The International Journal Of Engineering And Science (IJES)
||Volume||2 ||Issue|| 4 ||Pages|| 01-07||2013||
ISSN(e): 2319 – 1813 ISSN(p): 2319 – 1805

# Linear Prediction Analysis

[1,]Sai Bharadwaj B [2,] U.Vagdevi, Yadav Bineetha

[1,2,]*Department of electronics and communications,Vignan's Institute of Engineering for Women*

**Abstract**

*Most signals that we encounter in nature are random. A random signal has a predictable component and an unpredictable component. The aim of linear prediction is to use a linear model to model the mechanism that generates correlation among signal samples. In this project, we will use linear prediction to analyze the sound files(A bird chirping, A dog barking, The Girl speaking in English, Male speaking in Chinese) and to compare the predictability of sounds from animals and different languages . Levinson-Durbin algorithm also be applied as to implement the Linear Prediction. This project aims at Implementation of the Levinson-Durbin algorithm to analyze the four sound files, use a frame size of 20 msec to obtain the required result.*

***Keywords***: *Random signal, Linear prediction, Levinson-Durbin Algorithm*

Date Of Submission: 01 April 2013 ⟷ Date Of Publication: 30, April.2013

## I. INTRODUCTION

Speech coding is the operation of transforming the speech signal to a more compact form, which needs smaller bit-rate for transmitting over a communication channel. Generally, this transformation is lossy which means that the signal, reconstructed from the compact form is not exactly the same as the original signal.Considering the dynamics and limitations of the human ear, one can develop some compression schemes such that from the viewpoint of listener, the reconstructed signal is close enough to the original signal [1].

There are many coding techniques such as ADPCM, Adaptive Differential Pulse Code Modulation, is one of the most widely used speech compression. It falls into the category of waveform coding, which attempts to preserve the waveform of the original speech signal and achieves its compression improvements by taking advantage of the high correlation exhibited by successive speech samples. CELP, Code Excited Linear Prediction Code is one of the most effective coding methods at low bit-rates, which was proposed in the mid-eighties. This algorithm models speech by transmitting only parameters that are necessary to synthesize speech that is perceived to sound like the original. The basic idea of LPC is to transmit the prediction error (residue) instead of the speech signal. Since a linear predictor with properly chosen order can predict the signal with relatively small error variance, the power of residuals is effectively smaller than the power of the original signal.

This property of linear prediction enables us to use lower bit rate for transmitting the speech signal through a communication channel.[3]The traditional speech coding techniques like PCM will not compress the signal and ADPCM may not transmit signal at better quality. The extension coding techniques like CELP may transmit the signal with high quality but the rate of compression is less. Hence the proposed system is Linear Prediction of speech signal which results in best compression with better quality when reconstructed.[2]The basic idea of LPC is to transmit the prediction error (residue) instead of the speech signal. Since a linear predictor with properly chosen order can predict the signal with relatively small error variance, the power of residuals is effectively smaller than the power of the original signal. This property of linear prediction enables us to use lower bit rate for transmitting the speech signal through a communication channel.

## II. THEORITICAL ANALYSIS

Analysis of speech signals is made to obtain the spectral information of the speech signal. Analysis of speech signal is employed in variety of systems like voice recognition system and digital speech coding system. Accepted methods of analyzing the speech signals make use of linear predictive coding (LPC). Linear prediction is a good tool for the analysis of speech signals.

In linear prediction the human vocal tract is modeled as an infinite impulse response system for producing the speech signal. In LPC the current sample of a speech signal is estimated by the linear combination of a series of weighted past samples of the speech signal. The series of weights or coefficients represent the LPC coefficients which are used as filter coefficients in encoding and decoding process during coding. An alternate explanation is that linear prediction filters attempt to predict future values of the input signal based on past signals. LPC models speech as an autoregressive process, and sends the parameters of the process as opposed to sending the speech itself [1].

### 2.1. Choosing Prediction Order

Linear predictive coding is a time domain technique that models the speech signal as a linear combination of the weighted delayed past speech sample values. LPC order is an important parameter used in linear prediction, which will affect the quality of synthesized speech signal as the order determines how many number of weighted past samples are to be used to determine the current speech sample. The choice of prediction order should not be too low since key areas of resonance will be missed as there are insufficient poles to model them and if the prediction order is too high, source specific characteristics, e.g. harmonics, are determined, so there is need to select a optimum prediction order, by which the error is small enough and computation time is also reasonable.[3,5]Taking each frame as about 20ms, we can divide the signal of the sound wave into around 200 frames. This implies that we need to determine the prediction order and PARCO coefficients for each frame. We can also estimate that the characteristics of the signal depend on the particular frame that we choose for analysis. In such a case, it is difficult for us to analyze all these values to determine the predictabilities of these 4 sounds. As such, we have to know which one to choose so as to determine the predictability of these 4 input sound files. It is also known that if we were to choose just one frame out of each signal, the variance and error could be relatively high.

We basically go for a tradeoff between accuracy and computation speed. Therefore, in this project we choose the mean values of PARCO and ERROR coefficients of all the frames and at the same prediction order, and then use the defined criteria to determine the optimum prediction order for each sound file. As all four sounds are easily audible to humans, we have here used the maximum prediction order to be in the vicinity of the common linear prediction order that is recommended by researchers for human speech, i.e.; 10. So, we use a maximum prediction order of 30 for these four sound files. So we choose 30 as the upper boundary of the prediction order. Within 30, the optimum order will be determined when the change in the Mean Square Error satisfies the following criteria: The current order of Mean Square Error to the later one is less than 1% of the current MSE. [5].

i.e. $$\frac{\xi_f^{(m)} - \xi_f^{(m+1)}}{\xi_f^{(m)}} \leq 1\%$$

Where $\xi$ is mean square error order.

### 2.2. Linear predictive modeling of speech signals

Linear prediction is a technique of time series analysis that emerges from the examination of linear systems. Using linear prediction, the parameters of such system can be determined by analyzing the systems inputs and outputs. This section will review linear systems and, elaborating upon them, derives the mathematics of linear prediction.A linear system is such that produces its output as a linear combination of its current and previous inputs and its previous outputs [6]. It can be described as time-invariant if the system parameters do not change with time. Mathematically, linear time-invariant (LTI) systems can be represented by the following equation:

$$y(n) = \sum_{j=0}^{q} \mathrm{b_j}\, x(n-j) - \sum_{k=1}^{p} \mathrm{a_k}\, x(n-k) \qquad (1)$$

This is the general difference equation for any linear system, with output signal *y* and input signal *x*, and scalars *bj* and *ak*, for j = 0 . . . q and k = 1 . ..p where the maximum of p and q is the order of the system. By re-arranging equation (1) and transforming into the Z-domain, we can reveal the transfer function H (z) of such a system:

$$\mathrm{y}(n) + \sum_{k=1}^{p} \left(a_k\, \mathrm{y}(\mathrm{n}-\mathrm{k})\right) = \sum_{j=0}^{q} \left(b_j\, \mathrm{x}(\mathrm{n}-\mathrm{j})\right)$$

where $a_0$=1

$$\sum_{k=0}^{p} a_k z^{-k} X(z) = \sum_{j=0}^{q} b_j z^{-j} X(z)$$

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{j=0}^{q} b_j z^{-j} X(z)}{\sum_{k=0}^{p} a_k z^{-k} X(z)} \qquad (2)$$

The coefficients of the input and output signal samples in equation (1) reveal the poles and zeros of the transfer function. Linear prediction follows naturally from the general mathematics of linear systems. As the system output is defined as a linear combination of past samples, the system's future output can be predicted if the scaling coefficients *bj* and *ak* are known. These scalars are thus also known as the predictor coefficients of the system.The general linear system transfer function gives rise to three different types of linear model [5], dependent on the form of the transfer function H(z) given in equation (2) .

- When the numerator of the transfer function is constant, an all pole model is defined.
- When the denominator of transfer function is constant, an all zero model is defined.
- The third and most general case is the mixed pole model, where nothing is assumed about the transfer function.

The all pole model for linear prediction is the most widely studied and implemented of the three approaches, for a number of reasons. Firstly, the equations derived from the all pole model approach are relatively straight-forward to solve, where as the equations derived from all-zero modeling are non-linear equations. Finally, and perhaps the most important reason why all pole modeling is the preferred choice of engineers, many real world applications, including most types of speech production, can be faithfully modeled using the approach.Following from the linear system equation (1), one can formulate the equations necessary to determine the parameters of an all-pole linear system, the so-called linear prediction normal equations. First, following on from the all-pole model ỹ a linear prediction estimate at sample number n for the output signal y by a p[th] order prediction filter can be given by: (3)

$$\hat{y} = -\sum_{k=1}^{p} a_k\, y(n-k)$$

As in any practical system there will be slight difference between the actual output and the estimated output which nothing but error. The error or residue between the output signal and its estimate at sample n can then be expressed as the difference between the two signals.

$$e(n) = y(n) - \hat{y}(n) \qquad (4)$$

The primary objective of LP analysis is to compute the LP coefficients which minimizes the prediction error e(n). The popular method for computing the LP coefficients is by least squares auto correlation method. This is achieved by minimizing the total prediction error. The total prediction error can be represented as follows,

$$E = \sum_{n} [e(n)]^2 \qquad (5)$$

$$= \sum_{n} [y(n) - \hat{y}(n)]^2$$

$$= \sum_{n} [y(n)]^2 - 2y(n).\hat{y}(n) + [\hat{y}(n)]^2$$

Equation (5) gives a value indicative of the energy in the error signal. Obviously, it is desirous to choose the prediction coefficients so that the value of E is minimized over the unspecified interval. The optimal minimizing values can be determined through differential calculus, i.e. by obtaining the derivative of equation 5 with respect to each predictor coefficient and setting that value equal to zero.

$$\frac{\partial E}{\partial a_k} = 0 \qquad for\ 1 \le k \le p$$

$$=> \frac{\partial E}{\partial a_k}\left(\sum_n [y(n)]^2 - 2y(n).\hat{y}(n) + \hat{y}(n)^2\right) = 0$$

$$-2\sum_n y(n)\frac{\partial \hat{y}(n)}{\partial a_k} + 2\sum_n \hat{y}(n)\frac{\partial \hat{y}(n)}{\partial a_k} = 0$$

$$\sum_n y(n)\frac{\partial \hat{y}(n)}{\partial a_k} = \sum_n \hat{y}(n)\frac{\partial \hat{y}(n)}{\partial a_k}$$

$$\frac{\partial \hat{y}(n)}{\partial a_k} = -y(n-k) \qquad from\ eq(3)$$

$$=> \sum_n y(n).(-y(n-k)) = \sum_n \hat{y}(n).(-y(n-k))$$

$$-\sum_n y(n).y(n-k) = \sum_n \left(\sum_{i=1}^{p} a_i\, y(n-i).(-y(n-k))\right)$$

$$-\sum_n y(n).y(n-k) = \sum_{i=1}^{p} a_i \sum_n y(n-i).y(n-k)$$

For the sake of brevity and future utility, a correlation function Ø is defined. The expansion of this summation describes what will be called the correlation matrix.

$$\emptyset(i,k) = \sum_n y(n-i).y(n-k) \tag{6}$$

Substituting the correlation function into equation (6) allows it to be written more compactly:

$$-\emptyset(0,k) = \sum_{i=1}^{p} a_i\ \emptyset(i,k) \tag{7}$$

The derived set of equations is called the normal equations of linear prediction. The solutions to the normal equations are done by two methods namely the autocorrelation method and the covariance method[5].

### 2.3. Determination of LPC Coefficients

Two approaches can obtain the LPC coefficients $a_k$ characterizing an all-pole H(z) model. The two approaches are either by least square autocorrelation method or by covariance method. A question may arise as to whether to use the autocorrelation method or the covariance method in estimating the predictor parameters. The covariance method is quite general and can be used with no restrictions. The only problem is that of stability of the resulting filter. In the autocorrelation method on the other hand, the filter is guaranteed to be stable, but problems of the parameter accuracy can arise because of the necessity of the windowing (truncating) the time signal. This is usually a problem if the signal is a portion of an impulse response.[3] For example, if the impulse response of an all-pole filter is analyzed by covariance method, the filter parameters can be computed accurately from only a finite number of samples of the signal. Using the autocorrelation method, one cannot obtain the exact parameters values unless the whole infinite impulse response is used in the analysis. However, in practice, very good approximations can be obtained by truncating the impulse response at a point where most of the decay of the response has already occurred. In this thesis the LPC coefficients are determined using the autocorrelation method because of the greater redundancies in the matrix defined by the autocorrelation method, it is slightly easier to compute and experimental evidence indicates that the covariance method is more accurate for periodic speech sounds, while the autocorrelation method performs better for fricative sounds.[2] If the signal is non-zero from 0 to N −1, then the resulting error signal will be non-zero from 0 to N − 1 + p. Windowing a signal by multiplication with an appropriate function is shown in fig 2.1.

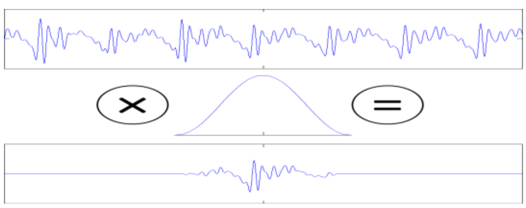


Fig 2.1 Autocorrelation method

Thus, summing the total energy over this interval is mathematically equivalent to summing over all time.

$$E = \sum_{-\infty}^{\infty}[e(n)]^2$$
$$= \sum_{n=0}^{N-1+p} [e(n)]^2 \qquad (8)$$

When these limits are applied to equation (7), a useful property emerges. Because the error signal is zero outside the analysis interval, the correlation function of the normal equations can be identically expressed in a more convenient form.

$$\boldsymbol{\varnothing_{auto}(i,k)} = \sum_{n=0}^{N-1+p} y(n-i).y(n-k) \qquad 1 \le i \le p\,,\ 1 \le k \le p$$
$$= \sum_{n=0}^{N-1+(i-k)} y(n).y(n+(i-k)) \qquad 1 \le i \le p\,,\ 1 \le k \le p \qquad (9)$$

This form of the correlation function is simply the short-time autocorrelation function of the signal, evaluated with a lag of (i − k) samples. This fact gives this method of solving the normal equations its name. The implication of this convenience is such that the correlation matrix defined by the normal equations exhibits a double-symmetry that can exploit by a computer algorithm.

$$\mathbf{R} = \begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_{p-1} \\ r_1 & r_0 & r_1 & \cdots & r_{p-2} \\ r_2 & r_1 & r_0 & \cdots & r_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & r_{p-3} & \cdots & r_0 \end{bmatrix}$$

The above matrix is a symmetric matrix and all the elements along a diagonal are equal i.e., the matrix is a Toeplitz matrix[4]. Equation (9) is solved by taking the inversion of p x p matrix but this result in computational errors. So the solution to equation (9) is to exploit the Toeplitz characteristics and to use efficient recursive procedures. The most widely used recursive procedure is the Levinson-Durbin's recursion.

Assuming we have already found out the autocorrelation functions for the input process x(n), we need to solve the M XM system of equations to get the desired coefficients ($a_k$). One can use standard methods for solving linear equations, but the structure of the R matrix gives us some very unique advantages. The Levinson-Durbin Recursion, named in recognition of its use first by Levinson (1947) and then its independent reformulation at a later date by Durbin (1960), is a direct recursive method for solving for the coefficients of the prediction filter. It makes particular use of the Toeplitz structure of the matrix R.The Levinson-Durbin (L-D) Algorithm is a recursive algorithm that is considered very computationally efficient since it takes advantage of the properties of R when determining the filter coefficients. During the process of computing the filter coefficients ($a_i$) a set of coefficients ($k_i$) called reflection coefficients or partial correlation coefficients (PARCOR) are generated. These coefficients are used to solve potential problems in transmitting the filter coefficients.[5]

The key to the Levinson Durbin method of solution that exploits the Toeplitz property of the matrix is to proceed recursively, beginning with a predictor of the order m=1 (one coefficient) and to increase the order recursively, using the lower order solutions to obtain the solution to the next higher order. Thus the solution to the first order predictor obtained by solving the equation is

$$a_1(1) = -R_{xx}(1)/R_{xx}(0)$$

The next step is to solve for the coefficients $a_2(1)$ and $a_2(2)$ of the second order predictor and express in the following equation.

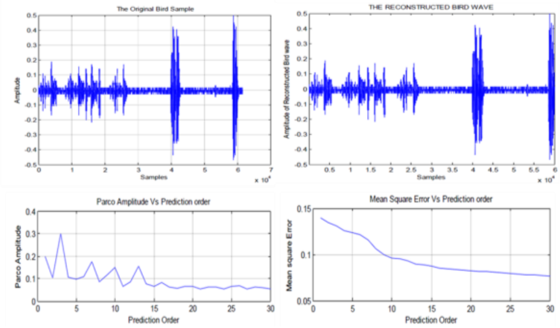
$$a_2(1)R_{xx}(0) + a_2(2)R_{xx}(1) = -R_{xx}(1)$$

The important virtue of the Levinson-Durbin algorithm is its computational efficiency, in that its use results in a big saving in the number of operations. The Levinson-Durbin recursion algorithm requires multiplications and additions to go from stage m to stage m+1. Therefore, for p stages it will take on the order of 1+2+3+…+p = p(p+1)/2 .[6]

## III. FIGURES AND TABLES

In our implementation if we enlarge the results waveforms of the reconstructed wave the Linear dependence of the implemented error signal with the original can be seen to approximately to be same. Thus we can say that the linear prediction Model holds good to predict the waveform.

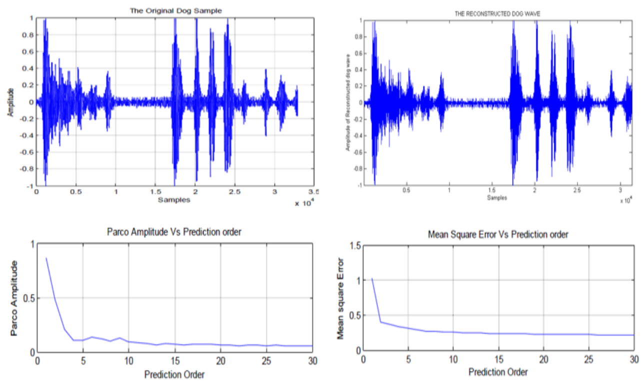
### 3.1 Simulation Graphs of Bird Signal



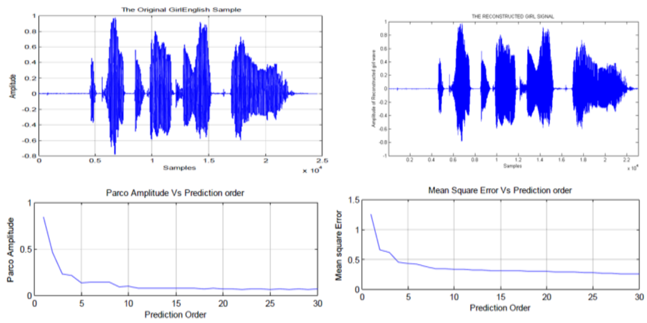

The above figure represents original bird signal and reconstructed signals in which X-axis denotes the number of samples present in the signal and the Y-axis denotes the amplitude of the signal. Here we can clearly observe that with the increase in prediction order, the prediction error decreases. The PARCOR 'K' of the chirping bird decreases with fluctuations with the increase in prediction order. The errors is larger for low correlated signal which is the sound of the bird.

### 3.2 Simulation Graphs of dog Signal

The below figure represents original dog signal and reconstructed signals in which X-axis denotes the number of samples present in the signal and the Y-axis denotes the amplitude of the signal. Here we can clearly observe that with the increase in prediction order, the prediction error decreases. The PARCOR 'K' of the chirping bird decreases with fluctuations with the increase in prediction order. The errors is larger for low correlated signal which is the sound of the dog.

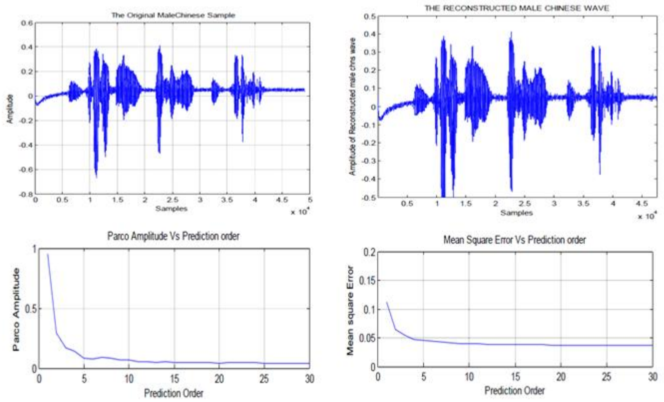



### 3.3 Simulation Graphs of Girl Signal




The above figure represents original girl signal and reconstructed signals in which X-axis denotes the number of samples present in the signal and the Y-axis denotes the amplitude of the signal. Here we can clearly observe that with the increase in prediction order, the prediction error decreases. The PARCOR 'K' of the chirping bird decreases with fluctuations with the increase in prediction order. The error is average for medium correlated signal which is the sound of the girl.

3.3 Simulation Graphs of male Signal



The above figure represents original male signal and reconstructed signals in which X-axis denotes the number of samples present in the signal and the Y-axis denotes the amplitude of the signal. Here we can clearly observe that with the increase in prediction order, the prediction error decreases. The PARCOR 'K' of the chirping bird decreases with fluctuations with the increase in prediction order. The error is low for highly correlated signal which is the sound of the male.

## IV. CONCLUSION

In our implementation if we enlarge the results waveforms of the reconstructed wave the Linear dependence of the implemented error signal with the original can be seen to approximately to be same. Thus we can say that the linear prediction Model holds good to predict the waveform.However, the rates of decreasing of different signals are different. The signal of male speaking Chinese decreases very fast, so it can be better predicted. The PARCOR K of the chirping bird decreases with fluctuating when prediction order increases, the K value of the Chinese speaking male decreases as well but more smoothly and rapidly at the beginning of order increasing. We can estimate that as the prediction order increases, the predict error becomes smaller and smaller, so the k value becomes closer to 0. So we can conclude that, higher the predictability level, the lower the prediction errors as **MALE>DOG>GIRL>BIRD**.

| SIGNAL | PREDICTION ORDER |
|---|---|
| MALE | 10 |
| DOG | 12 |
| GIRL | 14 |
| BIRD | 16 |

Table 4.1 Comparing predictions of different signals

## REFERENCES

[1] Arpana Mishra and Javed Asraf,Speech Compression with Linear Prediction, in IJETAE(International Journal for Emerging Technology and Advanced Engineering) vol.2,issue 6,june2012.
[2] A. Krishnamurthy and D.Childers,Two channel speech analysis and signal processing in IEEE Transactions on signal processing vol.4 issue 34, 1986.
[3] J. Makhoul. Linear prediction: A tutorial review. Proceedings of the IEEE, 63(4) 561–580, 1975.
[4] A.V. Oppenheim R. W. Schafer. Digital Signal Processing. Prentice Hall, Englewood Cliffs, New Jersey, 1975
[5] M. H. Hayes, Statistical Digital Signal Processing and Modeling, John Wiley & Sons, Inc., 1996.
[6] T. W. Parsons. Voice and speech processing. McGraw-Hill New York, 1987.