ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

Comparative Analysis of Pitch and Formant Frequency

Prof. Vinod.H.Patil[1], Aarushi Kaushal[2], Tushar Gupta[3], Vikas Lalotra[4]

Assistant Professor, Dept. of E&TC, BVUCOE, Pune, Maharashtra, India. [1]

Student, Dept. of E&TC, BVUCOE, Pune, Maharashtra, India[2]

Student, Dept. of E&TC, BVUCOE, Pune, Maharashtra, India[3]

Student, Dept. of E&TC, BVUCOE, Pune, Maharashtra, India[4]

ABSTRACT: The study of speech signal and the dealing with methods of these signals is called speech processing. The speech signal is basically created at the vocal cords, travels through the vocal tract, and produced at speaker's mouth. The signals are usually processed in two forms, one is pitch frequency and another is formant frequency. Emotion of human being can be identified by analysing pitch frequency and formant frequency. Commonly, the quality that makes the possibility of judging sounds as "higher" and "lower" in the sense associated with musical melodies is called as pitch where as the concentration of acoustic energy around a particular frequency in a speech wave is called formant.

KEYWORDS: Pitch; Formant; Acoustic energy.

I. INTRODUCTION

The main objective of the project is to analyse the pitch frequency and formant frequency of different age groups and gender. The pitch frequency and formant frequency can be used to describe the emotions of human beings. By creating a database of speech signals of males and females of different age groups one can easily analyse the variation in pitch and formant frequency. This project can be further used for Speaker Recognition System by using Feature Extraction (like Mel-frequency Cepstrum Coefficient or Linear Predictive code) and Feature Matching (like Dynamic time Warping or Vector Quantization).

The speech signal has been recorded by using MATLAB, for the number of seconds specified by length. The recorded audio signal is converted into "wav" file. The conversion is done because it can read only .wave file. After the conversion, the noise is been removed from the recorded wave signals. In this process of noise removal the recorded signal is analyzed frame by frame. Break the signal into the frames of 0.1sec and set the threshold value according to the recorded signal. Noise is the unwanted part of the signal which should be removed for more accurate results. The noise removed signal has been used for calculating the pitch frequency and formant frequency.

II. LITERATURE REVIEW

Generation of Human speech is complex and it consists of lungs, the vocal folds and vocal tracts. Production of speech begins with message formulation and then acoustic signal generates through vocal tract system. The vocal tract system of human is similar to the electronics parts power supply, oscillator and resonator. Lungs, ribcage and abdominal muscles are responsible for the source of speech. The combinations of these structures are responsible for generation of air-stream between the vocal folds which leads to expansion in chest cavity and contracts to force air from lungs and vibration of the vocal cords occurs which leads to generation of sound.

A pitch of a sound is basically refers to the sensation of a frequency. A high frequency sound wave corresponds to high pitch sound while a low frequency sound wave corresponds to a low pitch sound. A congregation of acoustic energy across a particular frequency in the speech wave is called formants. Each at different frequency, there are several formants. Formants rarely exist at 1000Hz intervals. In the vocal tract, each formants is analogous to a resonance in the speech signal.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

MATLAB (Matrix Laboratory) is a programming language developed by MathWorks which computes multi-pattern numerical including matrix manipulations. It is also used for plotting various types of functions and data. It has capability of interfacing with programs written in languages like C, C++, Java and Python. Praat is also an optional toolbox for analysis of speech signal. Pitch and formant frequency can also be calculated using praat.

III. CURRENT THEORY

Currently pitch and formant frequency are used for gender recognition. By referring the database (Table 1) we can analyse that usually Male have lower pitch frequency in comparison to female. Pitch frequency is also used in Voice Gender Conversion. In order to apply voice gender conversion initially the voice gender should be detected. If the recognized speech signal is female voice then the voice will be assigned using Voice Gender Conversion system and in case of male voice VGC will be neglected because the required spectral properties are already available.
Formant frequencies have rarely been used as auditory features for speech recognition system. During signal processing we can evaluate vowel formant frequency with the help of Linear Predictive Coding technique. Pitch frequency and formant frequency are also used for emotion recognition whether an individual is happy, sad etc. Every vowel show different formant frequency, so we always consider the average value of formant frequency.

IV. PROPOSED WORK

While considering the above, it will be apparent that, the pitch and formant frequency of different age group should be analysed as this helps in gender recognition, voice gender conversion. This project is basically analysis of pitch and formant frequency with the help of MATLAB code.
In this project we are supposed to create a database of different age group i.e. recording of audio signal. The time for recording can be adjusted by the user. The recorded audio signal consists of noise and that noise is removed by using frame by frame method. Then the filtered signal is converted into WAV file. Because of wave file conversion the sound wave will convert in the matrix form. Then the signal will be further processed for the pitch and formant estimation. The plots of the processed signals like the recorded signal, noise removed signals and formants will be there.

S. No.	GENDER	AGE GROUP	PITCH FREQUENCY	FORMANT FREQUENCY
11	Male 1	21-22 years	135.77 Hz	251.3 Hz
2	Male 2	21-22years	119.43 Hz	234.2 Hz
3	Male 3	21-22 years	124.27 Hz	269.8 Hz
4	Female 1	21-22 years	224.93 Hz	348.4 Hz
5	Female 2	21-22 years	376.50 Hz	377.8 Hz
6	Female 3	21-22 years	282.42 Hz	481.70 Hz
7	Female 4	21-22 years	255.89 Hz	350.8 Hz
8	Female 5	21-22 years	237.57 Hz	294.0 Hz
9	Female 6	30-40 years	173.65 Hz	332.4 Hz
10	Male 4	30-40 years	110.79 Hz	216.7 Hz
	Male 5	30-40 years	109.40 Hz	204.0 Hz

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

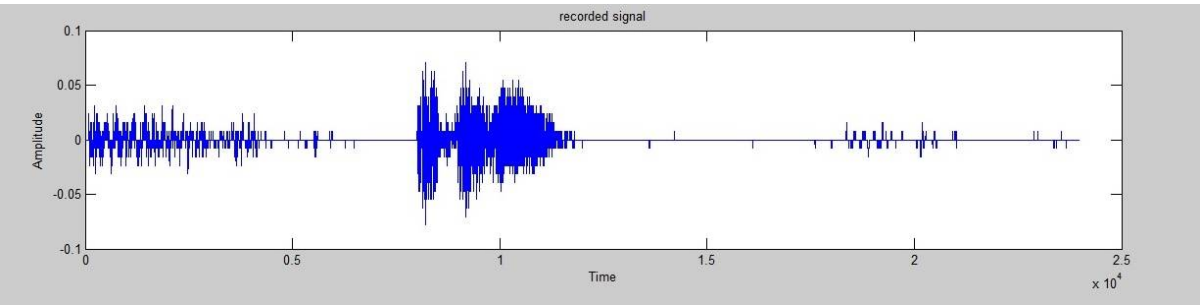
(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

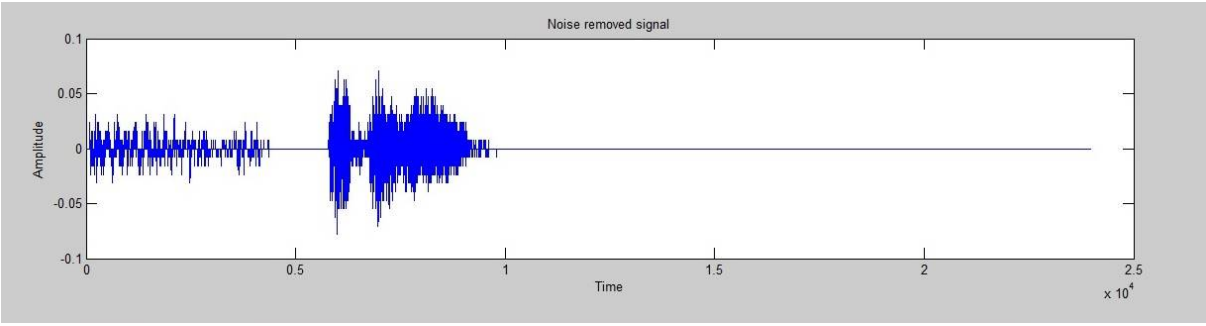
11				
1 12	Male 6	30-40 years	93.48 Hz	212.4 Hz
1 13	Male 7	30-40 years	93.79 Hz	216.7 Hz
1 14	Female 7	30-40 years	163.65 Hz	356.4 Hz
1 15	Male 8	50 years	90.57 Hz	190.7 Hz

TABLE 1

From the above database we can easily conclude that the male of age group of 22years have 130 Hz approx. pitch frequency and 245 Hz approx. formant frequency where as in case of female of age group 22years the pitch frequency is approx. 270 Hz and formant frequency is approx. 350 Hz. Further, the pitch frequency of male of age group 30-40years is 100 Hz and formant frequency is about 200 Hz while female of age group 30-40years have pitch frequency approx. 170 Hz and formant frequency is about 340 Hz.



RECORDED SIGNAL



NOISE REMOVED SIGNAL

V. RESULTS

The results shows that the male (gender) of age group of 22 years have 130 Hz approx. pitch frequency (avg.) and 245 Hz approx. formant frequency (avg.) where as in case of female (gender) of age group 22 years the pitch frequency (avg.) is approx. 270 Hz and formant frequency (avg.) is approx. 350 Hz. Further, the pitch frequency (avg.) of male of age group 30-40 years is 100 Hz and formant frequency (avg.) is about 200 Hz while female of age group 30-40 years have pitch frequency (avg.) approx. 170 Hz and formant frequency (avg.) is about 340 Hz.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

S. No.	GENDER	AGE GROUP	PITCH FREQUENCY (avg.)	FORMANT FREQUENCY (avg.)
1	Male	21-22 years	130 Hz	245 Hz
2	Female	21-22 years	270 Hz	350 Hz
3	Male	30-40 years	100 Hz	200 Hz
4	Female	30-40 years	170 Hz	340 Hz

VI. CONCLUSION

Hence we will be concluding that as the age of male and female increases the pitch and formant frequency decreases and the pitch and formant frequency of female is higher than that of male. This comparative data can be used in gender recognition, emotion detection and various other applications.

VII. FUTURE SCOPE

The comparative analysis of pitch and formant frequency will be very helpful in speech recognition in future, as the noise removed signal is further processed.

REFERENCES

1. *Juang, B. H., Rabiner, Lawrence R.,* "Automatic speech recognition–a brief history of the technology development" *(PDF): 6.* Retrieved 17 January 2015.
2. *Schwartz, David A., Dale Purves . "Pitch Is Determined by Naturally Occurring Periodic Sounds". Hearing Research* ***194****: 31–46, 2004.*
3. Titze, I.R., "Principles of Voice Production", Prentice Hall, ISBN 978-0-13-717893-3, 1994.
4. Ladefoged, Peter," Vowels and Consonants: An Introduction to the Sounds of Language", Maldern, MA: Blackwell, p. 40. ISBN 0-631-21412-7, 2001.
5. *Bakshi, K.A., A.V. Bakshi, U.A. Bakshi, "* Electronic Measurement Systems*", US: Technical Publications. pp. 4–14.* ISBN 978-81-8431-206-5, *2008.*
6. Roland," Introductory Signal Processing" World Scientific. p. 1. ISBN 9971509199, 1991.
7. T. H. Wilmshurst ," Signal Recovery from Noise in Electronic Instrumentation (2nd ed.)", CRC Press. pp. 11 ff. ISBN 0750300582, 1991.

BIOGRAPHY

Mr. Vinod. H. Patil received his BE degree in Electronics and telecommunication with distinction from Shivaji University, Kolhapur, India in the year 2008 and currently working as a Asst.Prof. In BVCOE, Pune (India).

Ms. Aarushi Kaushal has Pursuing B.Tech .in Electronics & Telecommunication in 2015 from the BVCOE Pune(India).

Mr. Tushar Gupta has Pursuing B.Tech .in Electronics & Telecommunication in 2015 from the BVCOE Pune(India).

Mr. Vikas Lalotra has Pursuing B.Tech .in Electronics & Telecommunication in 2015 from the BVCOE Pune(India).