Classification and Detection Model for Emotional States

Dr. Yogesh Haridas Gulhane

Assistant Professor, Department of Electronics & Telecommunication Engineering, Sipna College of Engineering & Technology, Amaravati

***Abstract* - Basic nature of features in speech under different emotional situations are different. Features are varied by place to place and gender and age also can reflect the variation in features of speech. In this research paper proposes to classify the types of emotions and impact of it on performance. Also, this research took a look at variation in features with respective genders, age and places. Implementation cases used data from three subjects. As part of the real input from a microphone, we recorded the voice of different subjects. The subjects were asked to express certain emotions when their speech was recorded. The subjects were studied Mongolian, Indians and they spoke English sentences under different emotional states. A micro-phone was used to record the speech and was kept at a distance about 15cms away from the mouth. The experiments were conducted in an ordinary classroom having an area of 25m2. For extracting features from the recorded speech segments, MATLAB functions were used. Success ratio of the model is 99%. Confusion Matrix is use to get the unidentified signals.**

***Index Terms* - signal Processing, Classification, Detection, SVM, Emotion Analysis.**

I.INTRODUCTION

The analysis of the recorded speech signals was done in a MATLAB environment which provides several graphical visualizations for analyzing a signal.

There are several reasons why changes in F0 with time are potentially capable of providing information concerning the emotional state of a speaker. Examples are enjoyment of learning; hope and pride related to success; and anxiety and shame related to failure. Achievement emotions are pervasive in academic settings, especially so when the importance of success and failure is made clear to students. First, considerable latitude is possible in the variations of FO, since only certain aspects of the F0 contour carry information [willum 1972].

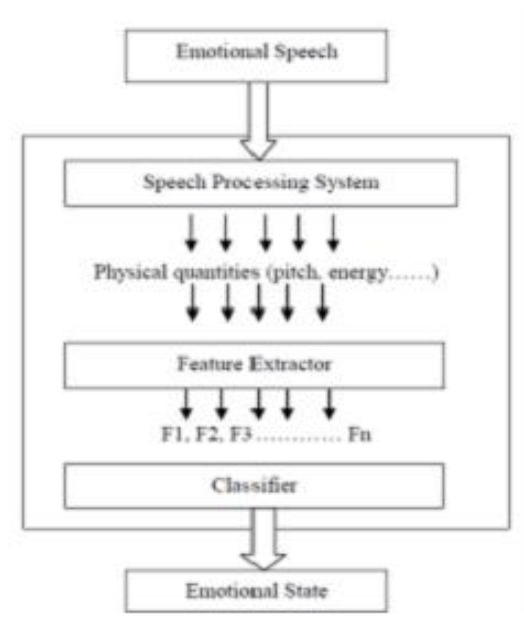


Figure 1.1- System Architecture

The Mel-frequency cepstral coefficients (MFCC) are widely used in audio classification experiments due to its good performance. It extracts and representational features of the speech signal. The Mel-cepstra takes short-time spectral shape with important data about the quality of voice and production effects [1]. A simple algorithm is used for estimating the reparability of the audio classes. In other words, a measure that describes how "easily" the features will be classified. In the case of a multi-class classification problem, the measure is calculated for each class opposed to all other classes, i.e. a measure value for each class is computed.

Designation	Frequency	Wavelength
ELF	Extremely low frequency	3Hz to 30Hz 100'000km to 10'000 km
SLF	superlow frequency	30Hz to 300Hz 10'000km to 1'000km
ULF	ultralow frequency	300Hz to 3000Hz 1'000km to 100km
VLF	very low frequency	3kHz to 30kHz 100km to 10km
LF	low frequency	30kHz to 300kHz 10km to 1km

MF	medium frequency	300kHz to 3000kHz 1km to 100m
HF	high frequency	3MHz to 30MHz 100m to 10m
VHF	very high frequency	30MHz to 300MHz 10m to 1m
UHF	ultrahigh frequency	300MHz to 3000MHz 1m to 10cm
SHF	super high frequency	3GHz to 30GHz 10cm to 1cm
EHF	extremely high frequency	30GHz to 300GHz 1cm to 1mm

There are dozens of emotions. They include anger, contempt, enthusiasm, envy, fear, frustration, disappointment, embarrassment, disgust, happiness, hate, hope, jealousy, joy, love, pride, surprise, and sadness. There have been numerous research efforts to limit and define the dozens of emotions into a fundamental or basic set of emotions. But some researchers argue that it makes no sense to think of basic emotions because even emotions we rarely experience, such as shock, can have a powerful effect on us.

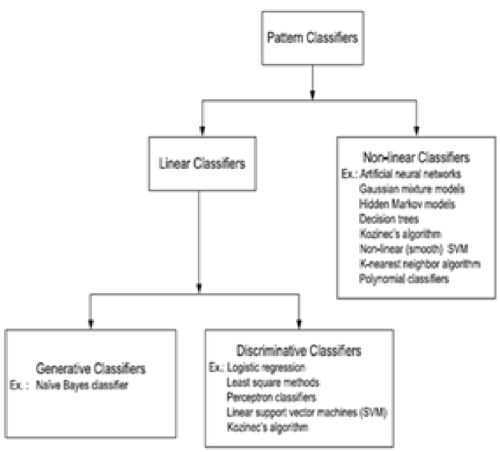


Figure 1: classifier Classification

P. Gangamohan et al,[14,15] Researchers reviewed and analyzed methods used for emotional speech and summarized overview of emotional speech data collection in their contribution towards the Analysis of Emotional Speech. This data collection is developed by different research groups. A hierarchical binary decision tree approach was used to accomplish the goal of research. Researchers characterizing the emotions as deviations in their work. In [16] researcher compare four ways to extend binary support vector machines (SVMs) to multiclass classification and shows that work done with SVM gives result accuracy of 75 to 80%. In general, pattern recognizers used for speech emotion classification can be categorized into two broad types namely

II.RESEARCH METHODOLOGY

Topic emotions pertain to the topics presented in the lessons. Examples are empathetic with the fate of one of the characters portrayed in a novel, anxiety and disgust when dealing with medical issues, or enjoyment of a painting discussed in an art course. Both positive and negative topic emotions can trigger students' interest in learning the material. In this research work, we performed four main stages. The first stage is called the speech analysis. The second stage prepares feature vector composed of static and dynamic features with MFCC feature extraction, third stage is classification using SVM. Finally, the last stage analysis of student performance under stress. Social emotion relate to teachers and peers in the classroom, such as love, sympathy, compassion, admiration, contempt, envy, anger or social anxiety. These emotions are especially important in teacher/student interaction and in group learning the Human e-motion at the time of Social Activity.

The speech signal is first divided into time frames consisting of an arbitrary number of samples .The filter coefficients w (n) of a Hamming window of length n are computed according to the formula:

$$W(n) = 0.54 - 0.46\cos(\frac{2\pi n}{N-1}), 0 \le n \le N-1$$

,Where N is total number of sample and n is current sample.

The relation between frequency of speech and Mel scale can be established as:

Frequency (Mel Scaled) =>

$$M(f) = 2595\log(1+\frac{f}{700}) \tag{2}$$

To go from Mels back to frequency:

$$M^{-1}(m) = 700(\exp(\frac{m}{1125})-1) \tag{3}$$

As don't have the frequency resolution required to put filters at the exact points calculated above, so we need to round those frequencies to the nearest FFT bin. This process does not affect the accuracy of the features. To convert the frequencies to fft bin numbers we need to know the FFT size and the sample rate,

f(i) = floor((nfft+1)*h(i)/samplerate)
This results in the following sequence:
f(i) = 9, 16, 25, 35, 47, 63, 81, 104, 132, 165, 206, 256
We can see that the final filterbank finishes at bin 256, which corresponds to 8kHz with a 512 point FFT size. MFCCs use Mel-scale filter bank where the higher frequency filters have greater bandwidth than the lower frequency filters, but their temporal resolutions are the same.
Each frequency range has a band designator and each range of frequencies behaves differently and performs different functions. The frequency spectrum is shared by civil, government, and military users of all nations according to International Telecommunications Union (ITU) radio regulations. For communications purposes, the usable frequency spectrum now extends from about 3Hz to about 300GHz. There are also some experiments at about 100THz where research on laser communications is taking place but we won't discuss this now. This range from 3Hz to 300GHz has been split into regions.
A formula for calculating these is as follows:

$$H_m(k) = \begin{cases} 0 & k < f(m-1) \\ \dfrac{k - f(m-1)}{f(m) - f(m-1)} & f(m-1) \le k \le f(m) \\ \dfrac{f(m+1) - k}{f(m+1) - f(m)} & f(m) \le k \le f(m+1) \\ 0 & k > f(m+1) \end{cases} \quad (4)$$

where M is the number of filters we want, and $f()$ is the list of M+2 Mel-spaced frequencies.

III.DISCUSSION

Almost all nationalities were having their uniqueness geographically and culturally. It is observed that in case of emotion, vibration the range of very high frequency is 30Mhz to 300MHz with the wave length 1m to 10m which get captured in aggressive stress and the range of very low frequency is 3KHz to 30KHz with the wavelength 100km to 10Km which are captured in the depressive stress. During the testing it has been seen that for the Mongolian dataset high frequency 1.0 classified as the aggressive stress in early stage on the other hand same frequency 1.0 indicate the normal speech in case if Indian dataset. From this observation, we can easily understand that the extent of F0 -excursions in speech increases slightly with age and also geography and cultural affects the FO. While comparing the approach with dynamic programming, we obtain the better satisfactory results. Studies have found that stress affects the performance of the student at an average percentage 10% as compared to normal performance. This system can extend to the medical application in case of physiological testing.
Emotions involve subjective experiences that vary between individuals. Different students can experience different emotions, even in the same situation, there performance variation obtained as shown in figure 7.1. For example, one student may be excited when doing today’s homework assignment in mathematics, whereas another student feels frustrated. These individual differences can relate to culture, ethnicity, gender, school membership, and class membership. For example, research has shown that average test anxiety is relatively high in students from some East Asian and Arab countries, as compared with students from Western countries. It has also been shown that average test anxiety is higher in female than in male students.
However, the differences in emotions experienced by different students within one culture are larger than the differences between cultures. Similarly, the differences among female students, and the differences among male students, are larger than the differences between the two genders. The same is true for ethnicity, school membership and class membership. Most of the differences between students are due to the uniqueness of students’ individual emotions and cannot be explained by group membership Students can also differ in how they react emotionally to different school subjects. For example, one student may enjoy mathematics but be bored by language instruction, whereas another student may be the opposite–bored by mathematics but enjoy languages. The emotions experienced in similar subjects (such as mathematics and science) are often similar, but the emotions experienced in dissimilar subjects (such as mathematics versus languages) can be quite different. The differences between emotions in different school subjects become larger as students progress in education and are most evident in high-school students. The reason for these differences is that students’ self confidence and interests often vary across different subjects. Therefore, emotions that are influenced by self-confidence and interest, such as

enjoyment of learning or anxiety, can vary as well. Finally, emotions can change over time. Emotional stability over time also differs between students. For example, some students tend to always enjoy mathematics instruction, whereas others are more variable in their emotional reactions.

The research findings imply that positive emotions can have profoundly positive effects on students' learning. However, this need not be true for all positive emotions. Specifically, positive task-related emotions, such as enjoyment of learning, focus students' attention on learning, promote their motivation to learn, and facilitate use of deep learning strategies and self-regulation of learning. Overall, you can expect these emotions to have positive effects on students' achievement. By contrast, positive emotions that do not relate to learning can draw attention away and lower performance, such as a student falling in love reducing his/her academic effort. Similarly, deactivating positive emotions, such as relief and relaxation, do not necessarily have positive effects.

The research evidence implies that negative emotions can strongly obstruct students' learning. Test anxiety, achievement-related hopelessness or boredom during lessons can lead students to with draw attention, avoid effort, procrastinate in doing assignments, fail exams, and drop out of school. Negative emotions are a major factor explaining why many students do not live up to their potential and fail to pursue the educational career that would correspond to their abilities and interests. Moreover, these emotions also jeopardize students' personal development and health, and contribute to the high numbers of suicides among youth in many countries–both unsuccessful and successful.

IV.RESULTS AND CONCLUSION

The available evidence for effects of emotions other than anxiety is limited. For positive emotions such as enjoyment of learning, positive correlations with academic achievement have been reported. For anger and shame, the findings suggest that overall relationships with students' achievement are negative. As with anxiety, however, the effects need not uniformly be negative.

Fundamental frequency was calculated using cepstral deconvolution method. When the log of the Fourier transform of the voiced signal is taken a distinguishable peak is obtained corresponding to the pitch period. A peak picking algorithm gives the fundamental frequency. The average and standard deviation of fundamental frequency were evaluated.

Classification: For the classification purpose SVM classifier is used. To classify the stress from the detected negative energy speech signal. SVM has general optimistic performance and hence it has led in the race of classifiers.

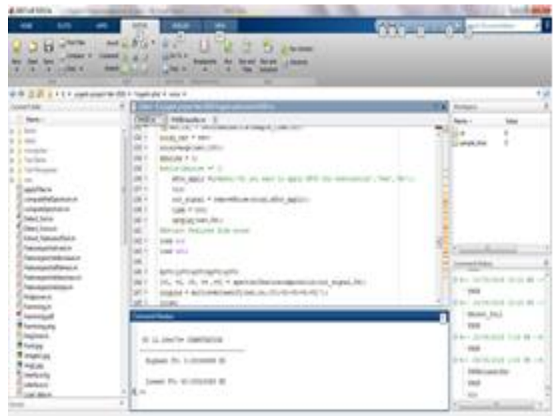


Figure 1Mongolian DB Natural testing

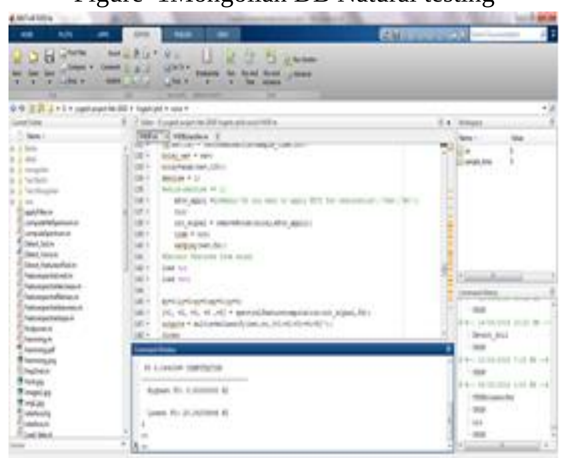
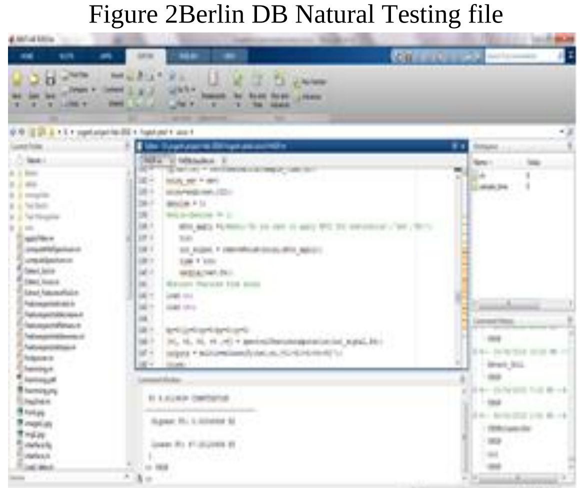


Figure 2Berlin DB Natural Testing file



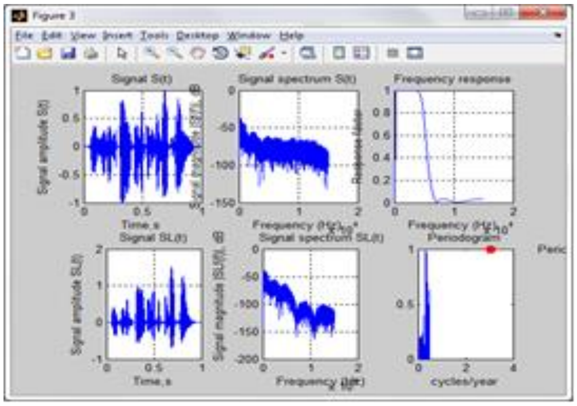


Figure b3 Indian DB Natural Testing

Spectrogram of Mongolian DB Testing for Mood Natural - Natural Spectrogram of Indian DB Testing for Mood- Natural Mood- Natural

V.FUTURE WORK

Implementation experience a great variety of emotions that can have profound effects on their learning, personality development and health. The effects of these emotions can be complex. Positive Emotions do not always benefit learning, and unpleasant emotions do not always impede learning. However, for the vast majority of students and academic learning tasks, enjoyment of learning is beneficial, whereas anxiety, shame, hopelessness and boredom are detrimental. Moreover, emotions are core elements of students' identity and well-being, implying that emotions are also important in and of themselves, beyond their functions for academic learning. For all these reasons, educators should attend to students' emotions.

REFERNCES

[1] K.R. Scherer. The effects of stress on the fundamental frequency of the voice. Journal of the Acoustical Society of America, 1977.

[2] L.J.M. Rothkrantz ,et.al " Voice Stress Analysis" Data and Knowledge Systems Group Delft University of Technology Mekelweg 4, 2628 CD Delft, The Netherlands.

[3] Analysis of Stress in Speech Using Empirical Mode Decomposition Nyaga Mbitiru, Peter Tay, James Z. Zhang, Robert D. Adams Department of Engineering and Technology Western Carolina University

[4] J.R. Stroop. Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 28:643–662, 1935.

[5] Milan Boštík "VOICE STRESS RECOGNITION METHODS" Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Radio Electronics, Purkyňova 118, 61200 Brno, Czech Republic.