ISSN No: - 2456- 2165

Gender Recognition Using Voice Signal Analysis

Mohd Aqib Dar
M tech
Uttaranchal University, Chandanwari
Pream Nagar, Dehradun, India

Devender Singh
HOD, ECE Department
Uttaranchal University, Chandanwari
Pream Nagar, ,Dehradun ,India

Abstract: This paper may be a few reasonable examination on speaking signs to plot a sex classifier. Sex cataloguing by speaking investigation primarily aims to predict the sex of the talker by investigating altogether totally unlike constraints of the speech trial. This relative examination within the main focuses on short time analysis of the speech signals. The analysis includes comparison of short-time average magnitude and short-time energy standards of masculine and feminine voice trials. This quantifiable assessment is completed through LAB-VIEW program design. A data entailing of vocal sound trials composed from many students, every masculine and feminine, of our institute was made. The short-time exploration was made on all the poised vocal sound trials and conjointly the constraints were equated to determine a operational principle for the sexual characteristics from speaking.

***Keywords**: Speech investigation, Speech acknowledgement, Speech handling, Gender detection, Detection procedures*.

I. INTRODUCTION

The intention of this paper is to recognize the gender of a speaker supported the voice of the speaker victimization certain voice process methods in real time victimization science lab read. Gender-based variations in human voice square measure unit because due to the physiological variations like fold thickness or vocal tract length and partially owing to variations in voice vogue. In view of these alteration square measures mirrored within the speech signal, we anticipate to use these properties to mechanically categorize a spokesperson as male or feminine.

A. Proposed Solution

In acquiring the gender of someone we've got used audio findings from the audio supply, to seek out the basic frequency (F0) or pitch severally. It is acquainted that F0 values for male voices area unit lower as a result of elongated and bushier vocal folding. F0 for fully developed men is often about a 120 Hz, whereas F0 for full-grown females is around 200 Hz. The band of voice generated by humans is delimited between fifty Hz to 3400 Hz and also the frequencies from twenty Hz to twenty kHz of the loud vary, .Additional fully developed men reveal lesser formant frequencies than fully developed women as a result of vocal tract length variations. The techniques won't to method voice signals which will be roughly speaking categorized as one or the other time-field or frequency-field examination. In timed-field examination, the computations area unit executed straight on the speaking sign to mine data. Whereas in frequency-field examination, the knowledge is taken out once the frequency at ease of the speaking sign calculated to make the band.

II. LITERATURE SURVEY

The aim of speech analysis techniques is to investigate the speech signal and estimate the parameters necessary for speech process applications. To derive these parameters, frequency domain illustration of speech signals is required. Numerous speech process techniques area unit then utilized. The challenge here is to search out the foremost economical methodology of speech process. There are, 3 most ordinarily used process ways are: Fourier analysis, cepstral analysis and linear prediction analysis

A. Fourier analysis

Fourier analysis is that the ancient technique of computing the phase and amplitude spectra of speech signal. It employs quick Fourier rework (FFT) rule and window functions to compute short-time Fourier values 3,4.we all know that the speech signal is comprised of excitation-source and linear signal spectra. Fourier analysis cannot estimate the 2spectral parts individually .so as to beat this limitation, speech analysis technique ought to perform the short time analysis of the speech signal in such some way that the two parts of speech are individually on the market for further process. Two such techniques are cepstral analysis and linear prediction analysis

B. Cepstral Analysis

Many applications demand the separation of excitation source and signal spectra. The cepstral analysis technique provides a good method of separating the 2 parts by reworking the merchandise of the 2 within the frequency domain into their total and additionally creating use of the very fact that the two signals have totally different spectral characteristics. This technique, thus, gives a method for transposition spectral envelope and pitch quality.

C. Linear Prediction Analysis

Like the cepstral analysis technique, linear prediction analysis method put together provides some way to separate the excitation-source and thus the signal spectra of the speech model. To comprehend this, it assumes associate all-pole model for the linear system. Excitation to the present model is either one impulse or a white random noise sequence.

III. MATERIALS AND METHODS

The following figure1 shows the idea of a gender recognition system. Notwithstanding the sort of the task (classification or verification), gender recognition system works in 2 modes: training and recognition modes.

Within the coaching mode, a replacement gender person's voice is recorded and analyzed for 2 seconds knowledge. The popularity mode, AN unknown gender person offers a speech input and therefore the system makes a choice regarding the speaker's identity. Each the coaching and conjointly the popularity modes embrace feature extraction, usually referred to as the front-end of the system. The feature extractor-converts the digital speech signal into a sequence of numerical descriptors, named as feature vectors. The alternatives offer much stable, robust, and compact illustration than the raw signal. Feature extraction is also thought-about as an info reduction methodology that tries to capture the essential characteristics of the speaker with a bit rate.

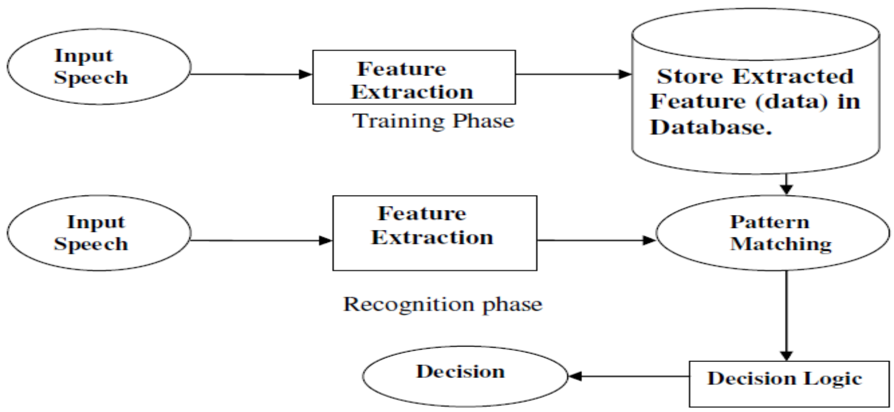

Figure 1. Block diagram of the Gender Recognition System

Feature extraction is that the strategy of fixing the initial speech signal to a continuing quantity illustration that provides a collection of purposeful options helpful for recognition. Feature extractions is that the combination of some signal process steps together with the computation of drive knowledge from wave sound, computation of fast Fourier transform (FFT), Power spectrum, the sample point at most power and eventually calculate the frequency[7].

A. Data Acquisition system

The data faller of the planned system is created up with the assistance of constitutional electro-acoustic transducer of the laptop having the data of 8 bit and sampling frequency of 11025 Hz. This is interfaced with the LABVIEW and recorded for the duration of 2 seconds as shown in figure2. The acquired data is further preprocessed to remove the unwanted signal using low pass digital filter of 3.5KHz [8],[9].

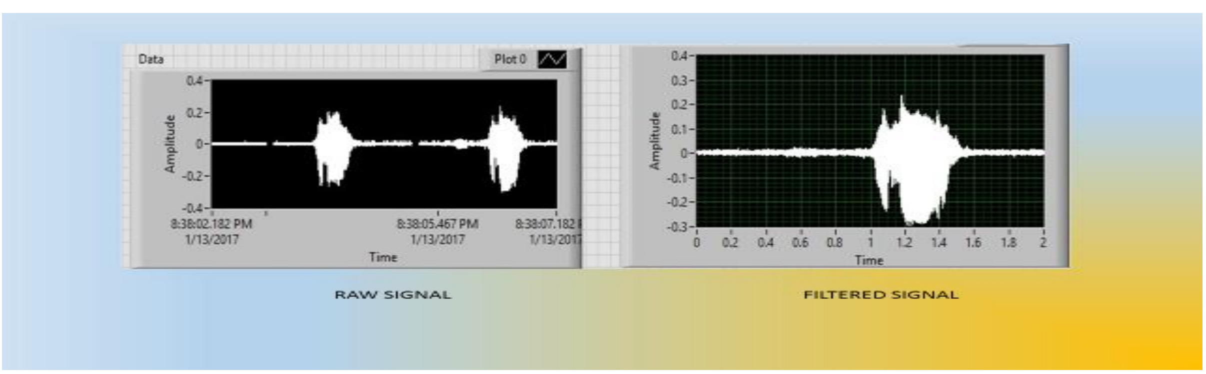

Figure 2. showing the raw signal and filtered signal of the word "hello"

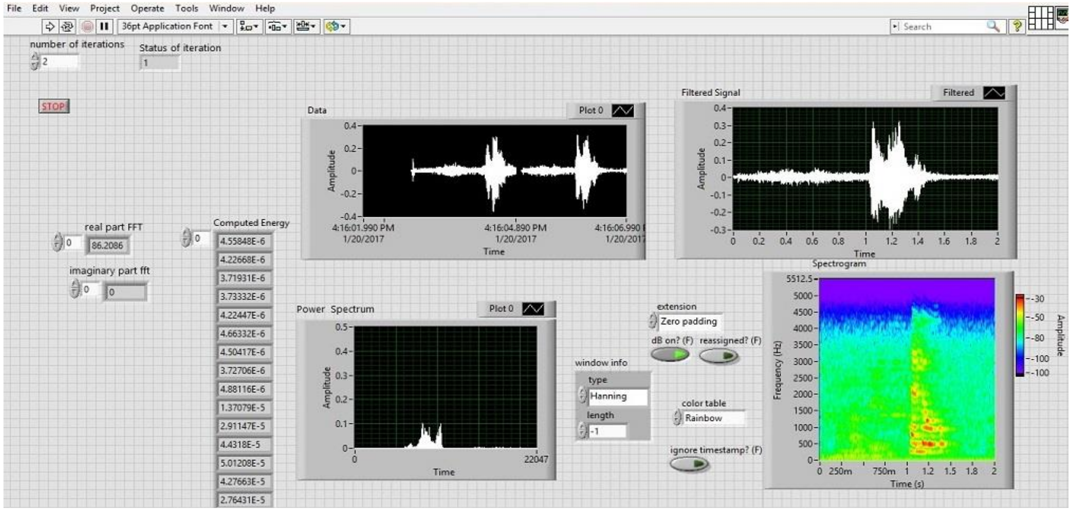

Figure 3. Front Panel of LABVIEW

Further it is processed to compute the power and energy spectrum of the signal of the voice as shown in the figure3.

IV. RESULTS AND DISCUSSION

Threshold price detection of the voice energy is completed by the exploitation STFT. (Short time Fourier transform) technique of 10 subjects (5 male and 5 female) of reciting the word hello as shown in figure4.

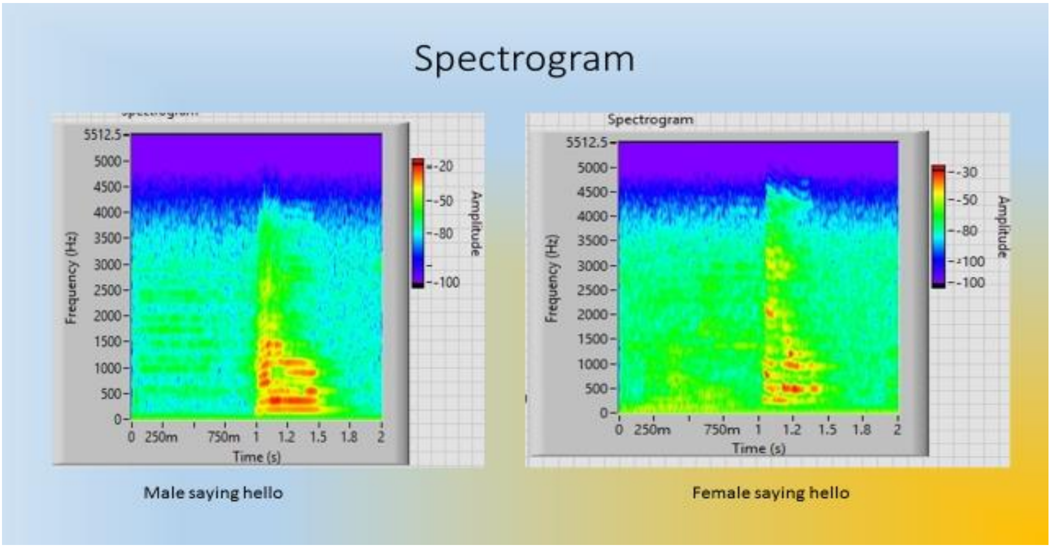

Figure 4. showing spectrogram of the word “hello” by male and female

Short-Time Energy (STE) is that the energy related to the signal in time domain.

Short-Time Energy of a speech signal is given by:

$$E_n = \sum_{m=-\infty}^{\infty} [x(m)w(n-m)]^2$$

Where,

E_n = Short-Time Energy

x = Speech Signal

w =Window

Based on the STE technique the vary of voice energy of various subjects has been determined .As shown within the

ISSN No: - 2456- 2165

figure5, the energy spectrum of various male and feminine subjects has been depicted. the info has been collected for 10 completely different person for two seconds. With reference of figure5, it may be inferred that the energy spectrum worth varies with person to person thus taking the typical value of all energy it's been ended that the typical energy for male and feminine ranges from a pair of.6227E^05volts to zero.000387744 volts as shown in figure6.

Based on these values the person Gender may be categorized as

if(x>=2.62E-05&&x<4E-04){

a=1;

b=0;

}

else if(x>=4E-04)

{a=0;b=1;}

else if(x<2.62E-05)

{a=0;b=0;} where x=energy values and a and b are the digital outputs to indicate the gender of the voice These threshold values area unit fed in to micro controller through Arduino cryptography and therefore the persons stress condition is displayed on the device as shown. in the figure7.

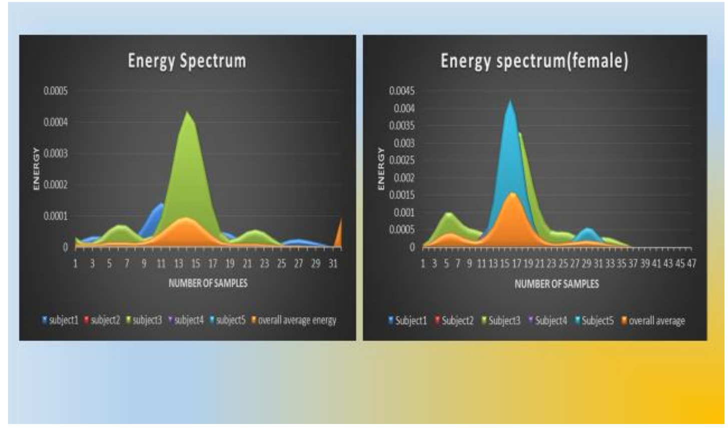

Figure 5. Energy Spectrum of different Male/female subjects

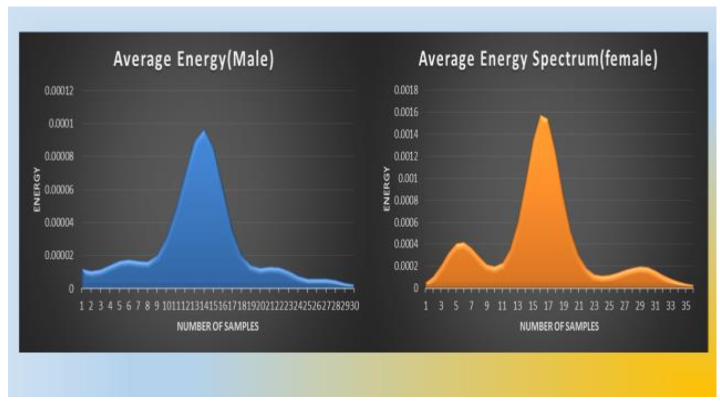

Figure 6. Average Energy Spectrum of Male/female subjects

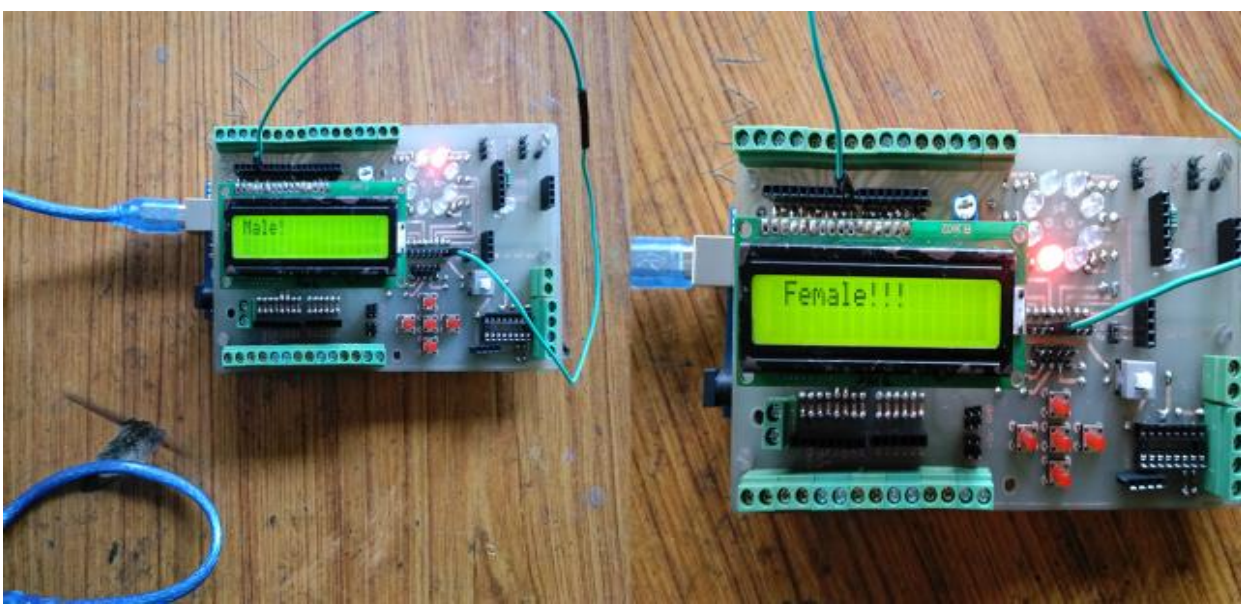

Figure 7. Hardware showing Gender recognition

V. CONCLUSION

Based on the study and observation of the STE (Short-Time Energy) technique it's all over that the common energy price is two.62E-05 for male and four.0E-04 for females. Females have higher energy values in their voice than the common men. These values are fed in to the microcontroller and programmed to research the gender condition. The more analysis and a lot of complicated algorithmic rule like ANN or formal logic to discover the pitch and frequency will increase the dependability of the system.

REFERENCES

1. Eric Keller, “Fundamentals Of Speech Synthesis And Speech Recognition”.
2. Lawrence Rabiner, “Fundamentals of Speech Recognition”.
3. John G. Proakis, Dimitris G. Manolakis, D.Sharma, “Digital signal Processing, Principles,Algorithms and Applications”
4. H. A. Murthy, "Algorithms for Processing Fourier Transform Phase of Signals," Speech and Vision Laboratory, Indian Institute of Technology, Madras, 1991.
5. M. S. S. A. M. M. a. A. A. A. Eslam Mansour Mohammed, "LPC and MFCC Performance Evaluation with Artificial Neural Network for Spoken Langu Identification," International Journal of Signal Processing, Im Processing and Pattern Recognition, vol. 6, no. 3, 2013.
6. R. P. Bishnu Prasad Das, "Recognition of Isolated Words using Features based on LPC, MFCC, ZCR and STE, with Neural Network Classifiers," International Journal of Modern Engineering Research (IJMER), vol. 2, no. 3, 2012.
7. John R. Deller, John G Proakis and John H. L. Hansen, “Discrete- Time Processing of SpeechSignals”.
8. H. Erokyar, "Age and Gender Recognition for Speech Applications based on Support Vector Machines," Scholar Commons, Tampa, 2014.
9. S. A. Hadei, "A Family of Adaptive Filter Algorithms in Noise Cancellation for Speech Enhancement," International Journal of Computer and Electrical Engineering, vol. 2, no. 2, pp. 67-78, 2010.
10. J.S Chitode, Anuradha S. Nigade" Throat Microphone Signals for Isolated Word Recognition Using LPC " International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 8, August 2012. ISSN: 2277 128X.