Basic Fundamental Recognition of Voiced Unvoiced and Silence Region of A Speech

B.M.S. Rani, A. Jhansi Rani, T. Ravi, M. Divya Sree

ABSTRACT:- Speech which is a natural and very easy way of exchanging the information if used as a medium to interact with the computer and can solve all these problems. Speech recognition technology has made it possible for computers to follow human voice commands and understand human languages. The main goal of speech recognition area is to develop techniques and systems for speech as input to machine.Objective is (a) Load, display and manipulation of speech signals.(b) Study and understand the time and frequency domain characteristics of voiced speech.(c) Classification of the voiced/unvoiced/silence features of speech signals in both time domain and frequency domain.(d)auto-correlation sequence of voiced,unvoiced,silence features of thespeech signals.

Keywords: voiced,un voiced,silence,time domain,frequency domain,auto correlation.

I. INTRODUCTION

Sound is variations in air pressure. The creation of sound is the process of setting the air in rapid vibration.Speech signal is an acoustic signal produced from a speech production system.The system characteristics depend on the design of the system. For the case of LTI system, this is completely characterized in terms its impulse response. However, the nature of output depends on the type of input excitation to the system. For instance, we have impulse response, sinusoidal response and so on for a given system. Each of these output responses are used to understand the behavior of the system under different conditions. A similar phenomenon happens in the production of speech also. Based on the input excitation phenomenon, the speech production can be broadly classified into three activities. The first case where the input is nearly periodic in nature, the second case where the input excitation is random noise-like in nature and third case where there is no excitation to the system. Accordingly, the speech signal can be broadly classified into three regions. The study of these regions is the aim of this paper. Considering speech signals in short ranges for plotting their waveforms and spectra. The lengths include 10-30 msec. Waveforms and their spectra for segments selected from the word *speech* recorded using 16 kHz f_s (sampling frequency) and 16 bit resolution

Manuscript Received November, 2014.

B.M.S. Rani, Assistant Professor,Vignan Nirula Institute and science for women Approved by A.I.C.T.E & Affiliated to Jntuk, Kakinada Pedapalakaluru, Guntur-522 005, Andhra Pradesh, India

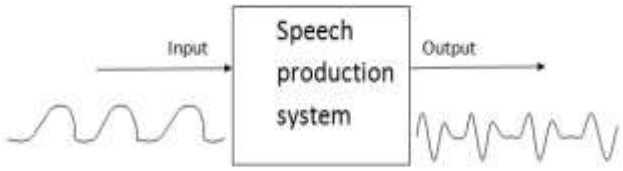
Dr A. Jhansi Rani, Professor, V.R. Siddhartha Engineering College, Andhra Pradesh - 520007, India

T. Ravi, Associate Professor, Koneru Lakshmaiah Education Foundation, Governorpet, Vijayawada. Andhra Pradesh, India.

M. Divya Sree, Assistant Professor, Vignan Nirula Institute of Technology and science for women Approved by A.I.C.T.E & Affiliated to Jntuk, Kakinada Pedapalakaluru, Guntur-522 005, Andhra Pradesh, India

II. Voiced Speech

If the input excitation to is a system is nearly periodic impulse sequence, then the corresponding speech looks visually nearly periodic and is termed as voiced speech.

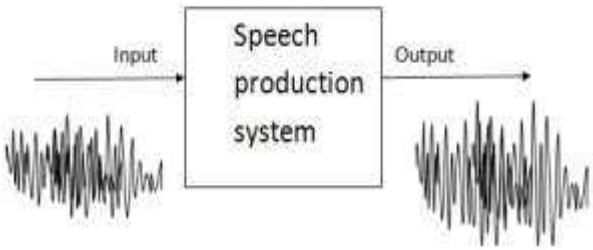


Block diagram representation of voiced speech production

The periodicity associated with the voiced speech can be measured by the autocorrelation analysis. This period is more commonly termed as pitchperiod. For male:T≈8ms⇒pitch≈125Hz; female:T≈4ms⇒pitch≈250Hz

III. Unvoiced Speech

If the excitation is random noise-like, then the resulting speech signal will also be random noise-like without any periodic nature and is termed as Unvoiced Speech. The typical nature of excitation and resulting unvoiced speech are shown in figure itself. As it can be seen, the unvoiced speech will be non- periodic in nature. This will be the main difference between voiced and unvoiced speech. The non- periodicity of unvoiced speech can also be observed by the autocorrelation analysis.



Block diagram representation of unvoiced speech production Silence Region:

The speech production process consists of generation of voiced and unvoiced speech in succession, separated by silence region. During silence region, there is no excitation supplied to the vocal tract and hence no speech output. However, silence is a part of speech signal. Without the presence of silence region between voiced and unvoiced speech, the speech will not be complete. The duration of silence along with other voiced or unvoiced speech is an indicator of certain types of sounds. Even though from amplitude/energy point of view, silence region is unimportant, but its duration is very essential for intelligible speech.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

IV. Voiced/Unvoiced/Silence Classification of Speech

Above discussion gave a feel about the production of voiced/unvoiced speech and also significance of silence region. Now the next question is how to identify these regions of speech? First by visual perception and next by automatic approach. If the speech signal waveform looks periodic in nature, then it may be marked as voiced speech, otherwise, as unvoiced/silence region based on the associated energy. If the signal amplitude is low or negligible, then it can be marked as silence, otherwise as unvoiced region.

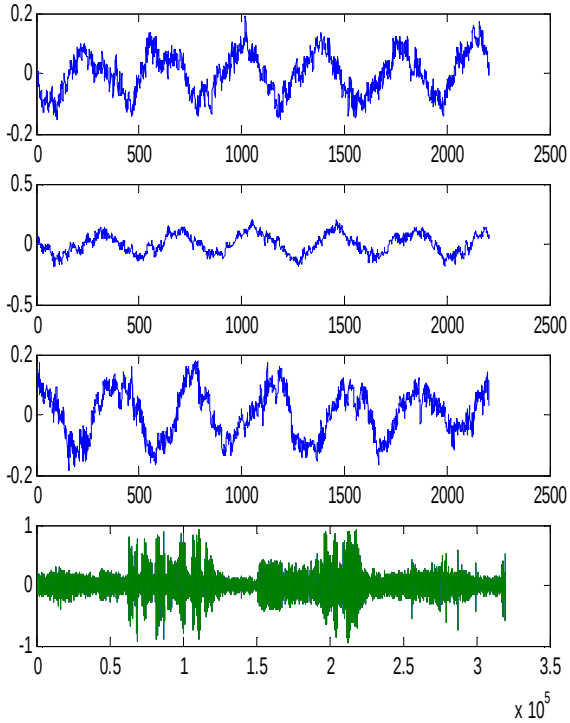


Figure 1: Time Domain Analysis of Voiced Speech Signal

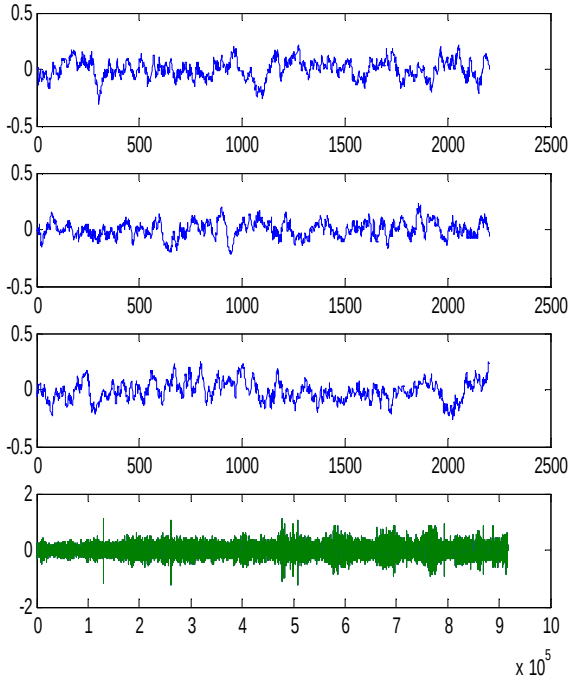


Figure 2: Time Domain Analysis of Unvoiced Speech Signal

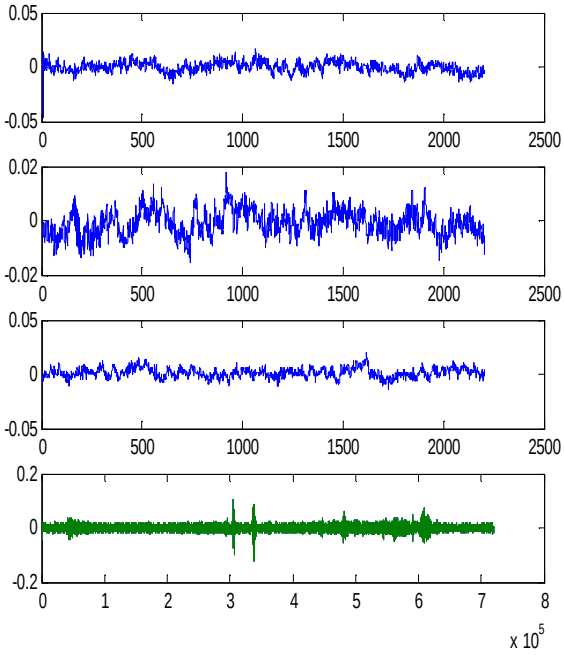


Figure 3:Time Domain Analysis of Silence Region of Speech Signalz

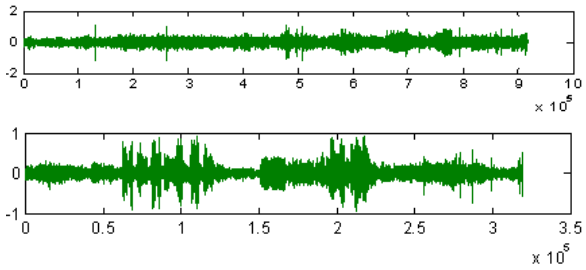


Figure 4:difference between voiced and unvoiced features in time domain analysis.

Frequency Domain Analysis:

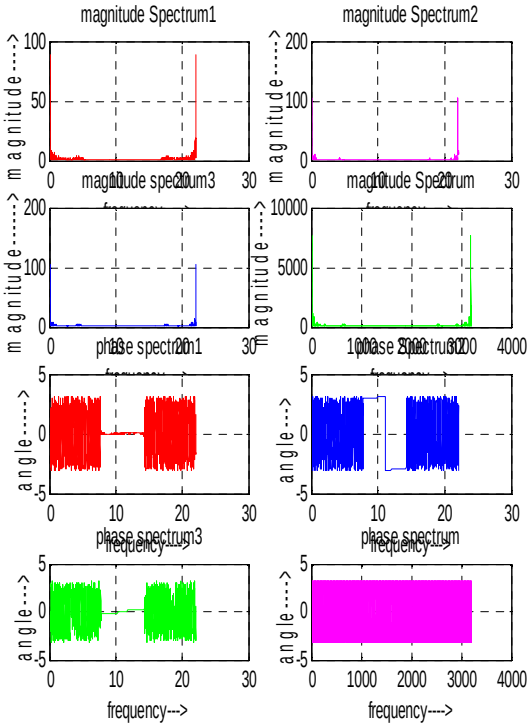


Figure 5:Frequency Domain Analysis of Voiced Speech Signal

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

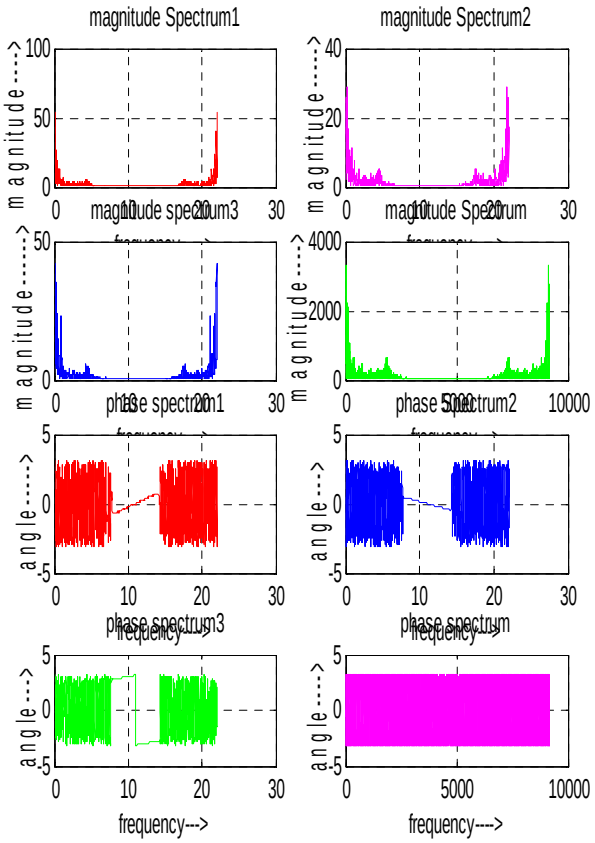


Figure 6:Frequency Domain Analysis of Unvoiced Speech Signal

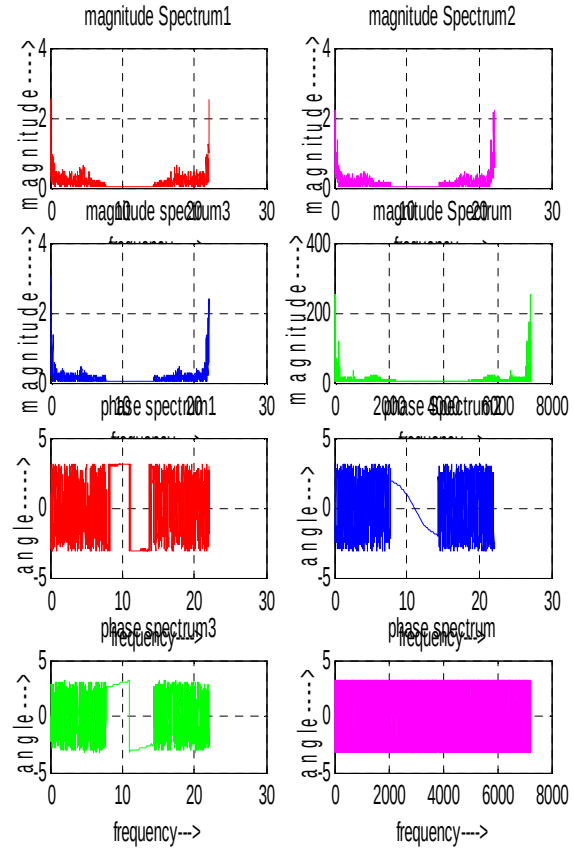


Figure 7:Freguency Time Domain Analysis of Silence Region of Speech Signal

Loading a speech signal, Plotting the spectral waveforms and auto-correlation sequence:

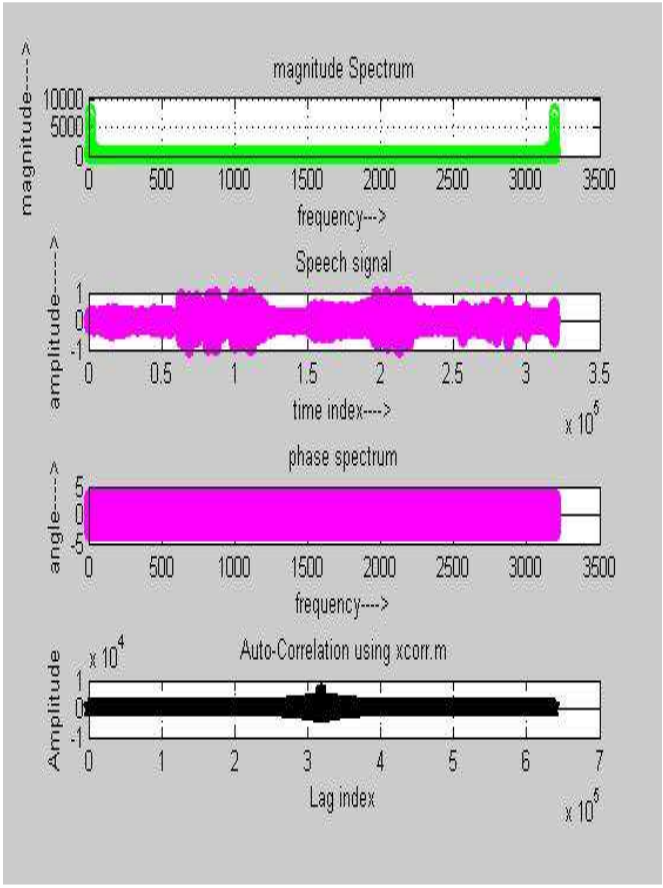


Figure 8:Auto-Correlation sequence of Voiced Speech Signal

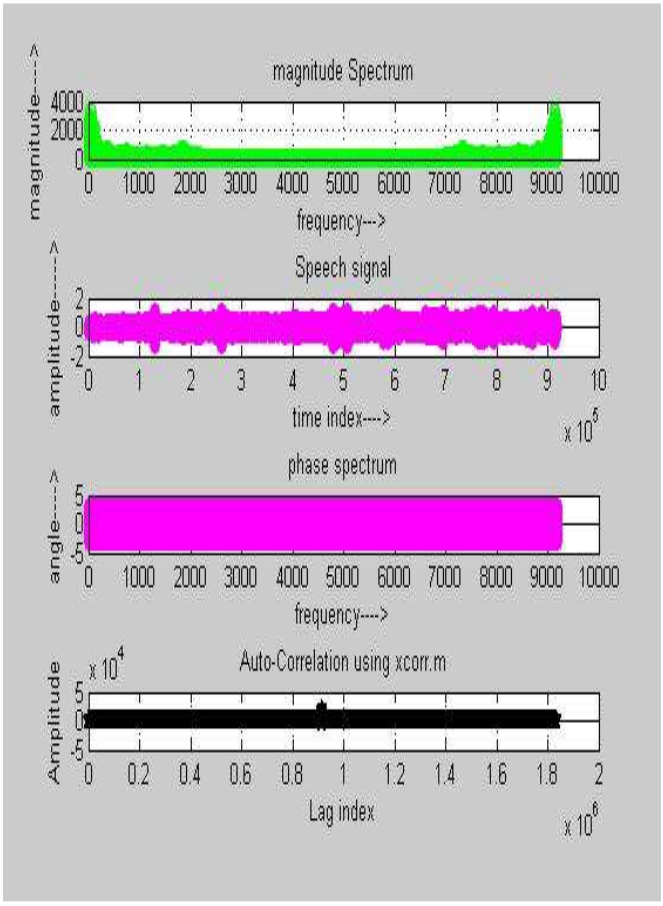


Figure 9:Auto-Correlation sequence of Unvoiced Speech Signal

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

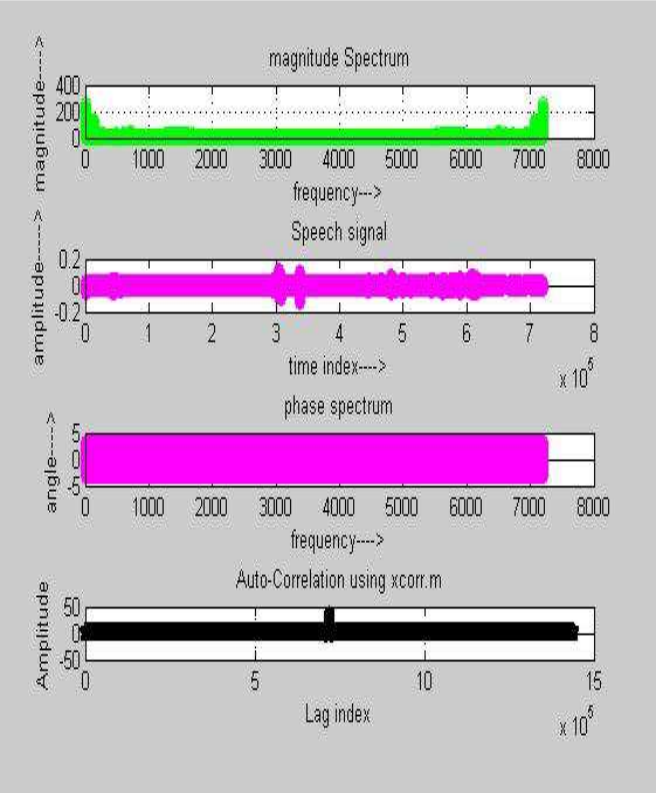


Figure 10: Auto-Correlation of Silence Region of Speech Signal

V. CONCLUSION AND FUTURE SCOPE

Speech Recognition is in research for many years. After years of research and development the accuracy of Automatic Speech Recognition still remains one of the important challenges. The design of Recognition systems requires careful attention to the issues like speech representation, preprocessing, feature extraction etc.

The classification of speech as voiced, unvoiced and silence is one of the most fundamental and difficult problem encountered in speech processing and if it is detected accurately it can increase the efficiency and accuracy of the recognition system to a great extent.

REFERENCES

1. Agarwal, A., Jain, A. and Prakash, N., 2010. Word Boundary Detection in Continuous Speech based on Suprasegmental Features for Hindi Language. 2nd International Conference on Signal Processing Systems, vol. 2, pp. V2-591-V2-594.
2. Anusuya, M. A. and Katti, S. K., 2009. Speech Recognition by machine: A review. International journal of Computer science and information security, vol. 6, no. 3, pp. 181-205.
3. Jong Kwan Lee, Chang D. Yoo, "Wavelet speech enhancement based on oiced/unvoiced decision",Korea Advanced Institute of Science and Technology The 32nd International Congress and Exposition on Noise Control Engineering, Jeju International Convention Center, Seogwipo, Korea ,August 25-28, 2003.
4. B. Atal, and L. Rabiner, "A Pattern Recognition Approach to Voiced-Unvoiced-Silence Classification with Applications to Speech Recognition," IEEE Trans. On ASSP, vol. ASSP-24, pp. 201-212, 1976.
5. S. Ahmadi, and A.S. Spanias, "Cepstrum-Based Pitch Detection using a New Statistical V/UV Classification Algorithm," IEEE Trans. Speech Audio Processing, vol. 7 No. 3, pp. 333-338, 1999.
6. Y. Qi, and B.R. Hunt, "Voiced-Unvoiced-Silence Classifications of Speech using Hybrid Features and a Network Classifier," IEEE Trans. Speech Audio Processing, vol. 1 No. 2, pp. 250-255, 1993.
7. L. Siegel, "A Procedure for using Pattern Classification Techniques to obtain a Voiced/Unvoiced Classifier", IEEE Trans. on ASSP, vol. ASSP-27, pp. 83- 88, 1979.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.