SPEECH CODING BY LINEAR PREDICTIVE CODING

Vivek Kothari, C.Chandra Vamshi, Vamshi Sai Teja
Student, Student, Student
Electronics and Communication Engineering ,
Gandhi Institute of Technology and Management, Hyderabad, India.

Abstract: **Linear predictive coding (LPC) is a method used mostly in audio signal processing and speech processing for representing the spectral envelope of a digital signal of speech in compressed form, using the information of a linear predictive model. It is one of the most powerful speech analysis techniques, and one of the most useful methods for encoding good quality speech at a low bit rate and provides highly accurate estimates of speech parameters. LPC is the most widely used method in speech coding and speech synthesis.**

Index Terms **– Linear predictive coding, Speech processing, Speech synthesis.**

1. INTRODUCTION

Speech is one of the most important mediums by which a communication can take place. With the invention and widespread use of mobiles, telephones, data storage devices etc. has provided a major help in setting up of speech communication and its analyzing. The term and the basic concept of speech identification was beginning in the early 1960`s with exploration into voiceprint analysis which was somewhat similar to fingerprint concept. Nowadays, with further growth & advancement in the field of speech recognition, the humans who are physically challenged such as blind and deaf can easily communicate with the machines. So, in biological terms a voice that is being generated through trachea will be decoded by brain.Speech coding uses speech-specific parameter estimation using audio signal processing techniques to model the speech signal, combined with generic data compression algorithms to represent the resulting modeled parameters in a compact bit stream. Some applications of speech coding are mobile telephony and voice over IP (VoIP). The most widely used speech coding technique in mobile telephony is linear predictive coding (LPC), while the most widely used in VoIP applications are the LPC and modified discrete cosine transform (MDCT) techniques.

2. SPEECH CODING

Speech is a very special type of signal for different reasons. The most preliminary of these is the fact that speech is a non-stationary signal. This makes the speech signal hard to analyze and model. The second reason is that factors like intelligibility, coherence and other such characteristics play a vital role in the analysis of the speech signals. The third reason in communication point of view is that the number of discrete values required to describe one second of speech signal corresponds to 8000 samples (at the minimum). As bandwidth is the parameter which affects the cost of processing, speech signals are subjected to compression before transmission.

Speech coding or compression is a process of obtaining a compact representation for the speech signals, for the purpose of efficient transmission over band limited wired or wireless channels and also for efficient storage. In recent day's speech coders became the essential components for telecommunications and multimedia as the utilization of the bandwidth affects the cost of transmission. The goal of speech coding is to represent the samples of a speech signal with a minimum number of bits without any reduction in the perceptual quality. Speech coding helps a telephone company to carry out more voice calls on a single fiber link or cable. Speech coding is very important in Mobile and Cellular communications where the data rates for a 2 particular user are limited, as lower the data rates for a voice call more services can be accommodated. Speech coding is also useful for Voice over IP, Video conferencing and Multimedia applications to reduce the bandwidth requirement over internet. In addition, most of the speech applications require minimum coding delay because long coding delays hinder the flow of the speech conversation.

A speech coder is one which converts a digitized speech signal into a coded representation and transmits it in the form of frames. At the receiving end the speech decoder receives the coded frames and performs synthesis to reconstruct the speech signal. The speech coders differ primarily in bit-rate, complexity, delay and perceptual quality of the synthesized speech . There exist two types of coding techniques, narrowband speech coding and wide band speech coding. Narrowband speech coding refers to coding of the speech signals whose bandwidth is between 300 to 3400 Hz (with 8 KHz sampling rate), while wide band speech coding refers to coding of the speech signals whose bandwidth is less than 5000 to 7000 Hz (with 14 –16 KHz sampling rate). Narrowband speech coding is more common than wide band speech coding because of the narrowband nature of the telephone channel lines (whose bandwidth lies between 300 to 3400 Hz) [6]. In recent days there is an increase in demand for wide band speech coding techniques in applications like video conferencing.

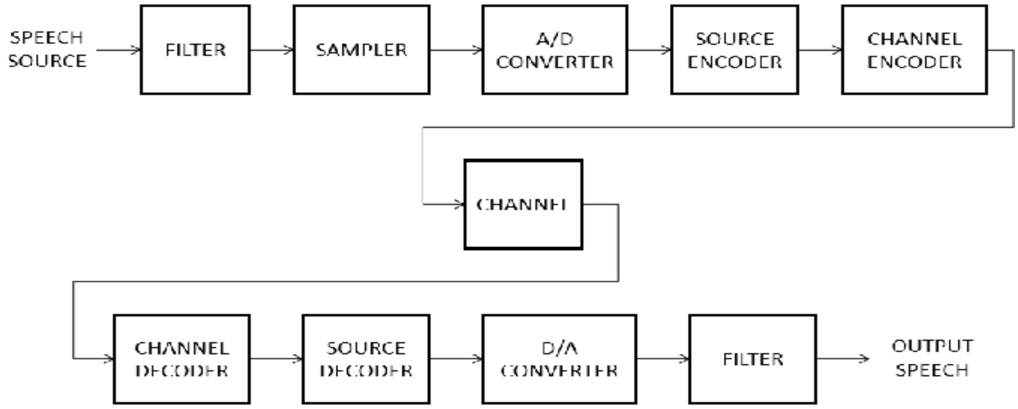


Figure 1 - General block diagram of speech coding system

3. SPEECH CODING TECHNIQUES

The techniques employed in speech coding are similar to those used in audio data compression and audio coding where knowledge in psychoacoustics is used to transmit only data that is relevant to the human auditory system. For example, in voice band speech coding, only information in the frequency band 400 Hz to 3500 Hz is transmitted but the reconstructed signal is still adequate for intelligibility.Speech coding differs from other forms of audio coding in that speech is a simpler signal than most other audio signals, and a lot more statistical information is available about the properties of speech. As a result, some auditory information which is relevant in audio coding can be unnecessary in the speech coding context. In speech coding, the most important criterion is preservation of intelligibility and "pleasantness" of speech, with a constrained amount of transmitted data. In addition, most speech applications require low coding delay, as long coding delays interfere with speech interaction. In order to compress the data different techniques are used. Speech coding techniques are classified based on bit-rate at which they produce output with reasonable quality and which are used for coding speech signal.

The coding algorithm used in this thesis for the reduction of the bit-rate is the linear predictive coding (LPC) algorithm. Using this algorithm, the bit rate is reduced to 1 Kbps, i.e., a reduction in bit-rate by 64 times with respect to the input. The output of the source encoder is given as an input to the channel encoder which provides error protection to the bit stream transmitted over the communication channel where noise and interference can reduce the reliability of the transmitted data. At the receiving end the channel decoder recovers the encoded data from the error protected data and will be fed to the source decoder so as to recover the original speech signal. Later the speech signal is fed to a digital to analog (D/A) converter to convert the speech signal from Channel A/D Converter Source Encoder Channel Encoder Filter Sampler Input Speech Channel Decoder Source Decoder D/A Converter Filter Output Speech 6 digital to analog format. Finally, the analog speech signal is fed to an anti-aliasing filter to prevent aliasing during the reconstruction of continuous speech signal from the speech samples, that again requires perfect stop-band rejection to guarantee zero aliasing.

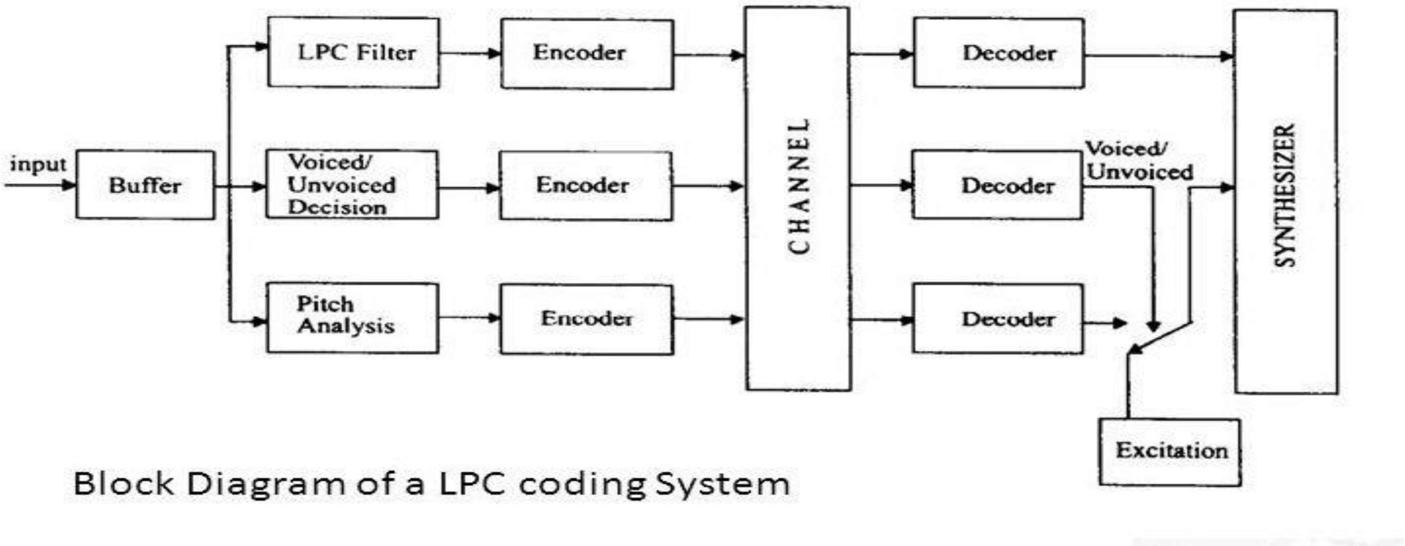


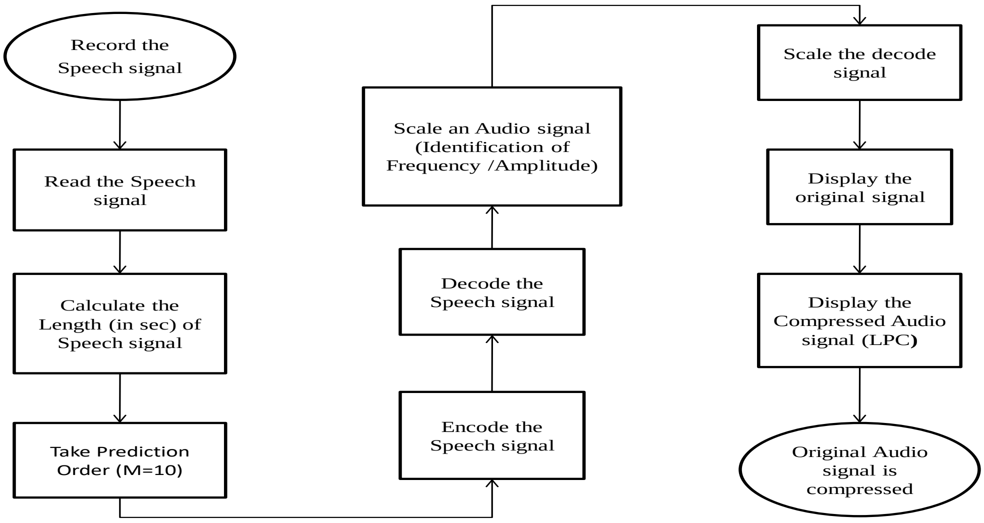
Figure 2 – Block diagram of linear predictive coding

.

4. ALGORITHM

- STEP 1: Record the audio signal.
- STEP 2: Read the audio signal.
- STEP 3: Calculate the length of audio signal (in seconds).
- STEP 4: Introduce Prediction order (M=10).
- STEP 5: Encode and decode the audio signal.
- STEP 6: Display the audio signal and compressed signal through LPC.

5. FLOWCHART



6. SIMULATION RESULTS

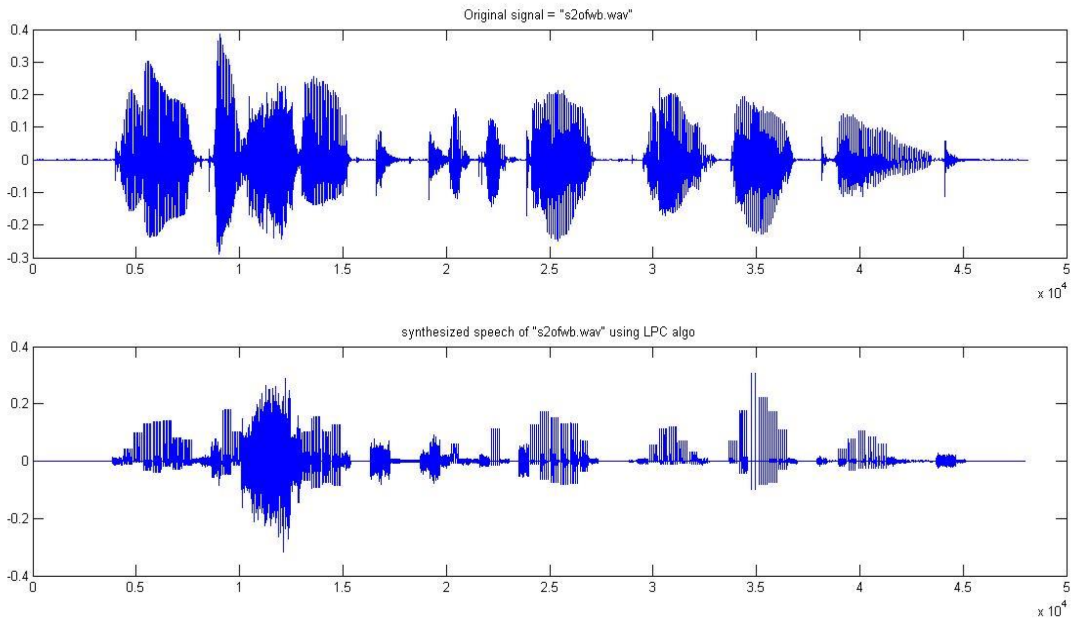


Figure 3 - Simulation result of speech coding using LPC

ACKNOWLEDGEMENT

We would like to thank our Pro Vice Chancellor- Dr.N Siva Prasad, Head of the Department- Dr.K Manjunathachari and our respected guide-Mrs.D.Anitha who have helped us in making this project. Without their support, this project would not have been made possible.

I would also like to thank the different resources and papers which acted as a guide in helping us do the project.

REFERENCES

[1] T. Fingschieldt, Suhadi, S. Stan. Environment-Optimized Speech Enhancement, IEEE transactions on Audio, Speech and Language processing 2008; vol. 16 (No4).

[2] P. Kabal. ITU-T G.723.1 Speech coder- MATLAB implementation, 2009.

[3] K. Satyapriya, Y. Dasari. Performance analysis of speech coding techniques, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, November 2013; vol.2; issue 11: 5725- 32.