International Journal of Scientific and Technical Advancements

ISSN: 2454-1532

Sound Compression Using Linear Predictive Coding Technique

Manoj Kumar[1], Sumer Khajuria[2]
[1, 2]Department of E&C Dept., GCET, Jammu, India
Email address: manoj963@rediffmail.com, sumeer.k@rediffmail.com

Abstract**—** Linear predictive analysis is based on the concept that the speech sample can be approximated as a linear combination of past samples. The Linear Prediction Model provides a robust, reliable and accurate method for estimating parameters that characterize the linear time varying system. In this project, we implement a voice excited LPC vocoder for low bit rate speech compression and its implementation in MATLAB.

Keywords**—** Speech compression, linear predictive coding (LPC), human speech production, LPC model.

I. INTRODUCTION

The human speech in its pristine form is acoustic signal. For the purpose of communication and storage, it is necessary to convert it into an electrical signal. This is accomplished with the help of certain instruments called 'transducers'. This electrical representation of speech has certain properties

- It is one-dimensional signal, with time as its independent variable.
- It is random in nature.
- It is non-stationary, i.e., the frequency spectrum is not constant with time.

Although human beings have an audible frequency range of 20 Hz – 20 KHz, the human speech has significant frequency components only upto 4 KHz, a property that is exploited in the compression of speech[1].

II. SPEECH COMPRESSION

Speech coding is a procedure to represent a digitized speech signal using as few bits as possible, maintaining at the same time a reasonable level of speech quality. It is also known as speech compression. In other words, it is concerned with obtaining compact digital representation of voice signals for the purpose of efficient transmission and storage the objective is to represent speech signal with minimum number of bits yet maintain the perceptual quality. Therefore, involves sampling and quantization [2].

The continuous time analog speech signal from a given source is digitized by a standard connection of filter which eliminates aliasing, sampler which converts into discrete-time, and analog-to-digital converter used for uniform quantization. The output is a discrete time speech signal whose sample values are also discrete. This signal is referred to as Digital speech. According to Nyquist theorem, the sampling frequency must be atleast twice the bandwidth of the continuous-time signal in order to avoid aliasing, a value of 8kHz is commonly selected as the standard sampling frequency for speech signals. So, Speech compression or speech coding is defined as a method for reducing the amount of information needed to represent a speech signal.

III. LINEAR PREDICTIVE CODING

Linear predictive coding (LPC) is defined as a digital method for encoding an analog signal in which a particular value is predicted by a linear function of the past values of the signal. Human speech is produced in the vocal tract which can be approximated as a variable diameter tube. The linear predictive coding (LPC) model is based on a mathematical approximation of the vocal tract represented by this tube of a varying diameter. At a particular time, t, the speech sample s (t) is represented as a linear sum of the p previous samples. Under normal circumstances, speech is sampled at 8000 samples/second with 8 bits used to represent each sample. This provides a rate of 64000 bits/second. LPC reduces this to 2400 bits/second [3]. At this reduced rate the speech has a distinctive synthetic sound and there is a noticeable loss of quality. However, the speech is still audible and it can still be easily understood. Since there is information loss, LPC is a lossy form of compression. Most forms of speech compression are achieved by modeling the process of speech production as a linear digital filter. The digital filter and its slow changing parameters are usually encoded to achieve compression from the speech signal. Linear Predictive Coding (LPC) models the process of speech compression as a linear sum of earlier samples using a digital filter inputting an excitement signal. Speech coding or compression is usually conducted with the use of voice coders or vocoders. LPC vocoders are considered model-based coders which means that LPC coding is lossy even if no quantization errors occur. Any bit rate below 64 kb/s is considered compression. LPC transmits speech at a bit rate of 2.4 kb/s, an excellent rate of compression. Linear predictive coding because of its high compression rate is very complex and involves executing millions of instructions per second. LPC often requires more than one processor to run in real time [4].

IV. THE HUMAN SPEECH PRODUCTION

The process of speech production in humans can be summarized as air being pushed from the lungs, through the vocal tract, and out through the mouth to generate speech. In this type of description the lungs can be thought of as the

Manoj kumar and Sumer Khajuria, "Sound compression using linear predictive coding technique," *International Journal of Scientific and Technical Advancements*, Volume 2, Issue 4, pp. 299-301, 2016.

source of the sound and the vocal tract can be thought of as a filter that produces the various types of sounds that make up speech.

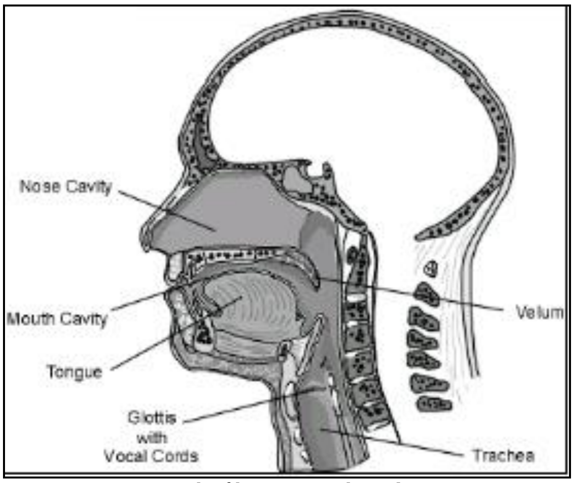

Fig.1. Path of human speech production.

Voiced sounds are usually vowels and often have high average energy levels and very distinct resonant or formant frequencies. Voiced sounds are generated by air from the lungs being forced over the vocal cords. As a result the vocal cords vibrate in a somewhat periodically pattern that produces a series of air pulses called glottal pulses. The rate at which the vocal cords vibrate is what determines the pitch of the sound produced. Unvoiced sounds are usually consonants and generally have less energy and higher frequencies than voiced sounds. The production of unvoiced sound involves air being forced through the vocal tract in a turbulent flow. During this process the vocal cords do not vibrate, instead, they stay open until the sound is produced. Pitch is an unimportant attribute of unvoiced speech since there is no vibration of the vocal cords and no glottal pulses. The other components that influence speech production are the shape of the vocal tract and the amount of air that originates in the lungs. The idea of the air from the lungs as a source and the vocal tract as a filter is called the source-filter model for sound production.

V. LPC MODEL

The particular source-filter model used in LPC is known as the Linear predictive coding model. It has two key components: analysis or encoding and synthesis or decoding. The analysis part of LPC involves examining the speech signal and breaking it down into segments or blocks. Each segment is then examined further to find whether the segment voiced or unvoiced, what is the pitch of the segment and what parameters are needed to build a filter that models the vocal tract for the current segment.

LPC analysis is usually conducted by a sender who answers these questions and usually transmits these answers onto a receiver. The receiver performs LPC synthesis by using the answers received to build a filter that when provided the correct input source will be able to accurately reproduce the original speech signal [5].

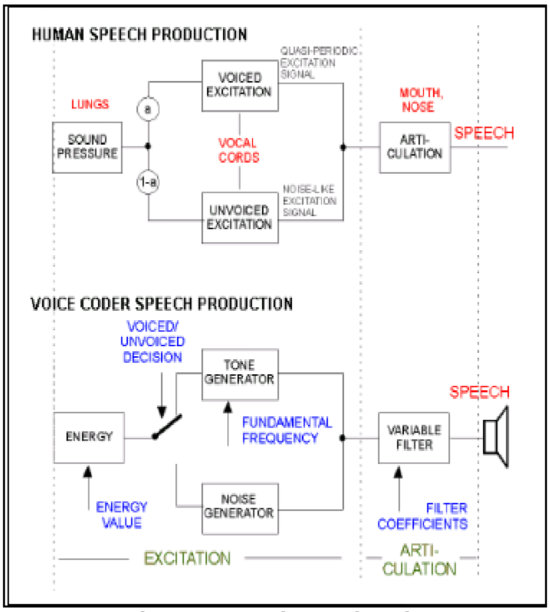

Fig. 2. human vs. vocoder speech production.

VI. RESULT

Waveforms of Original Speech Signal and LPC reconstructed speech signals gives the idea of quality of signals.

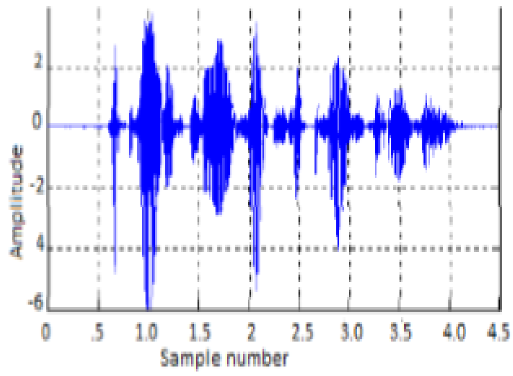

Fig. 3. waveform of original speech signal.

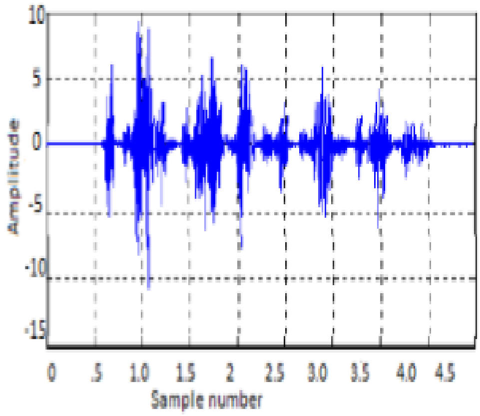

Fig. 4. Waveform of LPC reconstructed signal.

Manoj kumar and Sumer Khajuria, “Sound compression using linear predictive coding technique,” *International Journal of Scientific and Technical Advancements*, Volume 2, Issue 4, pp. 299-301, 2016.

VII. Conclusion

As can be seen from the waveforms that the voice excited waveform looks closer to the original sound but is not an exact replica. The synthesized speech is similar to the original speech signal because the power spectral density remains the same. Voice –discrimination tests indicate that voice identity is well preserved. Crucial factors influencing the reconstructed speech quality are accuracy of spectral flattening and the impulse response of the analyzer low pass filters.Thus, it can be concluded that using LPC Model, we can overcome the biggest constraints of speech analysis i.e., transmission and storage effectively.

References

[1] Hardman and O. Hodson. Internet/Mbone Audio (2000) 5-7.
[2] Mark Nelson and Jean-Loup Gailly. Speech Compression, The Data Compression Book (1995) 289-319.
[3] R. Sproat, and J. Olive. Text-to-Speech Synthesis, Digital Signal Processing Handbook (1999) 9-11.
[4] Richard V. Cox. Speech Coding (1999) 5-8.
[5] Khalid Sayood. Introduction to Data Compression (2000) 497-509.

Manoj kumar and Sumer Khajuria, "Sound compression using linear predictive coding technique," *International Journal of Scientific and Technical Advancements*, Volume 2, Issue 4, pp. 299-301, 2016.