IETE 46th Mid Term Symposium "Impact of Technology on Skill Development" MTS- 2015
Special Issue of International Journal of Electronics, Communication & Soft Computing Science and Engineering, ISSN: 2277-9477

# Analysis of Linear Predictive Coding

**Jennifer Jose**    **Smita Chopde**

*Abstract:* **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. The most important aspect of LPC is the linear predictive filter which allows the value of the next sample to be determined by a linear combination of previous samples. This paper describes how different parameters of speech signal are extracted for faithful reproduction of speech.**

**Key Words: Linear Predictive Coding (LPC), voiced sounds, pitch, vocal tract**

## I. INTRODUCTION

Linear Predictive Coding (LPC) is one of the methods of compression that models the process of speech production[1]. Specifically, LPC models this process 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. There are two types of voice coders: waveform-following coders and model-base coders. Waveform following coders will exactly reproduce the original speech signal if no quantization errors occur. Model based coders will never exactly reproduce the original speech signal, regardless of the presence of quantization errors, because they use a parametric model of speech production which involves encoding and transmitting the parameters not the signal. Lpc Vocoders are considered model-based coders which means that LPC coding is lossy even if no quantization errors occur. The general algorithm for linear predictive coding involves an analysis or encoding part and a synthesis or decoding part. In the encoding, LPC takes the speech signal in blocks or frames of speech and determines the input signal and the coefficients of the filter that will be capable of reproducing the current block of speech. This information is quantized and transmitted. In the decoding, LPC rebuilds the filter based on the coefficients received. The filter can be thought of as a tube which, when given an input signal, attempts to output speech. Additional information about the original speech signal is used by the decoder to determine the input or excitation signal that is sent to the filter for synthesis.

## II. 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.

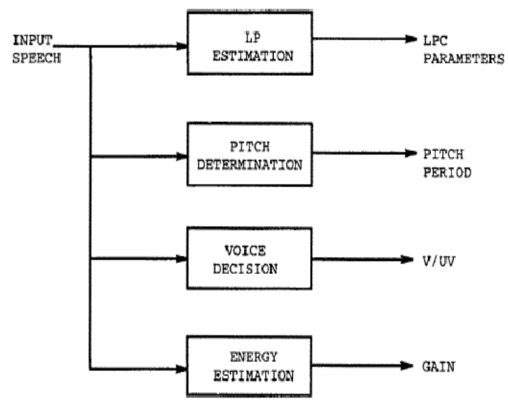


**Figure 1. Block diagram for the LPC encoder**

Figure 1 demonstrates parts of the LPC encoder receiver correspond to what parts in the human anatomy The analysis part of LPC involves examining the speech signal and breaking it down into segments or blocks. Each segment is than examined further to find the answers to several key questions:

- Is the segment voiced or unvoiced?
- What is the pitch of the segment?
- 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. Essentially, LPC synthesis tries to imitate human speech production.

## III. LPC ANALYSIS/ENCODING

**Input speech**

First the input signal is sampled at a rate of 8000 samples per second. The input signal is then broken up into segments or blocks which are each analysed and transmitted to the receiver. The 8000 samples in each second of speech. signal are broken into 180 sample segments. This means that each segment represents 22.5 milliseconds of the input speech signal.

**Voiced /Unvoiced Determination**

According to LPC-10 standards, before a speech segment is determined as being voiced or unvoiced it is first passed



Journal Impact Factor-2.02
Indexed by- ProQuest, DOAJ and many International Bodies.

through a low-pass filter with a bandwidth of 1 kHz. Determining if a segment is voiced or unvoiced is important because voiced sounds have a different waveform then unvoiced sounds. The differences in the two waveforms creates a need for the use of two different input signals for the LPC filter in the synthesis or decoding. One input signal is for voiced sounds and the other is for unvoiced. The LPC encoder notifies the decoder if a signal segment is voiced or unvoiced by sending a single bit[7].

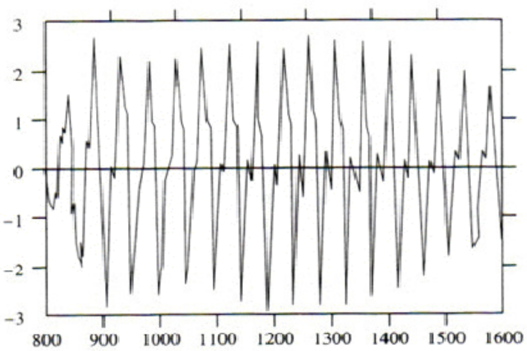


**Figure 2: Voiced sound – Letter ‘e’**

Voiced sounds are usually vowels and can be considered as a pulse that is similar to periodic waveforms. These sounds have high average energy levels which means that they have very large amplitudes. Voiced sounds also have distinct resonant or formant frequencies.

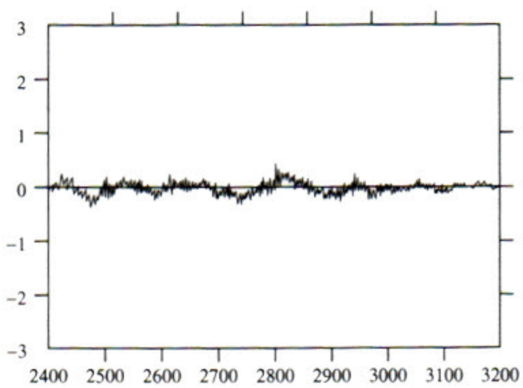


**Figure 3: Unvoiced sound – Letter ‘s’**

Unvoiced sounds are usually non-vowel or consonants sounds and often have very chaotic and random waveforms.

The following are the steps in the process of determining if a speech segment is voiced or unvoiced. The first step is to look at the amplitude of the signal, also known as the energy in the segment. If the amplitude levels are large then the segment is classified as voiced and if they are small then the segment is considered unvoiced. This determination requires a preconceived notion about the range of amplitude values and energy levels associated with the two types of sound[2].

**Pitch Period Estimation**

Determining if a segment is a voiced or unvoiced sound is not all of the information that is needed by the LPC decoder to accurately reproduce a speech signal. In order to produce an input signal for the LPC filter the decoder also needs another attribute of the current speech segment known as the pitch period. The period for any wave, including speech signals, can be defined as the time required for one wave cycle to completely pass a fixed position[7]. For speech signals, the pitch period can be thought of as the period of the vocal cord vibration that occurs during the production of voiced speech. Therefore, the pitch period is only needed for the decoding of voiced segments and is not required for unvoiced segments since they are produced by turbulent air flow not vocal cord vibrations. It is very computationally intensive to determine the pitch period for a given segment of speech. There are several different types of algorithms that could be used[4]. One type of algorithm takes advantage of the fact that the autocorrelation of a period function, $R_n(k)$, will have a maximum when k is equivalent to the pitch period. These algorithms usually detect a maximum value by checking the autocorrelation value against a threshold value.

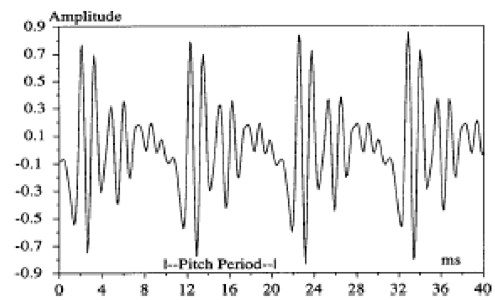


**Figure 4 : Time-domain waveform of a short segment of voiced speech[3]**

Figure 4 displays a time waveform for a short (40 ms) segment of a voiced sound. The x axis is the time scale, numbered in ms. The y axis is the amplitude of the recorded sound pressure. The high amplitude values mark the beginning of the pitch pulse. The first pitch period runs from near 0 ms to about 10 ms, the second from near 10 ms to about 20 ms.

**Determination of the Gain**

Gain should be determined by matching the energy in the signal with the energy of the linear predicted samples[6].The gain can be expressed as

$$G^2 = R_n(0) - \sum_{k=1}^{p} \alpha_k R_n(k)$$

Where $R_n(k)$ is the auto-correlation function and p is the predictor coefficients

## IV. IMPLEMENTATION AND TEST RESULTS

The implementation was carried out in MATLAB software to determine the different parameters of the speech signal. The sampled speech signal was divided into frames using a hamming window. Here the frame length used was of 20ms.The pitch estimation was carried out by the method of autocorrelation. The results obtained from MATLAB a of a Vowel-Consonant-Vowel "aa-Sh-aa", which is a recording of a female speaker.

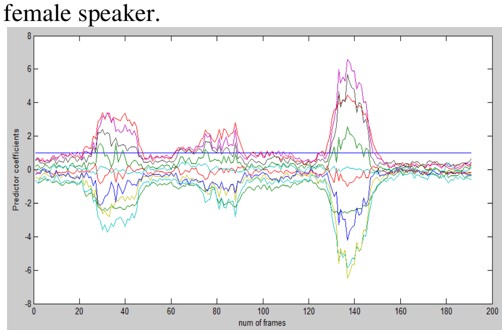


**Figure 5 : Plot of the predictor coefficients(p=10)**

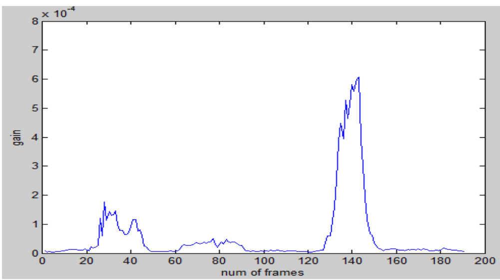


**Figure 6 : Plot of the Gain parameter**

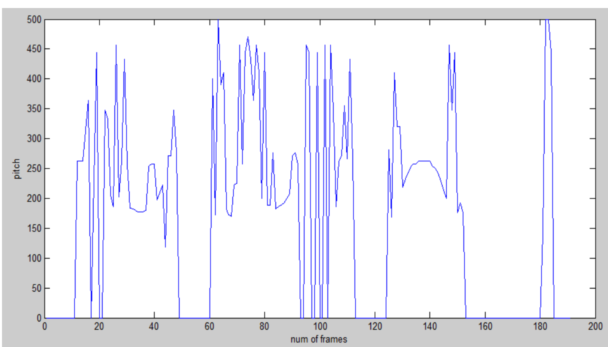


**Figure 7 : Plot of the pitch**

## REFERENCES

[1] D. O'Shaughnessy, Speech Communication- Human and Machine. Reading, MA: Addison-Wesley, 1987.

[2] M.H Johnson, Speech Coding: Fundamentals and Applications Wiley Encyclopedia of Telecommunications,2003

[3] A. S. Spanias, "Speech Coding: A Tutorial Review" Proc. IEEE. VOL.82, October 1994 Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.

[4] Simon Haykin " Communication systems" Wiley India.

[5] C.Xydeas."An Overview Of Speech Coding Techniques" IEEE Trans Vol No. 3, June 1981.

[6] L. R. Rabiner and R. W. Schafer, Digital Processing of Speech Signals. Englewood Cliffs,NJ: Prentice-Hall, 1978.

[7] R.G Goldberg, A Practical Handbook of Speech Coders. CRC Press 2000