CS571 PROJECT REPORT

NAME: RAHUL KUMAR
RALL NO: T21020

1 : SUMMARY

• The LP model has many interesting and useful properties

that follow from the structure of the Levinson-Durbin algorithms

• The different equivalent representations have different

properties under quantization

– polynomial coefficients (bad)

– polynomial roots (okay)

– PARCOR coefficients (okay)

– lossless tube areas (good)

– LSP root angles (good)

• Almost all LPC representations can be used with a range

of compression schemes and are all good candidates for

the technique of Vector Quantization

- Use Pth order linear predictor to predict (n^) from previous samples

- Minimize mean-squared error, E , over analysis window of duration L-samples
- Solution for optimum predictor coefficients, {αk}, is based on solving a matrix equation

=> Two solutions have evolved

1 Autocorrelation Method : Signal is windowed by a tapering window in order to minimize discontinuities at beginning (predicting speech from zero-valued samples)and end (predicting zero-valued samples from speech samples) of the interval; the matrix φ (i, k) is shown to be an autocorrelation function; the resulting autocorrelation matrix is Toeplitz and can be readily solved using standard matrix solutions

2 covariance method => The signal is extended by *p* samples outside the normal range
Of $0 \le m \le L-1$ to include P samples occurring prior to m=0; this eliminates large errors in computing the signal from values prior to m=0 (they are available) and eliminates the need for a tapering window; resulting matrix of correlations is symmetric but not Toeplitz

2 : INTRODUCTION

Linear prediction is one of the most important tool in speech analysis. The philosophy behind linear prediction is that a speech sample can be approximated as linear combination of past sample.

- LPC methods are the most widely used in speech coding, speech synthesis, speech recognition, speaker recognition and verification and for speech storage
  - LPC methods provide extremely accurate estimates of speech parameters, and does it extremely efficiently
  - basic idea of Linear Prediction: current speech sample can be closely approximated as a linear combination of past samples, i.e.,

$$s(n) = \sum_{k=1}^{p} \alpha_k s(n-k) \text{ for some value of } p, \alpha_k\text{'s}$$

3. SOLUTION

Solution for $\{\alpha_k\}$

- short-time average prediction squared-error is defined as

$$E_{\hat{n}} = \sum_m e_{\hat{n}}^2(m) = \sum_m \left(s_{\hat{n}}(m) - \tilde{s}_{\hat{n}}(m) \right)^2$$

$$= \sum_m \left(s_{\hat{n}}(m) - \sum_{k=1}^{p} \alpha_k s_{\hat{n}}(m-k) \right)^2$$

- select segment of speech $s_{\hat{n}}(m) = s(m+\hat{n})$ in the vicinity of sample $\hat{n}$
- the key issue to resolve is the range of m for summation (to be discussed later)

Solution for $\{\alpha_k\}$

- can find values of α_k that minimize $E_{\hat{n}}$ by setting:

$$\frac{\partial E_{\hat{n}}}{\partial \alpha_i} = 0, \qquad i = 1, 2, \ldots, p$$

- giving the set of equations

$$-2 \sum_m s_{\hat{n}}(m-i) [s_{\hat{n}}(m) - \sum_{k=1}^{p} \hat{\alpha}_k s_{\hat{n}}(m-k)] = 0, \quad 1 \le i \le p$$

$$-2 \sum_m s_{\hat{n}}(m-i)\, e_{\hat{n}}(m) = 0, \quad 1 \le i \le p$$

- where $\hat{\alpha}_k$ are the values of α_k that minimize $E_{\hat{n}}$ (from now on just use α_k rather than $\hat{\alpha}_k$ for the optimum values)
- prediction error ($e_{\hat{n}}(m)$) is orthogonal to signal ($s_{\hat{n}}(m-i)$) for delays (i) of 1 to p

3.1 : ASSUMPTION

- assumptions about excitation to solve for G
  - ***voiced speech***--$u(n) = \delta(n) \Rightarrow$ L order of a single pitch period; predictor order, p, large enough to model glottal pulse shape, vocal tract IR, and radiation
  - ***unvoiced speech***--$u(n)$-zero mean, unity variance, stationary white noise process

3.2 : ALGORITHM

Levinson-Durbin Algorithm 1

- Autocorrelation equations (at each frame $\hat{n}$):

$$\sum_{k=1}^{p} \alpha_k R[|i-k|] = R[i] \quad 1 \le i \le p$$

Rα = r

- R is a positive definite symmetric Toeplitz matrix
- The set of optimum predictor coefficients satisfy:

$$R[i] - \sum_{k=1}^{p} \alpha_k R[|i-k|] = 0, \quad 1 \le i \le p$$

- with minimum mean-squared prediction error of:

$$R[0] - \sum_{k=1}^{p} \alpha_k R[k] = E^{(p)}$$

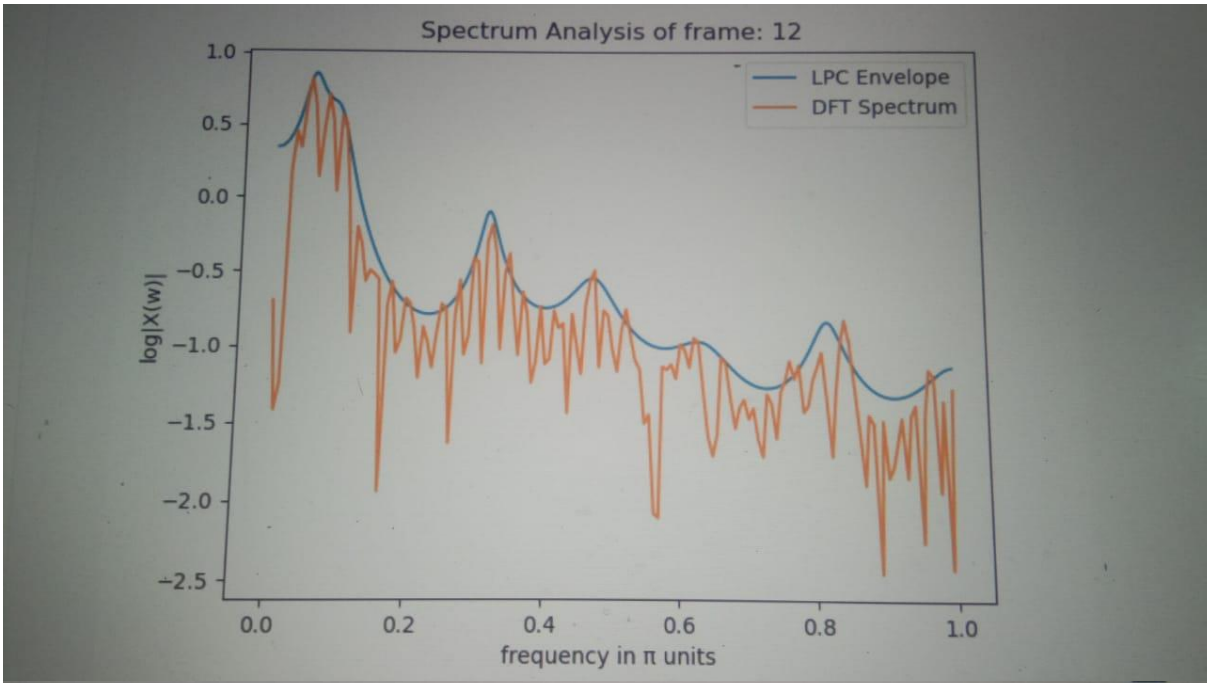
Levinson-Durbin Algorithm 2

- By combining the last two equations we get a larger matrix equation of the form:

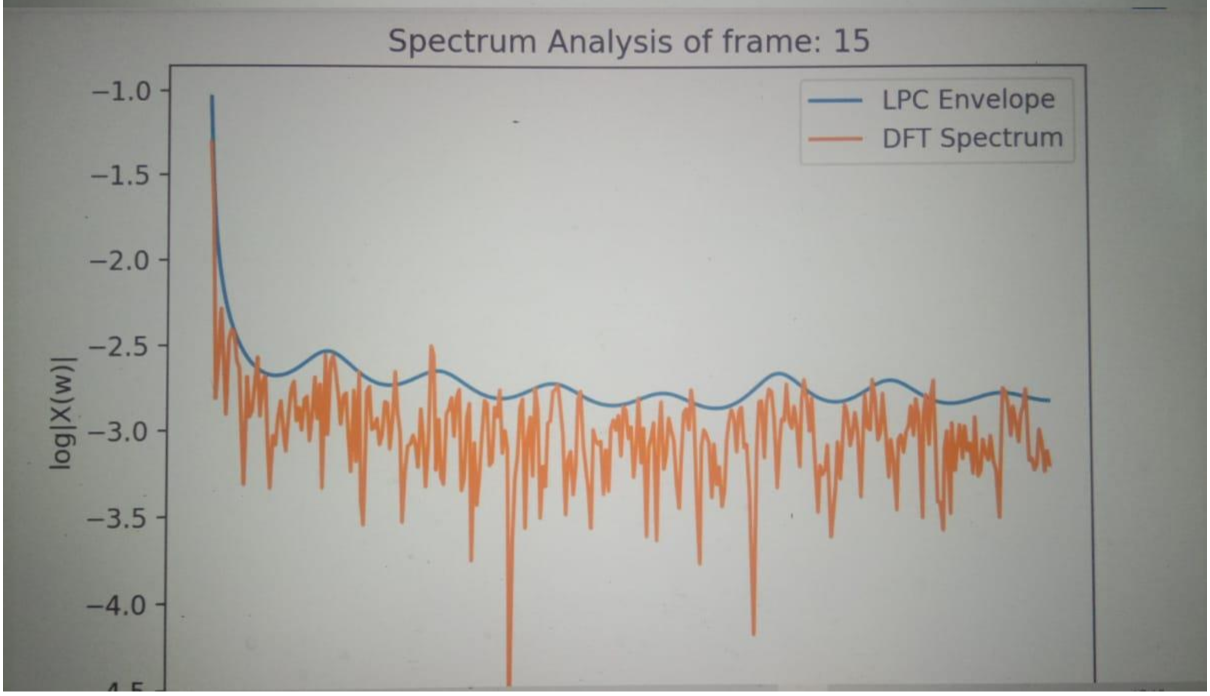
$$\begin{bmatrix} R[0] & R[1] & R[2] & \dots & R[p] \\ R[1] & R[0] & R[1] & \dots & R[p-1] \\ R[2] & R[1] & R[0] & \dots & R[p-2] \\ . & . & . & . & . \\ R[p] & R[p-1] & R[p-2] & \dots & R[0] \end{bmatrix} \begin{bmatrix} 1 \\ -\alpha_1^{(p)} \\ -\alpha_2^{(p)} \\ . \\ -\alpha_p^{(p)} \end{bmatrix} = \begin{bmatrix} E^{(p)} \\ 0 \\ 0 \\ . \\ 0 \end{bmatrix}$$

- expanded $(p+1)$x$(p+1)$ matrix is still Toeplitz and can be solved iteratively by incorporating new correlation value at each iteration and solving for next higher order predictor in terms of new correlation value and previous predictor

4 : RESULT



Spectrum Analysis of frame: 12
LPC Envelope
DFT Spectrum
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
log|X(w)|
0.0
0.2
0.4
0.6
0.8
1.0
frequency in π units



Spectrum Analysis of frame: 15
LPC Envelope
DFT Spectrum
−1.0
−1.5
−2.0
−2.5
−3.0
−3.5
−4.0
log|X(w)|

ANALYSIS

100 A Linear Prediction Analysis of Speech

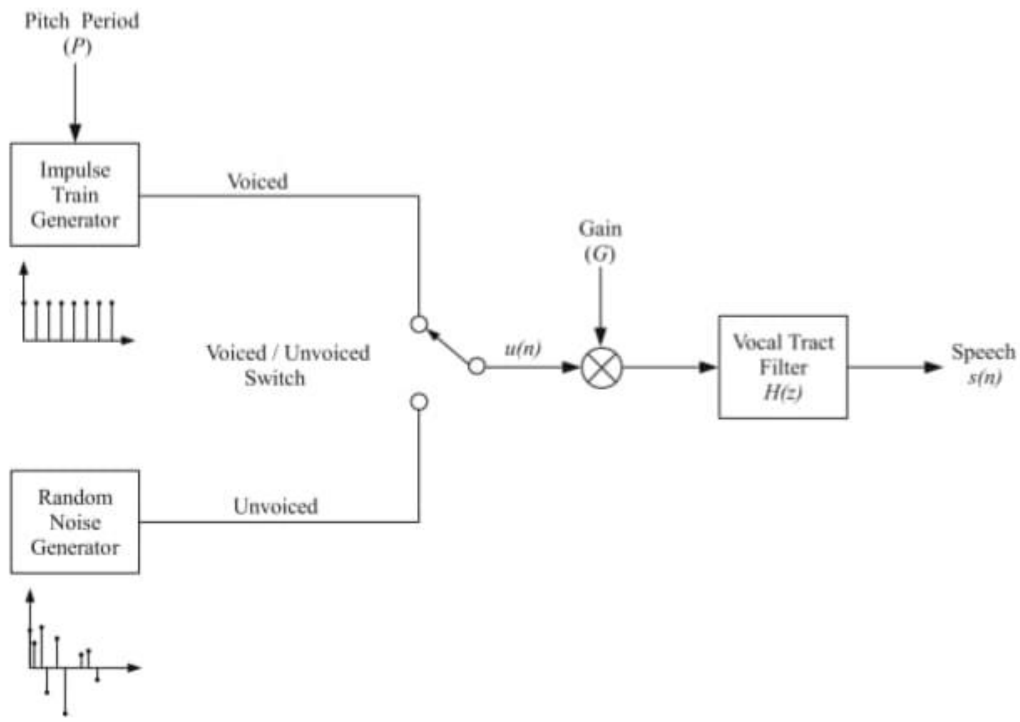


Fig. A.1 Model of speech production for LP analysis

5 : CONCLSION

The implementation of speech comperession technique using linear prediction coding. The implementation used the DSP ststem toolbox functionality available at the camand line. The code involves only calling of the successive system objects with appropriate input arguments

6 : PROJECT GITHUB PAGE :

https://github.com/t21020/LPC-PROJECT.git