ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 1, January 2018

Bandwidth Extension of Speech Signal: A Review

[1] Janki Patel, [2] Nikunj V. Tahilramani, [3] Ninad Bhatt, [4] Vandana Talreja
[1] PG student, [2] Assistant Prof. & Head, [3] Professor & Head, [4] Lecturer,
Department of E&C
[1][2] Silver Oak College of Engineering & Technology Ahmedabad, Gujarat, India.
C.K. Pithawala College of Engineering & Technology Surat, Gujarat, India.
Government Polytechnic, Kheda, Ahmedabad, Gujarat, India.

Abstract: - **The narrowband frequency range from about 300 Hz to 3.4 kHz. So, the audio quality of telephone network is restricted. Bandwidth extension(BWE) is used to extend narrowband(NB) frequency range to wideband(WB) frequency range which is 50 Hz to 7 kHz and wideband frequency range to super wideband(SWB) frequency range which is 50 Hz to 14 kHz. This bandwidth extension can be realized with or without side information. The Bandwidth extension is performed by adding missing frequencies artificially at the receiver using information contained in narrowband signal or either using some low bit rate side information. The basic principles of bandwidth extension address by this review paper and discuss different methods of bandwidth extension from narrowband to wideband and narrowband to super wideband.**

Key words: - **Bandwidth extension; narrowband; wideband; super wideband.**

I. INTRODUCTION

The main aspect of digital telecommunication systems is quality, clarity and naturalness of the speech. The telephone frequency range is 300 Hz-3400 Hz. Due to this limited bandwidth the speech quality is degraded. To improved speech quality and intelligibility many noise reduction and error cancellation techniques are used still it may sound unnatural and muffled. Human speech includes more frequency components compared to narrowband telephony speech[1]. For that bandwidth extension is used, which extend speech signal from narrowband to wideband, wideband to super wideband and narrowband to super wideband. Where narrow band frequency range is 300 Hz to 3.4 kHz, wideband frequency range is 50 Hz to 7 kHz and super wideband frequency is 50 Hz to 14 kHz. The bandwidth extension is performed by with or without side information. At user end terminal, the absent frequencies of speech is added artificially using information contained in NB speech. Otherwise, side information is used for bandwidth extension. Artificial bandwidth extension is evaluated by pattern recognition methods, which is Hidden Markov Models (HMMs) and Gaussian Mixture Models (GMMs). The telephony frequency range has been used in PSTN and also in 2G wireless communication system such as GSM. The narrow band speech quality degradation is mainly due to some information is lossed in 50 Hz to 300 Hz and 3.4 kHz to 8 kHz[3]. Which creates muffled voice and reduced the speech quality and clarity. These all problems solution is explain in next sections. Now in paper, next section A is an implementation of wideband coder. In section B, implementation of narrowband coder with basic two approaches. And in last super wideband extension and its comparison with wideband.

A. IMPLEMENTATION OF WIDEBAND CODER

At receiver, wideband coder is used. Its implementation gives good quality higher signal which is used in many applications. In this section, wideband coders are listed. In 1985, The first wideband speech codec is G.722 It was defined by CCITT (now ITU-T) in 1985 with three bit rates which is of 64, 56 and 48 Kbit/s. Which is used for ISDN and tele-conferencing applications. The second wideband codec is G.722.1. It was specified by ITU-T in 1999. It gives good speech quality at low bit rates. The latest wideband codec is adaptive multi-rate wideband (AMR-WB) codec, which was standardized by ETSI and 3GPP and used for CDMA cellular networks e.g. UMTS. The second generation G.722.2 is AMR-WB codec[11]. It was also invented by ITU-T This wideband codec is used in fixed network applications. All AMR-WB wideband codec has nine bit rates from 6.6 to 23.85 Kbit/s. After that AMR-WB+ codec was specified by ITU-T with bit rates from 6.6

All Rights Reserved © 2018 IJERECE

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 1, January 2018

and 32 bit/s and frequency range from 7 kHz to more than 16 kHz.

B. IMPLEMENTATION OF BWE IN NARROWBAND CODER

The Bandwidth extension is perform by adding missing frequencies artificially at the receiver[3,4] using NB speech contents or either using some low bit rate side information. At user end terminal, natural speech is generated by BWE. There are main two methods for BWE, which is distinguished as:

- BWE without side information
- BWE with side information

Bandwidth extension is applied for receive better quality speech. There is a no requirement of modification at the transmitter and also in network. For reverse process the NB encoder has been replace by WB encoder. Normally, all of the systems described below to generate NB to WB speech signal. This WB speech can be applied another time to achieve super-wideband(SWB) speech signal.

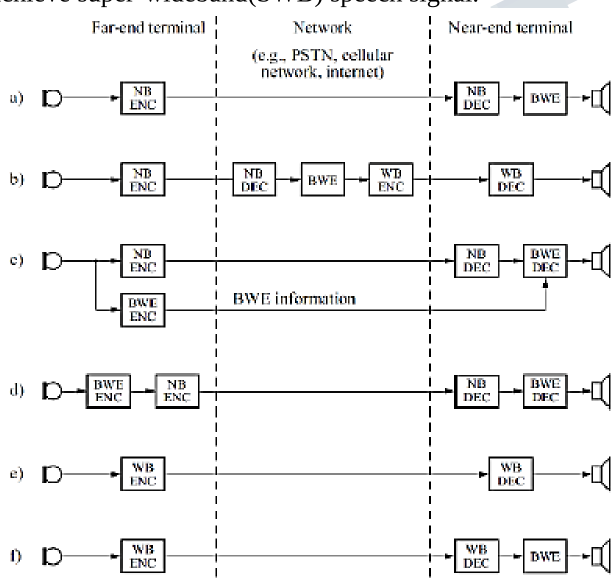

Fig.1 Steps from NB speech signal to WB speech signal[3]

Fig. 1 shows step by step as:
a) NB speech signal transmission and bandwidth extension in the user end terminal
b) NB transmission and BWE in the network
c) Parallel transmission of BWE information for BWE
d) BWE information embedding within narrowband signal
e) WB codec for speech transmission
f) For SWB speech, WB codec and bandwidth extension.

Different methods for the evaluate the missing frequency components are done and good outcome have been acquired by all of them.

II. BWE WITHOUT SIDE INFORMATION

BWE without side information means signal bandwidth is widening artificially, by producing missing frequencies at receiver using narrowband signal.

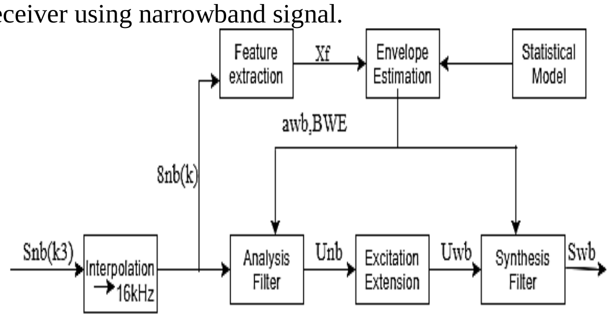

Fig.2 Basic block diagram of BWE without side information[6]

First, from NB signal feature vector has been extracted and wideband AR- coefficients are determined using this feature vector Xf. By this BWE can be split into main two sub parts described below:

- WB spectral envelope estimation
- NB excitation extension

1) WB spectral envelope estimation:
There are number of methods for WB spectral envelope estimation, which are as below:
a) Codebook mapping method:
This method is used for estimation of high band spectral envelope[13]. Codebook is trained using WB speech dataset. This pre-trained codebook is used to predict the high band envelope. Wide band envelope in the codebook is equated with collected narrow band envelope. All entries are compared with received narrow band envelope and closest entry is selected. This is called spectral envelope estimation. Codebook mapping is a basic method and compared with other methods. In the sense of spectral distortion, this methods gives good results. One limitation is that two wide band spectral envelope having same narrow band spectral envelope have same high band spectral envelope. There is a some delay because of comparison of narrow band envelope with all codebook codeword for getting best matching.
b)Adaptive codebook mapping method:
The limitation of above codebook mapping method is overcome by adaptive codebook mapping. There is an N closest codewords are selected. Then wide band code vector is generated using this N closest codewords by weighted average.

All Rights Reserved © 2018 IJERECE

One other modification is that different codebook is generated for voiced and unvoiced speech. This two different codebook are trained individually using voicing detection. This voicing information is used for expansion process.

c)Gaussian mixture model(GMM):

GMM evades the quantization because this approach is based on continuously mapping and soft clustering[9]. Codebook mapping is models as discretely and allowing hard clustering. Bandwidth extension in GMM providing high conversion accuracy with Minimum Mean Square Error (MMSE).

d)Hidden markov models(HMM):

HMM describes a time varying process. The advantage of HMM for envelope estimation is its capacity of completely utilizing data from the previous signal frames to enhance the estimation quality. The change of a state the HMM describe the envelope change of speech because each state of the HMM is correspond to a certain speech sound. Compare to GMM with higher order, HMM provides superior results with lower order. HMM model[7] uses Expectation Maximization (EM) trained GMM to estimate the observation probability PDFs. It makes a cepstral estimate of wide band coefficient. Using true state sequence, state and state transition probabilities are evaluated. The disadvantage of this method is that needs appreciable quantity of data to good estimation of state models and transition probability.

2)NB excitation extension:

In this excitation extension, the important feature is the regeneration of the HB /WB excitation signal. It doubles the sampling rate (8 kHz to 16 kHz).which is the main aim of the excitation extension and also the whole spectrum should be kept flat. The harmonics contained in the NB excitation should also be transferred in the WB excitation. There are some known technologies which are used to generating HB excitation signal from the given input NB residual signal. Which are spectral folding, spectral translation, non linear distortion, noise modulation, full wave rectification[14].

III. BWE WITH SIDE INFORMATION

All the techniques studied in above chapter are bandwidth extension without side information . These all methods are used to estimate high band features using only narrow band speech features. So the drawback is that an accurate estimation of envelope requires a speaker dependent training which is not applicable for real time application. BWE without side information is not enough for better quality wide band signal reconstruction at end terminal. These all limitation can be control by sending side information at the transmitter and at the receiver wideband speech is retrieve back. High band feature parameters are estimated and encoded at transmitter as a side information of original signal. At receiver side, these parameters are transmitted and estimate high band envelope. In traditional method is providing side information transmission which is not compatible with network. For that solution data hiding scheme is used. The side information is transmitted as a hidden data inside narrow band speech signal. At the user end terminal this hidden data is decoded and extracted. Which gives estimated the wide band speech and speech quality is improved.

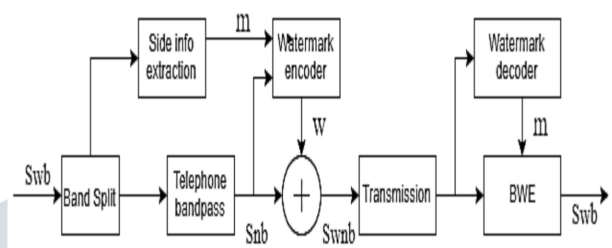

Fig.3 BWE with side information algorithm[2]

The algorithm of bandwidth extension with side information is shown in figure 3. At the first original speech(wide band speech) is band spitted and separate narrow band and high band components. This narrow band speech given to narrow band encoder to encode the speech. And high band features are encoded with side information. These encoded speech is watermarked inside the narrowband speech in narrow band coder. At the end terminal this hidden data is reproduced and bandwidth extension is applied to produce WB speech signal.

IV. WIDEBAND V/S SUPER WIDEBAND

The wideband speech is defined by its acoustic frequency range of 50 Hz to 7 kHz, where as super wideband speech provides a roughly doubled the frequency range of wideband Speech which is 50 Hz to 14 kHz. The lower cutoff frequency of 50 Hz is usually considered sufficient for a natural reproduction of speech signals The speech coding recommends improving the quality and increasing the bandwidth of the coded signal instead of increasing the absolute compression efficiency. For high quality voice calls , super wideband extension is used to implement with both speech and audio signals. G.729.1 and G.718 are scalable codec which is recently invented by ITU-T, with bit rates 32 kbps and below. A connection SWB and stereo extension for these codecs is on the ITU-T roadmap for standardization in 2009. When extending wideband codec to super wideband codec, many facets have to be considered, mainly in telecommunications applications in which the delay and

All Rights Reserved © 2018 IJERECE

complexity both should be conserved at a acceptable level. And at that time the codec should ideally be scalable. In audio coding applications for coding the high frequencies different bandwidth extension techniques[15,16] are commonly used. In these techniques the low frequencies is coded and coded this low frequency content is used for coding the high frequencies. Although, these techniques may not encounter the delay requirements, and in inclusion the scalability is not completely operational as the coded high frequency band usually cannot be enhanced with supplementary layer because of these techniques are not waveform coders, i.e., they do not retain the initial waveform shape of the high frequency spectrum.

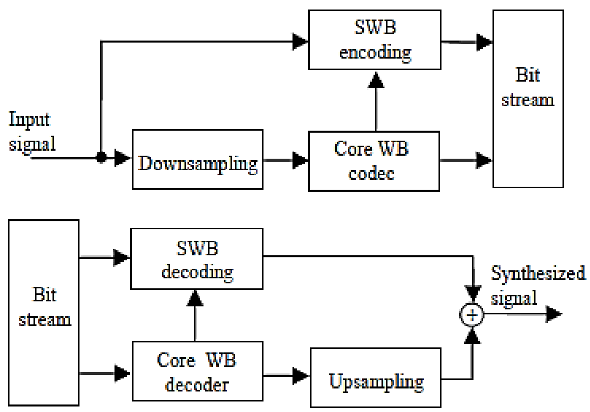

Fig.4 Block diagram of the Super wideband extension[12]

V. SUPER WIDEBAND CODEC

The block diagram of the proposed super wideband extension is proposed in Figure 4. This system uses the coded low frequency content. To coverts between WB and SWB speech signals, resampling (down sampling and up sampling) is used. In this system, the G.718 codec is used as the core codec which is modern in wideband speech and audio coding. G.718 is an embedded scalable speech and audio codec. Which is standardized for error susceptible communications channels and a wide area of applications comprises high quality audio conferencing and video conferencing, packetized voice, 3G and future wireless communication systems, and multimedia streaming. G.718 codec includes five from L1 (core level) to L5. Level L1 and level L2 are based on the ACELP technology. Levels L3 to L5 use modified discrete cosine transform (MDCT) to coding the error from level L2. The frame size of this codec is 20 ms. And for wideband signals the algorithmic delay is 42.875 ms. But the original baseline codec EV-VBR, which has 54.75 ms total algorithmic delay. This algorithmic delay difference is obtained, because of the window length utilized in the OLA operation for the transform coding in coder.

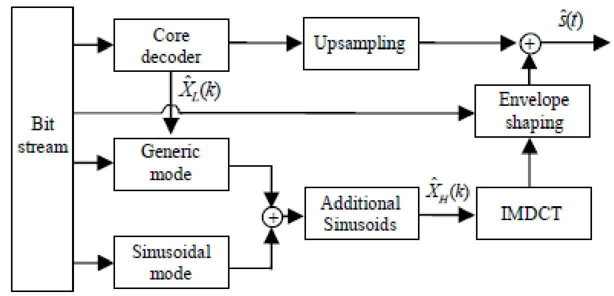

Fig.5 SWB decoder[12]

On top of the G.729.1 core, this super wideband extension was performed and further increase the quality, and reduction in delay and complexity.

VI. CONCLUSION

In ordinary telephone networks, the speech quality is limited because of NB frequency range (300 Hz to 3.4kHz). So, the bandwidth extension algorithm is applied to generate WB speech signal at receiver side. So, the quality and naturalness of the speech can be improved. By applying BWE algorithm, the missing higher frequency components of speech can be added at the end terminal to produce WB speech. This quality increment is produced without increasing the bit rate, i.e. no additonal bits are required for transmission but the high band features are inserted with narrowband speech signal. Bandwidth extension without side information is not good for real time application because of the high band speech which is produced at the receiver is only dependent on narrowband speech signal. For that solution, the high band features is embedded as a side information of narrow band speech. This watermarked data is retrieve at the receiver and embedded with narrow band speech. So bandwidth extension of speech signal is produced. Further improvement of the speech quality in some application is required. For that super wideband extension is used for that bandwidth extension with wide band transmission gives super wideband speech quality.

REFERENCES

[1] P. Jax and P. Vary, ―On artificial bandwidth extension of telephone speech,‖ Signal Process., vol. 83, no. 8, pp. 1707–1719, 2003.

[2] Chauhan, Tejal, Ninad Bhatt, and Shraddha Singh. "Artificial Bandwidth Extension Method of telephony Speech in Mobile Terminal: A Review."

[3] P. Jax and P. Vary, ―Bandwidth extension of speech signals: A catalyst for the introduction of wideband speech

All Rights Reserved © 2018 IJERECE

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 1, January 2018

coding?,‖ IEEE Commun.Mag., vol. 44, no. 5, pp. 106–111, 2006.

[4] Peter Jax ―‖Artificial Bandwidth Extension of Speech Signals‖ (IND) Aachen University (RWTH), Germany.

[5] Murali Mohan D, Dileep B. Karpur, Manoj Narayan, Kishore J ―Artificial bandwidth extension of narrowband Speech using Gaussian Mixture Model‖, BNM Institute of Technology,IEEE 2011,pp. 410-412.

] P. Jax and P. Vary, ―Wideband extension of telephone speech using a hidden Markov model,‖ in IEEE Speech Coding Worksh., Delavan,WI, USA, Sept. 2000, pp. 133–135.

[7] P. jax and P.vary ―Artificial bandwidth extension of speech signals Using MMSE estimation based on a hidden markov model‖ Institute of communication systems and data processing (ind), Aachen University (RWTH), 52056 Aachen, Germany,IEEE(ICASSP) 2003.

[8] P. Jax, ―Enhancement of bandlimited speech signals: Algorithms and theoretical bounds‖, Aachen University (RWTH), Aachen, Germany, vol. 15 of P. Vary , 2002.

[9] W. fujitsuruy, H. sekimotoy, ―Bandwidth extension of cellular phone speech based on maximum Likelihood estimation with GMM‖ International Workshop on Nonlinear Circuits and Signal Processing NCSP'08, Gold Coast, Australia, March 6-8, 2008.

[10] J. Makhoul and M. Berouti, ―High-frequency regeneration in speech coding systems,‖ in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), 1979, vol. 4, pp. 428–431.

[11] ITU-T Rec. G.722.2, ―Wideband Coding of Speech at around 16 kb/s Using Adaptive Multirate Wideband (AMR-WB),‖ Jan. 2002.

[12]Tammi, Mikko, et al. "Scalable superwideband extension for wideband coding." Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on. IEEE, 2009.

[13] Qian, Y. & Kabal, P. ―Wideband speech recovery from narrowband speech using Classified codebook mapping‖,(mcgill university) Proceedings of the 9th Australian International Conference on Speech Science & Technology Melbourne, December 2 to 5, 2002.

[14]Bhatt Ninad, and Yogeshwar Kosta. "A novel approach for artificial bandwidth extension of speech signals by LPC technique over proposed GSM FR NB coder using high band feature extraction and various extension of excitation methods." International Journal of Speech Technology 18.1 (2015): 57-64.

[15] J. Mäkinen et al., "AMR-WB+: A New Audio Coding Standard for 3rd Generation Mobile Audio Services," in Proc. ICASSP, Philadelphia, PA, USA, pp. 1109-1112, March 2005.

[16] M. Dietz, L. Liljeryd, K. Kjörling, and O. Kunz, "Spectral band replication, a novel approach in audio coding," in Proc. AES 112th Convention, Munich, Germany, Paper 5553, May 2002

All Rights Reserved © 2018 IJERECE