IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)
Volume 6, Issue 3, Ver. II (May. -Jun. 2016), PP 65-69
e-ISSN: 2319 – 4200, p-ISSN No. : 2319 – 4197
www.iosrjournals.org

Voice Excited Lpc for Speech Compression by V/Uv Classification

Veena T K[1], Dr. D. Geetha[2]
[1]M.Tech 4Th Sem Reva Institute of Technology Bengaluru, Karnataka, India
[2]Professor, Dept of ECE Reva Institute of Technology Bengaluru, Karnataka, India

Abstract: *Speech coding is an important application of speech processing. Linear predictive coding (LPC) is the powerful speech coding technique used for encoding speech signals at a low bit rate. This method provides accurate estimation of parameters with less complexity. In this paper we discuss the implementation of plain linear predictive coding (LPC) voice coder and voice excited linear predictive (VELP) voice coder. Both of these voice coders are based on the principle of linear prediction where the current sample is predicted by the linear function of past values. VELP is an improved version of plain LPC voice coder. It is implemented by using DCT for coefficients to improve quality of speech. Simulation results of plain LPC and VELP are compared and we find that VELP produces better quality of signal than LPC.*

Keywords: *Linear Predictive Coding, Levinson Durbin Recursion, Quantization, Autocorrelation.*

I. Introduction

Speech coding is the act of transformation of speech signal in more compact form or representation of speech signal by using lesser number of bits. Speech compression is greatly implied in Long-distance communication, cellular technology, high quality speech storage and message encryption [1]

Speech compression or coding is achieved with the use of voice coders or vocoders. All vocoders use the fact that speech production occurs through slow anatomical movements and the speech produced has the limited frequency range of around 300 Hz to 3400 Hz. There are two main classes of voice coders: Waveform approximation coders and Model based coders. Waveform approximation coders involve exact reproduction of the original speech signal if there is no quantization error. Model based coders will never exactly synthesizes the original speech signal regardless of the presence of quantization errors because Model based coders uses parametric model of speech production , this method involves encoding and transmitting only the parameters but not the whole signal. The LPC vocoder using in this paper is also a Model based coder, which is lossy even if there is no quantization errors [1].

All voice coders have four main constraints: Bit rate, Delay, Complexity, Quality. Any voice coders will have to make trade off's between these four constraints regardless of the algorithm. Speech coders are analyzed using both Subjective and it uses Objective analysis. Subjective analysis involves making a judgment on the speech quality by listening to the output speech signal; the quality of the output speech signal will depend fully on listeners. Objective analysis involves measurement of speech quality by using different techniques and it minimizes human interference [2].

II. Speech Production Model

This section gives the summary of speech production and classification of speech. Understanding the process of speech production facilitates us in developing a kind of model for speech production. Having a speech production model we are going to extract the parameters of speech production model and use this parameters in order to have a proper synthesis of speech at decoder end. Fig 2.1 shows the human speech production model and voice coder speech production model [3].

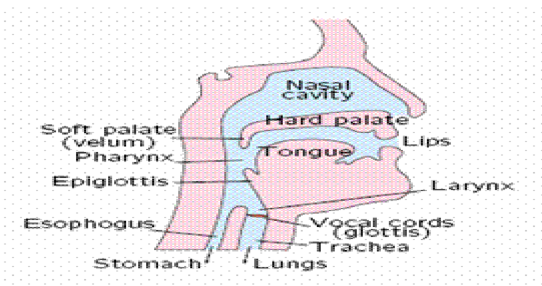

Fig.2.1 Human speech production

The sound is produced as the air from the lungs rush through the vocal tract and comes out of the mouth. Two kinds of sounds that are produced:

- Voiced speech is produced when the vocal cords vibrates. These are quasi-periodic like pattern. Ex: Voiced sounds /a/, /e/, /i/.
- Unvoiced speech is produced when vocal cords are kept open. These are noise like signal. Ex: Unvoiced sounds /s/, /f/ [4].

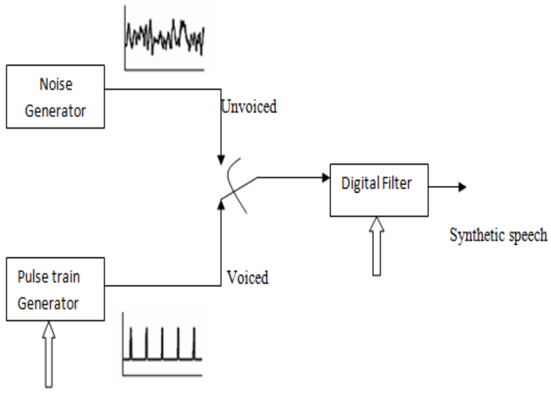

Fig.2.2 Speech production model

The vocal tract of human is modeled as all pole filter having the transfer function H(Z) as in equation (1) with gain G and q denotes the number of poles and a denotes the filter coefficients [1] and the speech production model is shown in figure 2.2:

$$H(Z) = \frac{G}{1-\sum_{k=1}^{q} a_k z^{-k}} \qquad (1)$$

III METHODOLOGY

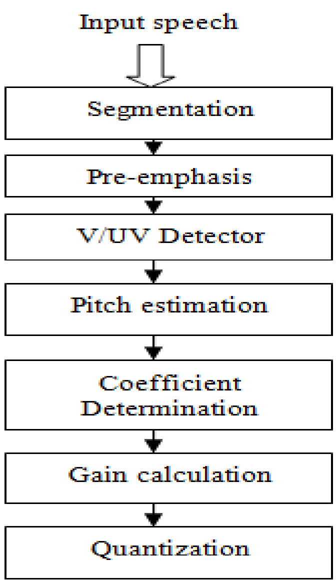

Fig.3.1 LPC Encoder

The flow of the encoding process of plain LPC voice coder is given in the figure 3.1. It consists of various functional block each one is explained in this section.

➢ Segmentation: The input speech signal which is in digital form having sampling rate 8000 samples/sec according to sampling theorem is first broken down into number of frames of definite intervals of 10 to 30 ms. In this project 20 ms interval is used. Windowing is performed by using rectangular window of length 30 ms. LPC process the signal frame by frame.

- Pre-emphasis: After windowing and framing the signal pre-emphasis filter is applied for each frame. Pre-emphasis filter is a high pass filter having the system function as in equation (2):

$$H(Z) = 1 - bZ - 1 \qquad (2)$$

Applying this filters flattens the signal spectrum and decreases the computational precision. In LPC the value of b is taken as 0.9378, it boosts the higher frequency.

- Voiced/Unvoiced Detection: The detection whether the frame is voiced or unvoiced is done by various methods. In this paper two features are used for this purpose that is short time energy and zero crossing count (ZCC). Short time energy is maximum for voiced frame and less for the unvoiced frame. A threshold is set for this purpose. The equation for short time energy is given in (3):

$$E_n = \sum_{m=-\infty}^{\infty} [x(m)W(n-m)]^2 \qquad (3)$$

Other feature is ZCC, it is the number of times the amplitude of the signal crossing x-axis. ZCC is more for unvoiced sound compared to voiced sound [4].

- Pitch Detection: After detecting the frame as V/UV the pitch of the voiced frame is calculated. LPC-10 uses average magnitude difference function (AMDF) for the calculation of pitch. The equation (4) is given for AMDF [2]:

$$AMDF(P) = \frac{1}{N} \sum_{i=k0+1}^{k0+N} |x_i + x_{i-p}| \qquad (4)$$

- Coefficient Determination: LPC finds the coefficients of the all pole filter by minimizing the squared difference between the present sample and the linear combination of previous samples. For the accurate estimation of the coefficients by maintaining stability of the filter LPC-10 uses Levinson-Durbin recursion algorithm.
- Gain Calculation: The gain of the coefficients obtained from levinson-durbin algorithm can be calculated by the equation (5):

$$G^2 = R(0) - \sum_{k=1}^{q} a_k R(k) \qquad (5)$$

- Quantization: The coefficients and speech parameters obtained are quantized, but instead of coding the filter coefficients the reflection coefficients obtained from the L-D algorithm are coded to ensure stability and coded into bit streams and transmitted to the decoder end where the decoder extracts the parameters and synthesizes the speech signal.

B) **Voice Excited LPC Voice Coder implementation**

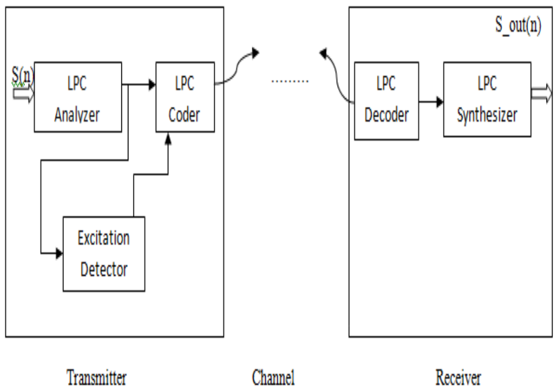

Fig.3.2 Voice Excited LPC Vocoder

The VELP voice coder block diagram shown in figure 3.2 is developed to increase the output speech quality as the plain LPC voice coder gives poor speech quality. The main difference between those two is the excitation, VELP uses a train of pulses while synthesizing the speech. Thus the Input speech signal in each frame is filtered with the estimated transfer function of LPC analyzer. The filtered signal is called residual. For a good reconstruction of the excitation only the low frequencies of the residual signals are coded. DCT is applied for the residual signal to get good quality of speech only the coefficients having high energy are transmitted to decoder in this case.

IV RESULTS

A.THROUGHPUT

MATLAB is used for the implementation and the results are obtained .The different parameters obtained are shown in the table1 and the number of bits allocated for each parameters is shown in table 2.

Table 1

Parameters	values
Sampling rate	8kHz
Frame size	180 samples
Frame rate	44.44 frames/second
Window size	30 ms
gain	RMS value
Assigned bit /frame	54
Order	10
Pitch	5 to 35
Bit rate	2400 bits/second

Table 2

Parameters	No of bits allocated
Pitch	7
Coefficients	41
Gain	5
Voicing decision	1
Total	54 bits/frame
Bit rate	54*44.44=2.4kbps

B. WAVEFORM ANALYSIS

The proposed algorithm is implemented in MATLAB tool. Waveforms of the original input speech signal is given in figure 4.1, the compressed signal of plain LPC method and VELP method are shown in figure 4.2 and 4.3 respectively, the comparison plot of gain, pitch and stream for orders 8, 10 and 12 for VELP vocoder is shown in figure 4.4, 4.5 and 4.6 respectively.

1.**Original input Speech Signal**

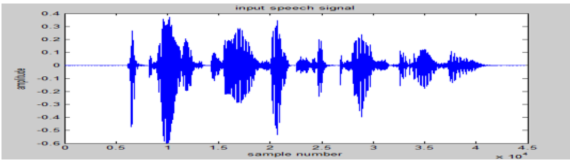

Fig.4.1 Original input signal

2. Synthesized Speech Signal of plain LPC vocoder

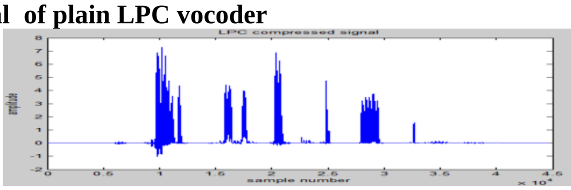

Fig.4.2 Waveform of LPC synthesized speech signal

3. Synthesized speech signal of VELP vocoder

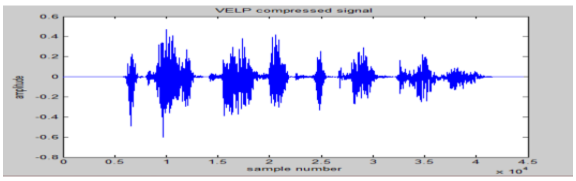

Fig.4.3 Waveform of Voice excited LPC synthesized speech signal

4. The gain for different orders 8,10 and 12 of VELP vocoder are obtained are shown below ;

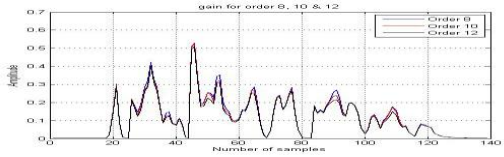

Fig 4.4 Gain in dB for different orders

5. Pitch for different orders 8, 10 and 12 of VELP are shown below:

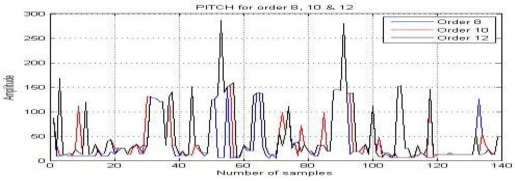

Fig 4.5 Pitch for different orders

5. Stream for different orders 8, 10 and 12 of VELP vocoder are shown below:

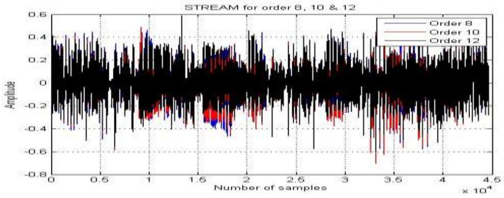

Fig 4.6 Stream for different orders

V. CONCLUSION

In this paper the plain LPC-10 and voice excited LPC voice coders are implemented. The results of which shows that the output obtained from plain LPC has poor sound quality and the output sounds more synthetic. To improve the speech quality we have developed VELP voice coder which makes precise estimation of pitch and uses train of pulses to synthesis the signal. DCT is used to increase the quality of output signal but at the cost of increase in bit rate which increases the bandwidth. Other coding methods like Huffmann coding can be used to get better results.

REFERENCES

[1] L R Rabiner and R W Schafer , "Digital Processing Of Speech Signal", Prentice-Hall, Eaglewood Cliffs ,NJ, 1978.

[2] Nikhil Sharma, Niharika Mehta, "Advanced Speech Compression VIA Voice Excited Linear Predictive Coding Using Discrete Cosine Transform(DCT)*"*, International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-2, Issue-3, pp.144-148, February 2013.

[3]. Mahmoud.A.Osman, Nasser Al, Hussein M. Magboub and S.AvAlfandi, "Speech compression using LPC and Wavelet", Volume 7, pp.92-99, IEEE transaction, 2010.

[4] Ykhlef Faycal and Messaoud Bensebti, "Comparative Performance Study of Several Features for Voiced/ Non-Voiced Classification", The International Arab Journal of Information Technology, Vol. 11, No. 3, May 2014.

[5] Landy Goldbarg and Lance Reik , "A practical handbook of Speech coding " , chapter 4,1-4 .

[6] K.Goyal, "Multiple description coding :compression meets the network,signal processing magazine", IEEE transaction , vol 18 no5, pp. 7493, 2001.

[7] W.Griffin and J.S.Lim, "Multiband excitation vocoder, Acoustic speech and signal processing", IEEE transactions on vol 36 no8, pp.1223-1235, 1998.

[8] Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of human perception", Proceedings of the IEEE, vol. 81, no. 10, pp. 1385–1422, 1993.

[9] S. Atal, "The history of linear prediction", IEEE Signal Processing Magazine, vol. 23, no. 2, pp.154–161. 2006.

[10] Aizawa and T. S. Huang, "Model-based image coding advanced video coding techniques for very low bit-rate applications", Proceedings of the IEEE, vol. 83, no. 2, pp. 259–271, 1995.

[11] L. Besacier, S. Grassi, A. Dufaux, M. Ansorge1, F.Pellandini, "GSM Speech Coding and Speaker Recognition", Proceedings of the International Conference on Acoustics, Speech, and SignalProcessing (ICASSP2000) 2, 1085-1088, pp. 1-3, 2000.

[12] Amol R Madane, Zalak Shah, Raina Shah, Sanket Thakur, "Speech Compression Using Linear Predictive Coding", International Workshop on Machine Intelligence Research, pp. 1-3, 2009.

[13] Johnson J., "Discrete Wavelet Transform Techniques in Speech Processing," in Proceedings *of* IEEE TENCON, Digital Signal Processing Applications, Australia, pp. 514-519, 1996.

[14] C. Hardwick and J. S. Lim, "A 4.8 kbps multi-band excitation speech coder", In Proceedings IEEE Int. Conf. on Acoustics, *Speech, Signal Processing*, pp. 374-377, New York, 1988.

[15] B. Almeida and J. M. Tribolet. "Harmonic coding: a low bit-rate good quality speech coding technique", In Proc. IEEE Int. Conf. on Acoustics, Speech, Signal Processing, pp. 1664-1667, Paris, 1982.