LPC ANALYSIS OF SPEECH SIGNAL

Lalima Singh1* and Rakesh Kumar Garg1

*Corresponding Author: Lalima Singh, Lalima775@gmail.com

Linear Predictive Coding is most powerful signal analysis technique. LPC is used to compress the spectral information for its efficient storage and transmission. In this paper the LPC is implemented using simulink model in matlab. LPC model consist of two parts analysis and synthesis. In analysis section, the reflection coefficient from input speech signal is extracted and use it to compute residual coefficient. In synthesis section, the original signal is reconstructed using reflection and residual coefficient. The input speech signal of 8 kHz is taken in ‘wav’ format.

Keywords: LPC, Source filter model, Autocorrelation, L-D algorithm, Simulink

INTRODUCTION

Speech is primary mode of communication among human being. It is most efficient and natural form of exchanging information among human. Speech sounds are sensation of air pressure vibration produced by air exhaled from lungs through the vibrating vocal cords and vocal tract and out from lips and noise airways. Speech sounds have a rich and multilayered temporal spectral vibration which conveys words, intention, intonation, expression, accent, speaker identity, gender, age, style of speaking, emotion and state health of speaker. The human speech production system is shown in Figure1.

The various organs responsible for producing speech are labelled. It consists of

1 Department of Electronics and Communication Engineering, Kurukshetra Institute of Management and Technology, Kurukshetra, India.
The lungs, larynx, vocal tract cavity, nasal cavity, connecting tubes, teeth. The lungs act as power supply and provide airflow to the larynx. The larynxes modulate the airflow and provide either a periodic puff like or noisy. The vocal tract [6][7] consist of oral, nasal and pharynx cavities, giving the modulated air flow its colour by spectrally shaping the source. The variation of air pressure at lips result in travelling sound wave that listener perceive as speech. The combined voice production mechanism produces variety of vibration and spectral temporal composition that from different speech sound.

**LINEAR PREDICTIVE CODING**

LPC [1][2][3] stands for linear predictive coding, it is most powerful signal analysis technique. LPC method is used to compress the spectral information for its efficient storage and transmission. The basic idea behind linear predictive analysis is that specific speech sample at the current time can be approximated as linear combination of past speech samples.

\[ E(n) = A(z)S(z) \]

where \( A(z) \) is

\[ A(z) = 1 + \sum_{k=1}^{p} a_k z^{-k} \]

The vocal transfer function \( H(z) \) is

\[ H(z) = 1/A(z) \]

The LPC method is based on source filter model. In this model the source signal is produced by the oscillation of the vocal fold and it is modified by resonance determined by the morphology of vocal tract and vocal cavities acting as a filter.

**IMPLEMENTATION**

To implement the linear predictive coding a Simulink [1] model is built. The simulink model
consists of two parts analysis and synthesis. The simulink model of LPC analysis is as shown in following figure.

**Pre-Emphasis Filter**

The pre-emphasis filter is used to boost the high frequency component of speech which is lost during the speech production. The speech is undergone a spectral tilt of -6db/oct. The pre-emphasis filter of following form is used:

\[ y(n) = x(n) - a . x(n-1) \]

**Framing**

The input speech signal is divided into frame of N samples. The adjacent frames are separated from each other. The starting frame consists of the first N samples of speech. The second frame begins M samples after the first samples. Similarly the third frame begins 2M samples after the first sample(or M samples after the second frame) and overlaps it by N-2M samples. This process is continues until all speech is accounted for within one or more frame. Thus, each frame overlaps with two other subsequent frames. This method is called framing of signal.

**Windowing**

To minimize the signal discontinuities at the beginning and end of each signal each individual frame is window using hamming window. The hamming window is defined by following equation:

\[ w(n) = 0.54 - 0.46 \cos \left( \frac{2\pi n}{N-1} \right), \quad 0 < n < N - 1 \]

where N is number of samples in each frame.

**Autocorrelation**

The encoder attempt to minimize the mean squared error in order to find the filter coefficient that best match the current segment being analysed. For this minimization of error criteria leads to following equation:

\[ \sum_{k=1}^{p} r_{|i-k|} a_k = r_i \quad 1 \leq i \leq p \]

where \( r_k \) is the kth autocorrelation coefficient of the windowed speech signal and is given by

\[ r_k = \frac{1}{N} \sum_{n=k}^{N} w_n s_n w_{n-k} s_{n-k} \]

where \((w)\) is windowed function of duration N samples. This equation is also called Yule-Walker equation. To obtain \( p \) LPC coefficient this equation is used. In matrix form:

\[ R_a = -r \]

\[ R = \begin{bmatrix} r_0 & r_1 & r_2 & \ldots & r_{p-1} \\ r_1 & r_0 & r_1 & \ldots & r_{p-2} \\ r_2 & r_3 & r_0 & \ldots & r_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & r_{p-3} & \ldots & r_0 \end{bmatrix} \]

\[ a = [a_1 \quad a_2 \quad a_3 \ldots \quad a_p]^T \]

\[ r = [r_1 \quad r_2 \quad r_3 \ldots \quad r_p]^T \]

The matrix \( R \) is called autocorrelation matrix.

**Levinson Durbin Algorithm**

In order to determine the LPC coefficient the equation \( R_a = -r \) must be solved. This can be done by first computing the matrix \( R \). the computation becomes easy when \( R \) is symmetric and diagonal consists of same element. This type of matrix is called Toeplitz matrix and can be easily inverted. The L-D
algorithm [1] [2] is recursive and it take the advantage of properties of $R$. When determining filter coefficient the filter order is denoted with a superscript $\{a_{ij}\}$ for $j$th order filter. The average mean squared error of $j$th order filter is denoted by $E_j$ instead of $E[en]$. When applied to $M^{th}$ order filter L-D algorithm computes all filters of order $N$ where $N = 1, \ldots, M-1$. Following steps are involved in L-D algorithm:

1. Set $E_0 = R_{yy}, i = 0$

   while ($i < M$)

2. $i++$

3. Calculate $k_i = \left[ \sum a_j (i-1) R_{yy}(i-j+1) - R_{yy}(i) \right] / E_i - 1$

4. Set $a_j (i) = k_i$

5. Calculate $a_j (i) = a_j (i-1) + k_i a_i - a_i (i-1), v_j = 1 \ldots i - 1$

6. Calculate $E_i = (1 - k_i^2) E_i - 1$

$R_{yy}$ denotes the estimate autocorrelation function of speech sample. During the process of computing the filter coefficient $\{a_i\}$, a set of coefficient $\{k_i\}$ is generated which is known as reflection coefficient.

**Time Varying Analysis Filter**

Time varying analysis filter is all zero filter through which the pre-emphasis voice is passed to create the residual or excitation signal. The bandwidth of residual signal is less than original signal.

**LPC SYNTHESIS**

The LPC synthesis process is inverse of LPC analysis process. Each segment is decoded individually and sequence of reproduced sound segments are joined together to represent the entire input speech signal. The following figure shows the simulink model of LPC synthesis.

**Time Varying Synthesis Filter**

Time varying synthesis filter used to recreate the original input signal based on set of coefficient. During encoding these coefficient is extracted from the original signal and are transmitting to the receiver for use in decoding. Each speech signal has different filter coefficient or parameter that it uses to recreate the original signal.

**De-Emphasis Filter**

De-emphasis filter is inverse of pre-emphasis filter. It is digital filter used to restore the high frequency components boosted for transmission. The de-emphasis filter is in form of:

$$y(n) = x(n) + a \cdot x(n-1)$$

**RESULTS AND DISCUSSION**

For LPC analysis and synthesis of speech signal firstly, a speech signal is taken in `.wav`
format whose original sampling frequency is 8 kHz. The LPC order of prediction is 15. For the input male speech signal uttering ‘kurukshetra’ the residual coefficient, reflection coefficient and LPC spectrum as shown in following figures:

**CONCLUSION**

The LPC is most powerful technique of features extraction of input speech signal. In this paper the LPC analysis and synthesis of speech signal is briefly explained. The LPC is simulated and implemented in real time. The reconstructed LPC signal is more spoken and less whispered. The output reconstructed signal is much easier to understand.

**REFERENCES**


