. Speech recognition technology has made it possible for computer to follow human voice commands and understand human languages. The main goal of speech recognition area is to develop techniques and systems for speech input to machine. For reasons ranging from technological curiosity about the mechanisms for mechanical realization of human speech capabilities to desire to automate simple tasks which necessitates human machine interactions and research in automatic speech recognition by machines has attracted a great deal of attention. Speech recognition technology was increasingly used within telephone networks to automate as well as to enhance the operator services [4].

The electroencephalogram (EEG) is a recording of the electrical activity of the brain from the scalp. The first recordings were made by Hans Berger in 1929 although similar studies had been carried out on animals in 1870. The EEG is used in the evaluation of brain disorders. Most commonly it is used to show the type and location of the activity in the brain during a seizure. It also is used to evaluate people who are having problems associated with brain function. These problems might include confusion, coma, tumors, long-term difficulties with thinking or memory, or weakening of specific parts of the body.

The paper is divided into five sections. Section 1 provides introduction about interrelated study in EEG, Emotion and Speech. Section 2 deal with the EEG brain signals and related techniques and Automatic speech recognition is presented in section 3. Section 4 covers the Emotion recognition. Section 5 gives the Conclusion followed by the Reference

## 2. EEG BRAIN SIGNALS AND RELATED TECHNIQUES

The waves recorded are thought to reflect the activity of the surface of the brain, the cortex. The nerve cells in the brain produce signals that are called action potentials. These action potentials move from one cell to another across a gap called the synapse. Special chemicals called neurotransmitters help the signals to move across the gap. There are two types of neurotransmitters; one will help the action potential to move to the next cell, the other will stop it moving to another nerve cell. The brain has to work hard to keep equal amount of each of these neurotransmitters in the brain.

EEG activity is quite small; it is measured in microvolts ( $\mu\text{V}$ ). The main frequency interest up to approximately 30 Hz. EEG signals are of different waveforms which are used. The waves like alpha ( $\alpha$ ), Beta ( $\beta$ ), Delta ( $\delta$ ), Theta ( $\theta$ ) as well as spikes associated with epilepsy. The Figure 1, describes the EEG signal is closely related to the level of consciousness of the person. As the activity increases, the EEG shifts to higher dominating frequency and lower amplitude. When the eyes are closed, the alpha waves begin to dominate the EEG. When the person falls asleep, the dominant EEG frequency decreases. In a certain phase of sleep, rapid eye movement called (REM) sleep, the person dreams and has active movements of the eyes, which can be seen as a characteristic EEG signal. In deep sleep, the EEG has large and slow deflections called delta waves. No cerebral activity can be detected from a patient with complete cerebral death.

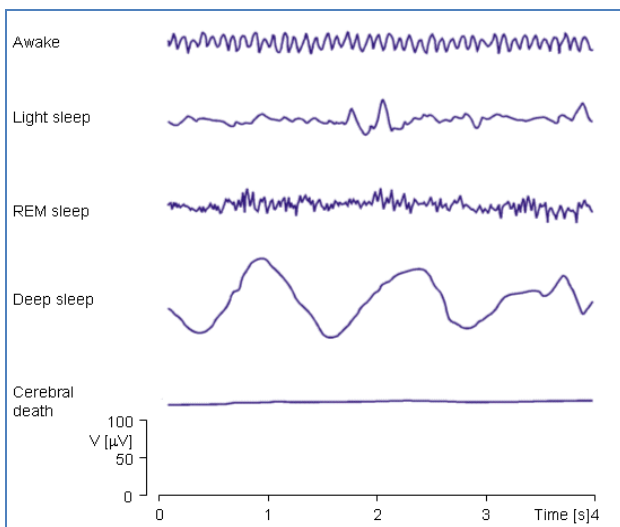


Figure 1: EEG activity dependent on the level of consciousness

We can elaborate various waves with their frequency as follows:

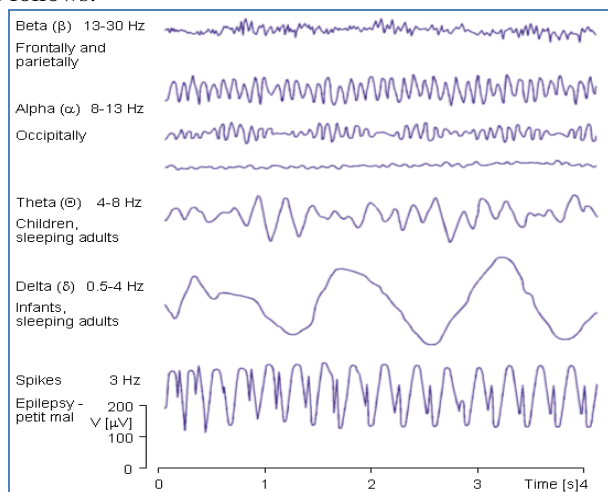


Figure 2: Examples of EEG waves

The figure 2 shows the frequency band of the beta waves is 13-30 Hz; these are detectable over the parietal and frontal lobes. The alpha waves have the frequency spectrum of 8-13 Hz and can be measured from the occipital region in an awake person when the eyes are closed. The theta waves have the frequency range of 4-8 Hz and are obtained from children and sleeping adults. The delta waves have the frequency range of 0.5-4 Hz and are detectable in infants and sleeping adults [5]. The EEG signals are basically and closely related to the level of consciousness of a person. When the eyes of the subject are closed then alpha waves begin to have a very strong influence over EEG [6].

Developing and understanding advanced signal processing techniques for the analysis of EEG signals which is crucial in the area of biomedical research. The sources project a unique topography on to the scalp i.e. "scalp map". The Independent Component Analysis (ICA) is a technique that attempts to reverse the superposition by separating the EEG into mutually independent scalp maps. A method called Polysomnogram (PSG) which records the signals more accurately overnight. [7], various classifications of EEG signals are performed by five mental states of brain. Wavelet packet transform (WPT) can be used for feature extraction of the relevant frequency bands from EEG signals. Various mental tasks like relaxed, arithmetic task, geometric figure rotation, visual counting, and mental letter composing task were performed by the subjects, the subjects were asked to perform the task without any movement of the body parts. To analyze, the frequency spectrum the FFT method is widely used. The FFT plot signals from all electrode pairs. The basic neuroscience knowledge that one hemisphere of the brain (left) is dominate for calculation while other hemisphere (right) of the brain is dominate for visual tasks [8]. Euclidean minimum-norm (MN) inverse is implemented in the study of single trial source localization of the scalp [9], The KNN method can be applicable to classify accuracy of the frequency with a K-nearest neighbor (KNN). It assigns new unclassified examples to the class to which the majority of its k-nearest neighbor belongs. The KNN method is based on Dempster-shafer theory. To extract features from the EEG autoregression (AR) models and wavelets decomposition are used. Dempster-Shafer theory and KNN classification method were used for classification of different mental tasks [12].

The feature extraction techniques used in the review are shown in table 1.

**Table 1: List of technique with their properties For Feature extraction**

Sr. No.	Method	Property	Procedure for Implementation
1	Independent Component Analysis (ICA)	Non linear feature extraction method, Linear map, iterative non-Gaussian	Blind course separation, used for de-mixing non-Gaussian distributed sources(features)
2	Mel-frequency cepstrum (MFCCs)	Power spectrum is computed by performing Fourier Analysis	This method is used for find our features
3	KNN method	It is used for finding nearest neighbor	This technique is useful for clustering implementation is depend on number of centroid.
4	Wavelet	Better time resolution than Fourier Transform and it is also known as Wavelet Packet Transform	It replaces the fixed bandwidth of Fourier transform with one proportional to frequency which allow better time resolution at high frequencies than Fourier Transform

### 3. AUTOMATIC SPEECH RECOGNITION SYSTEMS

After years of research and development the accuracy of automatic speech recognition remains one of the important research challenges. Speech Recognition is also known as Automatic Speech Recognition (ASR) or Computer Speech Recognition, processes are done on the speech by means of various algorithms implemented as a computer program. The main goal of speech recognition area is to develop techniques and systems for speech input to machine. There are various types of Speech Recognition:

- Isolated Words: It accepts single word or single utterances at a time.
- Connected Words: Is same as isolated word but allows a proper pause between the utterances.
- Continuous Speech: In these users speaks almost naturally and with the help of special methods the utterances is calculated.

Spontaneous Speech system handles variety of natural speech features various words sounding same run together[17]. The performance of Speech recognition systems is usually specified in terms of accuracy and speed. The accuracy is measures in means of Word Error Rate (WER) and speed is measured in terms of real time factor. The Mel-Frequency Cepstral Coefficients (MFCC) has been the most widely used front-end feature for speech recognition. Pitch is one of the most important features of speech. Here pitch refers to the fundamental frequency contour of successive frames, not the psychological one we feel. Roughly speaking, the so called prosody modeling for automatic speech recognition and understanding are mainly refer to pitch contour estimation and application. The pitch contour parameter has been widely used in the fields of speech transmission, compression and recognition, and it has been used in scene analysis and human emotion detection [18]. One of the first theories proposed was the James-Lange theory, in the end of the 19th century. This theory states that emotion is the experience of some changes in the body. For example: when one sees a bear, he will run away, and the running away will provide him the feeling of fear.

### 4. EMOTION RECOGNITION

Emotion is an important aspect in the interaction and communication between people. Even though emotions are intuitively known to everybody, it is hard to define emotion. Emotion is a complex set of interactions among subjective and objective factors, mediated by neural/hormonal systems, which can:

- Give rise to affective experiences such as feelings of arousal, pleasure/displeasure;
- Generate cognitive processes such as emotionally relevant perceptual effects, appraisals, labeling processes;
- Activate widespread physiological adjustments to the arousing conditions;
- Lead to behavior that is often, but not always, expressive, goal directed, and adaptive.

Various methods are described for emotion recognition which are shown in table 2.

**Table 2: Methods in emotion recognition**

Sr no.	Methods	Description
1	Emotion tracking	Emotion tracking is one of the most important aspects for EEG based emotion recognition. In this a suitable interface was designed in order to elicit the six basic emotions as defines previously in the line of a specified experiment.
2	Higher Order Crossings (HOC)	This concept was adopted in order to extract the feature vector classification in the six emotion classes.

3	Empirical Mode Decomposition (EMD)	This method considers the signal at their local oscillation scale, subtracts the faster oscillation, and iterates the residual [19].
---	------------------------------------	--

## 5. CONCLUSION

Emotion recognition by computers is becoming very popular. This paper attempts to present various techniques that can be used to recognize emotions using speech and EEG brain signals. Though emotion and speech recognition are two individual research areas, we are attempting to emerge with the new technique which will comprise of emotion and speech together for detecting emotions of human being.

## 6. ACKNOWLEDGMENTS

The authors would like to thank the Department of Computer Science and Information Technology, Dr. B.A.M. University authorities for providing infrastructure and the work is supported by University Grant Commission.

## 7. REFERENCES

- [1] Kazuhiko Takahashi, "Remarks on Emotion Recognition from Bio-Potential Signals" 2nd International Conference on Autonomous Robots and Agents December 13-15, 2004, Palmerston North, New Zealand, 2004.
- [2] Tsang-Long Pao, Yu-Ye Chen, Jun-Heng Yeh, Yuan-Hao Chang "Emotion Recognition and Evaluation of Mandarin Speech Using Weighted D-KNN Classification" Department of Computer Science and Engineering, Tatung University, Taipei.
- [3] Carlos Busso, Zhigang Deng, Serdar Yildirim, Murtaza Bulut, Chul Min Lee, Abe Kazemzadeh, Sungbok Lee, Ulrich Neumann, Shrikanth Narayanan, "Analysis of Emotion Recognition using Facial Expressions, Speech and Multimodal Information", Viterbi School of Engineering, University of Southern California, Los Angeles.
- [4] Victor Zue, Ron Cole, & Wayne Ward, Speech Recognition, Present at <http://cslu.cse.ogi.edu/HLTsurvey/ch1node4.html>
- [5] Electroencephalography, Present at <http://www.bem.fi/book/13/13.htm>
- [6] Rafal Bogacz, Urszula Markowska-Kaczmar, Androzej Kozik, "Blinking Artefact Recognition in EEG signal using Artificial Neural Network", Computer Science and Management Department, Wroclaw.
- [7] M.Ungureanu, C. Bigan, R. Strungaru, V.Lazarescu "Independent Component Analysis Applied in Biomedical Signal Processing". Measurement Science Review, Volume 4, section 2, 2004.
- [8] Vijay Khare, Jayashree Santosh and Sneh Anand, "Classification of EEG signals based on Neural Network to Discriminate Five Mental states" SPIT-IEEE colloquium and International Conference, Mumbai, India
- [9] Kaushik Majumdar, Member IEEE, "Constraining Minimum-Norm Inverse by Phase Synchronization and Signal power of the scalp EEG channels", IEEE Transactions of Biomedical engineering, Vol.56, No.4, April 2009.
- [10] David R. Hardoon and John Shawe-Taylor, "Signal Extraction for Brain-Computer Interface", Oct 9, 2003. M. Young, The Technical Writers Handbook. Mill Valley, CA: University Science, 1989
- [11] Filipe Silva and Pablo Dias, "Brain-Computer Interface for navigating in virtual environments", MAP-I doctoral programme, 2008-2009.3
- [12] Ashkan Yazdani, Touradj Ebrahimi, Ulrich Hoffmann, "Classification of EEG signals using Dempster Shafer Theory and a K- Nearest Neighbor Classifier", 4<sup>th</sup> International IEEE EMBS conference on Neural Engineering Turkey, April 29-May 2, 2009
- [13] Lester Ingber, "Statistical mechanics of Neocortical Interactions: EEG Eigenfunction of short-term memory", Chicago Mercantile Exchange Centre.
- [14] Srinath Vukkadala, Vijaylakshmi.S and Vijayapriya.S, "Automated Detection of Epileptic EEG using Approximate Entropy In Elman networks", International Journal of Recent Trends in Engineering", Vol1, No.1, May 2009.
- [15] Vitaly Schetinin, "Polynomial Neural Networks Learnt to classify EEG signals", NIMIA-SC2001-2001, Measurement and Related Industrial Applications: Studycases Crema, 9-20 Oct 2001
- [16] Rafal Bogacz, URSZULA Markowska-Kaczmar, Androzej Korik, "Blinking Artefact Recognition in EEG signal using Artificial Neural Network", Computer Science and Management Department, Wroclaw. G. R. Faulhaber, "Design of service systems with priority reservation," in *Conf. Rec. 1995 IEEE Int. Conf. Communications*, pp. 3-8.
- [17] M.A. Anusuya, S.K.Katti, "Speech Recognition BY machine: A Review", Department of Computer Science and Engineering, Mysore, International Journal of Computer Science and Information Society, Vol.6.No.3, 2009
- [18] WANG Dong, CHEN Yi-Ning, LIU Jia " An Algorithm for Voiced/ Unvoiced decision and pitch estimation in speech feature extraction" ,Department of Electronical Engineering, Beijing.
- [19] Robert Horlings, "Emotion recognition using brain activity", Man-Machine Interaction Group TU Delft, 27 March 2008.