



Emotion Classification and Detection

Yogesh Gulhane

Spine's College of Engineering & Technology,
Amravati, Maharashtra, India

Dr S. A Ladhake

Spine's College of Engineering & Technology,
Amravati, Maharashtra, India

ABSTRACT

Energy-in-Motion is symbol of living hoods. This Energy-in –Motion is also knows as emotion. Digitization and Globalization change our traditional ways of living .This change are not only in physical but also physiological and hence , new generation is in a dilemma of unfit and unequipped also this generation is has not a mental and physical strength to face and handle the emotional phases in for example stress or negative emotion . Now everyone is a part of race of success, from day one of school students enters in a world of in competition and stress. From the Kids to the youth are in stress of examination. Kids are missing there childhood. To bring their original age on their face is a need of understanding, Communication between the parents, teachers and friends. Right actions at right time can pull out them from the stress.. Hence researchers are working with this area of stress management. For our research work we simply tried to build a system 'Emotion Analysis using Rule base system.' Which can help to detect and classify the emotions?

Keywords: *Emotion, Rule base system, MFCC, Audio Frequency, Signal processing*

INTRODUCTION

In the fast track of life there is a interesting research areas in case of non-verbal communication. It is related to living hoods only, especially with youth living a fast and changing modern life and correlates of stress. Like in students while facing the race of competition no. of factors are important, in daily life one factor can see everywhere as student has to face is 'Stress of Study'. The non-verbal content or information is a silent factor about internal feeling and emotions [4,5]. This emotional feeling comes out from heart and express in the term of

sounds in speech. In this research we developed a system to classify the e-motion. We have shown result achievement of two input types of audio signals. For our research we have tested two of recordings alike we collect and test the recorded wav file and then compare with standard speech database for better result. The standard recordings were taken by recording equipment. The complete database is collected and evaluated with preserving their naturalness. The database can be accessed by the public via the internet (<http://www.expressive-speech.net/emodb>). In the second test type we tried to test real input /sound signals wav file and then compare with standard speech database for better result.

AIM

With the focus of student stress and emotional speech there is a need of implementation of automatic emotion detection system which can analysis the emotion of the students and classify according to mood and extract feature like energy. Different application has been implemented by the researchers with considering the area of audio signal processing and stress management. While implementing this system we had a focus on two steps for processing the voice.

- 1] Recording the audio through microphone or collect the available sample from sources.
- 2] Fundamental frequency evaluation in the speech signal.

EXPERIMENTAL SETUP

For studying real time input, to keep the naturalness in the speech signal under the different emotional

situations, we prepare a soundproof room to avoid the external barrier in the speech input. For recording the real input from the various students we use microphone connected to the system, we recorded variable input speech signals from the student having different mood with subjects of examination. These student were asked to express certain feeling about their exam or subject paper at the same time speech was recorded without aware them about recording to preserve the naturalness. Test is conducted for the Indian students and they spoke English or Marathi sentences under different emotional states. At the time of recording micro- phone was kept at a certain distance from the mouth. For feature extraction from the recorded speech segments, MATLAB functions were used.

MATHEMATICAL MODEL KEYS

Power and Energy content are used to calculated. Power = mean(x.^2) and energy = sum(x.^2) of the audio signal equation

$$E = T \sum_{n=0}^{N-1} x^2 [n] \text{-----} (1)$$

$$P = \frac{1}{N} \sum_{n=0}^{N-1} x^2 [n] \text{-----} (2)$$

1 and 2 shows the Energy E and power P respectively.

Where x(n) is the n:th sample within the frame and N is the length of frame in samples.

These parameter vectors can be described using GMM:

$$p(o | \lambda^s) = \sum_{i=1}^M w_i^s p_i^s(o),$$

Where M is components of class, w i, i = 1,..., M are weights of that sum of all weights is 1, and p means the probability . And others are mean value and covariance matrix C i.

Gaussian model can be defined by

$$\lambda^s = \{w_i^s, \mu_i^s, C_i^s\}, \quad i, \dots, M^s.$$

Using above factors we are able to detecting F0 detection in time domain, F0 plays an important role in frequency domain and F0 from cepstral coefficients. Popular autocorrelation function is used to determine the position of the first peak with the help of Pitch extraction concept . Simple formula is use for the final l calculation of the fundamental frequency as given bellow

$$F_0 = \frac{F_s}{k} .$$

RESULT

For a clean experimental setup everything except the issue under study is kept constant. Number of student speakers specks naturally with the emotions that they have to perform. Recordings are at high audio quality and without noise without which spectral measurements would not been possible. This experiment shows the emotion detection whose accuracy outperforms a Better than a number of papers Moreover, it achieves this in real-time, as opposed to previous work base on stored data. The novel application of the system for speech quality assessment also achieves high detection accuracies. Figure 3 Shows the Output classification result of Emotions and Table 1 shows the performance and the result.

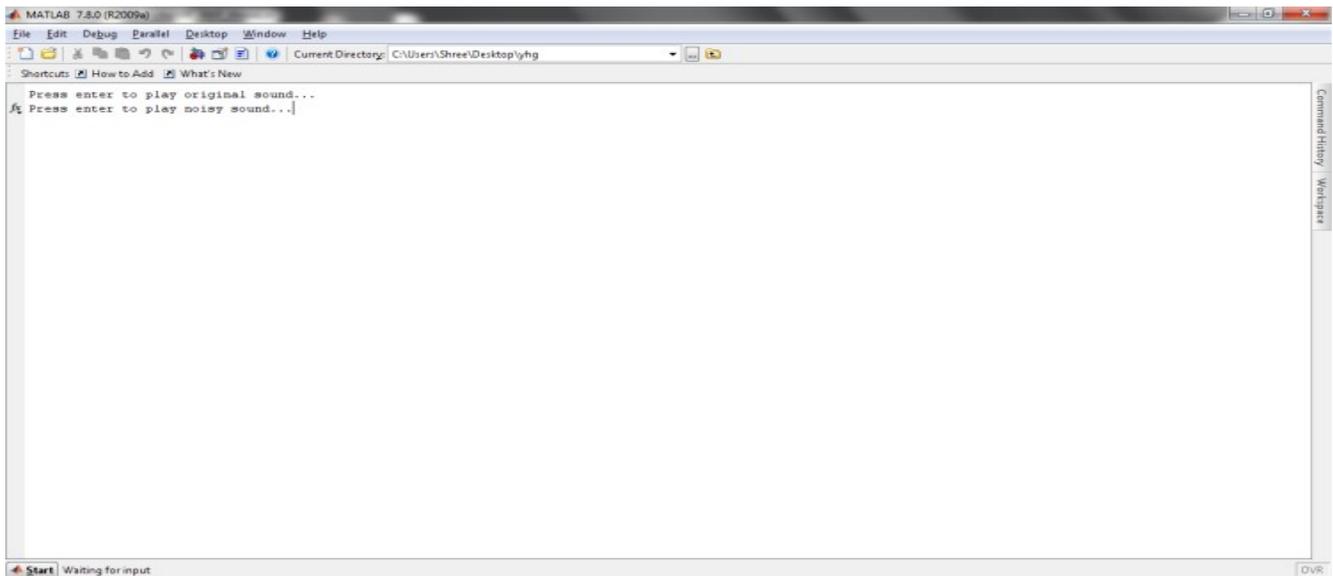


Figure A: Input for the testing.

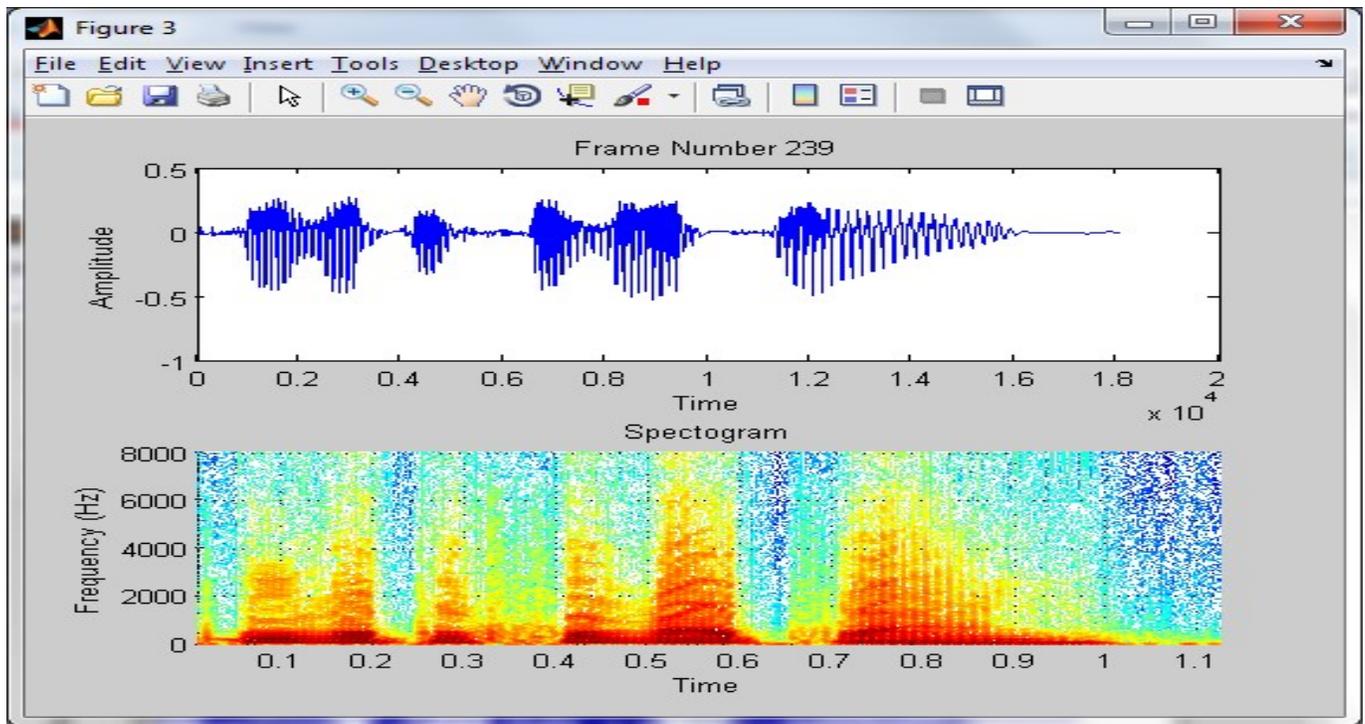


Figure B: Input and spectrogram of input sound

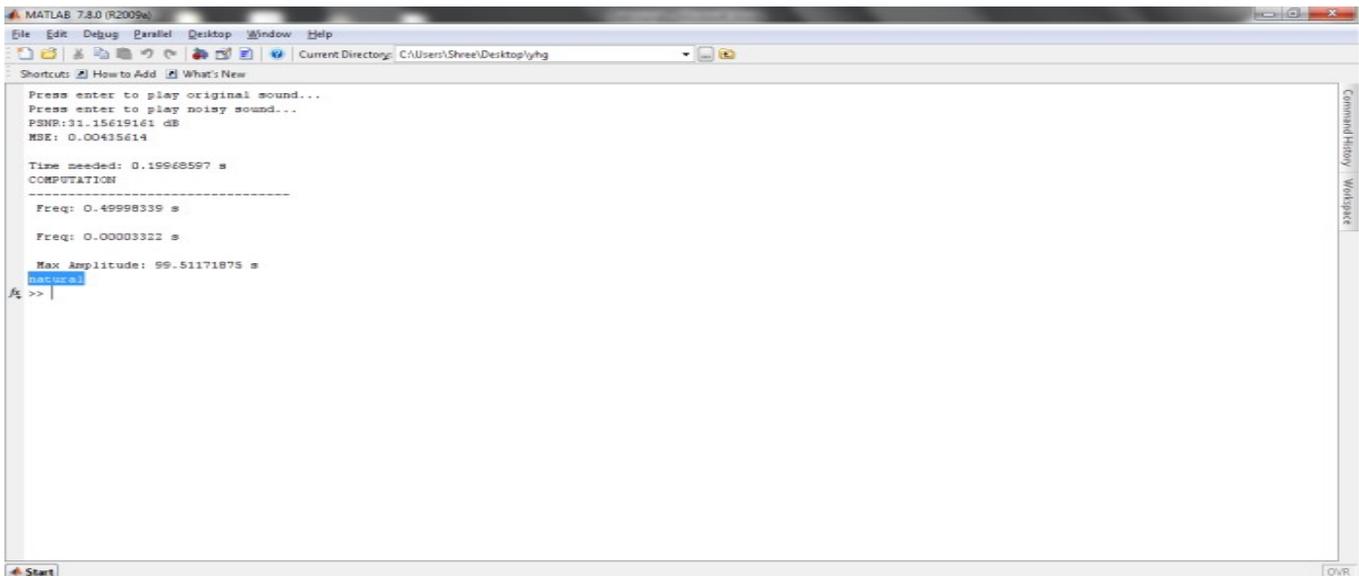


Figure 1: Features of the input Sound

Table 1: Classification of emotion

Category	Happy	Sad	Angry	Natural	Fear	boared	Error%
Happy	49	0	1	0	0	0	0.5
Fear	0	1	0	0	48	1	1
Natural	1	0	2	47	0	0	2
Boared	0	2	0	0		48	1
Angry	0	0	50	0	0	0	0
Sad	0	49	0	0	0	1	0.5

CONCLUSION

Earlier system shows there were a much more disadvantages without the MFCC features. Over a 20% drop in performance shows that overall the included features of the data set, it appears that the MFCC is the most important and also As it was found that not clustering data was advantageous to predicting the emotion, it was hypothesized that perhaps clustering provided some advantages to training time compared to the full feature set [1,6,7,8]. This experiment shows the emotion detection whose accuracy outperforms a Better than a number of papers Moreover, it achieves this in real-time, as opposed to previous work base on stored data. The novel application of the system for speech quality assessment also achieves high detection accuracies.

REFERENCES

1. Dominik Uhrin ; Zdenka Chmelikova ; Jaromir Tovarek ; Pavol Partila ; Miroslav Voznak "One approach to design of speech emotion database" *Proc. SPIE 9850, Machine Intelligence and Bio-inspired Computation: Theory and Applications X*, 98500B (May 12, 2016); doi:10.1117/12.2227067
2. Yixiong Pan, Peipei Shen and Liping Shen, "Speech Emotion Recognition Using Support Vector Machine", *International Journal of Smart Home*, Vol. 6, No. 2, April, 2012.
3. Ayadi M. E., Kamel M. S. and Karray F., "Survey on Speech Emotion Recognition: Features, Classification

4. *Schemes, and Databases', Pattern Recognition*, 44(16), 572-587, 2011.
5. Zhouy., Sun Y., Zhang J, Yan Y., "Speech Emotion Recognition using Both Spectral and Prosodic Features",IEEE, 23(5), 545-549, 2009.
6. Anurag Kumar, Parul Agarwal, Pranay Dighe1, "Speech Emotion Recognition by AdaBoost Algorithm and Feature Selection for Support Vector Machine".
7. T.-L. Pao, Y.-T. Chen, J.-H. Yeh, P.-J. Li, "Mandar in emotional speech recognition based on SVM and NN ", *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*, vol. 1, pp. 1096-1100,September 2006.
8. J.Lee and I Tashev. High-level feature representation using recurrent neural network for speech emotion recognition. In Interspeech 2015, 2015.
9. K.Han, D. Yu, and I. Tashev. Speech emotion recognition using deep neural network and extreme learning machine. In Interspeech 2014, 2014.
10. K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet *classification*. In ICCV 2015, 2015.