

IMPROVE THE PERFORMANCE OF VOICE-EXCITED LPC VOCODER IN SPEECH CODING APPLICATION

Assistant Lecturer: Awwab Qasim Jumaah Althahab

Assistant Lecturer: Ahmed Hussein Shatti Alisawi

Electrical Engineering Department, Collage of Engineering, University of Babylon

Email: eng.awwab.qasim@uobabylon.edu.iq

Email: Eng_Ahmed.Shatti2014@yahoo.com

Received 3 August 2015

Accepted 21 October 2015

ABSTRACT

One of the fundamental problems in the area of digital speech processing is a speech coding that has been studied for years. Speech coding simply transforms the speech signals as fewer numbers of binary digits as possible, which can be then transmitted through channels or stored in memory devices. Due to the fact that the bandwidth of the channels is not unlimited, speech compression is needed to let more space bandwidth; thereby more speech coded signals can be sent over same channel bandwidth. Linear Predictive Coding (LPC) that is based on linear prediction (LP) model, which is a method to represent and analyze human speech, is one of the most common speech coding techniques. It is used in compression the digital speech signals, resulting low bit rate. This method has become the dominant technique for determine the fundamental speech parameters such as pitch, formants, spectra, vocal tract area functions. However, the weakness of LPC is in estimating the fundamental speech parameters causes poor voice quality and performance. The aim of this paper is to build a system with precise detection of speech parameters for encoding a better speech quality at low bit rate. This can be done through proposing a modified version to the voice-excited LPC vocoder based on Discrete Cosine Transform (DCT) and quantization of residual error while retaining low bit rate; hence conserve the bandwidth. Segmental power signal to noise ratio (SEGPSNR) and mean square error (MSE) as an objective measure for speech signal quality are implemented for the proposed improvement through computer simulation using Matlab 11.

Keywords: Voice-Excited LPC vocoder, Levinson-Durbin recursion, IIR filters, Discrete Cosine Transform (DCT), Mean square error (MSE).

تحسين اداء الترميز الصوتي التنبؤي الخطي المثار في تطبيقات تشفير الكلام

مدرس مساعد: احمد حسين شاطي العيساوي

مدرس مساعد: اواب قاسم جمعة الذهب

قسم الهندسة الكهربائيه, كلية الهندسه, جامعه بابل

الخلاصة

واحدة من القضايا الرئيسية في مجال معالجة الكلام الرقمي هو تشفير الصوت الذي درس لسنوات. يمكن تعريف تشفير الكلام ببساطة على انه تحويل اشارات الكلام الى اقل عدد ممكن من الارقام الثنائية والتي تنقل بعدها عبر القنوات او تخزن في اجهزة

الذاكرة. بسبب الحقيقة بان عرض الحزمة للقنوات الناقلة للإشارة محدودة، لذلك نحتاج ضغط الكلام لكي نسمح فراغات أكثر في النطاق الترددي وبالتالي المزيد من الكلام المشفر ممكن إرساله عبر نفس قناة النطاق الترددي. الترميز الخطي التنبؤي (LPC)، الذي هو طريقة لتمثيل وتحليل كلام البشر، هو واحد من أكثر تقنيات تشفير الكلام شيوعاً والتي تستخدم في ضغط إشارات الصوت الرقمية، مما ينتج عن انخفاض معدل بت الترميز. هذه الطريقة أصبحت تقنية سائدة في حسابات خصائص الكلام الأساسية مثلًا النغمة، الأصوات، الطيف ودوال الجهاز الصوتي. ولكن الضعف في حسابات خصائص الكلام يجعل (LPC) نوع ضياعي من أنواع الترميز ويسبب ضعف في أداء وجودة الصوت. الهدف من هذا العمل هو بناء نظام مع الكشف الدقيق لخصائص الكلام لترميز أفضل جودة كلام في انخفاض معدل بت الترميز. ويمكن القيام بذلك من خلال اقتراح نسخة محسنة من الصوت المثار للترميز التنبؤي الخطي (VELPC) بالاعتماد على تحويلة (DCT) و تكميم الخطأ المتبقي مع الحفاظ على انخفاض معدل بت الترميز، وبالتالي الحفاظ على عرض النطاق الترددي. نسبة قدرة الإشارة الى الضوضاء (SEGPSNR) ومتوسط مربع الاخطاء (MSE) كمقياس موضوعي لجودة اشارة الكلام نفذت للتحسين المقترح من خلال محاكاة الكمبيوتر باستخدام ماتلاب 11.

الكلمات المفتاحية: الترميز الصوتي التنبؤي الخطي المثار، المرشحات ذو الاستجابة الترددية الانهائية، تحويلة (DCT) (MSE)، متوسط مربع الاخطاء

1. INTRODUCTION

The fundamental purpose of signal compression techniques is to reduce the number of bits required to represent a signal (speech, audio, image and video) while keeping an acceptable signal quality for the purpose of reaching the communication system target of low bit rate transmission or message encryption (Marcelo and Valdemar, 2005). Speech coding is the process of transforming the speech signal at hand, to a more compact form, which can then be transmitted with a considerably smaller memory. The motivation behind this is the fact that access to unlimited amount of bandwidth is not possible. Therefore, there is a need to code and compress speech signals. For example, in digital cellular technology, many users need to share the same frequency bandwidth. Utilizing speech compression makes it possible for more users to share the available system. Another example where speech compression needed is in digital voice storage. For a fixed amount of available memory, compression makes it possible to store longer messages.

One of the most powerful speech analysis techniques is the method of linear predictive analysis that is uses (LPC) vocoder (Daniele et al., 2014). The speech features such as pitch, formants, spectra and vocal tract transfer function can be all estimated using this type of technique. The good extracted parameters result good reconstructing the speech signal and then more intelligible speech. The weakest link in most LPC vocoders is estimations and representations of the excitation functions (especially the pitch period of excitation signal) (Marcelo and Valdemar, 2005). In this vein, (Yugandhar and Satyapriya, 2013) presented and implemented three coding techniques (LPC, Waveform and sub-band coding), checking their performance measures such as compression ratio and speech audible quality. In (Minal and Sonal, 2014), the authors proposed a system to implement a model based design by using the linear prediction coefficients of the encoded speech data and prove to be the promising method for speech compression. (Jingyun et al., 2014) presented a linear prediction model that is based on first order norm. They proposed a method based on linear programming to calculate the parameters of the model and analyze the performance of the first order norm.

A great challenge in digital coding of signals is the development of methods for assessing the quality of reconstructed signals. In this paper, for more accurate estimation of speech parameters, a modification on the voice-excited LPC vocoder based on Discrete Cosine Transform (DCT) is proposed for coding two male wideband speech signals. The measure used for assessing the quality of signals may be classified into two general groups: subjective quality and objective quality measures. For the purpose of this paper, the speech coder developed is evaluated using the objective measure, which is based on a direct mathematical comparison between the original and

processed signals. The objective analysis that will be performed consists of computing segmental power signal to noise ratio (*SEGPSNR*) and mean square error (*MSE*) between the original and the coded speech signals. Furthermore, the effect of quantizing the residual error on bit rate is studied.

The remaining of the paper is organized as follows: Section 2 demonstrates the speech production model and how speech can be represented as the output of a linear time varying system. In section 3, the proposed modification to find the optimum estimation of speech parameters is discussed and presented. In section 4, the performance of the proposed approach is evaluated using computer simulation in MATLAB 11. The paper is finished with some concluded remarks.

2. SPEECH PRODUCTION

Figure (1) depicts the simple speech production model. Speech is produced when an air being pushed from the lungs through the vocal tract, and out through the mouth to generate speech. In this type of description the lungs can be thought of as the source of the sound and the vocal tract can be thought of as a filter that produces the various types of sounds that make up speech (*Lawrence and Ronald, 2009*). Speech signals consist of several sequences of sounds, which can be classified into voiced and unvoiced. The fundamental difference between these two types of speech sounds comes from the way they are produced. Voiced sounds are produced by vibrating the vocal cords due to the air comes from the lungs. The rate at which the vocal cords vibrate dictates the pitch of the sound. However, unvoiced sounds do not rely on the vibration of the vocal cords. The unvoiced sounds are created by the constriction of the vocal tract that is modeled as a linear all pole filter (infinite impulse response filter). The vocal cords remain open, and the constrictions of the vocal tract force air out to produce the unvoiced sounds (*Lawrence and Ronald, 2009*). Now, speech can be modeled as the output of a linear time varying system (IIR filter), excited by either quasi-periodic impulse train or white noise to generate various components of speech.

During the production of a given speech signal, the encoding process of LPC analyzer uses to successfully predict and estimate a set of accurate parameters for modeling the vocal tract (all pole filter). The predictor parameters determined by minimizing the residual error, which is the sum of the squared differences between the actual speech signal and the linearly predicted one over frames of a finite duration, which is normally 20 ms long (*John et al., 1999*). Only the predictor coefficients and residual error are sent instead of sending original speech signals. Decoding process involves using the error and predicted parameters received to build a synthesized version of the original speech signal. The transfer function of the time-varying digital filter is given by (*Lawrence and Ronald, 2009*).

$$H(z) = \frac{G}{1 - \sum_{k=1}^P a_k z^{-k}} = \frac{S(z)}{U(z)} \quad (1)$$

where G, P, a_k are the gain, order and parameters of IIR filter. Only the first P^{th} coefficients are transmitted to the LPC synthesizer. The most common methods used to determine the coefficients are the covariance and the auto-correlation methods. For our implementation, the auto-correlation method will be used. The reason is that this method is superior to the covariance method in the sense that the roots of the polynomial in the denominator of the above equation is always guaranteed to be inside the unit circle; hence the system $H(z)$ is for sure stable (*Awwab, 2013; Carlo et al., 2009*). The Levinson Durbin algorithm will be used in our simulation to compute the required parameters for the auto-correlation method.

3. THE PROPOSED MODIFICATION ON VELPC MODEL

The classic approach to analyze human speech based on LPC showed poor sound quality, the voice excitation is the weakest portion of this method (Yi and Philipos, 2008). Therefore, voice-excited linear predictive coding (VELPC) is one approach to get better sound quality. A system of this type has been studied by (Thomas and Abeer, 2011). **Figure (2)** shows a block diagram of VELPC with excitation detector. The proposed modification on the model shown in **Figure (2)** and will be made in simulation is to use a pre-emphasis filter. It is used to make the spectrum as flat as possible by boosting the high frequencies in order to get a better result for estimation of the predictor parameters. Obviously, the predictor coefficients corresponding to higher frequencies can be better estimated. This kind of treatment is within the reconstruction part of the speech signal.

The input speech signal that is divided over finite duration of times (frames) is filtered by the estimated transfer function of linear predictive coding analyzer. The output of the analyzer is called the residual (error signal) that is sent with the predictor coefficients to the receiver. Consequently, a very good speech quality can be achieved. However, the trade off paid of this system is a high bit rate; therefore, one solution to reduce the bit rate to 16 kbits/sec is to use Discrete Cosine Transform (DCT) to the residual error. The fact behind the use of DCT is only the low frequencies of the residual signal are needed in order to maintain a good reconstruction of the excitation. The DCT concentrates most of the signal energy in the first few coefficients that will be then sent to achieve a high compression rate. Another process that will be executed in our simulations is shown that those DCT coefficients could be quantized using 4, 6 and 8 bits instead of 16 bits which is the original representation. The quantized process is based on the partial reflection coefficients (PRC), which are the average values during the calculation of the well-known Levinson-Durbin recursion. Finally, the receiver simply performs an inverse DCT and uses the resulting signal to excite the voice. From equation (1) and from the concept of speech production model, where current speech sample $s[n]$ is approximated as a linear combination of past samples:

$$s[n] = \sum_{k=1}^P \alpha_k s[n - k] + Gu[n] \tag{2}$$

where $u[n]$ is voiced or unvoiced sounds and n is sample index. A linear predictor with prediction coefficients α_k is define as a system whose output is

$$\tilde{s}[n] = \sum_{k=1}^P \alpha_k s[n - k] \tag{3}$$

The prediction error (excitation) is the difference between the observed and predicted signals, and it is assumed to be independent and identically distributed process (*i.i.d*) (John, 1975).

$$e[n] = s[n] - \tilde{s}[n] \tag{4}$$

Substituting (3) into (4); yields

$$e[n] = s[n] - \sum_{k=1}^P \alpha_k s[n - k] \tag{5}$$

Let now suppose that the prediction error filter can be represented as

$$A(z) = 1 - \sum_{k=1}^P \alpha_k z^{-k} \tag{6}$$

Then

$$E(z) = S(z) - S(z) \sum_{k=1}^P \alpha_k z^{-k} = S(z)[1 - \sum_{k=1}^P \alpha_k z^{-k}] = S(z)A(z) \tag{7}$$

If the coefficients of the predictor (α_k) converges exactly to (a_k), the error becomes

$$e[n] = Gu[n] \tag{8}$$

Or

$$E(z) = GU(z) \tag{9}$$

But from equation (7), $E(z) = S(z)A(z)$, substituting (7) into (9) and reallocate the terms, yields

$$A(z) = G \frac{U(z)}{S(z)} \tag{10}$$

From (1) and (10), we can conclude that

$$H(z) = \frac{G}{A(z)} \tag{11}$$

Equation (11) shows that the prediction error filter $A(z)$ (it is also called the analysis filter) is the inverse filter of the system $H(z)$, the synthesis filter. The optimization problem aims to find an estimate of the prediction coefficients from a set of observed real samples such that the prediction error is minimized (*Stephen and Lieven, 2004*). The resulting values are then assumed to be the parameters of the system function $H(z)$ which will be then used for synthesizing speech segments. To minimize the error, let $[x_1, x_2, \dots, x_m]$ set of past values of speech signal $s[n]$ are given; where m is the order of prediction error filter and equation (3) can be written as

$$\tilde{s}[n] = \sum_{k=1}^m \alpha_k s[n - k] = \alpha_1 s[n - 1] + \alpha_2 s[n - 2] + \dots + \alpha_m s[n - m] \tag{12}$$

$$\tilde{s} = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_m x_m = \sum_{i=1}^m \alpha_i x_i \tag{13}$$

According to (13), the prediction error is now written as:

$$err = s - \tilde{s} = s - \sum_{i=1}^m \alpha_i x_i \tag{14}$$

To find the predictor coefficients, the first order derivative is taken with respect to the predictor coefficient α_k to the mean squared error and equating the result to zero as in (15).

$$\frac{\partial MSE}{\partial x_k} = E\{2[s - \sum_{i=1}^m \alpha_i x_i](-x_k)\} = 0 \tag{15}$$

Rearenging (15), yields

$$E\{sx_k\} = E\{x_k \sum_{i=1}^m \alpha_i x_i\} \tag{16}$$

Setting $k = 1, 2, \dots, m$ and define covariance $R_{ji} = E[x_i x_j^*]$, and $R_{oj} = E[sx_j^*]$, (16) will be written as

$$\begin{bmatrix} R_{o1} \\ R_{o2} \\ \vdots \\ R_{om} \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1m} \\ R_{21} & R_{22} & \dots & R_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ R_{m1} & R_{m2} & \dots & R_{mm} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} \tag{17}$$

Equation (17) can be solved using the so-called Levinson-Durbin algorithm (*George, 1980*). In our simulations to the proposed scheme, objective measure of speech quality takes a mathematical criterion to analyze the performance and compare the origin with the reconstructed speech signals. Segmental Power SNR is calculated by first measure the SNR of each frame, then, take the average during the speech and it is defined in equation (18). Also, the mean square error (MSE) is calculated and defined by the equation (19).

$$SEGPSNR = \frac{10}{M} \sum_{k=1}^M \log_{10} \frac{\sum_{n=1}^N s^2(k,n)}{\sum_{n=1}^N err^2(k,n)} \quad (18)$$

and

$$MSE = \frac{1}{N} \sum_{n=1}^N |err(n)|^2 \quad (19)$$

where N is the frame length, M is the number of frames, $s(k, n)$ is the original speech of the n^{th} points of the k^{th} frame, and $err(k, n)$ is the residual error of the n^{th} points of the k^{th} frame (*Colin and Rainer, 2011*).

4. PERFORMANCE ANALYSIS AND SIMULATION RESULTS

The speech signals that will be coded are wideband signals. We utilize a modified version of voice-excited linear predictive coding (MVELPC) to code 2 males with the same phrase sentence (Welcome to The University of Babylon). The typical set of parameter values that have been used in our simulation results is as follows: the bandwidth of the speech signals is 8 kHz, the sampling frequency has to be at 16 kHz with a maximum end-to-end delay of 100 ms and length of each frame is 20 ms which results in 320 samples per frame. For perfect reconstruction, the overlapping length have to be 10 ms; hence the actual frame length is 30 ms which contains 480 samples, resulting 50 frames per second. The bit rate of original speech is at 250 kbps whereas the bit rate of synthesized speech is calculated in **Table (1)**.

Take into consideration the variation on the number of bits in quantization process and then multiply the result by 50 (number of frames per second); the bit rates are finally obtained. **Figure (3)** shows the original and the reconstructed speech signals based on MVELP vocoder with different quantized representation to the residual error (different No. of bits in quantization process). As can be seen from the figure, the reconstructed signal has a lower quality than the original signal when 4 bits used since there is a clear difference in shape between them. However, the similarity increases when the number of quantizing bits is raised, but does not sound exactly like the original speech signal.

Segmental PSNR and MSE are measured as can be seen in **Table (2)** and **(3)** for wave files (ahmed.wav) and (ali.wav), respectively. It is obvious from the values estimated that the reconstructed signals based on classic LPC have been very noisy since having very low SEGPSNR and high MSE. Meaning, the noise is stronger than the actual signal. However, the reconstructed speech signal that is based on the MVELPC vocoder sounds far better and its SEGPSNR is good enough. The SEGPSNR is increasing while MSE is decreasing when the number of bits is rising from 4 to 8 in the quantization process to the residual error whereas maintaining low bit rate, not exceeding 16 kbps. The VELPC vocoder is also implemented and the results obtained, as can be seen in **Table (2)** and **(3)**, are superior than both the classic LPC and MVELPC, but it demands a very high bit rate. This explains why the proposed system is desirable compared to what is achieved from other types of speech coding techniques such as VELPC, waveform and Subband Coders that require a very high bit rate for transmission.

A frame number 20 is selected from wave file (ahmed.wav), plus 240 points hamming window. The frequency response is estimated and plotted of the eight-pole LPC vocoder and MVELPC vocoder that is based on a pre-emphasis filter as shown in **Figure (4)**. Also, at the same figure and take 512 Fast Fourier Transform (512-FFT), the power spectral density (PSD) of the original speech signal is determined and plotted. As can be seen from **Figure (4)**, the spectrum of eight-pole MVELPC vocoder is good at peaks and troughs. The first, second, third, fourth and fifth peaks are good for fitting the original speech. However, the spectrum of eight-pole LPC vocoder shows poor performance for fitting the original speech at the peaks and troughs. **Figure (5)** illustrates the original speech signal with the reconstructed one for frame number 20 that is selected from wave file (ahmed.wav). The reconstructed one is shown to be very close and at most match the original signal.

5. CONCLUSION

In this paper, we presented a modified model vocoder that is based on voice-excited LPC to compress 2 male speech signals while maintaining low bit rate. The results have been achieved from the MVELPC are intelligible and desirable since the coder almost keeps perceptual relevant spectral characteristics of the speech signal. Also, high SEGPSNR and low MSE gained compared with the values obtained from classic LPC. The tradeoffs between speech quality on one side and bandwidth, the bit rate and complexity on the other side have analyzed and clearly appeared here. A better quality can be achieved by increasing the bit rate through an increase in bits used to quantize the DCT coefficients, causing larger bandwidth have to be used as shown in **Table (2)** and **(3)**. On the other hand, the classic LPC results are much poorer, and they are unintelligible and ineligible. Due to the fact that the MVELPC vocoder used gives pretty good results with the entire required limitations, particularly bit rate, the model can be more studied and improved.

REFERENCES

- [1] Awwab Q. Althahab, 2013 "*Performance Analysis of Adaptive Blind Equalization Algorithms for Noisy FIR and IIR Channels*" M.S. Thesis, University of Colorado.
- [2] Carlo Magi, Jouni Pohjalainen, Tom Backstrom and Paavo. Alku, 2009 "*Stabilized Weighted Linear Prediction*" *Speech Processing*, Vol. 51, No. 5, pp. 401-411.
- [3] Colin Breithaupt and Rainer Martin, 2011 "*Analysis of the Decision-Directed SNR Estimator for Speech Enhancement with Respect to Low-SNR and Transient Conditions*" *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 19, No. 2, pp. 277-289.
- [4] Daniele Giacobello, et al, 2014 "*Stable 1-Norm Error Minimization Based Linear Predictors for Speech Modeling*" *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 22, No. 5, pp. 912-922.
- [5] George Cybenko, 1980 "*The Numerical Stability of the Levinson-Durbin Algorithm for Toeplitz Systems of Equations*" *SIAM J. Sci. and Stat. Comput.*, Vol. 1, Issue 3, pp. 303-319.
- [6] Jingyun Xu, Xiaoqun Zhao, Qiao Wang and Digang Wang, 2014 "*Linear Prediction Analysis of Speech Signal Based on L_1 Norm*" *journal of Computational Information Systems*, Vol. 10, No. 17, pp. 7553-7560.

- [7] John Makhoul, 1975 "**Linear Prediction: A Tutorial Review**" Proc. IEEE, Vol. 63, No. 4, pp. 561-580.
- [8] John R. Deller, John G. Proakis and John H. L. Hansen, 1999 "**Discrete-Time Processing of Speech Signals**" Wiley-IEEE Press.
- [9] Lawrence Rabiner and Ronald W. Schafer, 2009 "**Theory and Application of Digital Speech Processing**" Prentice-Hall Inc.
- [10] Marcelo S. Alencar and Valdemar C. da Rocha Jr, 2005 "**Communication Systems**" Springer.
- [11] Minal Mulye and Sonal K. Jagtap, 2014 "**Speech Compression using Analysis by Synthesis**" IJECCE., Vol. 5, Issue 4, pp. 275-280.
- [12] Stephen Boyd and Lieven Vandenberghe, 2004 "**Convex Optimization**" Cambridge University Press.
- [13] Thomas Drugman and Abeer Alwan, 2011 "**Joint Robust Voicing Detection and Pitch Estimation Based on Residual Harmonics**" Proc. Interspeech, pp. 1973-1976.
- [14] Yi Hu and Philipos C. Loizou, 2008 "**Evaluation of Objective Quality Measures for Speech Enhancement**" IEEE Trans. On Speech, Audio and Language Processing, Vol. 16, No. 1, pp. 229-238.
- [15] Yugandhar Dasari and K. Satyapriya, 2013 "**Performance Analysis of Speech Coding Techniques**" IJAREEIE., Vol. 2, Issue 11, pp. 5725-5732.

Table (1): Bit rate calculation for the proposed model including DCT.

Parameters	Number of bits for each frame
Predictor coefficients=8	8*8=64 bits
Number of DCT coefficients=30	30*(number of bits in quantizing process of residual error)
Gain of the predictor	5 bits
Total number of bits for each frame	?

Table (2): Simulation results for speech file (ahmed.wav).

Type of vocoder	No. of predictor coefficients	No. of bits in quantization process of residual error	SEGPSNR in dB	MSE	Bit rate based on Table (1) (bits/sec)
Modified version of VELPC with 30 DCT coefficients	8	4	2.9403	0.0141	9450
	8	6	4.1106	0.0084	12450
	8	8	5.0021	0.0077	15450
VELPC	8	---	10.8479	0.0026	195450
Classic LPC	8	---	1.2687	0.1159	3800

Table (3): Simulation results for speech file (ali.wav).

Type of vocoder	No. of predictor coefficients	No. of bits in quantization process of residual error	SEGPSNR in dB	MSE	Bit rate based on Table(1) (bits/sec)
Modified version of VELPC with 30 DCT coefficients	8	4	2.2512	0.0164	9450
	8	6	4.1547	0.0081	12450
	8	8	5.0952	0.0074	15450
VELPC	8	---	12.6146	0.0014	195450
Classic LPC	8	---	1.2266	0.1564	3800

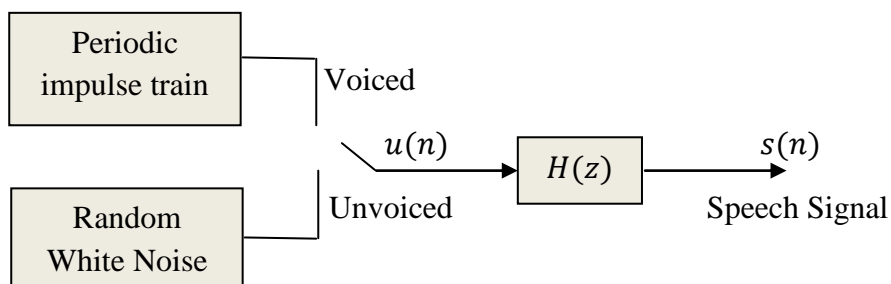


Figure (1): Simple speech production system

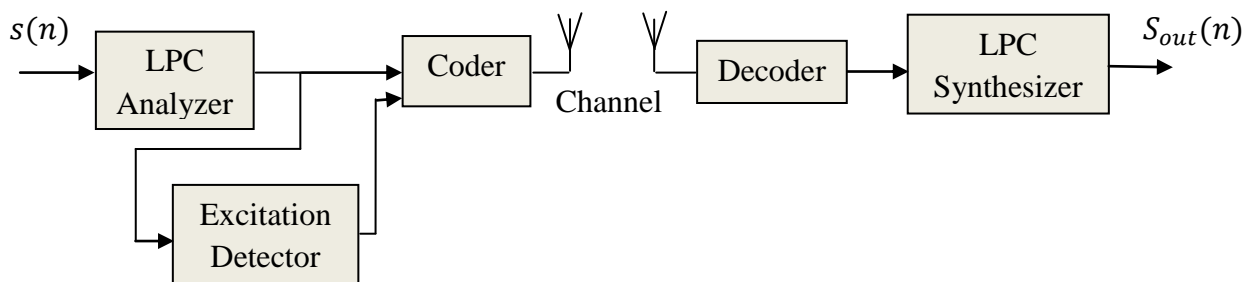


Figure (2): A voice-excited LPC system.

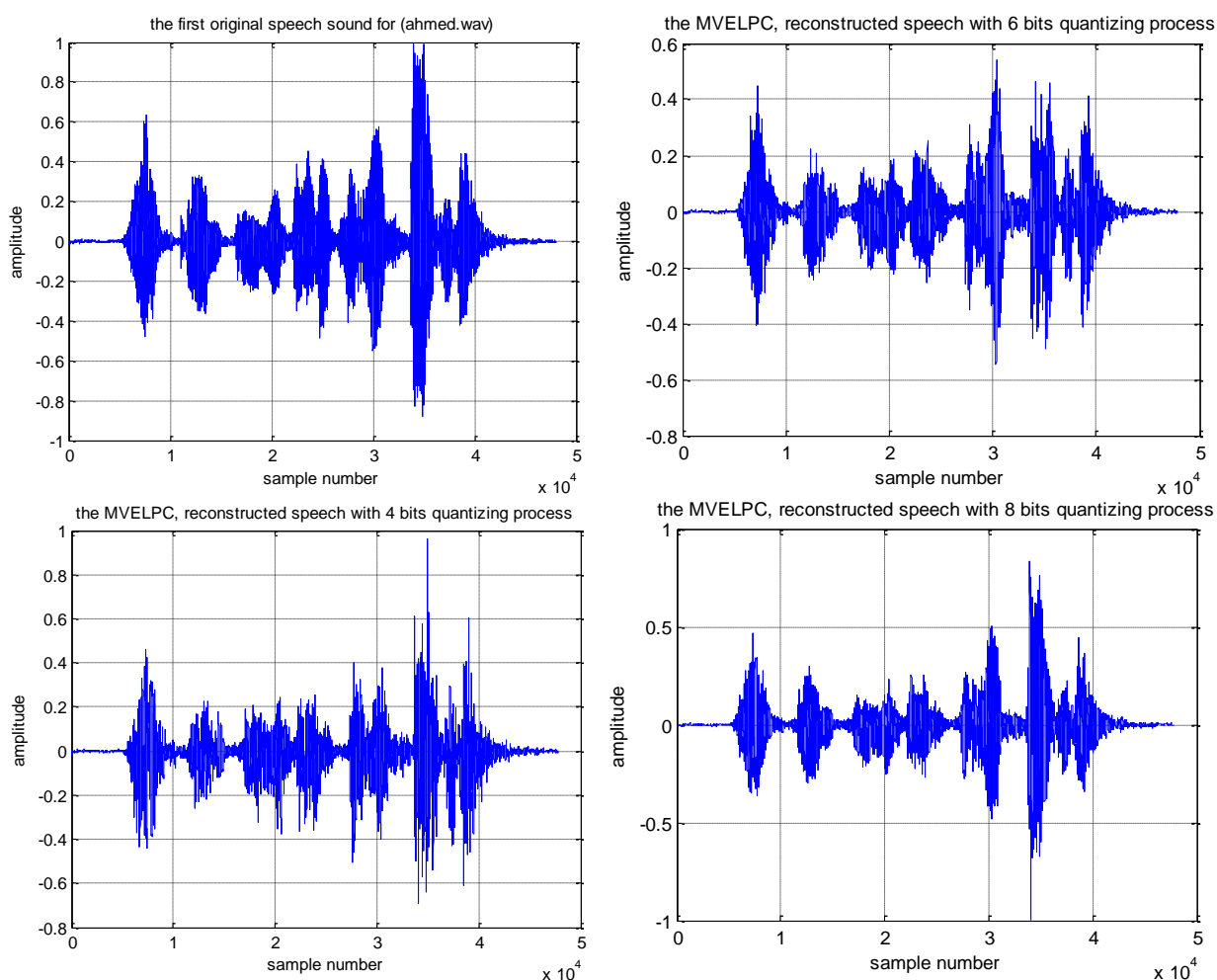


Figure (3): Waveform of the sentence "Welcome to the University of Babylon", the original and reconstructed speech signals.

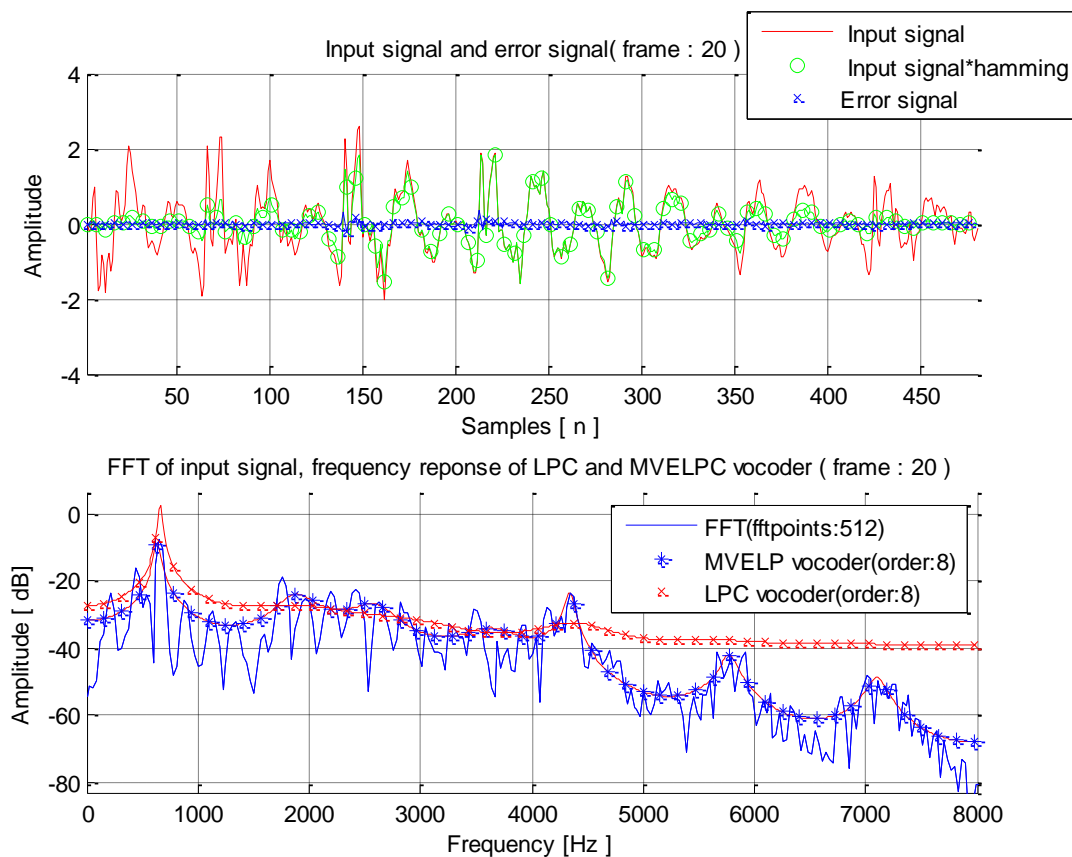


Figure (4): PSD of input speech, frequency response of LPC and MVELPC vocoder.

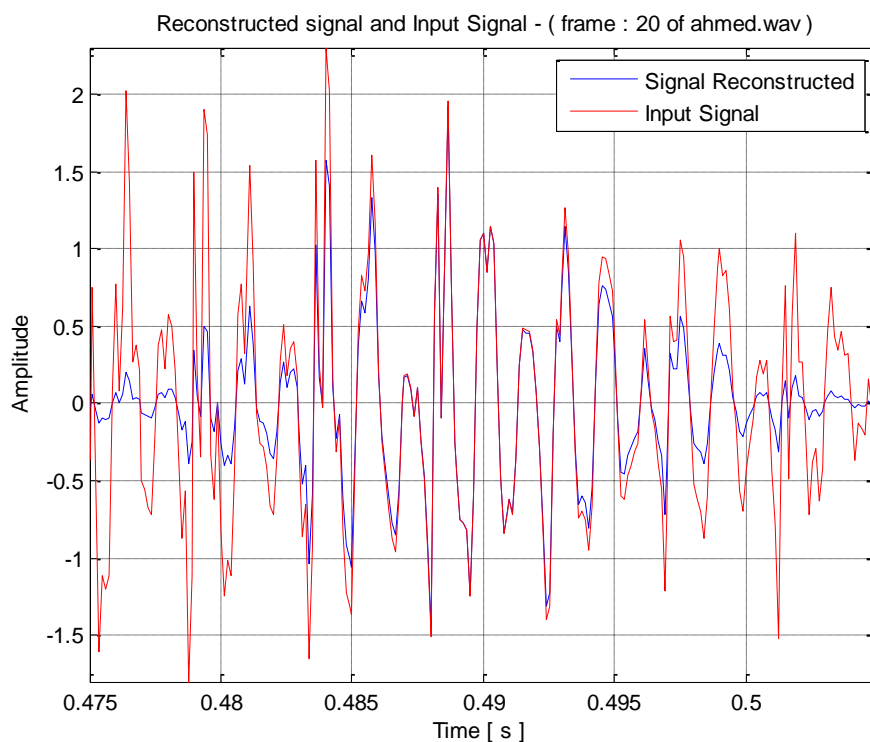


Figure (5): Original and reconstructed speech signal (frame No. 20 of ahmed.wav).