

# Experimental Analysis of Voice Ooze Algorithm using Voice-excited LPC Vocoder

Swati Raina  
 M. Tech. Student  
 Suresh Gyan Vihar University,  
 Jaipur, Rajasthan. India

Rohit Raina  
 Chief Technical Officer  
 MIET, Jammu,  
 J&K, India

Vinit Agarwal  
 Asst. Prof.  
 Suresh Gyan Vihar University,  
 Jaipur, Rajasthan. India

## ABSTRACT

This research paper examines the recent development in speech coding. Speech coding is a basic element of digital communications and constantly attracting attention due to the increase of power in telecommunication services and capabilities. This paper focuses on the design and implementation of a Voice-excited Linear Predictive coder at low bit-rate. And at last analyze result by using different voice examples. This research paper is divided into six divisions.

## General Terms

Voice-excited LPC Vocoder et. al.

## Keywords

Speech Coding, Communication, low-bit rate speech.

## 1. INTRODUCTION

In general, speech coding is a process to represent a digital speech signal with few bits, maintaining at the same time a reasonable level of speech quality. Speech coding plays a vital role in application area of signal processing. Now a day, speech communication and speech coding technology has received augmenting levels of interest from the research, standardization, and business communities.

Speech coding is performed using many different steps as an algorithm. In general, an algorithm is specified with a set of instructions. A computer or processor can execute them so as to complete the coding task. The instructions can also be translated to the structure of a digital circuit, carrying out the computation. The speech waveform is a voice generated from controlled movements of anatomical structures results into the human speech production system.

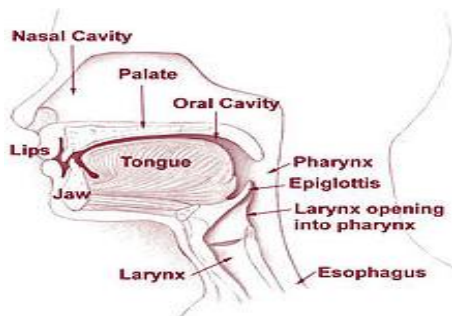


Figure 1: Human Speech Production System

Figure 1 shows the basic structure of speech production system [8]. Speech is produced when air is forced from the lungs through the vocal cord and along the vocal tract. Speech production in human can be crudely modeled as an excitation

signal driving a linear system. The filter co-efficient depend on shape of the vocal tract, which changes when producing different sound by changing the position of tongue and jaw. Vowels are produced when excitation is periodic pulse with different periods [4]. Consonants are generated when excitation signal noise like.

## 1.1 Speech Detection Technique

The system must know when a spoken word is input. Thus, a detection algorithm has to be devised. This can be done by continually computing the difference of the absolute average of two adjacent sound windows, and comparing it to a predefined threshold. The detector algorithm can be broken down as follows:

i. The absolute average  $w_1$  of a sound window of length  $W$  is computed from the sound samples is starting at  $s_a$  and ending at  $s_b$  as shown in Eq. 1.

$$w_1 = 1/W \sum_{i=s_a}^{s_b} |s_i| \text{ ----- (1)}$$

ii. The average of the second window  $w_2$  is computed from the sound samples  $s_i$  starting at  $s_b$  and ending at  $s_c$  as shown in Eq. 2.

$$w_2 = 1/W \sum_{i=s_b}^{s_c} |s_i| \text{ ----- (2)}$$

iii. The difference of  $w_1$  and  $w_2$  is then compared to the threshold value 'Th', if it is

$$\Sigma_i = a_i - b_i \text{ ----- (3)}$$

iv. The average of the oldest window ( $w_1$ ) is discarded, and replaced by  $w_2$ . Then, algorithm returns to step 2 again.

This value has been experimentally determined in the MATLAB implementation. Nevertheless, it may vary depending on the sound acquisition setup (i.e. position of the microphone, noise level, etc.). Finally, the length of the word is fixed to 1.024s for convenience [9].

## 2. VOICE-EXCITED LPC VOCODER

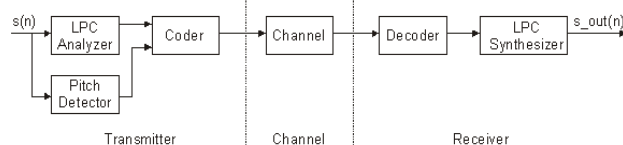


Figure 2: Block diagram of a voice-excited LPC vocoder.

The limitation of a plain LPC-10 vocoder [3] is based on methodology of voice excitation [1]. Here is one solution to improve the quality of the speech that the use of voice-excited LPC vocoders. Figure 2 shows a block diagram of a voice-excited LPC vocoder at low bit-rate [5]. The working of the

voice-excitation is to avoid the imprecise detection of the pitch and the use of an impulse train while synthesizing the speech [2]. One should rather try to come up with a better estimate of the excitation signal. Thus the input speech signal in each frame is filtered with the estimated transfer function of LPC analyzer.

### 3. PERFORMANCE ANALYSIS

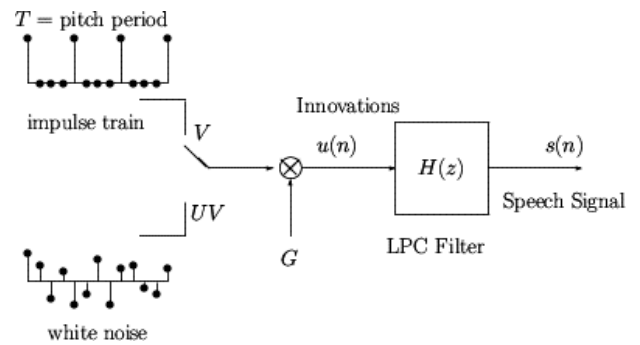
The segmental signal to noise ratio (SEGSNR) of the original speech file compared to the coded one and reconstructed speech file is measured using the provided Matlab function "segsnr". Then, the results which are obtained by this measurement are in Table 1: Scenario 1: A Female speaker saying: "It's easy to tell the depth of a well. Scenario 1: A Female speaker saying: "Glue the sheet to the dark blue background.

**Table 1. Plain LPC measures**

Vocoder type	SNR 1	SNR 2
Plain LPC	-24.85 dB	-23.94 dB

Note that, the calculation of the SNR(s) requires the signals to be normalized before the ratio can be calculated. A comparison of the original speech sentences against the LPC reconstructed speech and the voice-excited LPC methods were studied. In both cases, the reconstructed speech has a lower quality than the input speech sentences. Both of the reconstructed signals sound mechanized and noisy with the output of plain LPC vocoder being nearly unintelligible. The LPC reconstructed speech sounds guttural with a lower pitch than the original sound. The sound seems to be whispered. The noisy feeling is very strong. The voice-excited LPC reconstructed file sounds more spoken and less whispered. The guttural feeling is also less and the words are much easier to understand. Overall the speech that was reconstructed using voice-excited LPC sounded better, but still sounded muffled [6].

### 4. MATHEMATICAL MODEL OF SPEECH PRODUCTION



**Figure 3: Mathematical model of LPC Analysis.**

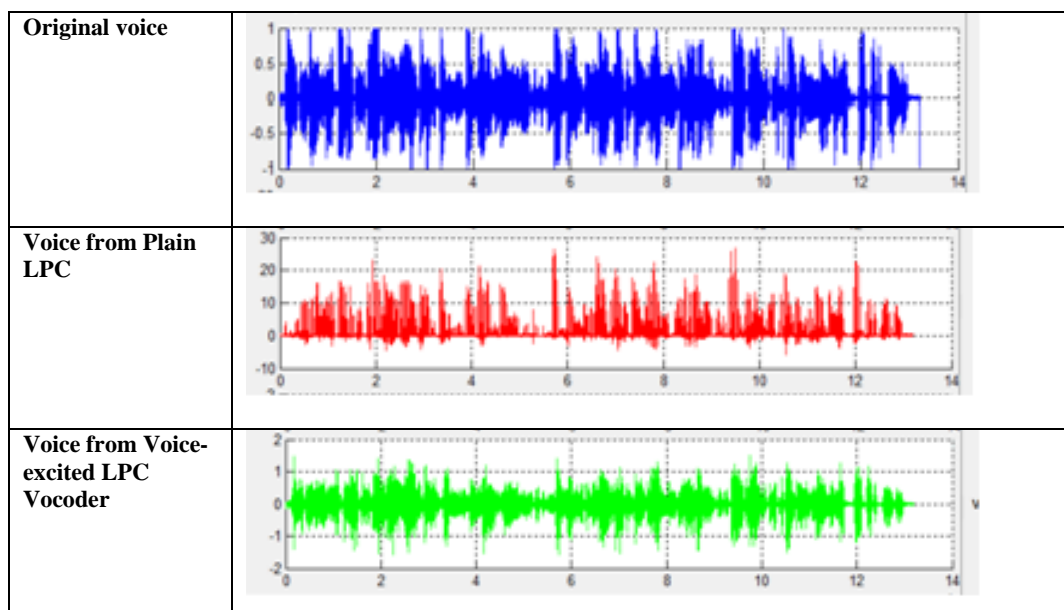
The mathematical model is basically used for different between voice and unvoiced frames. If frame is voiced an impulse train is generated with non zero (t) interval of the pitch period. If frame is unvoiced then white noise frame is generated and pitch period of zero. For estimation for pitch period the autocorrelation method is used. Impulse train is used to produce voiced sounds and the white noise sequence is used to produce unvoiced sounds.

### 5. EXPERIMENTAL ANALYSIS

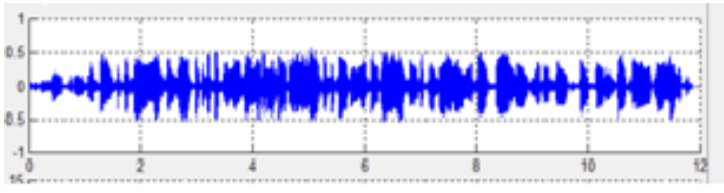
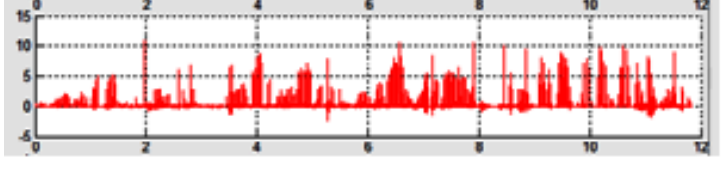
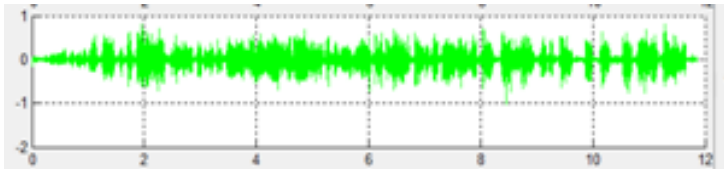
The results obtained from the algorithm built by us using voice excited LPC vocoder are pretty much clear and with very less distortion [5]. On the other side, the plain LPC are much poorer than that of this vocoder, accommodate signals other than speech such as music.

Some of the experimental observations with the different pattern of sounds using the built algorithm using voice-excited LPC vocoder are shown below:

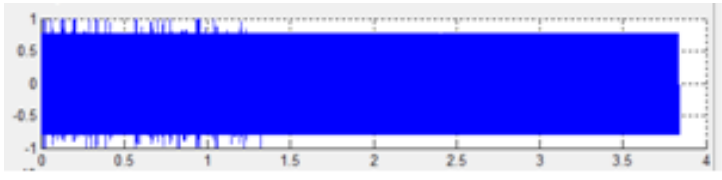
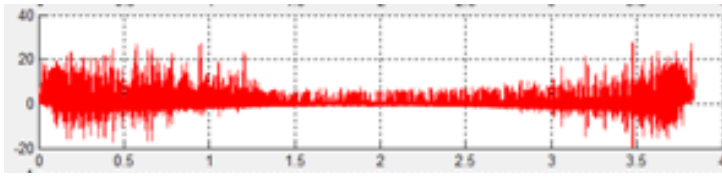
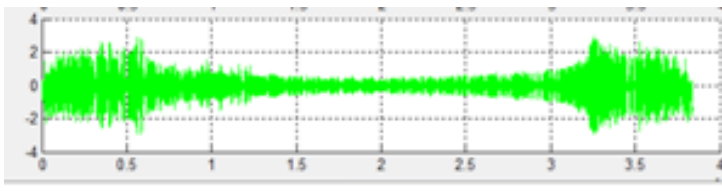
**Table 2: Male Voice**



**Table 3: Female mixed Voices**

<p><b>Original voice</b></p>	 <p>A blue waveform plot showing a complex, multi-frequency signal over a 12-second duration. The y-axis ranges from -1 to 1.</p>
<p><b>Voice from Plain LPC</b></p>	 <p>A red waveform plot showing a highly distorted and noisy signal over a 12-second duration. The y-axis ranges from -5 to 15.</p>
<p><b>Voice from Voice-excited LPC Vocoder</b></p>	 <p>A green waveform plot showing a clear, smooth signal over a 12-second duration, very similar to the original voice. The y-axis ranges from -2 to 1.</p>

**Table 4: Long Beep tone**

<p><b>Original voice</b></p>	 <p>A blue waveform plot showing a long, sustained, high-frequency tone over a 4-second duration. The y-axis ranges from -1 to 1.</p>
<p><b>Voice from Plain LPC</b></p>	 <p>A red waveform plot showing a highly distorted and noisy signal over a 4-second duration. The y-axis ranges from -20 to 40.</p>
<p><b>Voice from Voice-excited LPC Vocoder</b></p>	 <p>A green waveform plot showing a clear, smooth signal over a 4-second duration, very similar to the original voice. The y-axis ranges from -4 to 4.</p>

**6. CONCLUSION**

In Table 2, 3, and 4, the voices from a male vocal cord, mixed voices of two females, and a long beep tone respectively are used as to pass through the Plain LPC as well as through the algorithm using voice-excited LPC vocoder. From the above tables it is very much clear that the resultant output from Plain LPC is not clear and have many noise distortions added too. But the algorithm using voice-excited LPC vocoder give best

result with very much less distortion and is near about same as like the original voice [7].

From the experiments, it can also be examined that the output voice of long beep tone voice (like music) is not very much clear after passing it to the algorithm and this is the area where we need to work on in future.

## **7. REFERENCES**

- [1] David E. Schroeder M.R, Voice-excited Vocoder: IEEE Press, vol 8, Issue 5,1962.
- [2] Apeksha b Landge et al.,, “Wide band Speech Coding and It’s application,” in IJERST Vol 2 ,Issue 6 ,June 2013.
- [3] M.R. Schroeder “Vocoder: analysis and synthesis of speech”, proc. IEEE vol.54, no 3, pp 720-734, may 1966.
- [4] Prof. Dr. .Felbaurn., Human Speech Production Module.
- [5] B.S. Atal,” High Quality Speech at Low Bit-rate: multipulse and stochastically excited linear prediction codres” in proc. Int. Conf on Speech processing (Toronto, Canada).
- [6] Atal, B. S., Cuperman, V., Gersho, A., Advances in Speech Coding, Kluwer Academic Publishers, Boston
- [7] M. J. T. Smith and T. P. Barnwell, “Exact reconstruction techniques for tree-structured sub-band coders,” IEEE Trans Acousr., Speech, Signal Process., vol. ASSP-34, no. 3, p. 434, June 1986.
- [8] Klejin, W. B., Paliwal, K. K., Speech Coding and Synthesis, Elsevier, Amsterdam, 1995.
- [9] B.S. Atal and S. L. Hanauer,” Speech analysis and synthesis by linear prediction of speech” vol. 55, 1971.