Asian Journal of Convergence in Technology Volume III, Issue II
ISSN No.:2350-1146, I.F-2.71

Gender Classification by Speech Analysis

Shruti Nanda[1], Anjali Sahoo[2], Bhagyalaxmi Jena[3]
shruti.nanda.13@gmail.com , anjalisahoo1995@gmail.com , bjena@silicon.ac.in

Department of Electronics and Communication Engineering
Silicon Institute of Technology
Bhubaneswar, India

***Abstract*--Gender classification is widely used in automatic speech recognition systems to recognize a speaker speaking continuously in any language. This work aims at analysing speech signals based on some parameters so as to predict the gender of the speaker. This paper comprises of male and female voice samples which were collected to form a database. Parameters such as mean, variance and standard deviation were determined to help in classifying the gender of the speaker.**

Keywords - mean, standard deviation, variance.

I. INTRODUCTION

A speech signal consists of a sequence of basic sound units called phonemes. The same set of phonemes may contain same information but changes at acoustic level due to various physiological differences in the vocal tract of the speaker. Vocal tract is also known as the larynx. The main difference in the length of the vocal tract is noticeable between males and females. Adult male has 16.9 cm and adult female has 14.1 cm. The length and shape of the vocal tract along with other parts like tongue, teeth and dialect of the speaker helps in gender classification by speech analysis. [1].

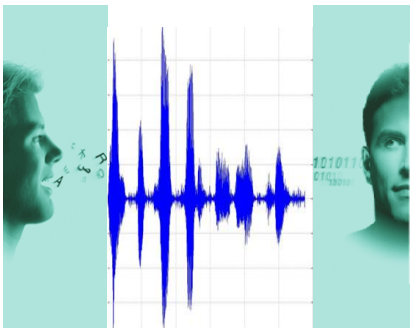

Fig.1. Generation, Transmission and Reception of Speech

II. SPEECH

Speech conveys an abundance of information with regards to speaker identity [5]. A speech or a normal speech can be defined as the expression of or the ability to express thoughts and feelings by articulate sounds.

A. Types of Speech Analysis

Speech can be analyzed in two different ways namely, time domain analysis and frequency domain analysis. Both the domains are interchangeable i.e. no information is lost during the change. It is achieved using various transforms like fourier transform. In time domain analysis "how long something takes to change" and in frequency domain "how fast or slow it is" can be measured respectively. In time domain analysis information can be extracted by directly measuring the signal whereas in frequency domain analysis information is extracted from the frequency content of the speech signal present in the spectrum.

B. Need of Gender Classifier

Gender classifier is the basic unit of any automatic speech recognition system. A gender classifier is embedded with programs which compare various speech parameters for determining the gender of the speaker. Speaker identification focuses on identifying the speaker rather than what the speaker speaks about. Some speech recognition applications are call routing, voice dialling, simple data entry and speech to text processing. These applications are implemented in fields like robotics, video games, voice password technology, high performance fighter aircrafts etc.

C .Pitch

The human voice has many components and is created through a myriad of muscle movement; pitch is an integral part of the human voice. The pitch of the voice is defined as the "rate of vibration of the vocal folds".

The sound of the voice changes according to rate of vibration. As the number of vibrations per second increases, pitch increases, which mean the voice sounds higher. Similarly, with the decrease in the number of vibrations per second, pitch decreases leading to deeper sound of voice.

The vibrations and the rate at which vocal folds vibrate depend on the length and thickness of the vocal cords along with the tightening and relaxation of the muscles surrounding them. Women tend to have higher voice than men because the former have shorter vocal cords than the latter. An adult male has pitch within a range of 85 to 180 Hz whereas an adult female has pitch between 165 to 255 Hz [2-3].

D. Database

This database contains the voice samples of students, both male and female, of our college using Wave-pad Sound Editor software. There were 05 male and 05 female students. The created database was named Gender Analysis. All the speakers were asked to utter the same sentence, "The weather is too hot today". Then the recorded utterances were processed and evaluated [4].

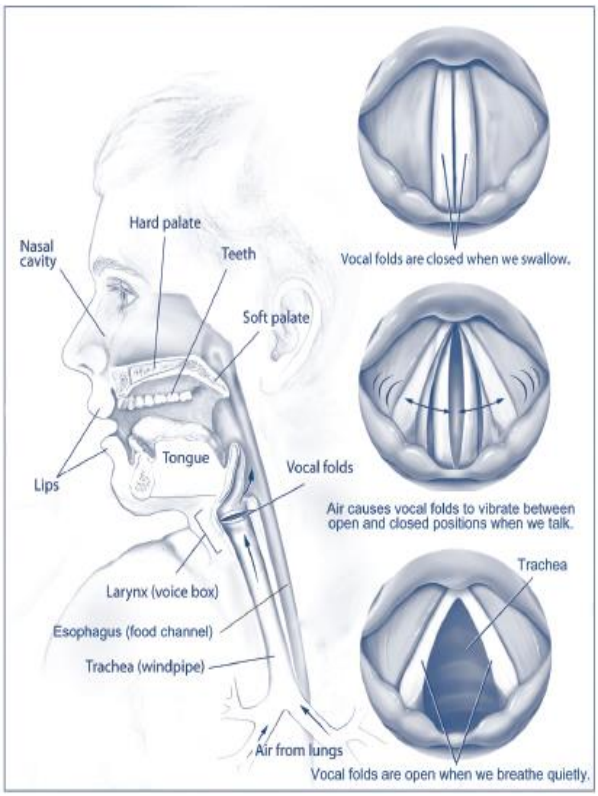

Fig. 2. Schematic Diagram of the Human Speech Production

III. MEAN

Mean is sum of the value in the signal x_i by letting the index i run from 0 to N-1.Then finish the calculation by dividing the sum by N. We get the average of the speech signal.

$$m = \frac{1}{N}\sum_{i=0}^{N-1} x_i \qquad (1)$$

m=Mean of the speech signal
N=Number of speech samples
x=Speech signal

IV. STANDARD DEVIATION

The average deviation of a signal is found by summing the deviation all individual samples, and then dividing by the number of samples, N. The average deviation provides a single number representing the typical distance that the samples are from the mean. The standard deviation is similar to the average deviation, except the averaging is done with power instead of amplitude [5-6].

$$\sigma^2 = \frac{1}{N-1}\sum_{i=0}^{N-1}(x_i - m)^2 \qquad (2)$$

σ^2=Variance
$\sqrt{\sigma^2}$=Standard Deviation
N=Number of speech samples
x= Speech signal
m=Mean of the speech signal

V. RESULTS

The result contains a table showing mean, variance and standard deviation of differing speech rate which helps in identifying the gender of the speaker [7-8].

A. Female

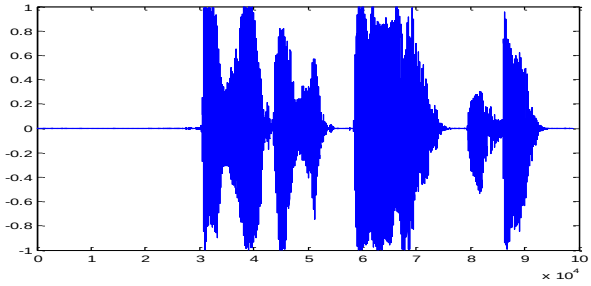

Fig. 3. Female

B. Male

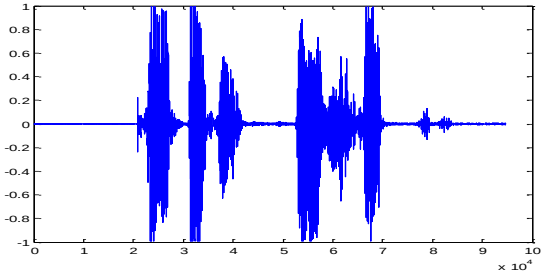

Fig. 4. Male

TABLE I

OBSERVATIONAL VALUES OF FEMALE VOICE SAMPLES

Speaker	Mean	Standard Deviation
F1	221.2667	13.06285
F2	212.6667	20.50668
F3	203.3333	13.76158

TABLE II

OBSERVATIONAL VALUES OF MALE VOICE SAMPLES

Speaker	Mean	Standard Deviation
M1	122.000	13.984
M2	103.4333	4.905
M3	121.0677	20.485

VI. OBSERVATIONS

From the calculations we observed that mean and standard deviation of female is found to be higher than that of male.

VII. CONCLUSION

At the end, when the results obtained from male and female voice samples of the corresponding parameters are compared, a competent difference in the values was observed. So these parameters can be used in gender classification by encoding them in a program and embedding in the gender classifier that can automatically predict the gender of the speaker.

REFERENCES

[1] M. Kepesi, and M. Sigmund, Automatic recognition of gender, *Proc. Radioelektronika'98, BUT, Brno, 1998, pp.200-203.*

[2] Milan Sigmund 2008 *Gender Distinction using Short Segments of Speech Signal*. Int. J. of Computer Science and Network Security, *Vol.8, No.10.*

[3] HESS, W. *Pitch Determination of Speech Signals*. New York: Springer Verlag , 1993.

[4] J. H. Hansen and S. E. Ghazale ,"Getting started with SUSAS,"*Proceedings of Eurospeech'97*. Rhodes, pp. 1743- 1746, 1997.

[5] M.Sigmund ,"Change in frequency spectrum of vowels due to psychological stress," In *Proceedings of the 20th International Conference Radioelecktronika.* Brno, (Czech Republic), 2010.

[6] Coleman, R. O. (1976). "A comparison of the contributions of two voice quality characteristics to the perception ofmaleness and femaleness in the voice". J. of Speech and Hearing Research 19, 168-180.

[7] Wu, K. and Childers, D.G. 1991 Gender recognition from speech. Part I: Coarse analysis. J.Acoust. Soc. of Am.,90(4):1828–1840.

[8] Rabiner Lawrence, Juang Bing-Hwang, *"Fundamentals of Speech Recognitions"*, Prentice Hall New Jersey, 1993, ISBN 0-13-015157-2.