International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 6 (2015) pp. 13975-13989

http://www.ripublication.com

Evaluation of Different Auto Regressive Modeling Methods For Source-Filter Model Based Artificial Bandwidth Extension Systems

G.Gandhimathi[1] , Dr.A.Hemalatha [2] Dr.S.P.K.Babu[3]
[1,2,3] *School of Architecture Engineering and Technology. Periyar Maniammai University Thanjavur, India*

Abstract

Artificial Band Extension (ABE) techniques are used to generate a wideband signal from the narrowband signal. Since most of the high frequency components and the fricative consonants were absent even in the narrowband representation of the sound, it is a challenging task to create those missing components in the wideband equivalent signal. Generally, there are different techniques of autoregressive (AR) modeling to find the filter coefficients which is important in source-filter based ABE. In this paper, the performance of five of autoregressive (AR) modeling methods namely 1. Autocorrelation Method (LPC), 2.Levinson-Durbin Recursion Method, 3.Yule-Walker Method, 4.Covariance Method and 5.Burg method were evaluated. Generally, these estimation methods lead to approximately the same results (same coefficients) for a particular autoregressive parameters. But, the small differences in such estimations will have a great impact on the quality of reproduced sound . In this work, we implemented a source-filter model based artificial speech bandwidth extension systems with the above five AR modeling methods and validated their performance with suitable metrics.

Introduction

Problem Specification

To store and keep the Digital speech achieves with best feature and high quality (Wide band speech 50 Hz to 8KHz) more memory is needed . Presently due to the efficient utilization of memory space all hand held devices can produce only low quality sound . We can achieve a best quality speech with less memory space by implementing the technique Artificial Bandwidth Extension (ABWE). The missing low frequency of NB component (> 3.4 KHz to 4 KHz) can be predicted , and the input speech signal’s high frequency component (4KHz to 7KHz) are artificially created at the receiving end . Creation of missing components between 3.4kHz. to 4.6kHz will be the exigent task

In this work we address these issues by designing a source-filter model based Artificial Bandwidth Extension system with different autoregressive (AR) modeling methods and assess its performance with ARCTIC sound database of Carnegie Mellon University. To find the filter coefficients generally, there are five different techniques of AR modeling namely 1.Autocorrelation Method (LPC), 2.Levinson-Durbin Recursion Method, 3.Yule-Walker Method, 4.Covariance Method and 5.Burg method is used in source-filter based ABE systems.

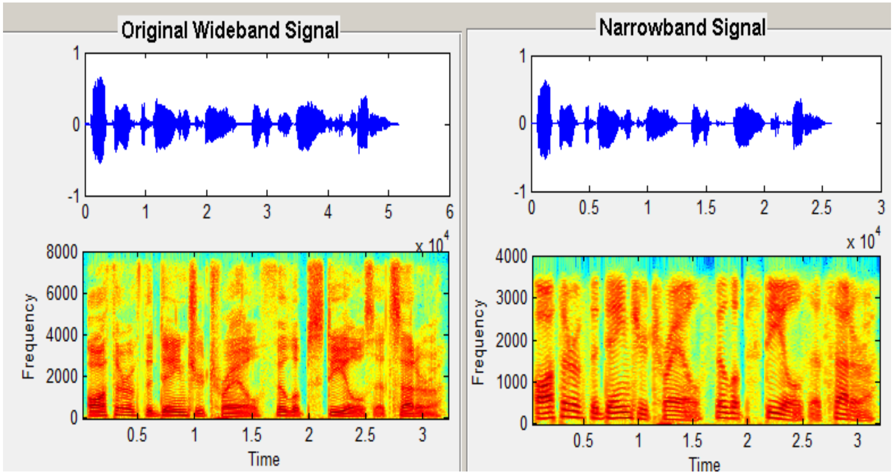


Fig 1: WB and NB speech

In this work, we will implemented a source-filter model based artificial speech bandwidth extension systems with the above five AR modeling methods and validated their performance with suitable metrics.

Previous Works

In our previous work, we designed and discussed about LPC based Source filter model and the ABWE of NB signal is analyzed by both Time Domain Interpolation method (TDI) and Analysis and Synthesis method using code book approach From that we had a conclusion that, for creating the missing components , the filter co-efficient played a dominant role and thus changes the reproduced sound quality . Gandhimathi [7] used different types of Auto recursive models [2012] to find the filter coefficients . In our present work we use all AR to find the potential features of speech and evaluate its performance for BWE application.

This paper organized as follows the section 2 explains basics of autoregressive (AR) modeling and the AR modeling algorithms under evaluation. Section 3 explains the design of ABE system with different AR modeling methods. In section 4 resents the results. Section 5 concludes with the findings.

Autoregressive Modeling

Autoregressive Modeling

Parametric methods can give up higher resolutions than nonparametric methods in cases when the signal length is short. The linear system model commonly used is the *all-pole model,* a filter with all of its zeroes at the origin in the *z*-plane. If the input of the filter is for white noise and an output of such a filter is an autoregressive (AR) process. Because of this reason, these methods are may referred as *AR methods* of spectral estimation. The AR methods tend to passably illustrate spectra of data that is "peaky," that is, data whose PSD is large at certain frequencies. The data in many practical applications (such as speech) tends to have "peaky spectra" so that AR models are may be useful. In addition, the AR models lead to a system of linear equations which is relatively simple to solve.

AR Modeling Procedure

In order to find the a(k)'s, we first construct forward and backward linear predictors are (n-p) to n and n to (n+p).

$$S_f(n) = -\sum_{k=1}^{p} a(k)s(n-k) \qquad \text{--------- (1)}$$

$$S_b(n) = -\sum_{k=1}^{p} a(k)s(n+k) \qquad \text{---------(2)}$$

Before measuring the prediction errors, we shift the indexes for the backward predictor by p as

$$S_b(n-p) = -\sum_{k=1}^{p} a(k)s(n+k-p) \qquad \text{---------} \quad (3)$$

This shifts the backward predictor's s arguments to be '(n-p)' to 'n'. The forward and backward prediction errors are respectively

$$r_f(n) = S(n) - S_f(n) \qquad \text{--------- (4)}$$

$$= \sum_{k=0}^{p} a(k)s(n-k) \qquad \text{--------- (5)}$$

$$r_b(n) = S(n-p) - S_b(n-p) \qquad \text{----------(6)}$$

$$= \sum_{k=0}^{p} a(k)s(n+k-p) \qquad \text{--------- (7)}$$

the forward and backward error energies are

$$V_f(a) = \sum_{n=n1}^{n2} \{rf(n)]^2 \qquad \text{---------} \quad (8)$$

$$V_b(a) = \sum_{n=n1}^{n2} \{rb(n)]^2 \qquad \text{---------} \quad (9)$$

In order to attempt to average out differences between the forward and backward models, we will minimize the combined prediction error function

$$V(a) = V_f(a)+V_b(a)$$

$$= \sum_{n=n1}^{n2} \{rf(n)\}^2 + \{rb(n)\}^2] \qquad \text{------(10)}$$

with respect to the p unknowns a(1) through a(p). There are two important cases, depending on how the limits n1 and n2 are chosen

Different Methods of Spectrum Estimation

Five methods of autoregressive-parameter estimation from these data samples shall be considered here

1. LPC (Autocorrelation Method)

LPC uses the autocorrelation method of 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.

2. Levinson-Durbin Recursion Method

The Levinson-Durbin recursion is an algorithm for finding an all-pole IIR filter with a prescribed deterministic autocorrelation sequence. It has applications in filter design, coding, and spectral estimation. The filter coefficients are ordered in descending powers of z.

3. Yule-Walker Method

The Yule-Walker AR method of spectral estimation computes the AR parameters by forming a biased estimate of the signal's autocorrelation function, and solving the least squares minimization of the forward prediction error. This results in the Yule-Walker equations.

The use of a biased estimate of the autocorrelation function ensures that the autocorrelation matrix above is positive definite. Hence, the matrix is invertible and a solution is guaranteed to exist. Moreover, the AR parameters thus computed always result in a stable all-pole model.

4. Covariance Method

The covariance method for AR spectral estimation is based on minimizing the forward prediction error. This method minimizes the forward prediction error in the least-squares sense. Vector a contains the normalized estimate of the AR system parameters, A(z), in descending powers of z.

5. Burg method.

The Burg method for AR spectral estimation is based on minimizing the forward and backward prediction errors while satisfying the Levinson-Durbin recursion. In contrast to other AR estimation methods, the Burg method avoids calculating the autocorrelation function, and instead estimates the reflection coefficients directly.

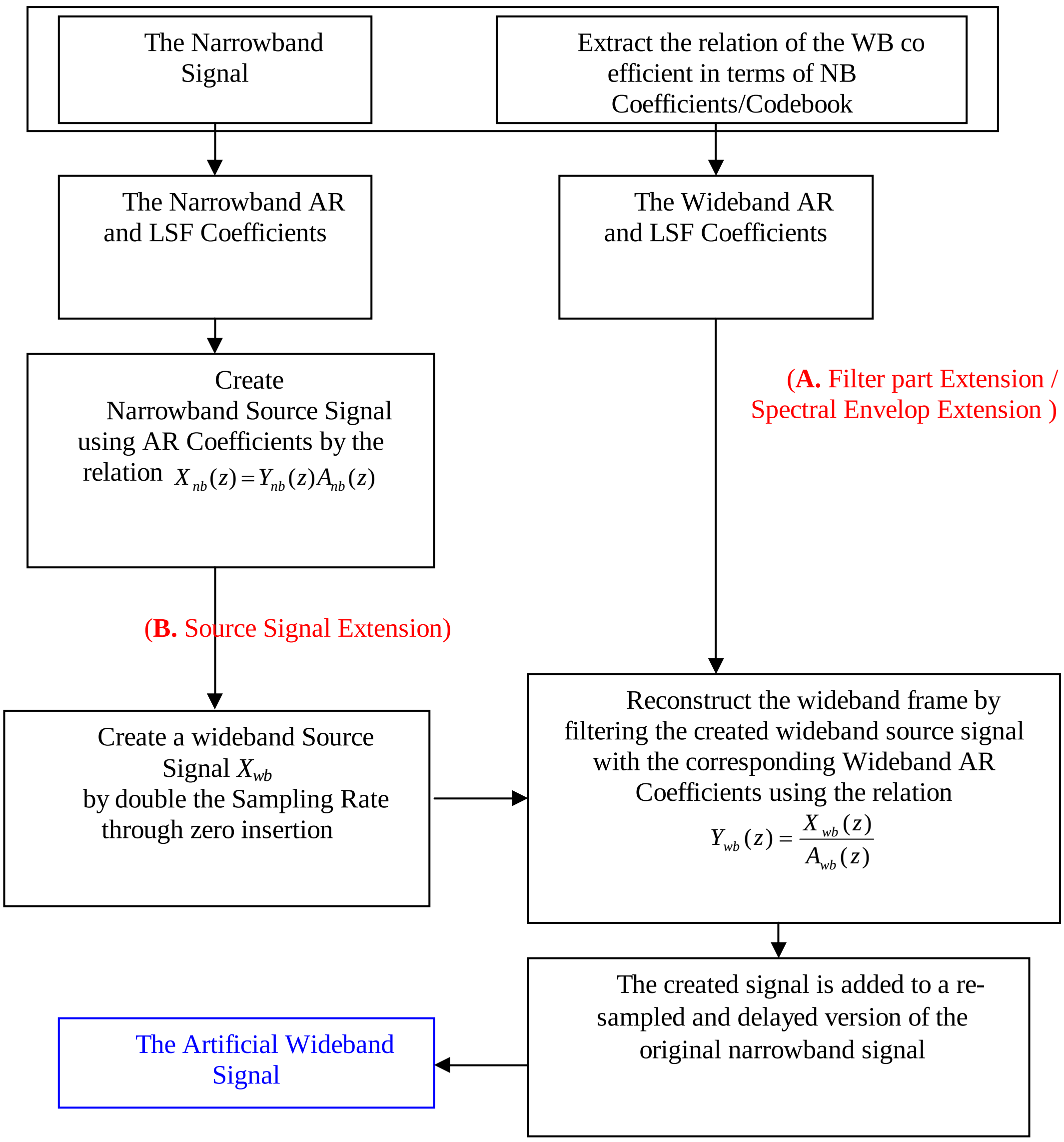


Fig 2 : ABWE System

ABE with different AR Modeling Methods

In this work, we assumed an simple and ideal code book. That is, for each narrowband LSF/AR coefficients, directly the corresponding wideband LSF/AR coefficients were mapped

Outline of Artificial Bandwidth Extention

Each narrowband signal frame is decomposed into a source part and a filter part and the parts are extended separately. The vocal tract is modeled as an all-pole filter and the filter coefficients are estimated using AR modeling. The model residual is used as a source signal. The vocal tract model is extended using the most suitable wideband model taken from a codebook or directly decoded from the file and the residual signal by time domain zero-insertion. The created signal is added to a re-sampled and delayed version of the original narrowband signal to form an artificial wideband signal.

During AR Modeling, the order of 10 is used in the case of narrowband signal and the order of 18 is used for wideband signal.

Results and Discussion

The Speech Database

ARCTIC database by SLT (CMU ARCTIC SLT 0.95) , Carnegie Mellon University includes 1132 utterances recording of the phonetically balanced US English speech by a female US English speaker. The speaker has experienced in building synthetic voices.

Performance Evaluation

We used three methods for evaluating the performance of the ABE techniques. The first is comparing the spectrograms of extended signal with the spectrograms original wideband signal. The second evaluation method is a objective distance measures scaled log spectral distance measure and the third evaluation method is by a subjective measure called mean opinion score.

1.Visualization using Spectrogram

Spectrogram, is a visual demonstration of sound. It is based on real measurements of the varying frequency component of a sound with time. Spectrogram provides more absolute and precise information. Fig 3. Spectrogram comparison for Sample Results with arctic_a0010.wav

2.Log Spectral Distance (LSD) measure.
As a objective distance measures we used Log Spectral Distance (LSD) measure. It is based upon difference between the logarithmic distance of the original spectra **P (□)** and the recreated wideband spectra **P(□).** Table 1 listed the Log Spectral Distances of ABE with LPC and Levinson-Durbin Recursion AR Models for 10 audio files from arctic_a0001.wav to arctic_a0010.wav

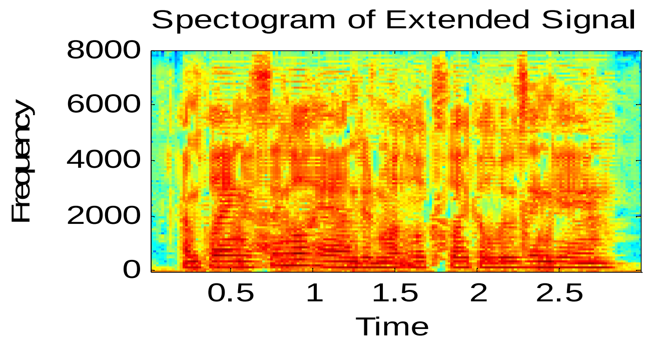
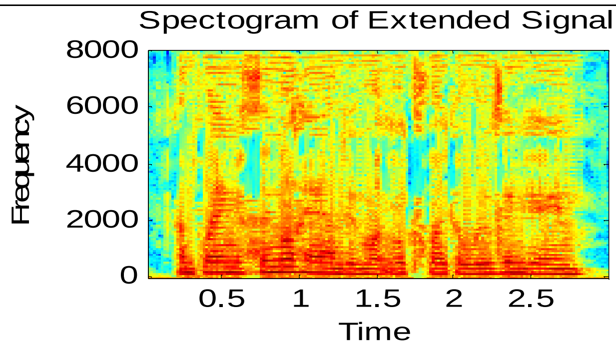
3.Mean Opinion Score
As a subjective measure, we conducted an MOS preference test between the Original Narrowband Signal, and the Bandwidth Extended version of signal. The objective of this test is to find the which signal has the more hearable or understandable high frequency component sounds such as "sh". Table 2 listed MOS of ABE with LPC and Levinson-Durbin Recursion AR Models for 10 audio files from arctic_a0001.wav to arctic_a0010.wav.

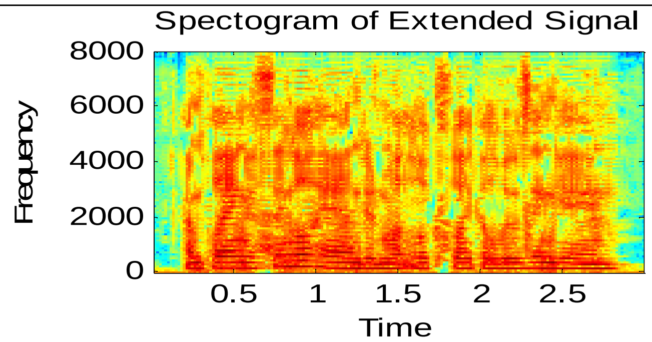
Comparison of Resultant Spectrogram

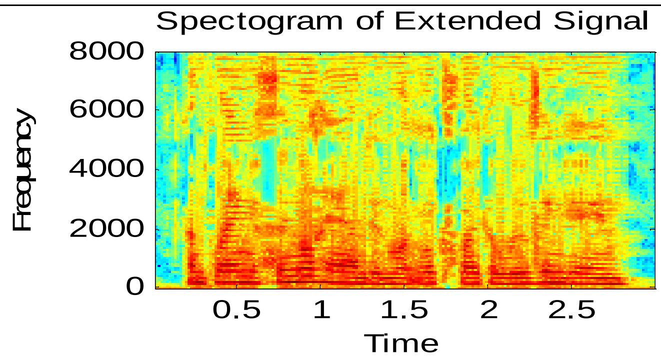
The spectrogram of Autocorrelation Method (LPC) and Yule-Walker Method. are almost like the spectrogram of reference wideband signal and both of them were almost similar to one another. The spectraogram of covariance method and Burg method were too much deviating from the spectrogram of original wideband reference singnal.
The spectrogram of Levinson-Durbin Recursion Method is verymuch similar to the spectrogram of reference wideband signal. So, in terms of spectrogram evaluations, the performance of Levinson-Durbin Recursion AR method based ABE system is significantly better than the other compared methods

A Sample Results with arctic_a0010.wav

ABWE Method	**Spectrogram and LSD (16bit LSF Coding)**

<table>
<tr><td>Original Wideband Signal</td><td>
</td></tr>
<tr><td>Autocorrelation Method (lpc)</td><td>

LSD:1.63</td></tr>
<tr><td>LD Recursion Method</td><td>

LSD: 1.41</td></tr>
<tr><td>Yule-Walker Method</td><td>

LSD: 1.63</td></tr>
</table>

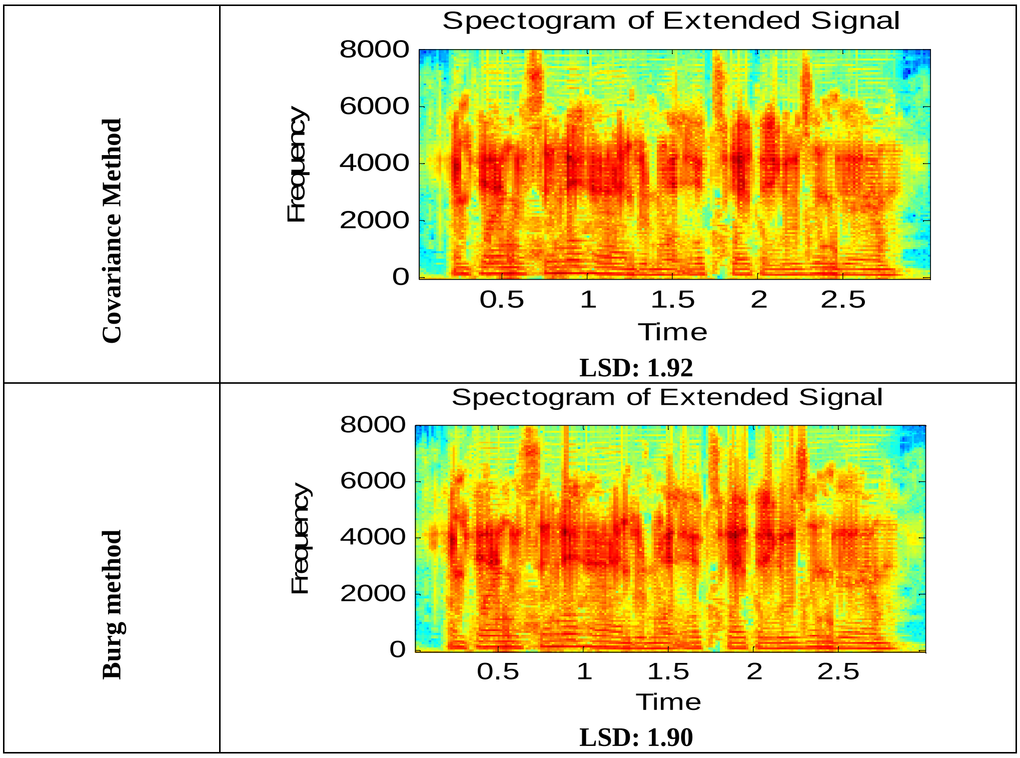


Fig 3: Spectrogram Comparison

Comparison of LOG Spectral Distances

The following Table: 1 illustrates Log Spectral Distances of ABE with different AR Models . Less LSD (**1.244)** implicates LD regenerated the WB with less distortion

Table 1. Log Spectral Distances of Diffetent Methods

Audio sample	Log Spectral Distances of ABE with different AR Models				
	Auto correlation (lpc)	**Levinson-Durbin Recursion**	**Yule-Walker**	**Covariance**	**Burg**
1	1.523	1.238	1.523	1.686	1.784
2	1.459	1.263	1.459	1.598	1.606
3	1.480	1.269	1.480	1.632	1.619
4	1.452	1.251	1.452	1.740	1.788
5	1.350	1.177	1.350	1.716	1.758
6	1.465	1.261	1.465	1.728	1.715
7	1.466	1.282	1.466	1.691	1.741
8	1.449	1.203	1.449	1.567	1.562

9	1.426	1.224	1.426	1.662	1.777
10	1.441	1.270	1.441	1.698	1.699
Avg.	**1.451**	**1.244**	**1.451**	**1.672**	**1.705**

The Comparison of MOS

The following Table 2 gives the Mean Opinion Score of different AR Models (based on the average values of scores from Different Listeners). The high MOS value in LD **(4.570)** perceived better Quality of reproduced speech

Table 2: Mean Opinion Score of Different Listeners

Audio sample	Mean Opinion Score of ABE with different AR Models				
	Auto correlation (lpc)	**Levinson-Durbin Recursion**	**Yule-Walker**	**Covariance**	**Burg**
1	4	4.5	4	2.5	2.3
2	4	4.25	4.5	2.5	2.75
3	4.5	4.75	4	2	2.0
4	4	4.5	4.25	2.5	2.75
5	4.25	4.75	4	2.5	2.75
6	4.75	4.75	4.5	2.5	2.75
7	4.5	4.65	4.75	2.5	2.75
8	4.5	4.75	4.25	2.15	2.75
9	4.25	4.55	4.25	2.35	2.3
10	4	4.25	4	2.5	2.15
Avg.	**4.275**	**4.570**	**4.250**	**2.400**	**2.525**

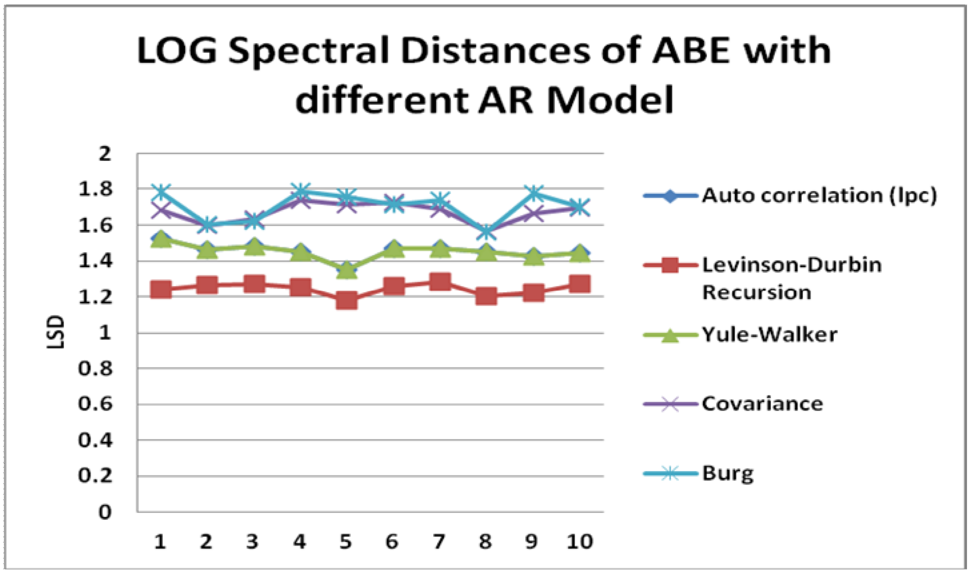


Fig 4: Evaluation of LSD for 10 audio files

The above Fig 4 shows the Evaluation of Log Spectral Distance LSD for 10 audio files. Following bar chart Fig 6 compares the average performance in terms of LSD., Levinson-Durbin Recursion Method outperformed well compare to all other methods.

The Fig 5 shows the Evaluation of Mean Opinion Score for 10 audio files. Following bar chart Fig 7 compares the average performance in terms of MOS., in both the cases Levinson-Durbin Recursion Method outperformed well compare to all other methods.

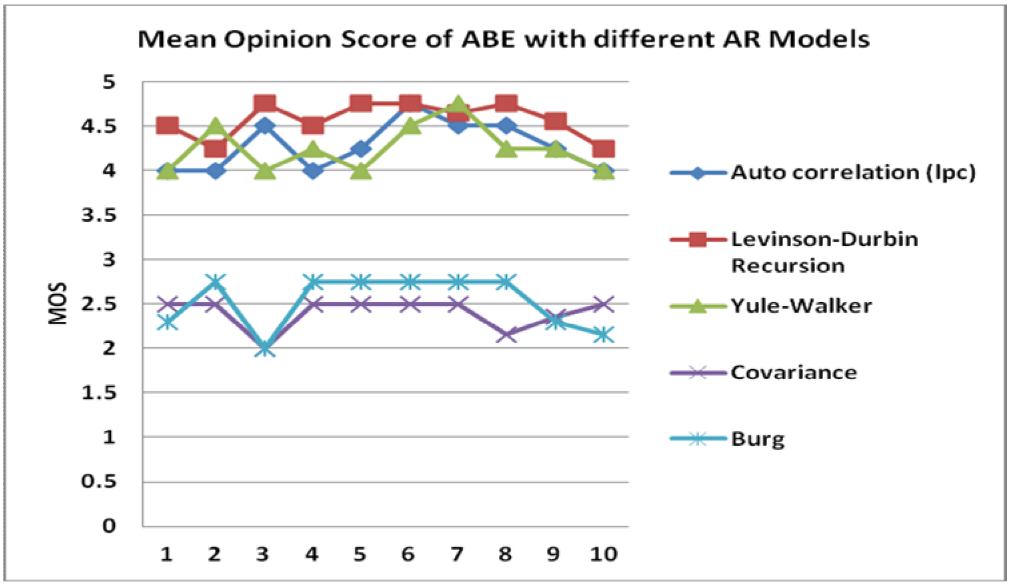


Fig 5: Performance of MOS for 10 audio files

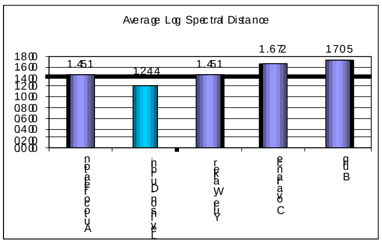


Fig 6: Average performance in terms of LSD

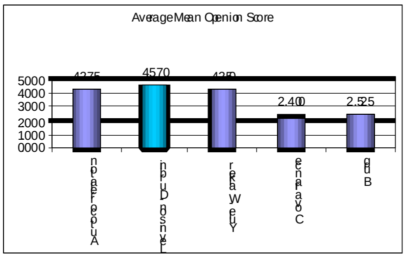


Fig 7: Average Performance In Terms Of MOS

Observations

- According to our abservations, the selectionf of narrowband pre-emphasis and and wideband pre-emphasis filters played an importand role in quality of the extended signal.
- Similarly, the post- emphasis filter also played an importand role in quality of the extended signal. It has great influence on changing the overall pitch of the original signal.
- Furher, the selection of these filter will depend upon the nature of the signal. For example, a set of filter used to extend a female voice may slightly create/extentd a male voice at different pitch.
- The window size used in analysis and synthesis also influences the quality of generated sound.
- The number of AR coefficients used to represent the wideband envelop also decides the quality of sound.
- Above all, the selection of AR model plays most important roll. Better the AR model, we will get better quality of wideband sound.

Conclusions

In this paper, we implemented a bandwidth extension systems using five different AR coefficients for scalable speech and audio coding. As a result, it was shown from the spectrogram comparison and the subjective and objective tests that the implemented Levinson-Durbin Recursion Method based BWE system provided better quality, especially for the voice signals which has high frequency component sounds such as 'sh" . The spectrogram of extended signals of the implemented Levinson-Durbin Recursion Method based BWE system shows the obvious creation of missing bands.

In most of the practical system, the wideband signal will be available at the transmitting end and before transmission, it will be converted in to narrowband signal to minimize transmission cost (narrowband channel is only available). In such cases, the wideband AR coefficients can be directly transmitted along with the narrowband signal. So that, at the receiving end, instead of maintaining a codebook, the bandwidth extension can be done by directly using the wideband AR coefficients.

So to model such fast and efficient system (without codebook) , we may use a custom audio encoding and decoding technique (to incorporate the transfer of wideband AR coefficients) at both transmitting and receiving end. Our future works will address the ways to find the relation between the WB co -efficient in terms of NB with less number of bits . In our future work by implementing Levinson-Durbin Recursion Method based ABWE we may expect better coding efficiency .

Acknowledgement

This work was supported by Periyar Maniammai University . This work has successfully completed by the active support of Prof. A.Hemalatha, and Prof. S.P.K.Babu , Periyar Maniammai University.

References

[1]. Carl,H., and Heute,U., 1994 “Bandwidth Enhancement of narrow band speech signals”, Signal Processing VII, Theories and applications, EUSIPCO, Vol 2, pp. 1178-1181

[2]. CCITT, 1988 ,“7 kHz Audio Coding Within 64 kBit/s”, Recommendation G.722, Vol. Fascile III.4 of Blue Book, Melbourne.

[3]. Cheng,Y.M., ’Shaughnessy,D.O, Mermelstein,P., 1994 “Statistical Recovery of Wideband Speech from Narrowband Speech”. IEEE Transactions on Speech and Audio Processing, vol.2, no4,pp.544– 548.

[4]. Chennoukh,S., Gerrits, A., Miet, G.,and Sluitjer, R., May 2001 "Speech enhancement via frequency bandwidth extension using line spectral frequencies," Pro. IEEE Int. Conf. On Acoustics, Speech, Signal Processing, vol. 1, pp. 665- 668

[5]. 3GPP TS 26.171, March 2001 “Speech Codec speech processing functions; AMRWB Speech Codec; General Description”. Version 1.

[6]. Gandhimathi,G., Narmadh,C., and Lakshmi,C. August 2010 ,” Simulation of Narrow Band Speech Signal using BPN Networks, International Journal of Computer Applications (0975 - 8887) Volume 5- No.9.

[7]. Gandhimathi.G., Jayakumar,S., 2012”, “ Efficient Method of Pitch Estimation for Speech Signal Using MATLAB , “IJCCT, ISSN (ONLINE) : 2231–0371, ISSN (PRINT) : 0975–7449, Volume- 3, Issue-1, pp. 107-112

[8]. Gandhimathi.G., Jayakumar,S., 2013” Speech enhancement Using Artificial Bandwidth Extension Algorithm in Multicast conferencing through Cloud services “, Information Technology Journal , ISSN 1812- 5638 pp.1-8.

[9]. D.A. Heide, G.S. Kang,, May 1998, "Speech enhancement for band limited speech", Proceedings of the ICASSP, Vol. 1, Seattle, WA, USA, pp. 393-396.

[10]. B. Iser and G. Schimidt, September 2003 "Neural Network Versus Codebooks in an Application for Bandwidth Extension of Speech Signal", in Proc. ROSPEECH'03, vol.1, pp. 565-568,.

[11]. P. Jax and P. Vary, May 2002., "An upper bound on the quality of artificial bandwidth extension of narrowband speech signals", Pro. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) vol. 1, pp 237-240, Orlando, FL, USA

[12]. P. Jax and P. Vary, May 2004 "Feature Selection for Improved Bandwidth Extension of Speech Signals", Pro. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) vol. 1, pp. 697-700, Montreal, Canada,.

[13]. G. Miet, A. Gerrits, and J. C. Valiere, Jun 2000, "Low-band extension of telephone-band speech," Pro. IEEE Int. Conf. On Acoustics, Speech, Signal Processing, vol. 3, pp. 1851-1854.

[14]. Y. Qian and P. Kabal, May 2004, "Combining Equalization and Estimation for Bandwidth Extension of Narrowband Speech", Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (Montreal, QC), pp. 1-713-1-716.

[15]. M.J.L. De Hoon, T.H.J.J. Van Der Hagen, H. Schoonewelle, And H. Van Dam "Why Yule-Walker Should not be Used for Autoregressive Modelling" Interfaculty Reactor Institute, Delft University of Technology, Mekelweg, Netherlands.