

A Simulation-Based Comparison on Code Excited Linear Prediction (CELP) Coder at Different Bit Rates

Swati Joshi, Hemant Purohit and Rita Choudhary

Abstract A better speech quality is necessary for transmission of speech in mobile communication systems in Digital Telephony. Nowadays, the conversion of analog speech signals into digital format is required for effective transmission of speech over different channels at far ends. The technique for conversion of speech in digital form is very old and ordinary one which is termed as pulse code modulation (PCM), but the bandwidth of the digitally converted data is too large, so a better level of compression is needed to reduce the bandwidth and enhance the capacity of channel. The compression of speech in nowadays is performed by a procedure called speech coding. In this paper, the code excited linear predictive (CELP) coding is summarized with different bit rates. The MATLAB R2016a version is used for simulating the 9.6 and 16 kbps CELP coder, and performance analysis is done with parameters MSE and SNR.

Keywords Speech coding · LPC · CELP

1 Introduction

The speech coding is a procedure to represent a digitized speech signal with its minimum bit format necessary to transmit it over different channels. Speech coding is a wide area of research from late 80s to present. The advancement in speech coding techniques is necessary due to rapid increase in users of mobile communication and limited bandwidth of channels. The speech coders are creating a

S. Joshi (✉)

Tula's Institute, Dehradun, India
e-mail: joshi92swati@gmail.com

H. Purohit · R. Choudhary

JJET Jodhpur, Jodhpur, Rajasthan, India
e-mail: hemant.purohit@jjetjodhpur.ac.in

R. Choudhary

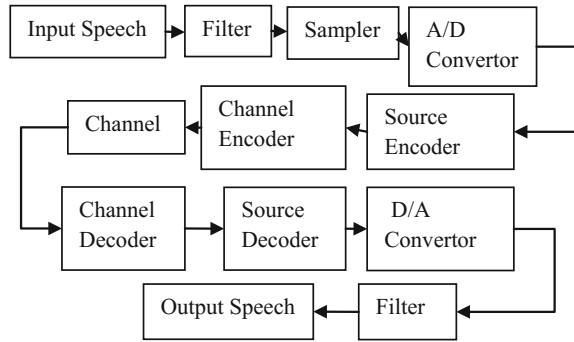
e-mail: rita.choudhary@jjetjodhpur.ac.in

© Springer Nature Singapore Pte Ltd. 2018

B. Tiwari et al. (eds.), *Proceedings of International Conference on Recent Advancement on Computer and Communication*, Lecture Notes in Networks and Systems 34, https://doi.org/10.1007/978-981-10-8198-9_31

297

Fig. 1 Block diagram of speech coding system



competitive environment between the telecom provider giants; better speech quality with lower bit rates is demanding one. The basic building blocks for speech coders are shown in Fig. 1. The low bit rate and high speech quality are the main requirements to design a speech coder [1, 2]. The speech coders are classified into two terms.

- (a) Waveform coders—pulse code modulation (PCM), differential pulse code modulation (DPCM), adaptive differential pulse code modulation (ADPCM).
- (b) Parametric coders—linear predictive coding (LPC), residual excited linear prediction, mixed excited linear prediction, code excited linear prediction (RELPC, MELPC, CELPC).

2 Literature Survey

The research on speech coding starts at Bell laboratories, and the first electronic voice synthesizer was invented by Homer W. Dudley in 1930 for secure voice transmission during world war. Motivation for speech coding research at that time is to design a system which is bandwidth efficient for telegraph cables. Dudley practically demonstrates the speech and figure out the redundancy in speech and finally set up with the new procedure of analysis by synthesis method for designing of speech coder [3].

The basic idea behind the first coder was to analyze speech in terms of its pitch spectrum by band-pass filter to analyze the periodic and random analysis of speech. The improvement on speech coders had been done during 1940s–1960s [4].

The early vocoder system is totally based upon the analog signal, and the digital representation gains interest because of its best encryption and better fidelity over long-range transmission. In 1940s, a new term introducing the speech coding named as pulse code modulation (PCM). PCM is a direct method for representation of discrete time and discrete amplitude of analog signals. The more advancement in this technique is started, and the best quantization capabilities were developed in

differential PCM, delta modulation, and adaptive DPCM were developed and speech coding in PCM with 64 kbps and with ADPCM 32 kbps become the standard of Consultative Committee for International Telephony and Telegraphy (CCITT) [2].

A great innovative invention was done by Prof. Fant in 1950s, the linear speech source system. The model consists of linear time-varying coefficient of speech signals excited by periodic impulse train for both voiced (speech) and unvoiced (noise) signal, and this model becomes the basic building block for new generation linear predictive speech coding [4, 5].

Theoretical and practical aspect on linear predictive speech coding is analyzed by Markel and Gray in 1970s. In between 1970s and 1980s, the rapid growth in speech coders was done because of drastic boom in VLSI technology.

In the duration on 1980s–1990s, the low rate high-quality speech coders were planned to design. The invention of code excited linear prediction coding was major improvement in speech coders, CELP is originally proposed by M. R. Schroeder and B. S. Atal in 1985. CELP was capable of producing low rate speech for communication purpose [6, 7].

The concept for hybrid code is finalized with the use of different structured codebooks in CELP. An 8 kbps hybrid coder was first hybrid coder which was selected for North American Digital Telephone Network. The hybrid coders are also selected for satellite systems.

The research on this field is still going on and researchers continuously working on to increase the capacity of systems at minimal bandwidth.

3 Linear Predictive Coding

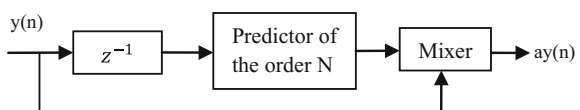
LPC technique is the most used technique in speech coding. LPC technique provides extremely accurate estimates of speech data sequence. Basic idea of linear prediction is that the current speech sample can be closely approximated as a linear combination of past samples, the block diagram for LPC filter is shown in Fig. 2.

The algorithm for LPC is given by the formula given below:

$$y(n) = \sum_{(i=1)}^N ay(n - i) \tag{1}$$

The efficient estimation of LP coefficients is based on the Levinson–Durbin algorithm which uses a forward and backward prediction for speech samples.

Fig. 2 Block diagram of linear predictive filter



The formula basically used for both forward and backward samples is:

$$r(i) = \sum_{n=0}^{N-1-i} s(n) \cdot s(n+i) \tag{2}$$

where r is a positive-definite matrix

$$r = [r(1)r(2) \dots r(n+1)]$$

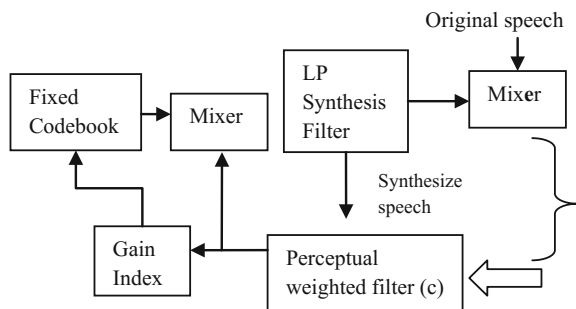
$$\begin{pmatrix} r1 & r2 & r(n) \\ r2 & r1 & r(n-1) \\ r(n) & r(n-1) & r1 \end{pmatrix} \begin{pmatrix} a1 \\ a2 \\ a(n+1) \end{pmatrix} = R(n) \tag{3}$$

4 Code Excited Linear Prediction (CELP)

Parametric coding is based on LP analysis, and CELP is standardized as parametric coder by CCITT in 1991 which is formally known as FS1016 CELP [8] and given the bit rate of 4.8 kbps. Advancement in this technique is observed in 1992, and another big achievement was successfully done by International Telecommunication Union (ITU), and another version of CELP is finalized as ITU-T G.728 LD-CELP with bit rate 16 kbps and abbreviated as low-delay code excited linear predictive coder (LD-CELP) which was designed to provide delay of less than 20 ms.

The block diagram in Fig. 3 shows the basic building block of CELP. In CELP, coder a fixed codebook is designed to provide initial code vectors for data bit comparison and hence the high quality of speech is attained at much lower bit rate than waveform coders; thus, the bandwidth is optimized as compared to waveform coders. The perceptual weighted filter is used to provide a fixed delay for each sample, and it is a constant value in between 0.1 and 0.9.

Fig. 3 Block diagram of CELP coder



5 Evaluation and Analysis

Analysis of 16 and 9.6 kbps CELP is done with the MATLAB simulating software version R2016a. The coder is designed to take audio speech samples at 8 kHz, and output is observed in 16 and 9.6 kbps. The ‘hello’ file is taken as input audio and ‘xhat1’ is decoded sound file in 16 kbps sampled format and ‘xhat2’ is 9.6 kbps decoded sound for CELP. Finally, the experiment is performed for a constant value of $c = 0.25$ and 0.65 . The original sound is standard sound in ITU-T test signal library. The signal sampled in 86,169 samples and the 50 LP coefficients are calculated randomly from 86,169 samples. Fig. 4 shows the LP estimation of original speech. Similarly, Figs. 5 and 6 show the comparison graph between 16 kbps CELP and 9.6 kbps CELP with $c = 0.25$.

6 Performance Comparison

SNR is abbreviated as signal to noise ratio, and MSE is abbreviated as mean square error estimation of speech signals. The two signals ‘xhat1’ (16 kbps) and ‘xhat2’ (9.6 kbps) CELP are compared with ‘hello’ original signal, and values for SNR and MSE are shown in Table 1.

Fig. 4 LP coefficient estimate

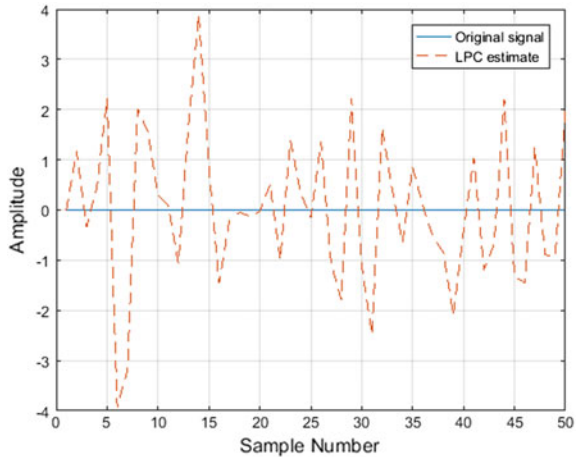


Fig. 5 Graph between 16 kbps CELP and original speech with $c = 0.25$

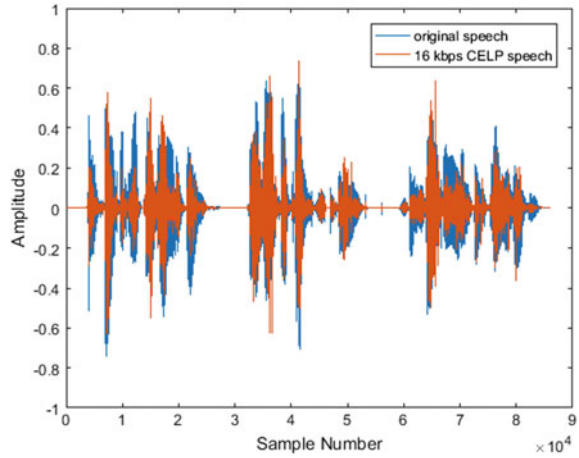


Fig. 6 Graph between 9.6 kbps CELP and original speech with $c = 0.25$

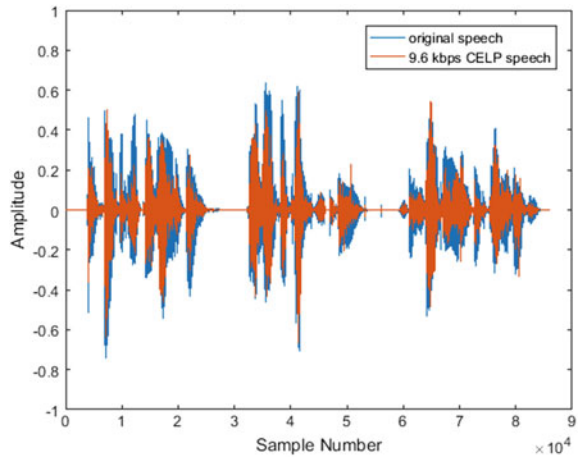


Table 1 SNR and MSE parameters for 16 kbps/9.6 kbps CELP coder

Perceptual weighted constant C	Different bit rates			
	16 kbps CELP		9.6 kbps CELP	
	SNR	MSE	SNR	MSE
0.65	1.2815	0.0047	0.2769	0.0175
0.25	4.2851	0.0016	3.5234	0.0132

Fig. 7 SNR of the 16 and 9.6 kbps CELP compared with original

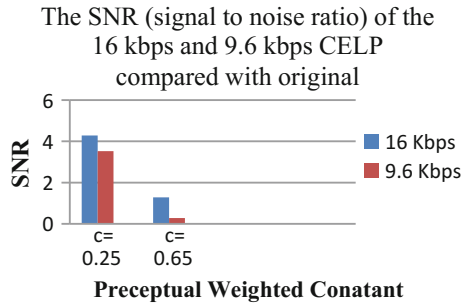
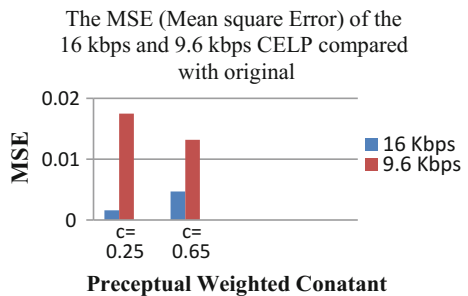


Fig. 8 MSE of the 16 and 9.6 kbps CELP compared with original



7 Result and Discussions

Here, a detail performance analysis of 16 kbps CELP and 9.6 kbps CELP with perceptual weighted value $c = 0.65$ and $c = 0.25$ is presented. This analysis is totally based on the value of perceptual weighted constant ‘ c .’ The factor ‘ c ’ is highly affecting factor for the quality of speech coders. These simulation-based comparative analyses illustrate the output speech quality in terms of signal to noise ratio (SNR) and mean square error (MSE), and from the comparison, it is clear that lower value of c is needed for better results. In Figs. 7 and 8, the graphical representation for both parameters is shown for various values of c .

8 Conclusion

From the above experimental test it is clearly analyzed that at lower the value of c better the SNR and similarly at lower values of c MSE is lesser. From both 16 and 9.6 kbps, the ratings for 16 kbps are better than 9.6 kbps. It is clearly shown that at $c = 0.25$ values for both SNR and MSE among both the better results are considered for 16 kbps CELP.

The 16 kbps CELP is a parametric coder and best for audio speech processing. The exponential growth in telecom field needs a better version of speech and video

processing both at same time for real-time implementation so research in this section is continuously growing and better quality of speech or video coder is implemented. The enhance voice services (EVS) and iLBC are example of it [9, 10].

References

1. Ogunfunmi, T., Narasimha, M.J.: Speech over VoIP networks: advanced signal processing and system implementation. In: IEEE Circuits and Systems Magazine, Second Quarter, pp. 35–55 (2012)
2. Ogunfunmi, T., Narasimha, M.J.: Principles of speech coding. CRC Press (2010)
3. Neuhoff, D.L., Philip, T.I.: Design of a CELP coder and analysis of various quantization techniques. Technical Report, University of Michigan EECS 651 (2005)
4. Ramo, A., Toukoma, H.: On comparing speech quality of various NB and WB speech coders. 0-7803-9243-4/05/, pp. 603–606, IEEE Multimedia Technologies laboratory, Nokia Research Center, Tampere, Finland (2005)
5. Edler, B., Schuller, G.: Perceptual audio coding using adaptive pre and post-filters and lossless compression. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), vol. 10, pp. 881–884 IEEE (2000)
6. Pettigrew, R., Cuperman, V.: Backward pitch prediction for low-delay speech coding. In: IEEE Global Tele-Communications Conference and Exhibition Communications Technology for the 1990s and beyond, (GLOBECOM), United State Patent, pp. 1–12, Patent Number: 4,963,034 (1989)
7. Pettigrew, R., Cuperman, V.: Robust low-complexity backward adaptive pitch predictor for low-delay speech coding. Proc. IEEE Commun. Speech Vis. **138**(4), 19–22 (1991)
8. Shetty, N., Gibson, J.D.: Perceptual pre-weighting and post-inverse weighting for speech coding. Technical Report, University of California (2007)
9. Dietz, M.: Overview of the EVS codec architecture. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP, 978-1-4673-6997-8/15/, pp. 5698–5602 (2015)
10. Qualcomm Inc., IWAENC.: Enhanced voice services. <http://www.iwaenc2014.org/>