International Journal of Electronics Communication and Computer Engineering
Volume 4, Issue (6) NCRTCST-2013, ISSN 2249–071X

National Conference on Recent Trends in Computer Science and Technology (NCRTCST)-2013

An Analysis on Stenography Based ABE System Using LPC Based Source Filter Model

G. Gandhimathi
Research Scholar,
School of Architecture Engineering and Technology
Periyar Maniammai University Thanjavur, India
Email: gandhimathiece@pmu.edu

Dr. S. Jayakumar
Asst. Dean of School of Computing Science
Periyar Maniammi University,
Thanjavur, India
Email: jk@mullivaikal.in

***Abstract* – As the limitation of the wireless transmission frequency band resources due to the complexity of the 3G mobile communications environment, in order to achieve efficient transmission and reconstruction of low bit rate with high-quality audio signal, we proposed a new wide band coding scheme based on any NB coder and wide band enhancement by few bits /frame of high band information. The Artificial bandwidth extension (ABE), extends the speech bandwidth by only information available in the narrowband speech signal. The proposed method is based on spectral envelope extension using classified codebook approach. The number of spectral peaks is used as the feature for classification. The results illustrate that the ABWE algorithm can achieve reliable perceived voice quality in comparison with the LPC source filter model with even less bit rate and evaluated its performance with suitable metrics.**

***Keywords* – Bitrates, Speech Codec, Speech Enhancement, LPC.**

I. Introduction

All digital hands-held communication devices, internet telephony need high feature good quality speech. To solve those problems the ITU standardizes the wideband (WB) speech codecs [2, 5] . An improved frequency bandwidth of speech signals contributes considerably to the apparent speech quality [15]. By using enhanced speech codecs WB speech communication needs a modification in the transmission link and increased bitrates.

Another approach for getting higher bandwidth is termed as Bandwidth Extension (BWE). The missing low frequency and high frequency of NB input speech signal are improved at the receiving end. By using a linear predictive model the redundancies in the frequency bands are discovered, then predict the missing components proposed in [3]. It can be done by two steps First, from the available narrow band (NB) speech the source model parameters are estimated then in second, from the combination of parameter can able to estimate the source model then add with the predicted missing components.

In this paper the BWE of speech for predicting high frequencies is examined. Here NB band speech of frequencies below 3.4 kHz is input signal and it is extended to 7 kHz artificially In this paper

Problem Specification

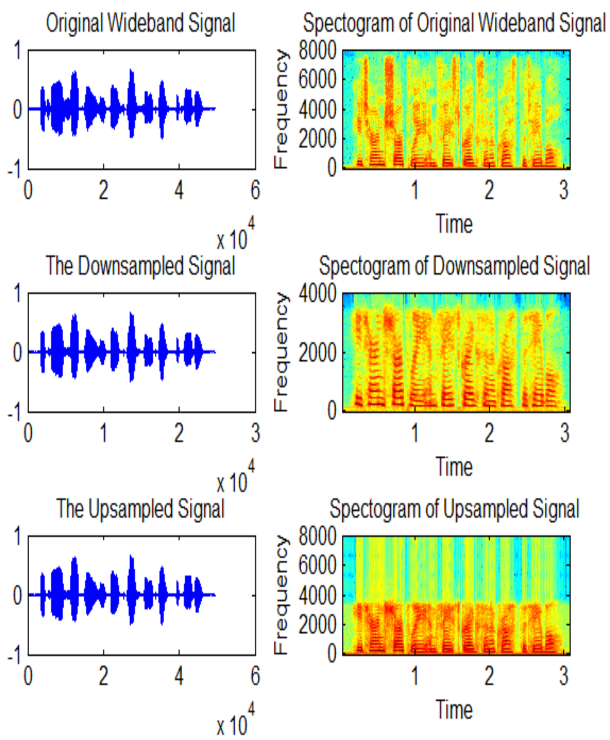


Fig 1. Missing Frequency components of WB and NB Signal

Fig.1 shows a WB signal (top) and its corresponding NB representation (middle) in which the whole high frequency components (in both the middle and bottom, above 3.4kHz) is missing. Even in the wideband signal the frequencies above 7khz is almost absent in the recording. It means that the first harmonic components of the band 3.4kHz. to 4kHz is also missing above 7kHz in the wideband spectrum. So, creating the missing frequency components between 3.4kHz. to 4.6kHz will be a challenging task.

So, at the receiving end, ABE techniques are used to construct a artificial wideband sound form the NB sound. Generally ABE techniques will create the missing frequency components from the learned mutual relationship between the lower and upper frequency regions of the spectrum. The source-filter model used in ABE systems will have some drawbacks and limitations.

International Journal of Electronics Communication and Computer Engineering
Volume 4, Issue (6) NCRTCST-2013, ISSN 2249–071X

This work will address these issues by developing a source-filter model based ABE system and evaluate its performance with Carnegie Mellon University's ARCTIC sound database [9].

Previous works

Miet et al (2000) splits the NB speech into short-term residual and spectral envelope . They are individually extended then recombined to produce a WB signal. Another novel approach is discussed by Qian and kabal (2004) to create a WB signal (3.4-7 KHz) by combination of equalization and estimation. Equalization technique is used in mid range (3.400-4KHz). Because of the small mutual dependency between the NB and WB parameters its performance may improved than statistical estimation procedures.

This paper structured as, in section II explores the proposed methodology. Section III discuss in detail about the design of the proposed source filter modeled ABE, Section IV elaborates the results, Section V concludes its performance and the future work.

II. Source-Filter Model of Speech Production

Model: Fig.2. shows generation of speech by source-filter model based on Vocal tract. (VT assumed as all pole filter).

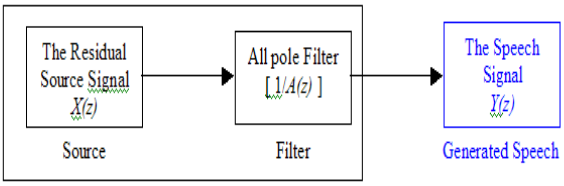


Fig.2. The Source-Filter Model of Vocal Tract

Consider VT be a pth order all-pole filter

$$\frac{1}{A(z)} = \frac{1}{1 + a_1 z^{-1} + \dots\dots\dots + a_p z^{-p}} \text{..........} \quad \text{.....(1)}$$

Where the filter coefficients a_1 ...a_p are estimated using linear prediction.Speech signal *Y(z)* (in frequency domain) is formed by filtering the residual/source signal *X(z)* by the all-pole flter 1/*A(z)*

Y(z) = X (z) / A(z) -----------(2)

In fact, from the recorded audio, we have only the speech signal *Y(z).* So that, the residual/source signal *X(z)* can be formed by filtering the speech singnal *Y(z)* by the VT inverse filter *A(z).* We can find the *A(z)* using the filter coefficients a_1 ...a_p that were estimated using linear prediction on *Y(z).*

X(z) = Y(z) A(z) ---------------- (3)

Fig 3a (Top) shows one frame of a given input signal (red) and its corresponding Error/source(blue) Signal calculated by the above relationship. It means, in the figure1, if we supply the error/source signal at the left hand side, then it will generate the speech at the right hand side. The source-filter model is related to linear prediction and it used in both speech analysis and speech synthesis,

Fig.3b (bottom) shows the frequency domain representation (FFT spectrum shown in blue) of the previously shown time domain signal frame and its corresponding LPC coefficients (red).

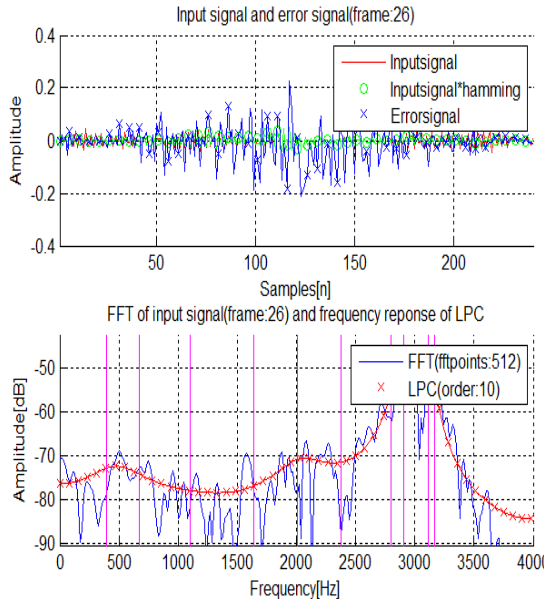


Fig.3.a. Input and error signal. 3b. FFT and LPC of Input signal

Implementation: For voiced speech, excitation signal often modeled as a periodic impulse train, or white noise for unvoiced speech. The VT filter is approximated by an all-pole filter where the coefficients are predicted by linear prediction. To get synthesized speech excitation signal and the filter response are convolved.

III. Modeling The Artificial Bandwidth Extention System

i) WB LPC hidden on NB

Various codebooks are designed corresponding to frames having same number of spectral peaks. In this scheme three codebook pairs are trained for spectral peaks ranging from one to five. 10th order LSP is used for narrow band speech and 18th order LSP is used for wideband speech. The LSP parameters are calculated every 10ms frame of NB and the WB speech for the same corresponding speech frames after the classification is done to the training set based on the number of spectral peaks in the envelope. (Using the same frame structure as in G.729). The narrow band training data are quantized by predetermined NB codebook, Each wideband LPC is retrieved from the corresponding narrowband LSP.

International Journal of Electronics Communication and Computer Engineering
Volume 4, Issue (6) NCRTCST-2013, ISSN 2249–071X

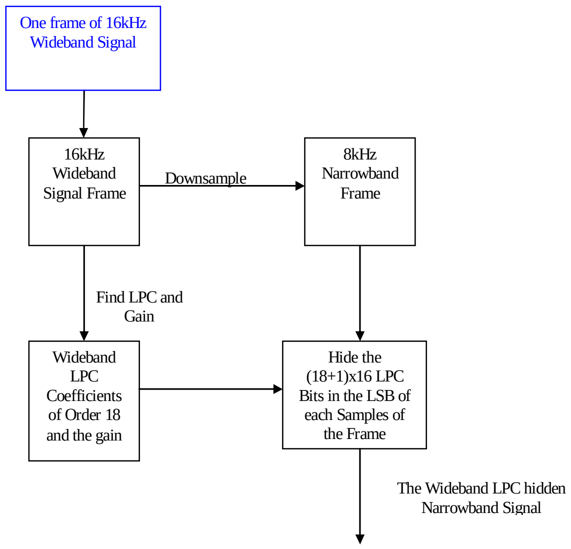


Fig.4. WB LPC hidden on NB signal

The following steps explain the construction codebook (Fig.4).

1. The wideband pre-emphasis filter was applied on the wideband training data wave files. In this implementation use a FIR filter H (z) = 1 - $0.95z^{-1}$ is used as wideband filter
2. The narrowband signals are formed by decimating (down sampling) the same wideband training data wave files and a suitable narrowband pre-emphasis filter was applied on the narrowband signals.
3. The wideband signal is decomposed in to frames of uniform 25ms window size with no overlapping between adjacent frames and calculate Wideband LPC Coefficients of Order 18 and the gain
4. The narrow band signal is decomposed in to frames of uniform 25ms window size with no overlapping between adjacent frames. Hide the (18+1)x16 LPC Bits in the LSB of each Samples of the Frame
5. Save the Wideband LPC hidden Narrowband Signal

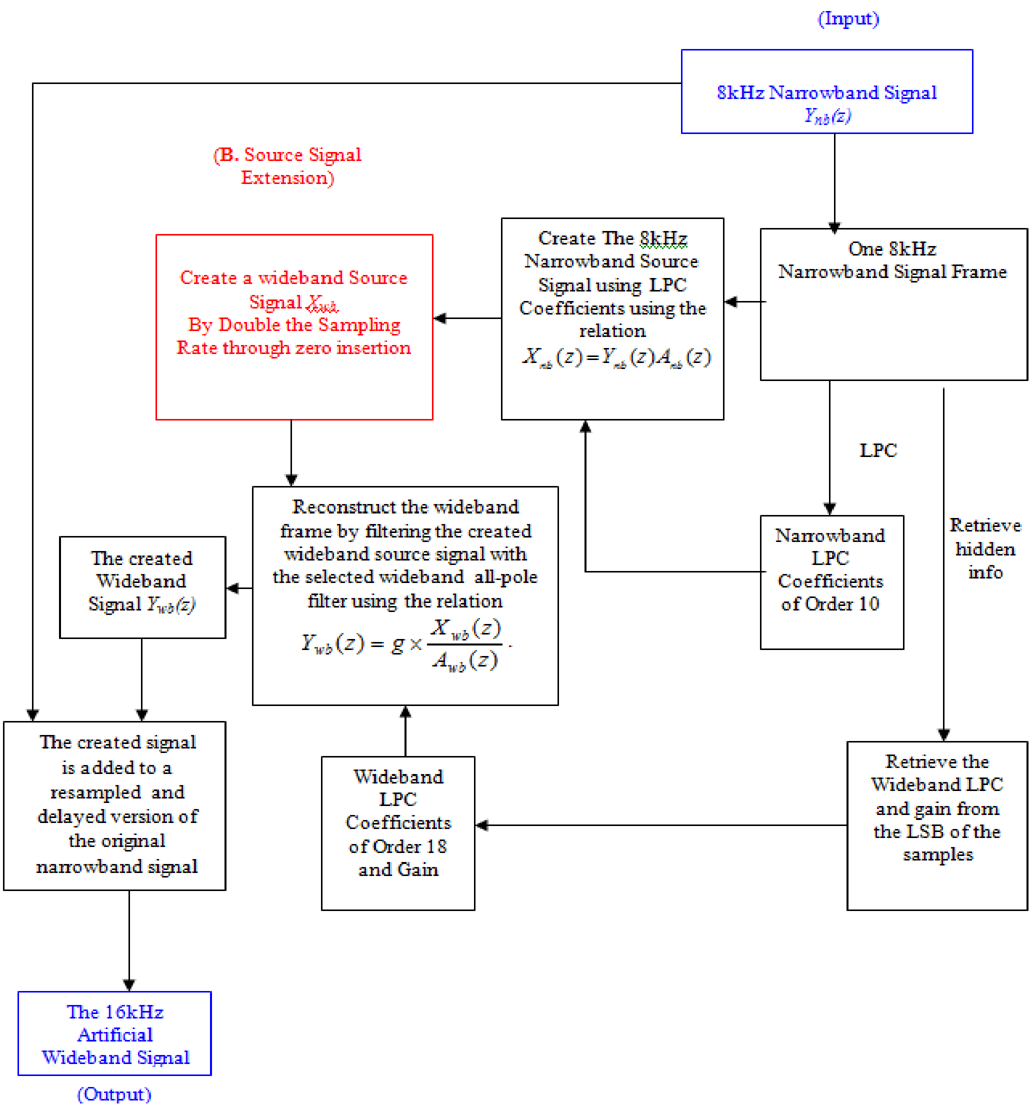


Fig.5. Artificial Bandwidth Extention (ABE) using Steganography based hidden Wideband Information.

International Journal of Electronics Communication and Computer Engineering
Volume 4, Issue (6) NCRTCST-2013, ISSN 2249–071X
National Conference on Recent Trends in Computer Science and Technology (NCRTCST)-2013

ii) Artificial Bandwidth Extention (ABE)

In ABWE, each NB signal frame is split into a filter part and source part .Each part is extended separately. using linear prediction (LP), the all pole filter (VT filter) the filter coefficients can be estimated .The model residual be a source signal. The VT replica is extended from a codebook using the most appropriate WB model and the residual signal by Time domain (TD) with zero-insertion. The produced signal is added with a delayed resampled version of original NB signal to produce an Artificial WB signal.

ABE Steps :

The basic idea is to create a signal that contains the frequencies that are missing from the original narrowband signal (Fig 5).

A. Spectral Envelope Extension / Filter part Extension

1. The NB signals which is to be extended is opened and a suitable NB pre-emphasis filter was applied on the NB signals.

$X_{nb}(z) = Y_{nb}(z)A_{nb}(z)$ ---------------- (6)

2. The sampling rate of NB signal is increased using TD with zero-insertion .This will build a signal of 4-8kHz and it is a mirror copy of the 0-4kHz
3. The LP coefficients of each NB frame is calculated and Convert all LPC coefficients .
4. Retrieve the hidden information from the LSB of the samples.
5. Retrieve WB LPC coefficients of order 18 and gain from the LSB of the samples

B. Source Signal Extension

1. Using the narrowband signal frames from (6) and its LPC coefficients calculate the source signal.
2. Extend the narrowband source signal using zero insertion technique (X_{wb}).
3. Using the extended source signal from (7), and the LPC coefficients from (5) calculate the output signal.
4. The extended frames from (8) are concatenated using overlap and add method. For analysis window 25ms and synthesis window 10ms are used. Both analysis and synthesis, between adjacent frames the time difference is 5ms.
5. To get the final synthetic WB signal, the signal from (9) and the delayed version of resampled NB signal were added/mixed. To synchronize the signals, the resampled NB signal must be delayed since the extension may causes a delay.

Fig 6 shows an example of NB LPC coefficients (top) of order 12 and its corresponding frequency response curve (bottom).

Fig.7 shows a original sample from (first) and its windowed version of signal (second). The third is the reconstructed NB signal.

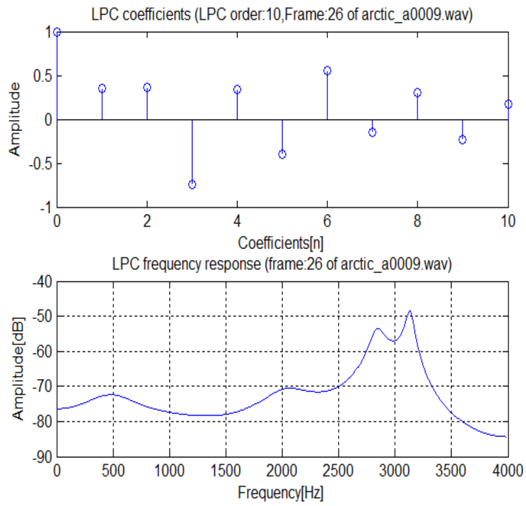


Fig.6. NB LPC coefficients and its frequency Response curve

Visualization using Spectrogram

For visual representation of sound we used spectrogram. It provides entire and accurate information of sound signal since it is based on change in frequency content of a sound wave over time. At every picky time, relative intensity of sound and its frequency are shown by the color at that point of spectrogram hearable or understandable high frequency component sounds such as “sh”. We selected 10 NB files in which the high frequency components are almost absent but present in the corresponding original WB signals.

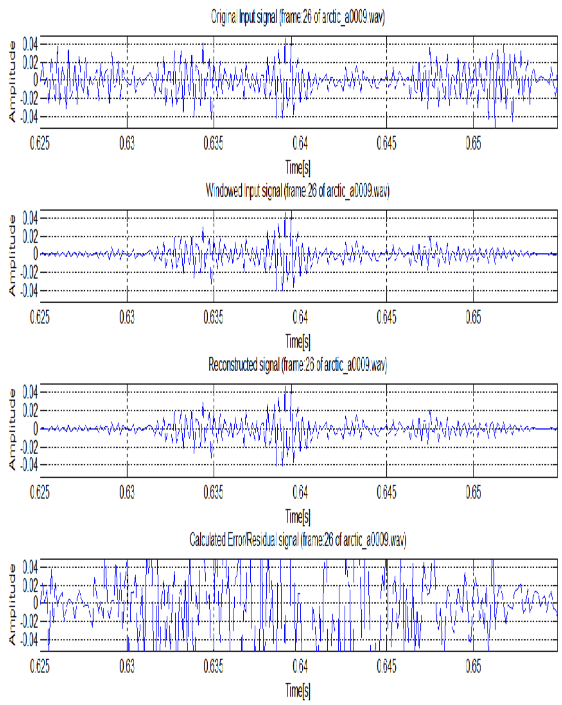
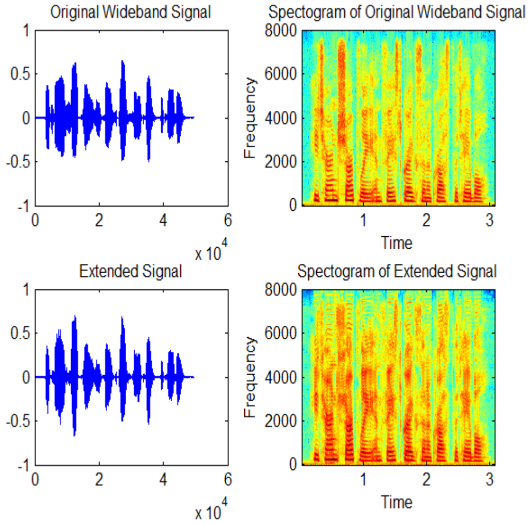


Fig.7. Different types of Time domain frames

International Journal of Electronics Communication and Computer Engineering
Volume 4, Issue (6) NCRTCST-2013, ISSN 2249–071X

As shown in Fig 8(b), the spectrogram of decoded signal by the conventional BWE algorithm had horizontal lines at round 3.6 kHz and 7.6 kHz, as denoted by a dotted box, which were caused by the gain mismatch. This mismatch, however, was mitigated by using the proposed BWE algorithm thus, there were no horizontal lines in Figure 8(c). These results implied that the proposed BWE algorithm could provide better quality than the conventional BWE algorithm

Comparison of Resultant Spectrogram



Observations :

According to our observations, the selection of NB pre-emphasis and and WB pre-emphasis filters 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.

The window size used in analysis and synthesis also influences the quality of generated sound.

CONCLUSION

In this paper, we employed a ABWE algorithm using LPC coefficients for scalable speech and audio coding. As a outcome, it was clearly exposed from the preference test and comparison of spectrogram, employed ABWE algorithm provided better speech quality, especially which has a high frequency component sounds like 'sh" . The spectrogram of extended signals shows the obvious creation of missing bands. The conducted MOS preference test between the Original NB Signal, the interpolated version of NB signal and the Bandwidth Extended version of signal shows that the ABE system almost created the original WB signal from the NB signal.

We use Steganography based data hiding technique to hide the LPC coefficients of each signal frame inside the corresponding signal frame itself.

Steganography based ABE system will not affect the existing transmission and reception methods very much. Even a standard receiving system may play the NB signal as it is without processing the hidden LPC coefficients. Steganography based ABE systems and it is fast and efficient

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 & 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; AMR WB 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., 2013" Speech enhancement Using Artificial Bandwidth Extension Algorithm in Multicast conferencing through Cloud services ", Information Technology Journal, ISSN 1812- 5638 pp.1-8.