INFANT SOUND BEHAVIOUR ANALYSIS

[1]Dr. S.M. Ali, [2]Samiksha Nagpure, [3]Ruqaiya Anjum, [4]Mubasshira Ali,

[1]Professor, [2]Student, [3]Student, [4]Student,

[1]Department of Electronics And Telecommunication,

Anjuman College Of Engineering And Technology, Nagpur, India.

Abstract : This paper present an infant emotion recognition system using audio information. Crying or making sound is not just an infant behaviour but a part of human system that assures survival.There are various reasons related to an infant crying such as hunger, anger ,happy, sad,cooing etc .There are many research paper have been written on analysis of infant by using various methods i.e, melody shape method, inverse filtering etc. ,In this paper we used feature extraction technique that includes power spectrum density, mean-average values, then covert it into mel-frequency which is quite simpler then other extraction methods. Back propagation neural network is used as classifier . The set of data is made from processed infant crying data ,there are two sets of data i.e, training data and testing data.GUI(Graphical user interface) is used to represent the information and make it easy to understand. In this research we study five reasons due to which the infant is crying .hopefully, in next research more reasons of crying can be added which will impact on the accuracy .

***Index Terms* – Feature extraction, Mel frequency, GUI.**

Introduction

Infant Cry is wholly associated with the respiratory system and the nervous system. It forms the most only means of communication immediately after baby's birth. In respiratory system vocal cord and vocal tract are in response to produce the cry sound, which has the fundamental frequency range of 250 – 600 Hz .We always wants to see a cute smiling baby rather than a Crying baby. And most of all have no reasons for which the baby cry's. At times mother can only understand their child's behavior. No two signals of an infant cry is found to be similar. The infant cry signal is always different for different emotional reasons like hungry, cooing, sad, happy, angry, etc.

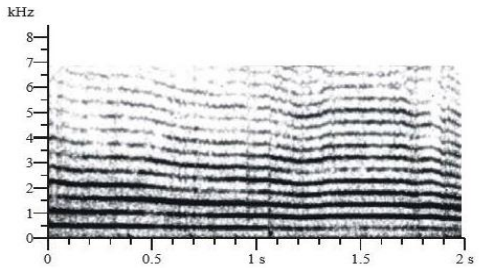


Figure 1 : A (digitally produced) spectrogram of an infant sound.

Infant crying is characterized by its periodic nature, i.e. alternating cry utterances and inspirations. By using a rapid flow of air through the larynx burst sound is produced ,because of that there is repeated opening and closing of the vocal folds, which in turn generates periodic excitation. This excitation is transferred through the vocal tract to produce the cry sound, which normally has a fundamental frequency (pitch) of 250–600 Hz. The acoustic signal of an infant's cry contains valuable information about their physical and physiological condition, such as health, weigh, identity, gender and emotions. Here for cry detection, zero crossingrate (ZCR) and fundamental frequency [1], Fast Fourier transforms coefficients, were determined and analyzed to detect the crying signals. In clinical settings one can assume noise-free conditions, and the research depends on finding subtle difference between cries that may be used for diagnostic purposes. In contrast to the clinical settings the detection problem does not assume noise-free conditions. In other words, the focus is on robustness in detecting crying signals in noisy and unpredictable environments.In this paper, we explain an analysis of infants' cry and present an algorithm for cry detection, which is aimed to alert parents in potential physical danger situations. The proposed algorithm is based on two main stages. The first stage involves feature extraction , in which pitch related parameters, are extracted from the signal. In the second stage, the signal is classified using Neural Network classifier and later verified as a cry signal

I. RESEARCH METHODOLOGY

The infant cry signals are stored in test folder which is are collected from the site (http://www.soundJay.com) and provided as an input voice to the simulating software MATLAB where it is compared with the database. Then the further analysis is computed and the results are classified according to the extracted feature by using their specified method.

In this present work, database folder contains 11 types of crying signal in which each signal has its own signal characteristics and defines the mood of the infant like hungry, sad, cooing, laughter, etc. In order to test the proposed algorithms a suitable database has to be collected or acquired. Here we have collected all the crying signals randomly from different databases available on internet. In this paper we have analyzed 5 different types of infant cry which states the following reasons,

1. **Hungry:** If the infant cries because of hunger, then the caretaker have to feed them to stop crying.

2. **Angry:** If the infant cries due to anger , then the caretaker have to deal with anger.

3. **happy:** If the infant is happy no need to worry.

4. **Sad:** If the infant cries out of sadness, then the caretaker should demand more attention towards their infant and has to make them feel safer.

5. Cooing: If the infant is cooing than there is no reason to worry

2.1 Data Collection

The data used for this study is the infant cries that are grouped into five cries of hungry baby, sleepy baby, like burping baby, in pain baby (in the stomach), and the uncomfortable baby. The data is taken from Dunstan Baby Language videos that has been processed.

The data is divided into two, training data and testing data.

II. FEATURE EXTRACTION METHODS

Of various methods, in this paper we have processed the infant cry signals using 3 different methods of feature extraction.

Pitch Information: As the infants cry has some normal frequency range, the fundamental frequency (f0) is considered as an classifier. Pitch and frequency

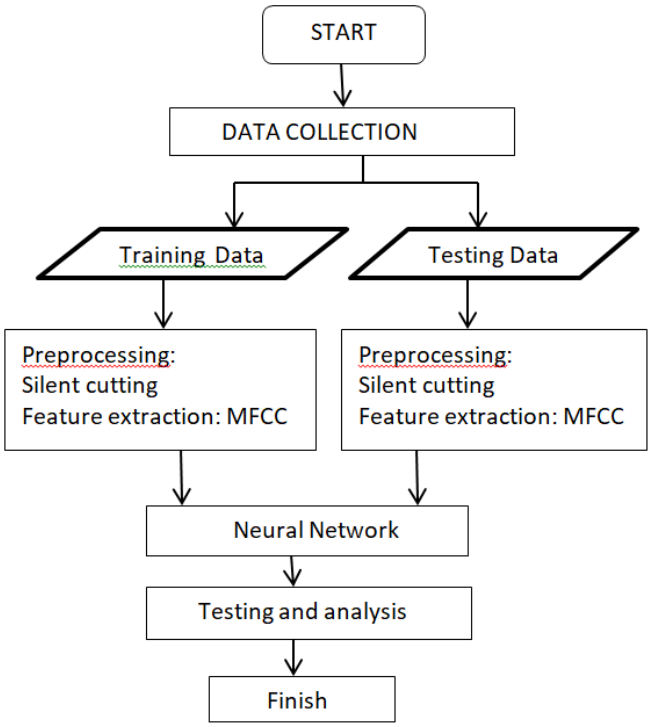


Figure 2: Flow chart for infant cry analysis.

(i) Pitch Information:

Pitch is an important attribute of voiced speech. It contains most specific information about the speaker. Thus estimation of fundamental frequency (f0) or pitch frequency is an important part in infant cry analysis. Initially, input voice push button is pressed and the test cry signal is loaded. Then pitch of the signal is detected if the fundamental frequency ranks between 100 to 400 Hz and with the cry duration between 1 to 1.5s. It is completely a time domain analysis.

(ii) Mel Frequency Cepstral Coefficient and Cross-correlation Method:

Pitch detection using time domain analysis has some limitations like that there may occur larger pitch signals which is not in the pitch period (T0). Thus we have implemented the same pitch detection algorithm in frequency domain by using the simplest cepstral analysis.

Pitch feature by using Mel-frequency cepstral coefficient method can function only in combination with Cross-correlation method.

What is the Mel scale?

The Mel scale relates perceived frequency, or pitch, of a pure tone to its actual measured frequency. Humans are much better at discerning small changes in pitch at low frequencies than they are at high frequencies. Incorporating this scale makes our features match more closely what humans hear.

The formula for converting from frequency to Mel scale is:

$$M(f) = 1125 \ln(1 + f/700)$$

To go from Mel back to frequency:

$$M^{-1}(m) = 700(\exp(m/1125) - 1)$$

(iii) NEURAL NETWORK- A neural network (NN), in the case of artificial neurons called artificial neural network (ANN) or simulated neural network (SNN), is an interconnected group of natural or artificial neurons that uses a mathematical or computational model for information processing based on a connectionistic approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network.

In more practical terms neural networks are non-linear statistical data modeling or decision making tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data.

A neural network (NN), in the case of artificial neurons called artificial neural network (ANN) or simulated neural network (SNN), is an interconnected group of natural or artificial neurons that uses a mathematical or computational model for information processing based on a connectionistic approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network.

In more practical terms neural networks are non-linear statistical data modeling or decision making tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data.

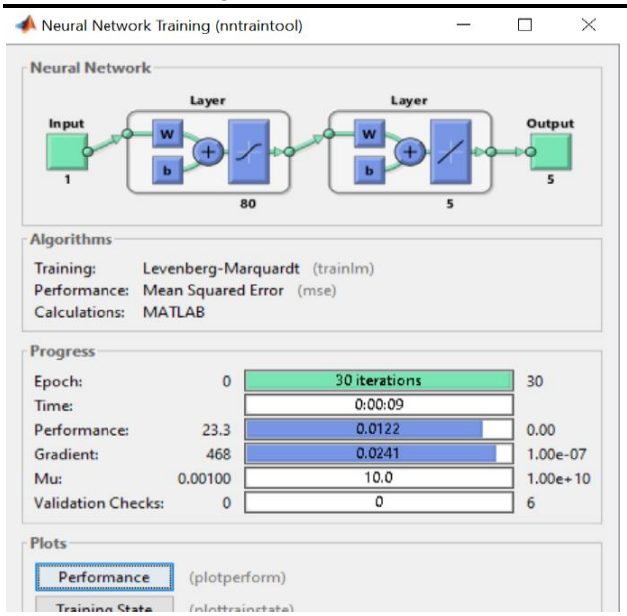


Figure 3- Neural network

IV. RESULTS AND DISCUSSION

Crying signals are stored in the test folder. Then one by one push button on GUI is pressed and the process starts. The comparison takes place by using cross correlation
algorithm and resampling those signals which are of different lengths the input signal will be loaded for further processing.

MATLAB 10.1 is used for coding. Codes are developed for representing cry signal in time & frequency domain and estimating fundamental frequency in time & Frequency domain. Also silenced region of cry signal has been removed for reducing the processing time. The results are as follows.

Outputs of testing samples-

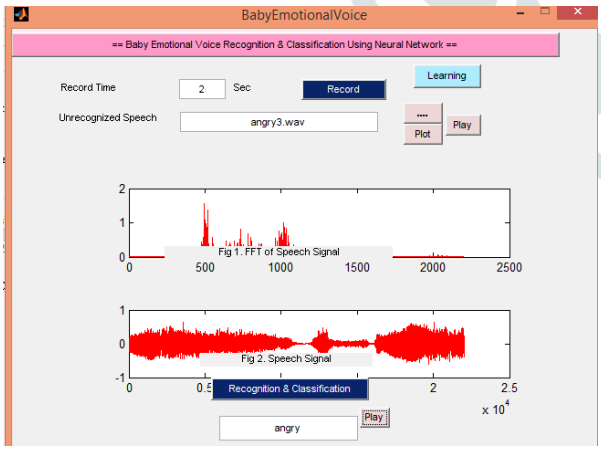
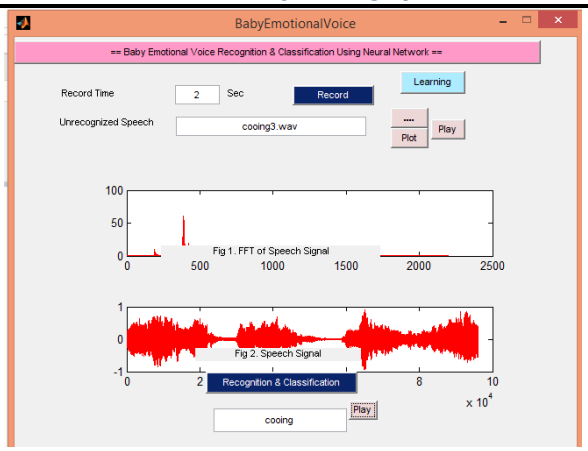
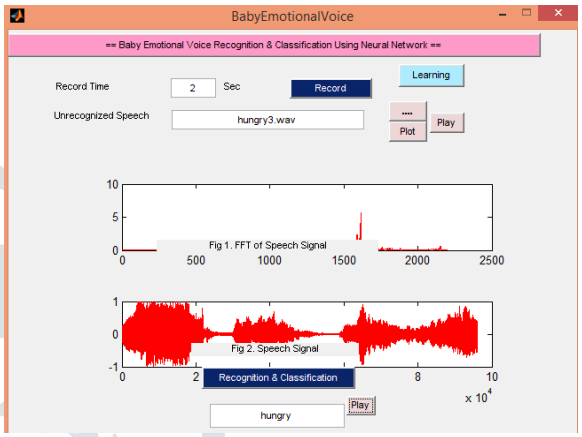


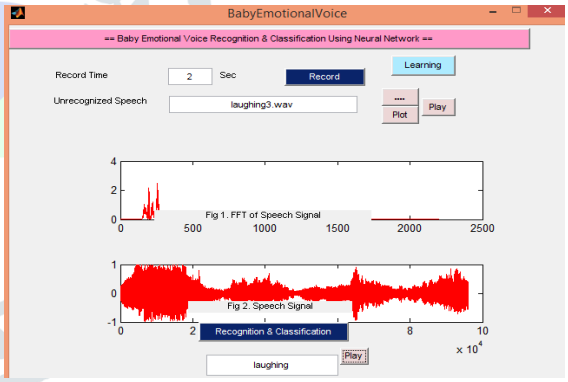
Fig.(4.a) Angry



Fig(4.b) Cooing



Fig(4.c) Hungry



Fig(4.d) Laughter

CONCLUSION:

The infant cry cannot only be treated as a mean of communication, or a sign of mood as it is very important acoustic signal, so it can be used to diagnose several diseases and accordingly the neonates can be treated in an effective manner. In this paper, we developed GUI and various databases and compared the crying signals, which are to be tested with the signals and the reason for crying is identified. This is done through cross-correlation algorithm. Additionally the harmonics, pitch, frequency range of the cry signals are classified in two different techniques that includes Mel-frequency cepstrum and linear prediction coding methods. The efficiency of the two techniques has also been differentiated in the statistical data using excel. Another advantage of the algorithm is its simplicity, it is based on a small number of features, which are relatively simple to implement.

References

［1］ Swarna laxmi Nagrajan: Infant Cry analysis ,Department of Biomedical Engineering, Coimatore-22
2017.
［2］ Medhanita Dewi Renti: Infant Cries Identification By Using Codebook As Feature Matching.

Department Of Computer Science, Bogor
Bogor Agriculture University.
［3］ Roopa S M, Smt. Anitha G, Dr. Mohamed Rafi.: Analysis of Infant
Cry through Weighted Linear Prediction Cepstral Coefficent and
Probabilistic Neural Network. Comp. Sc. Eng. IJCSMC, Vol. 4, Issue.4, Pg. 413−420, April 2015

[4] Patel K, Prasad RK. Speech Recognition and Verification Using MFCC & VQ. *International Journal of Emerging Science and Engineering* (IJESE) Vol. 1, No. 7, 2013. pp. 33-37.
[5] Dror Lederman ,"Estimation of Infants" Cry Fundamental Frequency Using a Modified SIFT algorithm", 2010 [4] Yasmina Kheddache, Chakib Tadj ,"Characterization of Pathologic Cries of Newborns Based on Fundamental Frequency Estimation" , October 2013
[6] Raina P. Daga and Anagha M. Panditrao, "Acoustical Analysis of Pain Cries" in Neonates: Fundamental Frequency", Special Issue of International Journal of Computer Applications on Electronics, Information and Communication Engineering - ICEICE No.3, Dec 2011
[7] Rami Cohen1, Yizhar Lavner, IEEE 27-th Convention of Electrical 2002