iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** XI **Month of publication:** November 2017

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue XI November 2017- Available at www.ijraset.com

Classification and Analysis of Emotion from Speech Signals

Renuka Vijayrao Kukade[1], Prof. G D Dalavi[2]
[1]*M.E. Student, EXTC Department, P.R.Pote (Patil) Welfare & Education Trust's Group of Institute, College of Engineering & Management, Amravati, Maharashtra, India.*
[2]*Professor, EXTC Department, P.R.Pote (Patil) Welfare & Education Trust's Group of Institute, College of Engineering & Management, Amravati, Maharashtra, India.*

Abstract: ***In recent years the workings which require human-machine interaction such as speech recognition, emotion recognition from speech recognition is increasing. Because emotion recognition from speech is an important area in research that represents human and computer interaction. Not only the speech recognition also the features during the conversation is studied like melody, emotion, pitch, emphasis. It has been proven with the research that it can be reached meaningful results using prosodic features of speech. In this paper we performed pre-processing necessary for emotion recognition from speech data. We extracted features from speech signal. To recognize emotion it has been extracted Mel Frequency Cepstral Coefficients (MFCC) from the signals. Speech emotion recognition results are also reviewed in this paper for the purpose of identifying the number of speakers, language used and emotion classification till date. This paper conducts an experimental study on recognizing emotions from human speech. The six emotions considered for the experiment includes neutral, anger, joy sadness, boredom, fear. The distinuguishability of emotional features in speech were studied first and then emotion classification is performed on a custom dataset.***

Keywords: Emotion Analysis, Emotion Classification, Speech Processing, Mel-Frequency Cepstral Coefficients, Human-Computer Interface

I. INTRODUCTION

one of the most natural form of communication between human and computer is the speech. Speech signal is one of the fastest methods of communications between humans. so the speech can be more efficient and fast method of inter-action between machine and human. Speech is a complex signal which contains information about the speaker, message, language and emotions. An emotion makes speech more effective and expressive. Different ways like laughing, yelling, teasing, crying, etc, are used by humans to express their emotions.

Emotion recognition through speech is an area which is increasingly attracting the attention the field of pattern recognition and speech signal processing in recent years. Automatic emotion recognition pays close attention to identify emotional state of speaker from voice signal. It is important medium of expressing humans perspective or fillings and his or hers mental state to others. Humans have natural ability to recognize emotions through speech information but the task of emotion recognition for machine using speech signal is very difficult since machine does not have sufficient intelligence to analyze emotions from the speech. Machine can detect who said and what is said by using speaker identification and speech recognition techniques but if we implied emotion recognition system through speech then machine can also detect how it is said. As emotions plays an important role in rational actions of human being there is a desirable requirement for intelligent machine human interfaces for better human machine communication and decision making. Emotion recognition through speech means detection of the emotional state of human through feature extracted from his or her voice signal(speech signal). Emotion recognition through Speech is particularly useful for applications in the field of human machine interaction to make better human machine interface. Speech emotion recognition has many more applications in our daily life.

Some of these applications include-

A. Most important application is for enhancing the human and machine interaction.
B. In psychiatric diagnosis, lies detection.
C. For analysing the behavioural study in call centre conversation between customer and employee.
D. In aircraft cockpits, for the better performance speech emotion recognition system trained for stress speech.
E. For understanding the criminals behaviour that would help for analysis in criminal department.

©IJRASET (UGC Approved Journal): All Rights are Reserved

F. Emotion recognition in robotic, it would be more realistic and enjoyable, if they understand and express the emotions like human.

II. DATABASE REVIEW

Emotion speech database collected for the variety in research. Speech corpora used for developing the recognition of emotion from speech system, it can be divided in three ways as follows:

A. Simulated based emotional speech database: Where the database is collected from actors, experienced and trained artists. In this artist express their natural words or sentence in different emotions. This one is the easier way to collect the database. More that 60% of databases are collected in this way.

B. Induced emotional speech database: This type of database is collected by simulating artificial emotional situation, without knowledge of the speaker. This database may be more natural as compare to simulated database, but there may be a problem when speaker know that they are being recorded, then they are not that much of expressive.

C. Natural emotional speech database: These type of emotions are sometimes difficult to recognize. Natural emotions also called underlying emotions. Natural database can be recorded form call centre conversation and emotional conversation between public places and so on.

III. WORKING AND BLOCK DIAGRAM

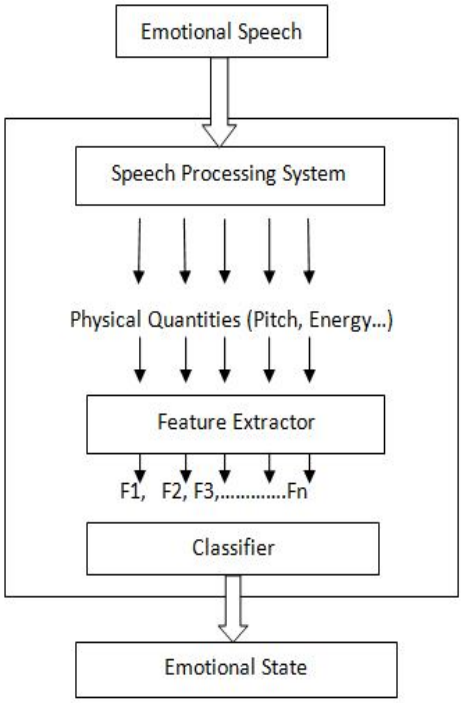


Fig -1: proposed system block diagram

MFCC is most widely used in audio classification experiments due to its good performance. It extracts and represent feature of speech signal. The Mel-cepstra takes short-time spectral shape with important data about the quality of voice and production effects . To calculate these coefficients the cosine transform of real logarithm of the short-term spectrum of energy must be done. Then it is performed in mel-frequency scale. After pre-emphasizing these speech segments are windowed . The simple window used for this process is a hamming window based on reduction of leakage effect. It smears energy from true signal frequency into neighboring ones thus negatively affecting the performance. It also contributes to avoiding the discontinuity of the speech signal in time domain that might occur during Fast Fourier Transform. FFT algorithm is used for converting the N samples from time domain to frequency

©IJRASET (UGC Approved Journal): All Rights are Reserved

domain. It is used to evaluate frequency spectrum of speech. The concept of windowing is based on multiplying the signal frames by window function. Windowing is used to window each individual frame in order to remove the discontinuities at the start and end of the frame. Hamming window is mostly used due to its relatively narrow main lobe width hence, removes distortion.

IV. FEATURE EXTRACTION

Speech is partitioned into small intervals known as frames. The process of partitioning speech into frames based on the information they are carrying about emotion is known as feature extraction. Feature extraction is the vital step in SER (speech emotion recognition) system. Some of the features that help to figure out emotions from speech are-

A. Pitch

It is the main component of any speech which is defined as the lowness or highness of a voice as identified by the human ears. Pitch is dependent on the vibrations per second. The value of pitch parameter is extracted by using cepstrum in the frequency domain. Pitch helps in identifying the neutral and angry emotions from speech sample.

B. ENERGY

Intensity of the speech defines the energy level of speech. Energy level for each frame is calculated as first the square of all sample amplitude is done and then summing up the values of all the squared sample amplitudes.

C. Pitch difference and energy difference

Difference between the values of pitch or energy level of neighbouring segments is use to categorised speech parameters into emotions. The more the fluctuation, more it is easier to reveal the lively emotions like happiness and anger.

D. Formants

Formants are governed by the shape of the vocal tract and are manipulated by different emotions for eg, the state of excitement results in obtaining the higher mean values of the first formant frequency. The fundamental frequency helps in identifying happy emotion from speech samples.

E. Mel-Frequency Cepstrum

1) *Coefficient (mfcc) :*MFCC is the most vital parameter in which best describes the emotional state by using simple calculations. MFCC also provides good frequency resolution when the speech frequency is low. MFCC based parameters show the energy migration in frequency domain and also helps in identifying phonetic characteristics of speech.

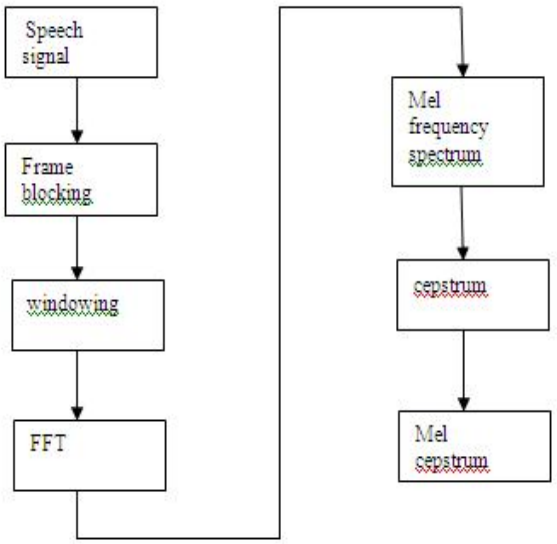


Fig -2: MFCC Block Diagram

V. RESULT

©IJRASET (UGC Approved Journal): All Rights are Reserved

In this project the data we need to give as an input is need to be in ".wav" format. T he steps of emotion recognition include selecting the input sounds , followed by adding the selected sound to the database then we can check the entire database information on the screen and the we can recognize the emotion of a person by giving input samples to check.

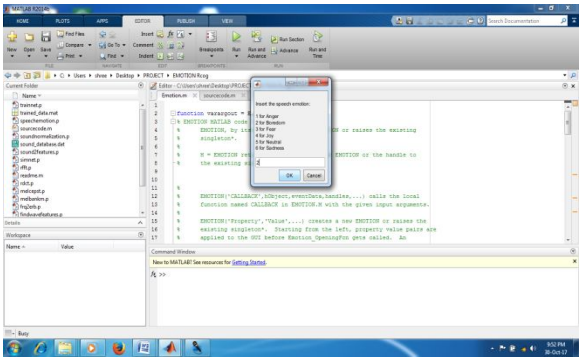


Fig -3: Add Input database

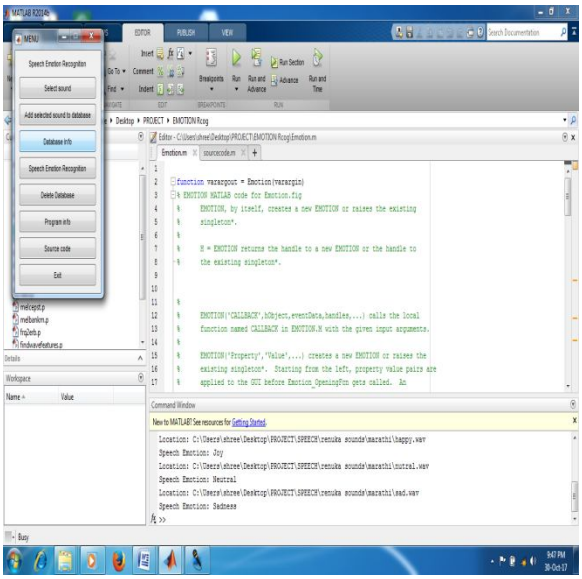


Fig -4: Input database information

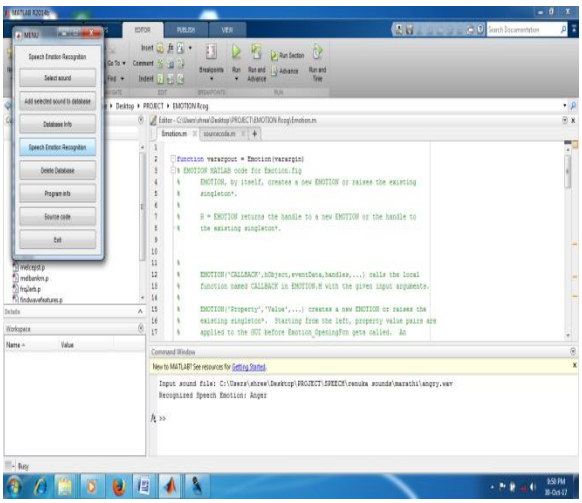


Fig -5: Speech emotion recognition

©IJRASET (UGC Approved Journal): All Rights are Reserved

VI. CONCLUSIONS

Emotion recognition system is an important research area in today's fields. There are the several applications where speech emotion recognition can be deployed. A properly and well designed database is essential for developing the emotion recognition system. This research paper covers the recent work of speech emotion recognition for filling some important re-search gaps. This paper contains the research of recent works in speech emotion recognition from the points of views of emotional databases, speech features, and classification model. The paper explores the idea of detecting the emotional state of a person by speech processing techniques. The study on words and letters under different emotional situations proved that the emotional state can alter the speech signal. The development of a software based agent for emotion detection and heart rate analysis can greatly improve telemedicine based systems can also be improved.

VII. ACKNOWLEDGMENT

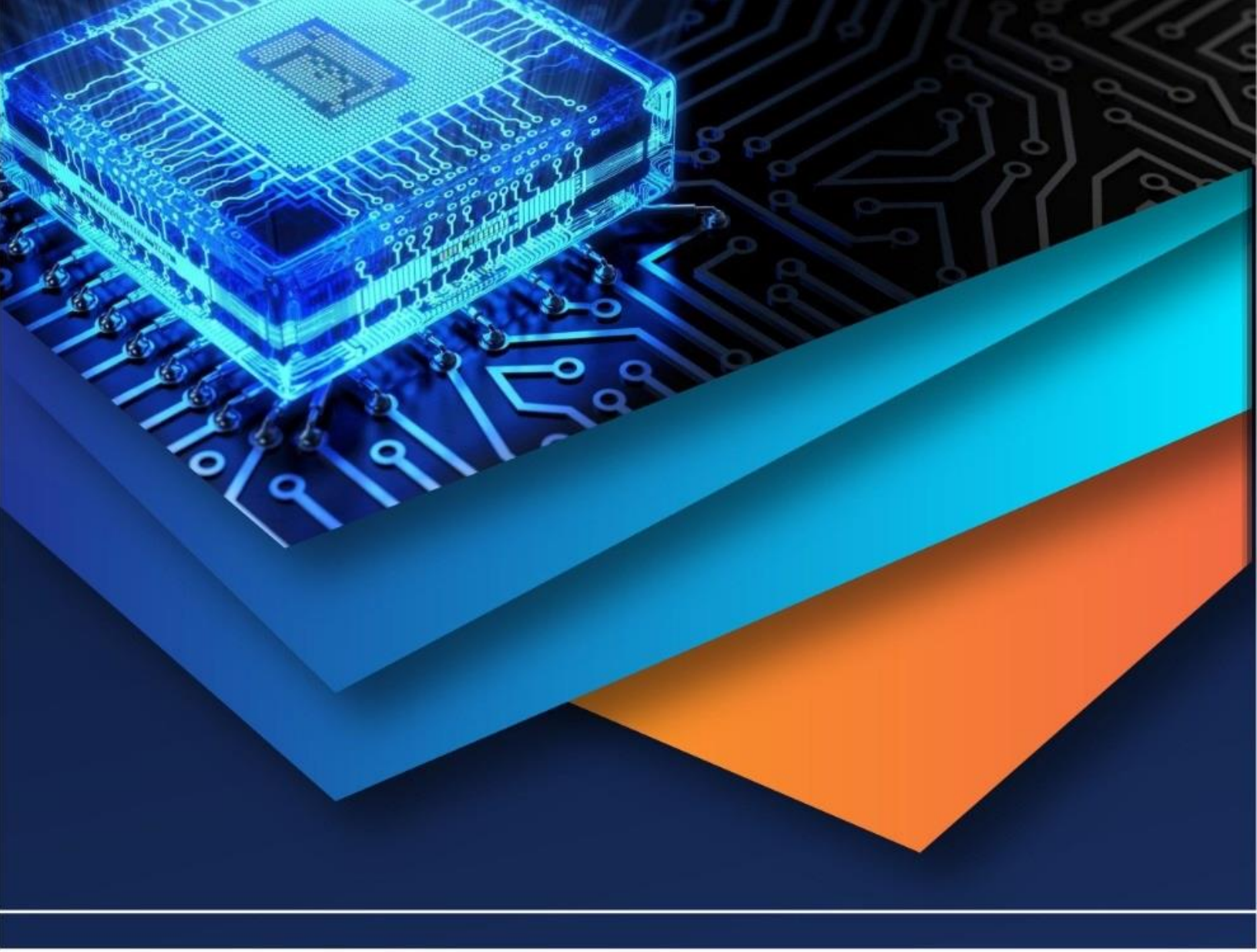
We would like to acknowledge the faculties of Electronics & Telecommunications Department, P. R. Pote (Patil) Welfare & Education Trust's college of Engineering & Management, Amravati for their valuable support.

I am Miss Renuka V. Kukade specially like to express my sincere thanks to my guide Prof. G. D. Dalvi and respective faculties for their motivation and useful suggestions which truly helped me in improving the project work.

REFERENCES

[1] "Analysis Of Emotion Recognition System Through Speech Signal Using KNN, GMM & SVM Classifier" by Chandra Praksah , Prof. V. B. Gaikwad in ijecs (June 2015)
[2] "Understanding and estimation of emotional expression using acoustic analysis of natural speech" by Nilima Salankar Fulmare ,Prasun Chakrabarti and Divakar Yadav in International Journal on Natural Language Computing (IJNLC) Vol. 2, No.4, October 201
[3] "Speech Emotion Recognition Based on SVM Using MATLAB by Ritu D.Shah1, Dr.Anil.C.Suthar2 in IJIRCCE (March 2016).
[4] "A Survey on Speech Emotion Recognition in Human Computer Interaction" by Sanju D. Garle, Angad Singh, Girish Kapse in IJEMT (January 2016).
[5] "Recognition of Human Emotion from a Speech Signal Based on Plutchik's Model" byDorota Kami ´nska and Adam Pelikant in jet intl(may 2012).
[6] "Feature Extraction from Speech Data for Emotion Recognition" by S. Demircan and H. Kahramanlı in Journal of Advances in Computer Networks, Vol. 2, No. 1(March 2014).
[7] "Emotion recognition in speech signal using emotion extracting binary decision trees Jarosław Cichosz, Krzysztof Ślot supported by the Polish State Fund for Research Support
[8] "Speaker Emotion Recognition Based on Speech Features and Classification Techniques" J. Sirisha Devi, Dr. Srinivas Yarramalle, Siva Prasad Nandyala in MECS(in June 2014).
[9] "Speech Emotion Classification using Machine Learning" by Pooja Yadav, Gaurav Aggarwal in International Journal of Computer Applications (May 2015).
[10] "Analysis of speech under stress using linear techniques and non-linear techniques for emotion recognition system" by A. Khulageand Prof. B. V. Pathak in csit (2012).
[11] " Detection of Speech under Stress: A Review" by Neha P. Dhole, Dr. Ajay A.Gurjar in IJEIT (April 2013).
[12] " detection of speech under stress using spectral analysis" by neha p. dhole, dr. ajay a.gurjar in ijret (Apr-2013).
[13] "Detection and Analysis of Emotion From Speech Signals" by Assel Davletcharova, Sherin Sugathan, Bibia Abraham, Alex Pappachen James in ScienceDirect(2015).
[14] "Voice Stress Analysis and Evaluation" by Darren M. Haddada, Roy Ratleyb in Proc. SPIE Vol. 4232(2014).
[15] " Emotion Recognition from Speech: A Survey" by Rani P. Gadhe,Shaikh Nilofer R. A., V. B.Waghmare, P. P. Shrishrimal, R. R. Deshmukh in International Journal of Scientific & Engineering Research, (April-2015).

©IJRASET (UGC Approved Journal): All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)