ISSN - 2250-1991 | IF : 5.215 | IC Value : 77.65

Research Paper Computer Science

Development of Noise free Marathi Speech database using various Filtering Techniques

Agnihotri P.P.	Assistant Professor, School of Technology, S.R.T.M. University, Nanded Sub-Centre,Latur, India
Shinde A.R.	Assistant Professor, School of Computational Sciences, Solapur University, Solapur, India
Khanale P.B.	Professor, Department of Computer Science, Dnyanopasak College, Parbhani, India

ABSTRACT

This paper describes the development of noise free speech database in Marathi language. Speech data was collected from 30 speakers from Marathwada Region of Maharashtra with the help of Laptop and unidirectional condenser Microphone. As a result, the acquired speech data captured large amount of noise. The noise from speech data was removed through filtering techniques like Pre-emphasis, Butterworth low pass, Butterworth high pass, Butterworth pass band and Butterworth stop band filtering techniques to get the actual signal. Band stop filtering technique is an efficient technique to remove the noise from speech signal. Templates of each spoken corpus of Marathi Language were constructed which helps to develop speaker independent system.

KEYWORDS Marathi Language, speech database, sampling frequency, and filtering techniques

Introduction

The Indian constitution identifies 22 languages, of which six languages (Hindi, Telugu, Tamil, Bengali, Marathi and Gujarati) are spoken by at least 50 million people within the boundaries of the country. Marathi has the fourth largest number of native speakers in India. In the 2001 Indian census, 73 million (73,000,000) people reported Marathi to be their native language[1] .

The script used for writing Marathi is called Devanagari derived from the Brahmi script of Ashoka. Devanagari was originally developed to write Sanskrit but was later adapted to write many other languages, including Marathi and Hindi. Devanagari has 65 consonants, 18 full vowel letters, 17 vowels symbols and 2 symbols for nasal sounds. Marathi Language uses only 36 consonants Syllables, 13 vowels and 2 symbols for nasal sounds. The main feature of the script is left-to-right and words are spelled phonetically [2, 3, and 4].

This paper is presenting the steps requires to create a noise free speech database or speech corpus for Marathi language Marathi Speech Database

2.1 Speech sampling

Speech signals, i.e., signals intended to carry only human speech can usually be sampled at a much lower rate. For most phonemes, almost all of the energy is contained in the 5Hz-4 kHz range, allowing a sampling rate of 8 KHz for telephone system [6]. In Speech Recognition System, speech is recorded at 16 KHz sampling rate [5] to get accuracy in speech recognition using Praat software[6].

Thirty speakers from three districts of Marathwada Region of Maharashtra i.e. Latur, Parbhani and Nanded of different age group, gender and socio linguistic background were selected for development of speech database of Marathi Language.

The Data acquisition process has been conducted at 16 kHz sampling frequency using Praat in the silent environment by using Laptop with unidirectional condenser Microphone.

Marathi language has 49 characters, so, 1470 samples were collected for development Speech database of Marathi language. While data acquisition process, there are many types and sources of noise or distortions are included in a speech signal.

Signal distortion is the term often used to describe systematic undesirable change in a signal and refers to changes in a signal from the non-ideal characteristics of the communication channel, signal fading reverberations, echo, and multipath reflections and missing samples [7].

3. Noise removal through filtering techniques

3.1 Pre emphasis filter

The high-frequency signal components are emphasized to produce a more equal modulation index for the transmitted frequency spectrum. In Pre-emphasis technique, speech signals are processed to increase the magnitude of higher frequencies within frequency band with respect to the magnitude frequencies in order to improve the overall signal-to-noise ratio by minimizing the adverse effects of such phenomena as attenuation distortion or saturation of recording media [7,8,9].

The frequency response is decided by special time constants. The cutoff frequency can be calculated from the constant value 0.95. Pre-emphasis Digital Filter is used to reduce differences in power of different components of the signal and it is defined as

$$y(n) = x(n) - a.x(n-1)$$

where: $x(n)$ = value of input signal at discrete time step n,

$y(n)$ =value of output signal at discrete time step n

a =constant = 0.95.

Pre-emphasis Algorithm

Step 1: Input the spoken vowels and consonants of Marathi Language recorded at 16 khz sampling frequency and store the signal in 'x'

Step 2: Set the constant value as 0.95 to the variable 'a' to determine the cut off frequency

Step 3: Set n =length(x)

Step 4: Set y(n) = x(n) - a.x(n-1) where x(n) is original signal and y(n) filter signal

Step 5: Restore the filtered vowels and consonants into corresponding wave file.

Step 6: Determine the accuracy of filtered spoken vowels and consonants recognition

3.2 Butterworth filter

The Butterworth filter is a kind of signal processing filter designed to have as flat a frequency response as possible in the pass band. It is also mentioned to as a maximally flat magnitude filter. Butterworth filters are normal in character and of different orders, the lowest order showing the excellent in the time domain, and the higher orders showing better in the frequency domain. The Butterworth filter has minimal phase shift over the filter's band pass when evaluate to other conventional filters [8,9].

3.2.1 Low pass filter

Low pass filter is used to remove useless higher frequency interference. Its function is to filter high-frequency component and increase the low-frequency component. Sometimes it is called high-cut filter. For different low-pass filter, there is different weakening of frequency in each signal. When using it in audio applications, sometimes it is called high-cut filter or treble cut filter [8].

3.2.2 High-pass filter

High-pass filter is a filter that removes useless low frequency interference. Its function is to filter low-frequency component and increase the high-frequency component. Sometimes it is called low-cut filter or bass-out filter. In addition, high-pass filter always appears with low-pass filter. No matter high-pass filter or low-pass filter, their function is sending the sound frequencies to the appropriate unit[9].

3.2.3 Band-pass filter

A band-pass filter is used to pass frequencies within a certain range and rejects frequencies outside that range. An example of band-pass filter is resistance-inductance-capacitance circuit [8, 9].

3.2.4 Band stop filter

In signal processing, a band-stop filter or band-rejection filter is a filter that passes most frequencies unaltered, but attenuates those in a specific range to very low levels. It is the opposite of a band-pass filter. A notch filter is a band-stop filter with a narrow stopband.

Band stop filter is also called as 'band limit filter', 'T-notch filter', 'band-elimination filter and 'band-reject filter [7,8,9].

Butterworth filter Algorithm:

Step 1: Input the spoken vowels and consonants of Marathi Language recorded at 16 khz sampling frequency and store the signal in 'y'

Step 2: a. If Low pass filter is used the set cut off frequency as 8000 Hz,

b. If High pass filter is used then set cut off frequency as 2000 Hz

c. If band pass and band stop filter is used the set the band pass and stop band as Frequency range as 1000hz and 2000hz

Step 3: Determine the appropriate filter order

Step 4: Determine the filter coefficient by assigning [b, a]=butter(filterOrder, cutOffFreq / (fs/2) ,'type') where type is low pass, high pass, band pass or Band stop

Step 5: Determine the frequency response and angular frequency from filter coefficient vector b and store the frequency response and angular frequency in the vector h and w respectively

Step 6: Apply the filter y1=filter (b,a,y)

Step 7: Restore the filtered vowels and consonants into corresponding wave file.

Step 8: Determine the accuracy of filtered spoken vowels and consonants recognition

4. Results and conclusion

Filtering techniques are applied on Marathi Language database which consists of Speech Corpora and results are obtained as shown in the figure 1.

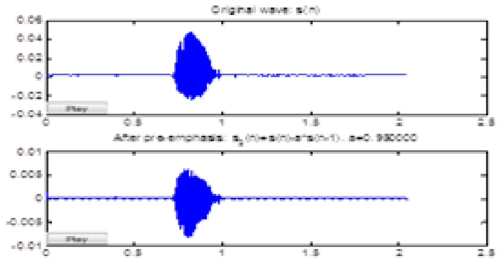

Fig. 1 (a) Pre-emphasis filter applied on vowel 'अ'

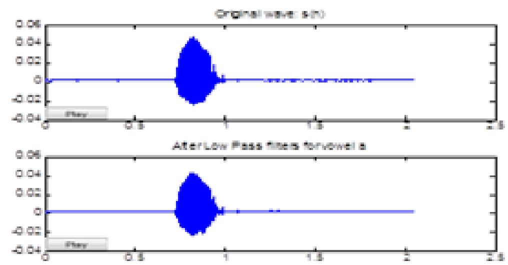

Fig. 1 (b) Low pass filter applied on vowel 'अ'

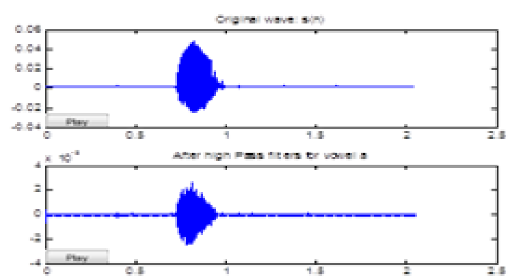

Fig. 1 (c) High pass filter applied on vowel 'अ'

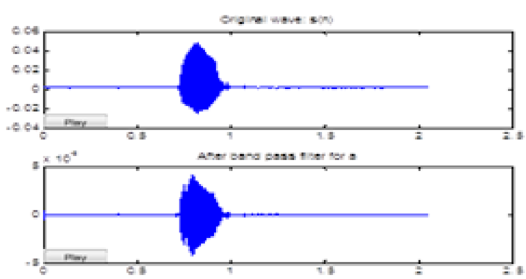

Fig. 1 (d) Band Pass filter applied on vowel 'अ'

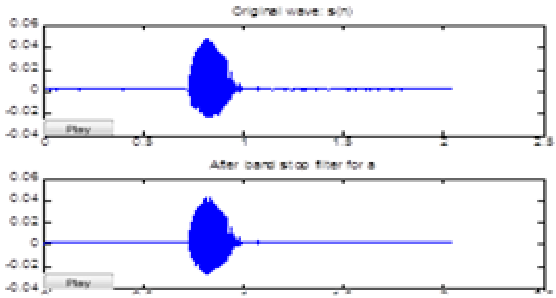

Fig. 1 (e) Band Stop filter applied on vowel 'अ'

The main objective of this study is to develop a noise free Marathi Language corpus database and to achieve this Marathi language corpus was recorded at 16 khz sampling frequency and noise removal techniques are applied to remove the noise from speech database. all types of noise removal techniques were implemented and accuracy of vowels and consonants recognitions are noted which is shown in the table 1.1 (a and b).

Table 1.1 (a) Comparison of average accuracy of recognition of vowels various filtering techniques.

	% of recognition					
Vowels	Original Signal	Pre-emphasis	Low pass	High pass	Band pass	Band stop
अ	93.25	94.75	94.25	94.45	95.20	98.15
आ	92.25	94.55	94.15	94.35	95.10	98.55
इ	93.25	94.75	94.25	94.45	95.20	98.15
ई	92.25	94.55	94.15	94.35	95.10	97.55
उ	93.25	94.75	94.25	94.45	95.20	98.15
ऊ	92.25	94.55	94.15	94.35	95.10	98.55
ए	93.25	94.75	94.25	94.45	95.20	98.15
ऐ	92.25	94.55	94.15	94.35	95.10	97.55
ओ	93.25	94.75	94.25	94.45	95.20	98.15
औ	92.25	94.55	94.15	94.35	95.10	98.55
अं	92.25	94.75	94.25	94.45	95.20	98.15
अः	93.25	94.55	94.15	94.35	95.10	97.55
रू	93.25	94.75	94.25	94.45	95.20	98.15

Table 1.1 (b) Comparison of average accuracy of recognition of Consonants various filtering techniques.

	% of recognition					
Consonants	Original Signal	Pre-emphasis	Low pass	High pass	Band pass	Band stop
क	93.25	94.75	94.25	94.75	95.45	97.88
ख	92.25	94.55	94.15	94.75	95.45	97.88
ग	93.25	94.75	94.25	94.75	95.45	97.88
ध	92.25	94.55	94.15	94.75	95.45	97.88
ङ	93.25	94.75	94.25	94.75	95.45	97.88
च	92.25	94.55	94.15	94.75	95.45	97.88
छ	93.25	94.75	94.25	94.75	95.45	97.88
झ	92.25	94.55	94.15	94.75	95.45	97.88
ञ	93.25	94.75	94.25	94.75	95.45	97.88
ट	92.25	94.55	94.15	94.75	95.45	97.88
ठ	92.25	94.75	94.25	94.75	95.45	97.88
ड	93.25	94.55	94.15	94.75	95.45	97.88
ढ	93.25	94.75	94.25	94.75	95.45	97.88
णं	92.25	94.75	94.25	94.75	95.45	97.88
त	93.25	94.55	94.15	94.75	95.45	97.88
थ	93.25	94.75	94.25	94.75	95.45	97.88
द	93.25	94.75	94.25	94.75	95.45	97.88
थ	92.25	94.55	94.15	94.75	95.45	97.88
न	92.25	94.55	94.15	94.75	95.45	97.88
च	93.25	94.75	94.25	94.75	95.45	97.88
प	92.25	94.55	94.15	94.75	95.45	97.88
फ	93.25	94.75	94.25	94.45	95.45	97.15
ब	92.25	94.55	94.15	94.35	95.45	97.15
भ	93.25	94.75	94.25	94.45	95.45	97.15
म	92.25	94.55	94.15	94.35	95.45	97.15
य	93.25	94.75	94.25	94.45	95.45	97.15
र	92.25	94.55	94.15	94.35	95.45	97.15
ल	92.25	94.75	94.25	94.45	95.45	97.15
व	93.25	94.55	94.15	94.35	95.45	97.15
श	93.25	94.75	94.25	94.45	95.45	97.15
ष	92.25	94.75	94.25	94.45	95.45	97.15
स	93.25	94.55	94.15	94.35	95.45	97.15
ह	93.25	94.75	94.25	94.45	95.45	97.15
ळ	93.25	94.75	94.25	94.45	95.45	97.15
क्ष	93.25	94.75	94.25	94.45	95.45	97.15
ज्ञ	93.25	94.75	94.25	94.45	95.45	97.15

The first conclusion of this research work is that band stop filtering technique is an efficient technique to remove noise from speech signal as compared to other noise removal techniques and second conclusion is that when noise removal techniques are applied on the speech corpus of thirty different speakers, it is found that average accuracy of recognition of speech corpus was 98.55 % which is helpful to develop an speaker independent system.

References

1. Anil Kumar Singh, 2009, Hyderabad "A Computational Phonetic Model for Indian Language Scripts", Digitizing the Legacy of Indian Languages-Book, ICFAI Books.
2. http://censusindia.gov.in/census_data_2001/census_data_online/language/statement6.aspx accessed on December 2013
3. Akshar Bharati, vinit Chaitanya, Rajiv Sangal, 2010, "Natural Language Processing : A Paninian perspective" Prentice Hall India (PHI), New Delhi Anil Kumar Singh, 2009, Hyderabad "A Computational Phonetic Model for Indian Language Scripts", Digitizing the Legacy of Indian Languages-Book, ICFAI Books.
4. Anil Kumar Singh, 2009, Hyderabad "A Computational Phonetic Model for Indian Language Scripts", Digitizing the Legacy of Indian Languages-Book, ICFAI Books.
5. http://www.censusindia.gov.in/Census_Data_2001/Census_Data_Online/Lan
6. guage/State ment 6.aspx accessed on Dec. 2013
7. G. Cardona, Panini: His Work and Its Traditions: Background and Introduction, Motilal Banarsidass,
8. Zue, V., Cole, R., Ward, W. ", "Speech Recognition. Survey of the State of the Artin Human http://www.censusindia.gov.in/Census_Data_2001/Census_Data_Online/Language/ Language Technology",", Kauii, Hawaii, USA
9. www.en.wikipedia.org/wiki/Sampling_(signal_processing) accessed on 2/08/2011
10. www.praat.org accessed on 12/06/2012
11. James F. Buckwalter,, Mounir Meghelli, Daniel J. Friedman and Ali Hajimiri "Phase and Amplitude Pre-Emphasis techniques for low power serial links" IEEE Journal Of Solid-State Circuits, VOL. 41, NO. 6
12. Prajoy Podder, Md. Mehedi Hasan and Md.Rafiqul Islam, 2014, "Design and Implementation of Butterworth, Chebyshev-I and Elliptic Filter for Speech Signal Analysis" International Journal of Computer Applications (ISSN NO: 0975 – 8887) Volume 98– No.7
13. Neelam Mehela, Anand,2013 "IIR Multiple Notch Filter Design for Power Line Interference Removal" International Journal of Emerging Science and Engineering (IJESE) ISSN: 2319–6378, Volume-1, Issue-10