

Performance Analysis of Optimization Tool for Speech Recognition Using LPC & DSK TMS3206711/13 Using Simulink & Matlab

¹Kadam V.K, ²Dr.R.C.Thool

¹ Associate Professor, Department of Electronics, P.E.S College of Engineering, Nagsenvan, Aurangabad-431002.

²Professor & Head, Department of Information Technology, SGGGS Institute of Engineering Technology, Nanded 431606.

Abstract:

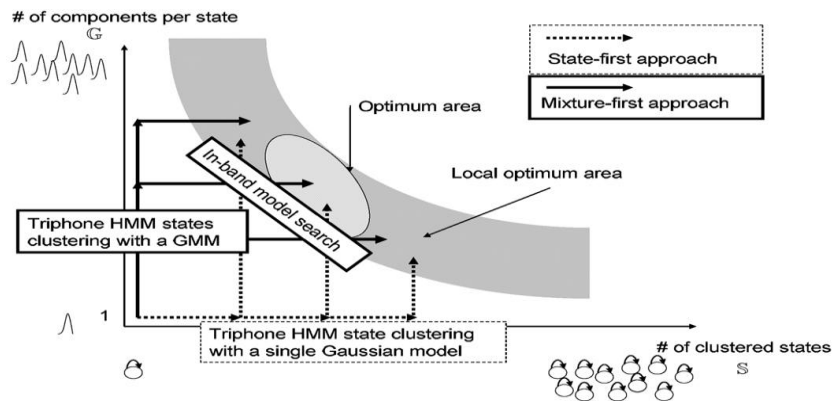
In this we have used various analysis tools like impulse response, frequency response, phase response etc. to get the performance analysis of said system. Here we implement a speech compression technique known as Linear Prediction Coding (LPC) using DSP System Toolbox from Matlab functionality available at the MATLAB command line. In this system we use the Levinson-Durbin and Time-Varying Lattice Filter blocks for low-bandwidth transmission of speech using linear predictive coding. We have changed the various order of the filter for various speeches signal & optimized that. We have received which is reflected in the figures given below.

Keywords: Introduction, System Implementation, Performance Analysis, Figures and Graphs, Conclusion, References, Lpc

1. Introduction

A number of noisy speech enhancement algorithms are experimentally compared in terms of linear predictive coding (LPC) perturbations. The enhancement algorithms considered are simple spectral subtraction, spectral over subtraction with use of a spectral floor, spectral subtraction with residual noise removal, and time-domain and frequency-domain adaptive minimum mean-square-error filtering. LPC perturbations considered are LPC cepstral distance, log likelihood r Linear Predictive Coding (LPC) has been used to compress and encode speech signals for digital transmission at a low bit rate. LPC determines a FIR system that predicts a speech sample from the past samples by minimizing the squared error between the actual occurrence and the estimated. The coefficients of the FIR system are encoded and sent. At the receiving end, the inverse system called AR model is excited by a random signal to reproduce the encoded speech. The use of LPC can be extended to speech recognition since the FIR coefficients are the condensed information of a speech signal of typically 10ms -30ms. PARCOR parameter associated with LPC that represents a vocal tract model based on a lattice filter structure is considered for speech recognition. The use of FIR coefficients and the frequency response of AR model were previously investigated. [1] In this we have taken the method to detect a limited number of phonemes from a continuous stream of speech. A system being developed slides a time window of 16 ms and calculates the PARCOR parameters continuously, feeding them to a classifier. A classifier is a supervised classifier that requires training. The classifier uses the Maximum Likelihood Decision Rule. The training uses TIMIT speech database, which contains the recordings of 20 speakers of 8 major dialects of American English. The classification results of some typical vowel and consonant phonemes segmented from the continuous speech are listed. The vowel and consonant correct classification rate are 65.22% and 93.51%. Overall, they indicate that the PARCOR parameters have the potential capability to characterize the phonemes. atios, and weighted likelihood ratio. [2] A communication system was built and tested to operate in the land mobile VHF band (150-174 MHz) at a channel separation of only 6 kHz. The audio source was digitally encoded at 2.4 kbits/s using linear predictive coding (LPC). The speech data stream was transmitted by frequency shift keying (FSK) which allowed the use of class-C transmitters and discriminator detection in the receiver. Baseband filtering of the NRZ data resulted in a narrow transmitter spectrum. The receiver had a 3 dB bandwidth of 2.4 kHz which allowed data transmission with minimal intersymbol interference and frequency offset degradation. A 58 percent eye opening was found. Bit error rate (BER) performance was measured with simulated Rayleigh fading at typical 150 MHz rates. Additional tests included capture, ignition noise susceptibility, adjacent channel protection, degradation from frequency offset, and bit error effects upon speech quality. A field test was conducted to compare the speech quality of the digital radio to that of a conventional 5 kHz deviation FM mobile radio. [3] In this, we try to use some part of propose a speech-model based method using the linear predictive (LP) residual of the speech signal and the maximum-likelihood (ML) estimator proposed in "Blind estimation of reverberation time," (R. Ratnam, J. Acoust. Soc. Amer., 2004) to blindly estimate the reverberation time (RT60). The input speech is passed through a low order linear predictive coding (LPC) filter to obtain the LP residual signal. It is proven that the unbiased autocorrelation function of this LP residual has the required properties to be used as an input to the ML estimator. It is shown that this method can successfully estimate the reverberation time with less data than existing blind methods. Experiments show that the proposed method can produce

better estimates of RT60, even in highly reverberant rooms. This is because the entire input speech data is used in the estimation process. The proposed method is not sensitive to the type of input data (voiced, unvoiced), number of gaps, or window length. In addition, evaluation using white Gaussian noise and recorded babble noise shows that it can estimate RT60 in the presence of (moderate) background noise. [4]



“Figure 1: Optimum model search for an acoustic model”

2. System Implementation

2.1. Introduction

In this paper, a comparative study between two speech coders have been reported, considering their performances in simulation and in real-time. The speech coders taken for study are Linear Predictive Coder (LPC) and Cepstral coder. The simulation models are made on SIMULINK® and the real-time models can be implemented on TMS320C6713® DSK. For simulation, a comparison between synthesized speech signals using both the speech coders is given. For real-time implementation, some important parameters like memory consumption and execution time for these coders have been calculated. [5] In this system we implement LPC analysis and synthesis (LPC coding) of a speech signal. This process consists of two steps; analysis and synthesis. In the analysis section, we extract the reflection coefficients from the signal and use it to compute the residual signal. In the synthesis section, we reconstruct the signal using the residual signal and reflection coefficients. The residual signal and reflection coefficients require less number of bits to code than the original speech signal. The block diagram below shows the system we will implement. In this simulation, the speech signal is divided into frames of size 320 samples, with an overlap of 160 samples. Each frame is windowed using a Hamming window. Twelfth-order autocorrelation coefficients are found, and then the reflection coefficients are calculated from the autocorrelation coefficients using the Levinson-Durbin algorithm. The original speech signal is passed through an analysis filter, which is an all-zero filter with coefficients as the reflection coefficients obtained above. The output of the filter is the residual signal. This residual signal is passed through a synthesis filter which is the inverse of the analysis filter. The output of the synthesis filter is the original optimized signal. [6], [7], [8], [9] The Optimum Reflection Coefficients for the Lattice Forward and Backward Predictors in Section we derived the set of linear equations which provide the predictor coefficients that minimize the mean-square value of the prediction error. In this section we consider the problem of optimizing the reflection coefficients in the lattice predictor and expressing the reflection coefficients in terms of the forward and backward prediction errors. The forward prediction error in the lattice -filter is expressed as the reflection coefficient K , yields the result

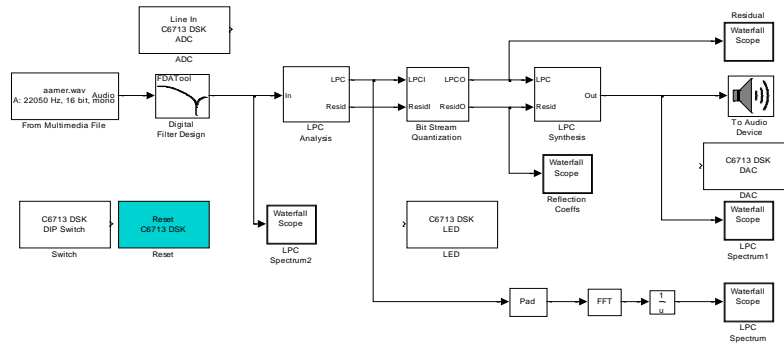
$$K_m = \frac{-E[f_m - 1(n)g^n m - 1(n - 1)]}{E[lgm - 1(n - 1)]^2} \quad (1)$$

We observe that the optimum choice of the reflection coefficients in the lattice predictor is the negative of the (normalized) cross correlation coefficients between the forward and backward errors in the lattice.' Since it is apparent from (1) that K , it follows that the minimum mean-square value of the prediction error, which can be expressed recursively as a monotonically decreasing sequence. [10] Here we initialize some of the variables like the frame size and also instantiate the System objects used in our processing. These objects also pre-compute any necessary variables or tables resulting in efficient processing calls later inside a loop. We create a buffer System object and set its properties such that we get an

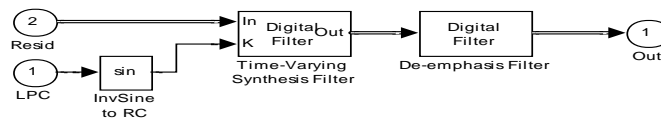
output of twice the length of the frame Size with an overlap length of frame Size. We also create a window System object. Here we will use the default window which is Hamming. By creating an autocorrelator System object and set its properties to compute the lags in the range [0:12] scaled by the length of input. We create a System object which computes the reflection coefficients from auto-correlation function using the Levinson-Durbin recursion. We configure it to output both polynomial coefficients and reflection coefficients. The polynomial coefficients are used to compute and plot the LPC spectrum. By creating an FIR digital filter System object used for analysis. Also create two all-pole digital filter System objects used for synthesis and de-emphasis.

2.2. Stream Processing Loop

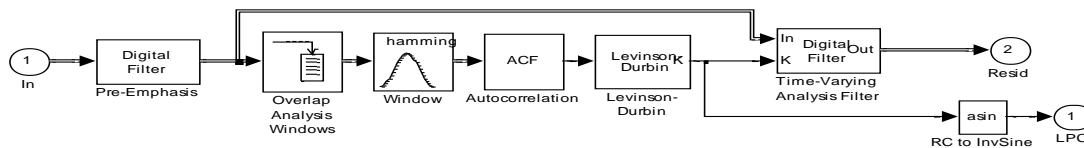
Here we call our processing loop where we do the LPC analysis and synthesis of the input audio signal using the System objects we have instantiated. The loop is stopped when we reach the end of the input file, which is detected by the AudioFileReader System object. Following fig shows the signal & LPC Spectrum.



“Figure 2: Block representation of system implementation using Simulink”



“Figure 3: LPC Analysis”

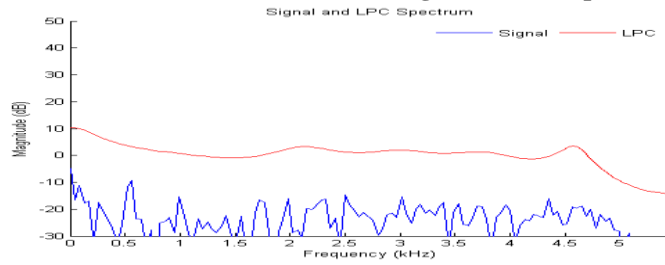


“Figure 4: LPC Synthesis”

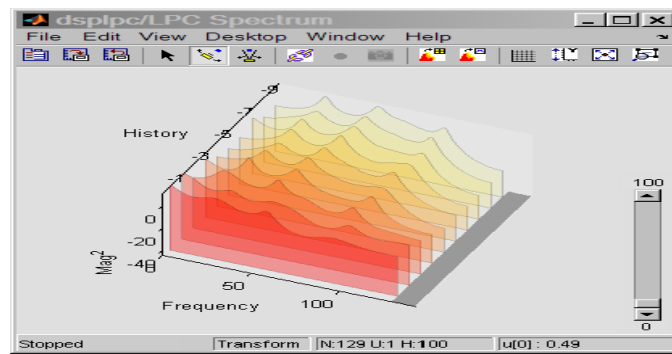
3. Performance Analysis

LPC determines the coefficients of a forward linear predictor by minimizing the prediction error in the least squares sense. It has applications in filter design and speech coding. $[a,g] = \text{lpc}(x,p)$ finds the coefficients of a p th-order linear predictor (FIR filter) that predicts the current value of the real-valued time series x based on past samples. p is the order of the prediction filter polynomial, $a = [1 a(2) \dots a(p+1)]$. If p is unspecified, lpc uses as a default $p = \text{length}(x)-1$. If x is a matrix containing a separate signal in each column, lpc returns a model estimate for each column in the rows of matrix a and a column vector of prediction error variances g . The length of p must be less than or equal to the length of x . Algorithms for lpc uses the autocorrelation method of autoregressive (AR) modeling to find the filter coefficients. The generated filter might not model the process exactly even if the data sequence is truly an AR process of the correct order. This is because the autocorrelation method implicitly windows the data, that is, it assumes that signal samples beyond the length of x are 0. [11]

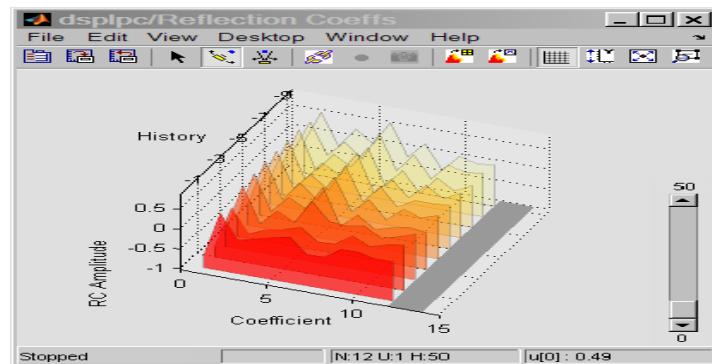
4. Figures and Graphs



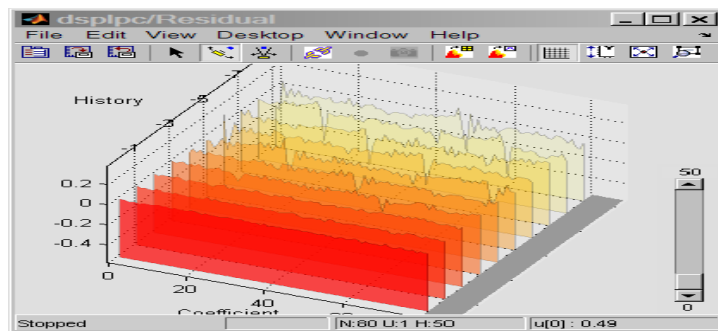
“Figure 5: Graph shows the the signal & LPC Signal”



“Figure 6: LPC Spectrum of signal”

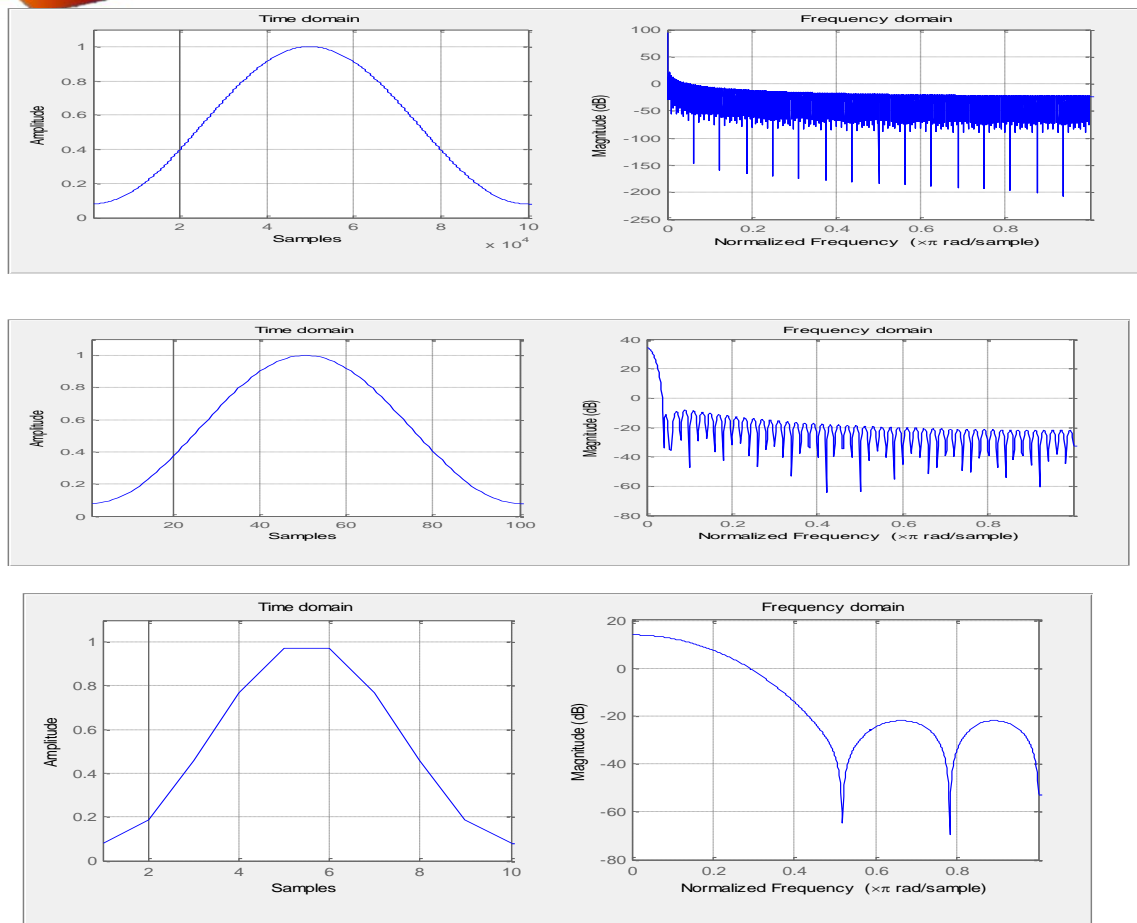


“Figure 7: Reflection coefficients of signal”



“Figure 8: Residual of Signal”

Following figures shows the analysis of the system for various orders of filters & windows



“Figure 9: Time & Frequency representation”

5. Conclusion

This analysis gives the performance of the system implemented using the DSK TMS3206711/13 for various values of the order of filters & various windows. We have seen here the implementation of speech compression technique using Linear Prediction Coding. The implementation used the DSP System Toolbox functionality available at the MATLAB command line. The code involves only calling of the successive System objects with appropriate input arguments. This involves no error prone manual state tracking which may be the case for instance for a MATLAB implementation of Buffer. From the performance it is observed that the optimized speech recognition can be archive

References

- [1] Ahmed, M.S. Dept. of Syst. Eng., King Fahd Univ. of Pet. & Miner., Dhahran ,Comparison of noisy speech nhancement algorithms in terms of LPC perturbation, Acoustics, Speech and Signal Processing, IEEE Transactions on Date of Publication: Jan 1989, Volume: 37 , Issue: 1 Page(s): 121 - 125
- [2] Ying Cui; Takaya ,Recognition of Phonemes In a Continuous Speech Stream By Means of PARCOR Parameter In LPCVocoder, K.Electrical and Computer Engineering, 2007. CCECE 2007. Canadian Conference on Digital Object Identifier:10.1109/CCECE.2007.402 Publication Year: 2007, Page(s): 1606 – 1609
- [3] Speech ‘McLaughlin, M.; Linder, D. Carney. S, Design and Test of a Spectrally Efficient Land Mobile Communications

System Using LPC, Selected Areas in Communications, and IEEE Journal on Volume: 2,

Issue: 4 Digital Object

Identifier: 10.1109/JSAC.1984.1146086 Publication Year: 1984, Page(s): 611 – 620

- [4] Keshavarz, A.; Mosayyebpour, S.; Biguesh, M.; Gulliver, T.A.; Esmaili M, Speech-Model Based Accurate Blind Reverberation Time Estimation Using an LPC Filter, Audio, Speech, and Language Processing, IEEE Transactions on
Volume: 20 , Issue: 6 Digital Object Identifier: 10.1109/TASL.2012.2191283 Publication Year: 2012 , Page(s): 1884 – 1893
- [5] Bhattacharya, S.; Singh, S.K.; Abhinav, T, Performance evaluation of LPC and cepstral speechcoder in simulation and in real-time Recent Advances in Information Technology (RAIT), 2012 1st International Conference on Digital Object Identifier:10.1109/RAIT.2012.6194531 Publication Year: 2012, Page(s): 826 - 831
- [6] Fliege, N.J., Multirate Digital Signal Processing (John Wiley and Sons, 1994).
- [7] Mitra, S.K., Digital Signal Processing (McGraw-Hill, 1998).
- [8] Orfanidis, S.J., Introduction to Signal Processing (Prentice-Hall, Inc., 1996).
- [9] Vaidyanathan, P.P., Multirate Systems and Filter Banks (Prentice-Hall, Inc., 1993).
- [10] Proakis, Digital Signal Processing (third edition pp. 863-64).
- [11] Jackson, L.B., Digital *Filters and Signal Processing* (Second Edition, Kluwer Academic Publishers, 1989. pp.255-257).



Kadam V.K¹

Associate Professor & Research Student,
Department of Electronics,
P.E.S College of Engineering
Nagsenvan, Aurangabad-431002 (M.S)
Dr.Babasaheb Ambedkar Marathwada University,
Aurangabad-431002 (MS)



Dr.R.C Thool²

Department of Information Technology
SGGS Institute of Engineering & Technology
Vishnupuri, Nanded 431606 (M.S)
(An autonomous institute set up and 100% funded by
Government of Maharashtra)