Feasibility Study, Implementation and Overall Evaluation of Performance of Audio Signals over Code Excited Linear Prediction based Adaptive Multi-Rate Wideband Coder

Ninad Bhatt[1] and Yogeshwar Kosta[2]

[1]*Research Scholar, Veer Narmad South Gujarat University, Surat, Gujarat, India*

[2]*Campus Director, Marwadi Education Foundation, Rajkot, Gujarat, India*

bhattninad@gmail.com, ypkosta@gmail.com

Abstract

This paper is a humble step towards investigating feasibility of implementation of Code Excited Linear Prediction algorithm on Adaptive Multi Rate Wideband coder over audio signals. Proposed coder offers inherent capability of adaptively change its bitrates with variation in received C/I ratio depending upon channel conditions. Though proposed WB coder exhibits nine different bitrate modes (between 6.6kbps and 23.85kbps), its practical implementation reveals that only two upper most bitrates (which are 23.05kbps and 23.85kbps) have been found suitable for audio transmission and recovery over WB link. An e-test bench using MATLAB is created to implement proposed WB coder and series of simulations are carried out to judge the performance of implemented coder for audio signal using Subjective (Mean Opinion Scores) and Objective (Perceptual Evaluation of Audio Quality- Objective Difference Grade) analysis. Simulation results clearly advocate that it is possible to reproduce audio signals with comparable quality when implemented with upper most two bitrates of AMR WB coder. It is also evident from the obtained simulation results that PEAQ and MOS for chosen audio signal are also quite comparable and satisfactory.

Keywords: *Audio coding, Adaptive Multi Rate Wide Band coder, Code Excited Linear Prediction, MATLAB, Subjective analysis, Objective analysis*

1. Introduction

Today, the major limitations in existing wireless communication systems are bandwidth and power for any signal transmission. It has become indeed a need for communication community to envisage the development of efficient and effective methods which maintains adequate quality of recovered signal at receiving terminal [8-9]. The role of the WB coder is to compensate the need of utilization of existing conventional Narrowband telephony and wireless networks. The inherent benefit in WB communication is that it provides listeners with the feeling of transparent communication and in turn also reduces the listening efforts made by listeners. As an outcome of such benefits WB coding systems have become revolutionary offering numerous wide area applications both in speech and in audio processing. Potential applications of WB coding include Voice over Internet Protocol (VoIP), high-quality audio conferencing, 3G wireless communications, audio streaming, downloading and uploading content etc. [2, 5, 6].

Adaptive Multi Rate system combines source and a channel coder which are triggered and controlled by signaling means. It aims at providing best recovered signal quality under the scenario of background noise as well as transmission errors thus mitigates the dependence of recovered signal quality over variable and unpredictable behavior of channel conditions. In this paper, AMR WB coder is implemented and simulated in MATLAB environment using Code Excited Linear Prediction algorithm over audio signal. Further subjective and objective evaluations are carried out on the proposed coder to judge its performance for audio signals and the same are compared with its counterpart speech signals [11].

The paper is organized as follows: In Section 2, CELP speech coder is described whose quantization bits are affected to be adjusted to AMR WB Codec mode bitrates. In Section 3 MATLAB simulation of proposed coder is described. Section 4 describes Subjective and Objective performance evaluation of proposed coder using set of tables and graphs. Finally the concluding remarks are given in Section 5.

2. Overview of CELP Coder

Introduction to Code Excited Linear Prediction coder block diagram is depicted in Fig. 1. Audio signal input *s(n)* is applied as an input to buffer and LPC process. Main role of buffering followed by Linear Prediction analysis is for the estimation of vocal impulse response system for each frame which then produces pitch delay and LP coefficient. The pitch delay will be used in pitch synthesis filter and the LP coefficient (a_i) at LP synthesis filter. Before the process of pitch synthesis filter, pitch filter coefficient *(b)* will be produced first from the computation using pitch delay *(P)*. The process of LP synthesis filter before the processing of LP coefficient in the block is carried out by converting LP coefficient (a_i) to produce reflection coefficient of LP [1].

In order to obtain gain parameter and codebook index, perceptual weighting filter process will be carried out and then error minimization process is performed. In error minimization section, gain and codebook index, which will be used in the next block, will be determined. After all parameters are obtained, audio coding process can be carried out to compress the audio signal. Functioning of each section is as demonstrated in Figure 1. Therefore, the compressed audio signal can be formed as an outcome and adapted to the AMR codec mode. Further, CELP parameters obtained from CELP analysis sections are eventually utilized to reproduce audio signal [10-11].

3. Simulation of Proposed Coder

AMR-WB coder exhibits nine bitrate modes in kbps i.e. 23.85, 23.05, 19.85, 18.25, 15.85, 14.25, 12.65, 8.85 & 6.60 kbps. The above codec modes are derived from ETSI standardization in 2001 [6].

In this research four audio (mono-left channel) wave files (each having sampling frequency f_s=44.1 kHz) have been utilized. The subset was chosen to include critical and varied material from mono instrumental (first three wave files like drumsA.wav, drumsB.wav and Chimes.wav) to complex polyphonic (song.wav). Here, for each wave files samples are coded by 16 bits/sample. Frame duration of 20ms have been taken for the vocal-tract analysis (320 samples) and block (sub-frame) duration of 5ms (80 samples) for determining the excitation [13].

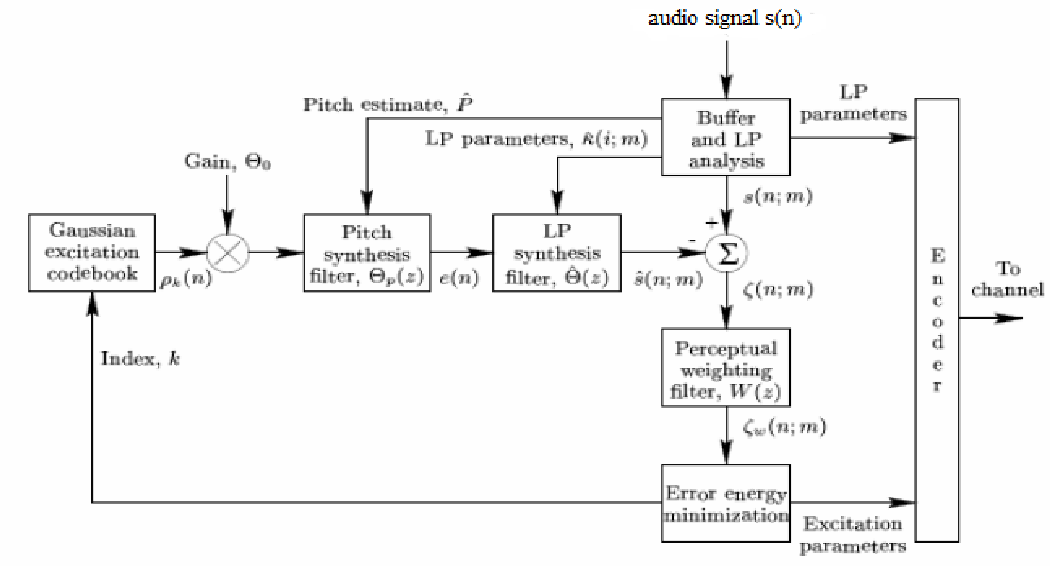

Figure 1. Structure of CELP Coder

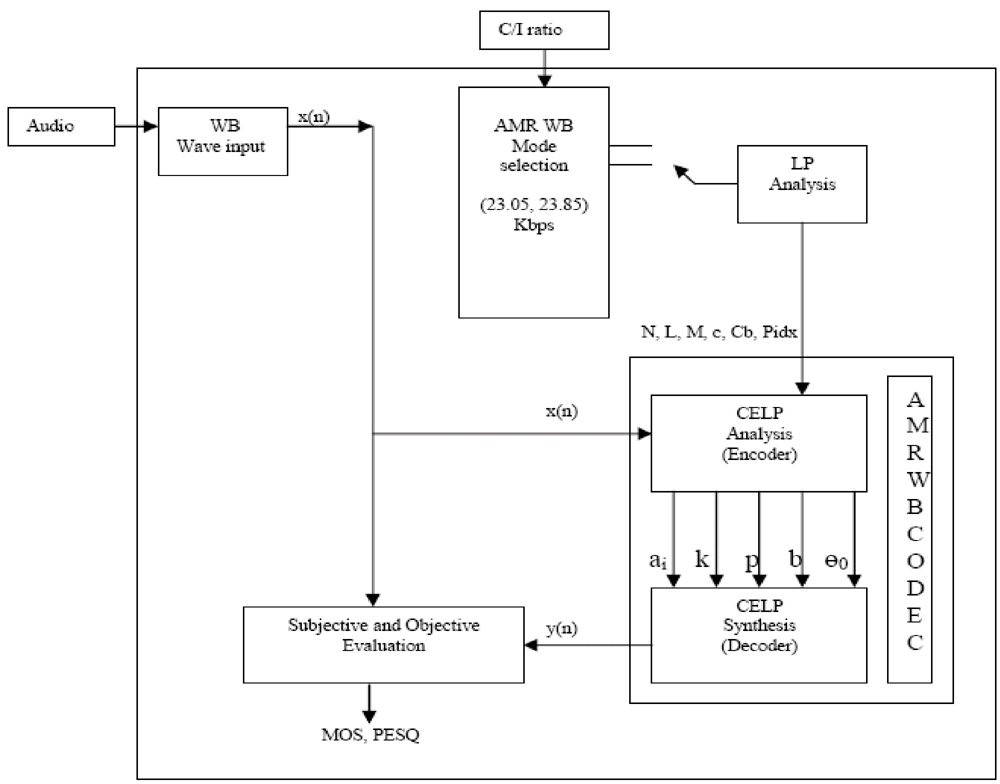

Figure 2. MATLAB Simulation of Proposed Coder

As illustrated in Figure 2, audio signals having wave file format have been supplied as an input to the block of proposed coder. Along with audio input, here C/I ratio has also been provided as the other input. Provision is made in MATLAB code to pick any specific bitrate mode of AMR WB coder based on the given C/I ratio input. Further, LPC analysis has been performed on given audio file for the selected mode of operation to produce the parameters like frame length *(N)*, Block length *(L)*, order of filter *(M)*, LP parameter *(c)*, codebook index (C_b) and Pitch index (P_{idx}) which are then supplied to CELP analysis block of proposed AMR WB coder. Information parameter as a result of CELP analysis related to proposed AMR WB codec include: linear prediction coefficient *(a)*, pitch lag *(p)*, codebook index *(k)*, gain (Θo) and pitch filter coefficient *(b)*. The bit allocations of all the five information parameters tuned as per AMR WB coder bitrate are highlighted in Table 1 [12]. Information parameters generated from CELP analysis section are then passed on to CELP synthesis section for the recovery of audio signal. Finally, for the purpose to judge the performance of proposed coder for feasibility of audio transmission and recovery, subjective (MOS) and objective (PEAQ – Objective Difference Grade) analysis have been conducted [3].

4. Overall Performance Analysis of Proposed Coder

This paper presents overall performance analysis of proposed coder using both objective and subjective analysis as described below.

4.1. Subjective Analysis

In this category of analysis, Mean Opinion Score rating has been conducted on four different clean audio files for listening test. MOS analysis is carried out in quiet environment and with the use of high quality headphones. To conduct MOS analysis, twenty un-trained listeners are chosen to participate in listening test. Out of which ten listeners are men and ten are women listeners. Each listener is offered with total of eight (four wave files, each of two bitrate modes) wave files. Ratings given by all twenty listeners (for each wave files and for each coder's bitrate mode) are then averaged to produce final MOS ratings [14].

As could be witnessed from Table 2 and Figure 3, obtain results for MOS scores advocate the performance of proposed WB coder for upper most two bitrate modes for given four audio wave files. It is evident from the results cited in table and bar graph that, 23.85kbps bitrate mode outperforms its counterpart 23.05kbps mode. Moreover it should be noted as this juncture of time that all other lower bitrate modes of proposed coder (seven bitrate modes between 6.6 kbps and 19.85 kbps) were not found suitable for audio coding and transmission because it rapidly degrades MOS score values with gradual reduction in bitrate modes. It could also be revealed from this implementation and analysis that obtained MOS values for chosen audio signals offer quite comparable values to MOS scores produced in [12] for given set of speech wave files.

4.2. Objective Analysis

In the category of Objective analysis, only perceptual based analyses based on Perceptual Evaluation of Audio Quality-Objective Difference Grade [13] have been conducted in this paper as per ITU-R 1387-1[3]. Here ODG has been considered to be output parameter for perceptual measurement of decoded audio signals. Typical range of ODG lies between 0 and -4 [3].

As can be demonstrated in Table 2 and Figure 4, with reduction in bitrate modes from 23.85kbps to 23.05kbps, small decrement in ODG ratings have been quite evident for all audio files. While comparing MOS and ODG scores, resultant values clearly advocate that both offer quite comparative results. ODG scores for given audio files are also found satisfactory and comparable to Perceptual Evaluation of Speech Quality (PESQ) scores obtained in [12] for different speech corpuses.

Table 1. AMR-WB Codec Mode Bit Allocation [12]

AMR (kbps)	*A*	*P*	*K*	Θo	*B*	Total bits
23.85	37	9,9,9,9	25,25,25,25	38,38,38,38	38,38,38,38	477
23.05	37	9,9,9,9	23,23,23,23	37,37,37,37	37,37,37,37	461
19.85	33	9,9,9,9	18,18,18,18	32,32,32,32	32,32,32,32	397
18.25	29	9,9,9,9	17,17,17,17	29,29,29,29	29,29,29,29	365
15.85	29	9,9,9,9	16,16,16,16	24,24,24,24	23,23,23,23	317
14.25	25	9,9,9,9	15,15,15,15	21,21,21,21	20,20,20,20	285
12.65	21	9,9,9,9	14,14,14,14	18,18,18,18	17,17,17,17	253
8.85	21	9,9,9,9	13,13,13,13	9,9,9,9	8,8,8,8	177
6.6	20	9,9,9,9	10,10,10,10	5,5,5,5	4,4,4,4	132

Table 2. Subjective and Objective Analysis of Proposed Coder

CELP based AMR WB codec	PEAQ (Objective Difference Grade)		MOS Score	
Audio wave file	23.05 Kbps	23.85 Kbps	23.05 Kbps	23.85 Kbps
drumA.wav	-1.455	-1.273	3.47	3.66
drumB.wav	-1.397	-1.196	3.54	3.78
Chimes.wav	-1.261	-1.174	3.72	3.80
Song.wav	-1.294	-1.184	3.69	3.75

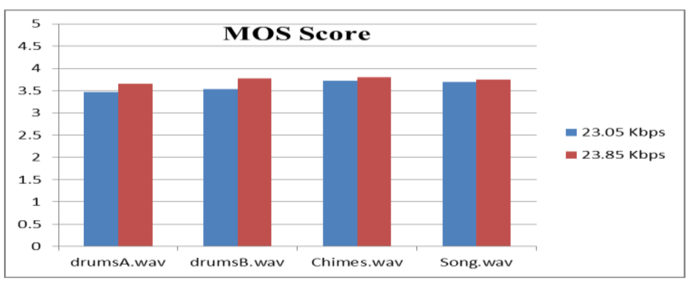

Figure 3. MOS Score Comparison between Two Bitrate Modes of Proposed Coder

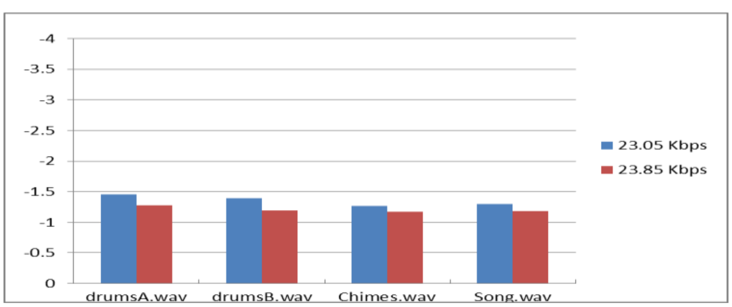

Figure 4. ODG Comparison between Two Bitrate Modes of Proposed Coder

5. Concluding Remarks

In the recent scenario of wireless communication, aggressive researches have been reported in the domain of wideband coding in last decade. In order to mitigate the dependence of quality of recovered signals at receiver from erroneous channel conditions, WB AMR is a technique which combines source and a channel coder so that codec modes are made adaptive to the variation in channel conditions (as per received C/I ratio). Here, CELP based AMR WB coder has been proposed for application of transmission of audio signal over wideband link.

The motive of this research is to comment on feasibility of coding and transmission of audio signals over AMR WB coder, and for the same, series of simulations have been carried out. The results obtained as an outcome reveal that out of nine available bitrate modes of AMR WB coder, only upper most two bitrates (which are 23.05kbps and 23.85kbps) have been found suitable for audio coding and transmission. As audio signals have broader spectra compared to speech and hence other lower bitrates of AMR WB coder exhibits degradation in the quality of recovered audio. Subjective (MOS) analysis and Objective (PEAQ-ODG) analysis demonstrates that proposed coder works well for audio signal for last two bitrate modes. Both MOS and PEAQ-ODG values for chosen audio are quite satisfactory and comparable to the values obtained for speech signals.

References

[1] X. Jianming, L. Xun and W. Lei, "Software simulationin GSM Environment of Federal Standard 1016 CELP Vocoder", International Conference on Communication Technology, Beijing, China, **(1998)** October 22-24.

[2] J. Schnitzler and P. Vary, "Trends and perspectives in wideband speech coding", Signal Processing, vol. 80, **(2000)**, pp. 2267-2281.

[3] ITU-R Recommendation BS 1387-1, "Methods for objective measurements of perceived audio quality", Geneva, Switzerland, **(1998-2001)**.

[4] R. Huber and B. Kollmeier, "PEMO-Qa new method for objective audio quality assessment using a model of auditory percep-tion", IEEE Trans. Audio, Speech, Lang. Process., vol. 14, no. 6, **(2006)** November, pp.1902-1911.

[5] R. Salami, B. Bessette, R. Lefebvre and M. Jelinek, "The Adaptive Multi-Rate Wideband Codec: History And Performance", IEEE Conference, **(2002)**.

[6] B. Bessette, R. Salami, R. Lefebvre and M. Jelinek, "The Adaptive Multi-Rate Wideband Speech Codec(AMR-WB)", IEEE Trans. on Speech and Audio Processing, **(2002)**.

[7] W. C. Chu, "Speech Coding Algorithms, Foundations & Evolution of Standardized coders", A John Wiley & Sons, Inc., Publications, **(2003)**.

[8] D. Malkovic, "Speech Coding Methods in Mobile Radio Communication Systems", 17th International Conference on Applied Electromagnetics and Communications, Croatia, **(2003)**.

[9] J. D. Gibson, "Speech Coding Methods, Standards and Applications", IEEE Circuits and Systems Magazine, IEEE Fourth Quarter, **(2005)**.

[10] E. Pryadi, K. Gandi and H. Kanalebe, "Speech Compression Using CELP Speech Coding Technique in GSM AMR", IEEE Conference, **(2008)**.

[11] N. Bhatt and Y. Kosta, "Overall Performance Evaluation of Adaptive Multi Rate 06.90 Speech Codec based on Code Excited Linear Prediction Algorithm using MATLAB", International Journal of Speech Technology, springer, vol. 15, no. 2, **(2012)** June, pp. 119-129.

[12] N. Bhatt, Y. Kosta and D. Bhagat, "Adaptive Multi Rate Wideband speech codec based on CELP algorithm: Architectural study, implementation & performance analysis", IEEE International conference on Communication Systems and Network Technologies, Rajkot, India, **(2012)** May, pp. 547-551.

[13] E. Ravelli, G. Richard and L. Daudet, "Union of MDCT Bases for Audio Coding", IEEE Transaction on Audio, Speech, and language processing, vol. 16, no. 8, **(2008)** November, pp. 1361-1372.

[14] F. Kurth, "An audio codec for multiple generations compression without loss of perceptual quality", Proceedings of the AES- 17th International Conference on High Quality Audio Coding, **(1999)**.

Corresponding Author

Ninad Bhatt. He has received his B.E. degree in Electronics Engineering from VNSGU, Surat, Gujarat in 2001and M.E. in Electronics and Communication Systems from DDU, Nadiad, Gujarat in 2007 and currently pursuing his Ph.D. in the area of Speech and Audio compression from VNSGU, Surat, Gujarat.

He is currently serving as an Associate Professor and Head of Electronics and Communication Department of C. K. Pithawalla College Of Engg. and Tech., Surat, Gujarat affiliated with Gujarat Technological University, Amdavad,, Gujarat. He has around twenty research publications in his domain in various national, international conferences including IEEE conferences and reputed international journals like springer.