

EMOTION DETECTION ANALYSIS THROUGH TONE OF USER: AN IMPLEMENTATION

Ankita Dutta Chowdhuri¹ and Sachin Bojewar²

Abstract- Speech emotion detection aims to analyze the speech emotion state based on the training and testing process along with classifiers to the database provided. The below framed work has been subdivided in three sections namely HAPPY ,SAD & FEAR. This reserch & implementation work has two major phases that is training and testing. The training has been done on the basis of wave files provided for every group. Features have been extracted for all groups and have been saved into the database. The testing section classifies the training set of data with the help of BACK PROPAGATION Method. The results of the BACK PROPAGATION Method have been found superior in terms of classification accuracy.

Keywords – Speech Emotion,Back propogation,Fear,Happy,Sad

I. INTRODUCTION

Emotion recognition with speech has now a day's getting attention of engineers in field of pattern recognition and speech signal processing. As computers is one of the important part now a days, so the requirement for communication between computer and humans. Through voice signals Automatic emotion recognition recognize the emotional state of the speaker. In humans life emotions plays a vital role. Humans show's their perspective or feelings or their mental state to others through emotions. Human have a natural gift to sense the emotion through speech on the other hand if we take machines to understand emotions through speech is bit difficult as it lacks the basic intelligence to monitor emotions. Understanding emotions in speech is a convoluted task as there is no obvious solution to what the right emotion for a given speech sample. Through speaker identification and speech recognition techniques machines can understand what is said but if we use emotion recognition system then machine can understand how it is said. As in human being emotions have a vital role for their different actions so there is a favorable requirement for a system interface through which we can have a good human machine communication and decision making. Emotion recognition through speech defines disclosure of emotional state of humans through feature extracted from his or her voice signal. Some other applications which require this interface such as Interactive movie, storytelling, and electronic machine pet, remote teach school & E-tutoring application Different intelligent systems have

¹ *Department of Computer Engineering ARMIET, Asangaon, Maharashtra, India*

² *Dept of Information Technology, Vidyalkar Institute of Technology, Mumbai University, India*

been developed on the basis of different emotions such as anger, happiness, sadness, surprise, neutral, disgust, fearful, stressed etc .

II. PROPOSED ALGORITHM

There are number of milestones that need to be achieved in order to reach the goal, so following are the sequential steps that will led to attain the signal.

1. Ready database.
2. Train Systems using the feature extraction method. The features extracted are Maximum Frequency, Average Frequency, Minimum Frequency, Roll off, Pitch, and Loudness etc.
3. Repeat step 2 for all categories.
4. Save data to db.
5. Upload a file to test.
6. Extract same features of the uploaded file.
7. Set Target values for the evaluation.
8. Apply BPNN algorithm for classification.
9. Compare the accuracy of SMO and BPNN

The main function of back propagation method is that it moves from input to output layer. In the back propagation neural network error is found at output layer, so to remove error at output layer, moves backward takes place , hence name is back propagation neural network.

There are two sections in our research work. The sections are explained as follows.

A. Training: The training section ensures that the database gets trained properly so that at the time of testing it produces wide-ranging results. The characteristics of the training are as follows.

□ Maximum Frequency: It is the rate which we get at the tip on a frequency chart. When so ever we place a voice sample over the time and frequency model, the maximum peak is known as the maximum frequency of the voice

illustration.

□ Minimum Frequency: It is the rate which we get at the tip on a frequency chart. When so ever we place a voice sample over the time and frequency model, the minimum peak is known as minimum frequency of the voice

illustration.

□ Average Frequency: The average frequency can be considered using two techniques. The first technique is to append all the frequency samples and then separate the entire sum with the total number of frequency. The second

method is an incredibly principled method in which we can insert the minimum frequency and the maximum frequency and then we can divide them by two.

Average Frequency = (Minimum frequency + Maximum Frequency)/2

□ Spectral Roll off: The spectral roll off is difference between the maximum frequency differences with the adjacent frequency.

□ Noise Level: The noise level is the additional amount of bits which has been supplementary into the voice sample.

1. Uniform Noise: Uniform noise is the noise which is concurrently equivalent all over the voice sample.

2. Non Uniform Noise: Non-uniform noise is the noise that does not remain constant all over the sample.

- Pitch: It is the standard value of the entire voice sample.
- Spectral Frequency: The spectral frequency is the frequency of the voice pitch subsequently to the highest voice sample.

B. Testing: In the testing phase, the trained network was computer-generated with unknown speech blueprint. It was observed that the trained network performs very well and more than ten words can be recognized by using the developed system

Database

A data base is the gathering of data. In our proposed work we have used speech samples of three actors for the database. The actors are male & female both with different age groups. In the database we find attributes of the speech signals and then we save them into the database. The query is that how we are going to achieve the target store hundreds of files in the database. The steps would be as follows. First of all we would fetch the attributes of the voice samples. All those properties which are required would be calculated and then it would be stored into a storage. The array would move on as the files would move. We would fetch the features and would take the average by the end and then store them into the database for each category of the voice which we have taken i.e. HAPPY, SAD AND FEAR.

BPA Algorithm

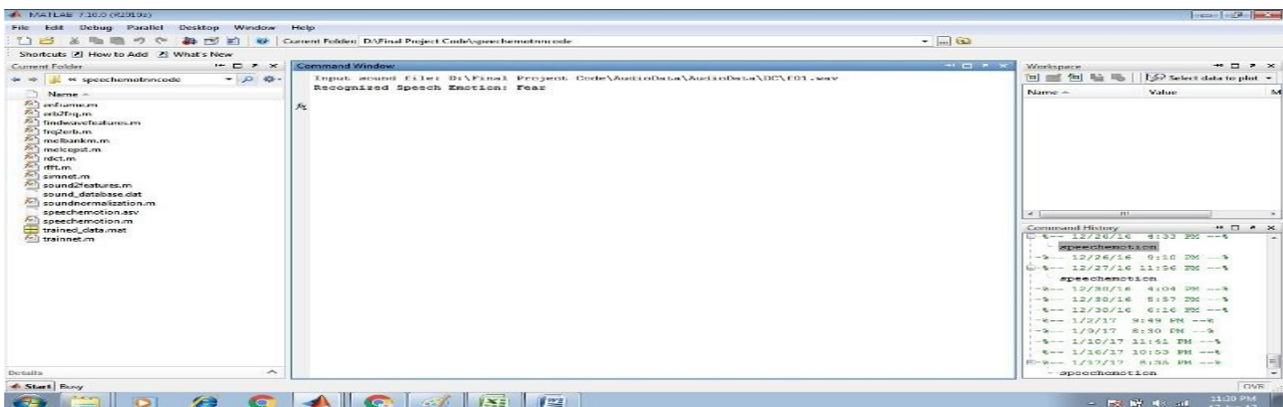
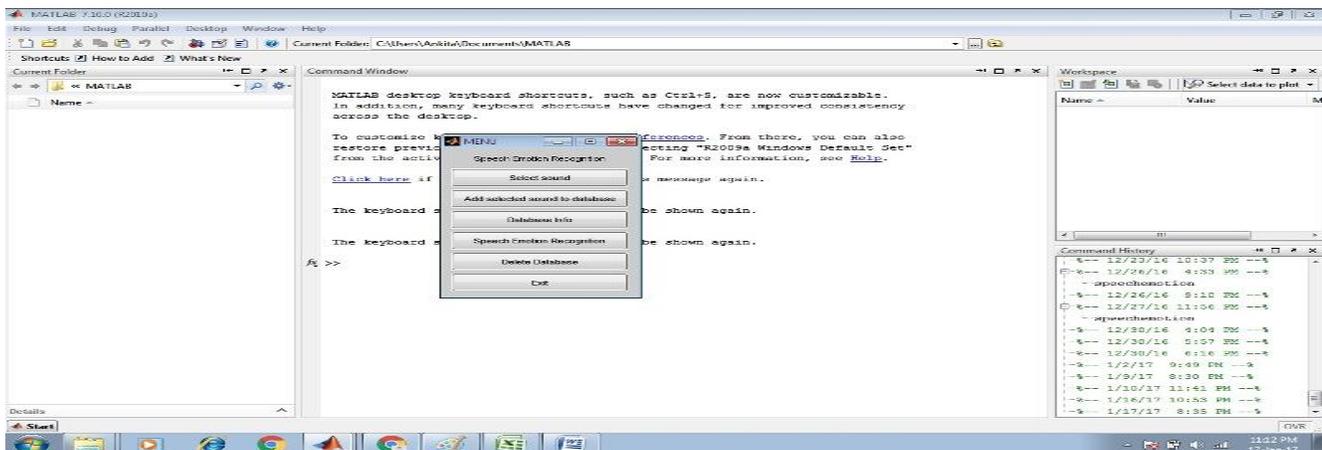
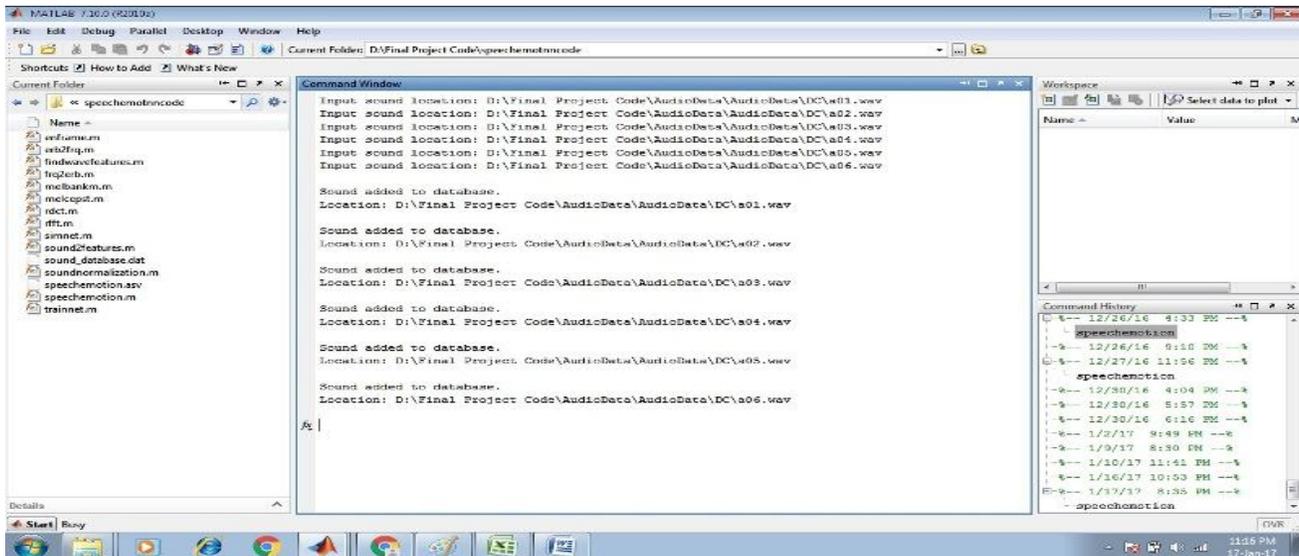
BPA algorithm is one of the best algorithms of Neural Network [10]. BP algorithm can be broken into four main steps:

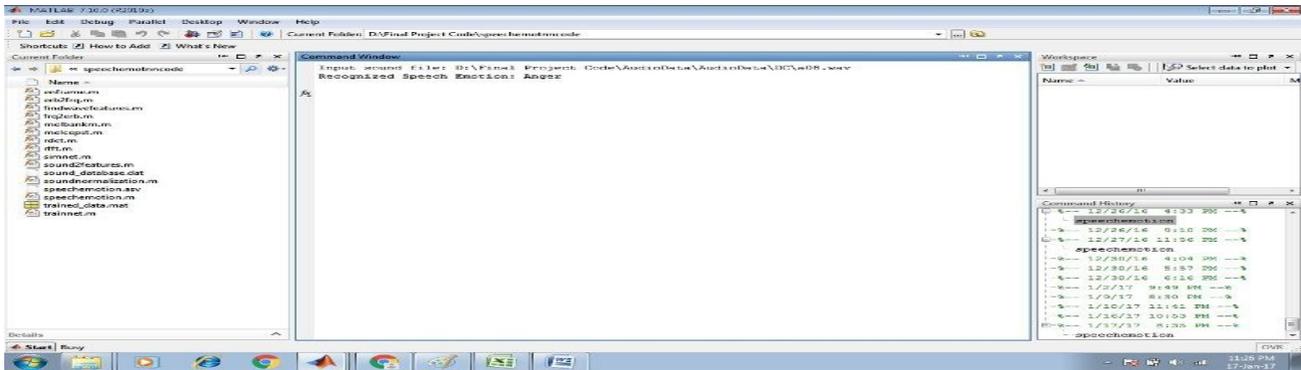
- i) Feed-forward calculation
- ii) Back propagation useful on output layer
- iii) Back propagation applied on hidden layer
- iv) Updating of weight performed

III. EXPERIMENT AND RESULT

Audio Samples	Age Group	JOY	ANGRY	FEAR
(Female speaker)	In-between (25-30)	9/10	10/10	7/10
(Male Speaker)	In-between (21-25)	10/10	9/10	6/10
(Female Speaker)	In-between (55-60)	10/10	8/10	9/10

Table 1: Output Results per 10 input audio samples





IV.CONCLUSION

Although it is impossible to accurately compare recognition accuracies from this study to other studies because of the different data sets used, the methods implemented here are extremely promising. The recognition accuracies obtained using back propagation are higher than any other study. Previous studies have neglected to separate out male and female speakers. This project shows that there is significant benefit in doing so. Our methods are reasonably accurate at recognizing emotions in female and all speakers. Our project shows that features derived from agitated emotions

REFERENCES

[1] Dr.Yousra F., Al-Irhaim Enaam Ghanem Saeed, "Arabic word recognition using wavelet neural network", *Scientific Conference in Information Technology*, November 2010.
 [2] Sonia Sunny, David Peter S, K Poulouse Jacob, "Design of a Novel Hybrid Algorithm for Improved Speech Recognition with Support vector Machines Classifier", *International Journal of Emerging Technology and Advanced Engineering*, vol.3, pp.249-254, June 2013.
 [3] Tingxiao Yang, "The Algorithms of Speech Recognition, Programming and Simulating in MATLAB", *University of Gavale*, pp.1-49, January 2012.

-
- [4] Scherer S, Hofmann H, Lampmann M, Steffenrhinow M, Mschwenker F, Untherpalm G. Emotion recognition from speech: Stress experiment. Germany; 2008. p. 1–6.
- [5] Shashidhar G, Koolagudi K, Rao S. Emotion recognition from speech: A review. *Int J Speech Technol.* 2012; 3(2):1–5.
- [6] ElAyadi MA, Mohamed SB, Karray BF. Survey on speech emotion recognition: features, classification schemes, and databases. *Pattern Recognition.* 2011; 44(3):572–87.
- [7] Kim EH, Hyun KH, Kim SH, Kwak YK. Improved emotion recognition with a novel speaker-independent feature. *IEEE/ASME Transactions on Mechatronics.* 2009; 14(3):317–25