Dirasat, Engineering Sciences, Volume **33**, No. **2**, 2006

A New Architecture Model for Multi Pulse Linear Predictive Coder for Low-Bit-Rate Speech Coding

*Ibrahim M. Mansour and Samer J. Al-Abed **

ABSTRACT

Speech coding is a very important area that finds civilian and military applications. It deals with the problem of reducing the bit rate required for the speech representation while preserving the quality of the speech which reconstructed from that representation.

In this paper, we focused on developing algorithms and methods for a waveform speech coder operating at low bit rate with high quality reconstructed speech signal. Therefore, a new model for linear predictive coding of speech that can be used to produce high quality speech at low data rate is introduced. In this model, the residual (excitation signal) encoding is based on selecting the important pulses in the residual signal rather than encoding all pulses. Hence, this vocoder forms an excitation sequence which consists of multiple non-uniformly spaced pulses. During analysis, both the amplitude and location of the pulses are determined. In addition, this paper involves new techniques in the process of modeling and encoding of the amplitude and location of each pulse as well as linear prediction parameters.

KEYWORDS: Linear Predictive Coding, speech analysis, speech coding.

1. INTRODUCTION

In model-based source coding, the source is modeled as a linear system (filter) that, when excited by an appropriate input signal, results in the observed source output. Instead of transmitting the samples of the source waveform to the receiver, the parameters of the linear system are transmitted along with an appropriate excitation signal. If the number of parameters is sufficiently small, the model-based methods provide a large compression of the data. The most widely used model-based coding method is called Linear Predictive Coding. In this, the sampled sequence, denoted by s(n), n=0, 1, …,N-1, is assumed to have been generated by an all-pole (discrete-time) filter.

Linear Predictive Coding (LPC) method is one of the most important speech coding techniques for low bit rate speech coding, since it has been so widely studied and applied (Alku and Bäckstrōm, 2004; Atal, 1982; Barnwell, 1980; Brinker et al., 2004; Viswanathan et al., 1982; Chong and Cox, 2003; Galand et al., 1986; Gray and Wong, 1980; Härmä, 2001; Hu Hwai-Tsu and Wultsi-Tsung, 2000; Johansson, 1986). It was initially developed in the late 1960's. LPC is a type of speech coding that can predict the signal from the past samples and depends on the autocorrelation function (Proakis and Manolakis, 1996), (Spanias, 1994: 1541-1582). Linear Predictive Coding method provides an efficient way of coding the vocal tract filter information. This method for speech analysis and synthesis is based on modeling the vocal tract by a time varying all poles model and manages to remove a lot of redundancy in the speech signal (Spanias, 1994: 1541-1582), (Papamichalis, 1987). The remaining signal must then be coded in some way and there are several methods that give varying quality but also different bit-rates.

To summarize, the whole model can be decomposed into the following two parts, the analysis part as shown in Fig. 1 (a) and the synthesis part as shown in Fig. 1 (b). At the encoder, the speech encoder must determine the filter coefficients and the excitation signal for each segment.

The speech signal is first filtered by the analysis filter $A(z)$ and the output is the error signal. Therefore, the encoder (analysis part) analyzes the speech signal and

* Department of Electrical Engineering, Faculty of Engineering and Technology, University of Jordan.
Received on 15/6/2005 and Accepted for Publication on 2/10/2006.

produces the error signal. At the receiver, the decoder (synthesis part) takes the error signal as an input. The input is filtered by the synthesis filter $1/A(z)$, and the output is the speech signal (CCITT Recommendation G. 721, 1984).The error signal $e(n)$ is sometimes called the residual signal or the excitation signal. If the error signal from the analysis part is not used in synthesis, or if the synthesis filter is not exactly the inverse of the analysis filter, the synthesized speech signal will not be the same as the original signal. To differentiate between the two signals, we use the notation $s'(n)$ for the synthesized speech signal.

2. SOURCE FILTER MODEL

Speech signals are non-stationary, and at best they can be considered as stationary over short segments (typically 5-40 ms). Hence, the human speech production process is characterized by two factors: the source excitation and the vocal tract shape. In order to model speech production we have to model these two factors (Hu Hwai-Tsu and Wu His-Tsung, 2000; Spanias, 1994; Papamichalis, 1987; Ingle and Proakis, 2000). The vocal tract is modeled as an all-pole transfer function H(z). The transfer function of the all-pole model is

$$H(z) = \frac{G}{[1 + \sum_{k=1}^{p} a_k(k).z^{-k}]} \qquad (1)$$

Where P is the order of the linear prediction filter or the number of poles, G is the filter gain and $\{a_k\}$ are the parameters that determine the poles (filter coefficients). Filter coefficients (a_k) are called the linear prediction coefficients.

The vocal tract model $H(z)$ is excited by a discrete time glottal excitation signal $e(n)$ to produce the speech signal $s(n)$ as shown in Fig. 1 (b).

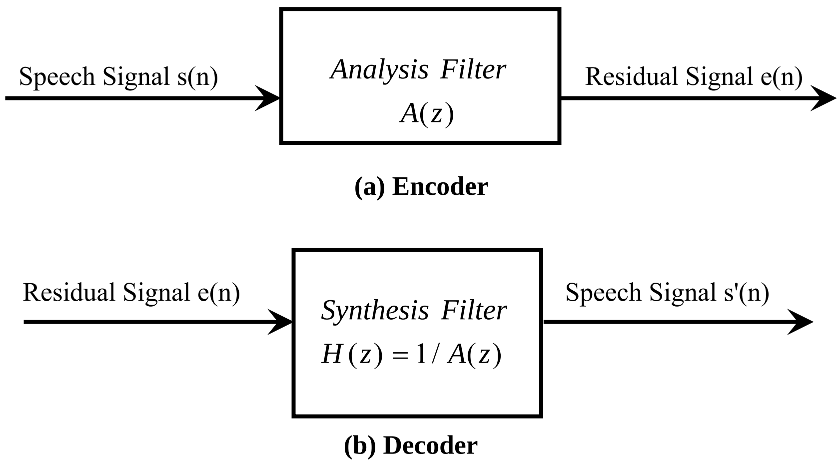

Figure 1. Linear Prediction Analysis and Synthesis Model (a) Encoder, (b) Decoder.

3. BASIC LINEAR PREDICTIVE MODEL

A simplified view of speech is that it consists of two types of sounds, voiced (vowels) and unvoiced (consonants). Consequently, Voiced speech is produced by exciting the vocal tract with quasi-periodic glottal air pulses generated by the vibrating vocal tract chords (Spanias, 1994: 1541-1582). The frequency of the periodic pulses is referred to as the fundamental frequency or pitch (Spanias, 1994: 1541-1582), (Papamichalis, 1987), (Hu et al., 2001). Unvoiced speech is produced by forcing air through a constriction in the vocal tract. In other words, this system assumes a two-state excitation (impulse-train for voiced and random noise for unvoiced speech). We can represent the previous description in engineering sense by a system in which the vocal tract is represented as a time-varying filter (Spanias, 1994: 1541-1582), (Papamichalis, 1987), (Proakis, 2001), (Ingle and Proakis, 2000). A block diagram can be obtained to represent the whole system; such block diagram is shown in Fig. 2.

4. ANALYSIS-SYNTHESIS MODEL

The ideal excitation for linear predictive coding synthesis is the prediction residual (the difference between estimated and exact signal):

$$e(n) = s(n) - \tilde{s}(n) = s(n) + \sum_{k=1}^{p} \alpha_k s(n-k) \qquad (2)$$

In classical linear predictive coding, this excitation $e(n)$ is modeled by a pitch periodic impulse sequence for voiced speech and a random noise sequence for unvoiced speech (Spanias, 1994: 1541-1582), (Papamichalis,1987), (Proakis,2001), (Ingle and Proakis, 2000). Both the periodic and noise source are scaled by an appropriate gain. The source filter model of speech production is shown in Fig. 2. A different approach is used in Residual Excited LP (RELP), where the baseband of the prediction residual is encoded. The Multi-Pulse Excited Linear Prediction (MPLP) algorithm uses Analysis-by-synthesis approach, and forms an excitation sequence which consists of multiple non-uniformly spaced pulses. The pulse locations and amplitudes are determined by minimizing the weighted mean-squared error created by the difference between the original and the LP synthesis filtered signal. In other words, during analysis both the amplitude and locations of the pulses are determined (sequentially), one pulse at a time such that the weighted mean squared error is minimized. The MPLP algorithm typically uses 4-6 pulses every 5 ms. (Spanias, 1994: 1541-1582).

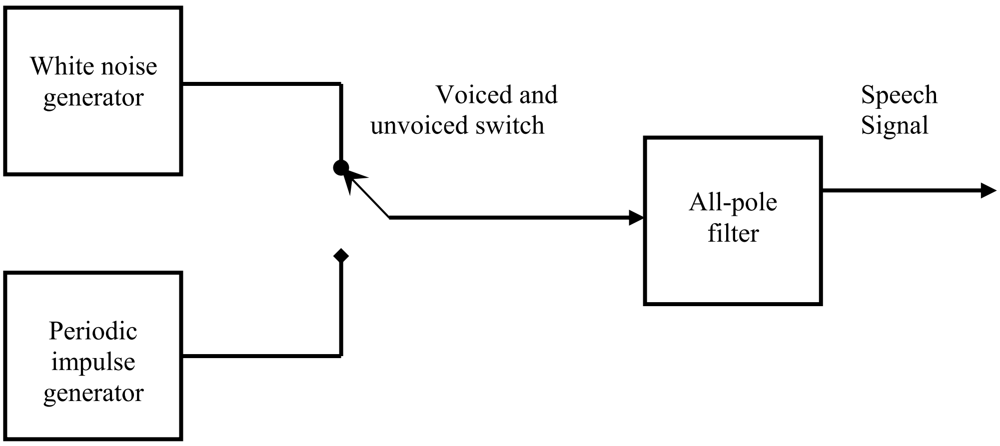

Figure 2. A Block Diagram Showing the Source Filter Model of Speech Production.

5. THE ENCODER STAGE

LPC removes the redundancies of a speech signal by modeling the speech signal as a linear all-pole filter, excited by a signal called the excitation signal (residual signal). Speech coders process a certain group of samples called a frame or a segment. The linear predictive coefficients (the filter coefficients) are selected to minimize the energy of the residual signal at the output of the filter for that frame (Proakis and Manolakis, 1996), (Proakis, 2001). This filter is called a linear prediction analysis filter.

The speech signal is first filtered through the linear prediction analysis filter. The resulting signal is called the residual signal for that particular frame.

In order to reduce the total bit rate, speech coders such as LPC-10 (Linear Predictive Coding with a 10^{th} order model) do not transmit the whole residual signal (Spanias, 1994: 1541-1582), (Papamichalis,1987), (Ingle and Proakis, 2000). In practical systems, the speech encoder at the transmitter must determine the filter coefficients and the proper excitation signal for each frame.

Our new model for linear predictive coding of speech is a vocoding technique that can be used to produce high quality speech at low data rate. This vocoder uses the linear predictive coding method with a 10^{th} order model. To generate the linear prediction parameters for each frame, the speech is sampled at 8kHz, and then the speech is analyzed by first segmenting the signal using a finite duration analysis window (e.g., a Hamming window of length N), then for each segment, the excitation parameters are determined. The excitation parameters consist of a voiced / unvoiced (v/uv) decision and the amplitude and location of each pulse (sample) in the residual signal. For the voicing decision, the algorithm uses

- The energy of the frame.
- The zero crossing count.

Since, unvoiced segments are associated with small energy and large number of zero crossings, voicing can be determined by energy and zero-crossing measurements. In many cases, voicing information is also provided by the pitch detection algorithm.

At the encoder, we use an analysis system that uses linear prediction to determine the filter coefficients $\{a_k\}$ and the excitation signal. It is believed that the Linear Prediction (LP) parameters are sufficient to represent the vocal tract and still producing high quality speech (Spanias, 1994: 1541-1582), (Papamichalis, 1987). We found that this is true not only in the voiced frames of the speech but also for unvoiced and transient frames. Hence, we used a 10^{th} order linear predictive coding model for voiced (un-silent) frames, but only a 6^{th} order model for unvoiced (silent) frames. The reason is that the speech spectrum of unvoiced sounds is described sufficiently well by the lower order model. Filter coefficients and excitation are usually determined every 32 ms (256 samples) for speech sampled at 8 kHz. The filter essentially represents the vocal tract.

Moreover, it is known that the residual signal is the perfect excitation for the all pole synthesis filter and a class of linear predictive vocoder depends on encoding this signal efficiently. In other words, the residual excitation carries all the information that has not been captured by linear prediction analysis, e.g. phase, pitch information or zeros due to nasal sounds… etc.

Although the concept of encoding the prediction residual is also utilized in Differential Pulse Code Modulation (DPCM) (Spanias, 1994: 1541-1582), (Papamichalis, 1987), (Proakis, 2001), (Ingle and Proakis, 2000), (Kim and Lee, 1999), our proposed method is different in that the residual encoding is based on selecting the most important pulses in the residual signal rather than encoding all pulses (samples).

This vocoder forms an excitation sequence which consists of non-uniformly spaced pulses. During analysis both the amplitude and location of the pulses are determined. This vocoder algorithm typically uses (8-16) pulses every 32 ms.

The block diagram of this vocoder which encodes the residual in time domain is shown in Fig. 3. In this system, the residual signal is computed and then the absolute values of the samples are obtained. The maximum amplitudes of the residual signal within the effective band and the linear prediction coefficients are encoded, quantized and transmitted.

6. THE DECODER STAGE

At the decoder, the inverse of the linear prediction analysis filter works as the linear prediction synthesis filter, while the residual signal acts as the excitation signal for the linear prediction synthesis filter.

At the receiver in our new model, the significant amplitudes of the residual signal plus a suitable random noise (AWGN) are used as the excitation signal. Hence, the synthesis system consists of an all pole filter that uses the coefficients obtained from the analysis system and the excitation signal to generate a synthetic speech signal as shown in Fig. 4.

The first step in this process is to find if voiced or unvoiced frame, if voiced frame, a 10^{th} order model and a large number of residual samples (amplitudes and locations) are used to generate the residual signal, then the LPC coefficients (a_k) are used to find the voiced speech frame. In unvoiced speech frames, a 6^{th} order model and less number of residual samples (amplitudes and locations) are used to generate the residual signal, then the LPC coefficients (a_k) are used to find the unvoiced speech frame.

In order to allow for efficient quantization, the linear prediction parameters are encoded as Line Spectrum Pairs (Spanias, 1994: 1541-1582), (Papamichalis, 1987), (Kim and Lee, 1999), (Kim et al., 2000), (Rothweiler, 1999), (Kang et al., 2004). In the next section, the quantization process of the amplitudes and locations of the pulses and the Line Spectrum Pairs (LSPs) is described.

7. MODELING AND ENCODING PROCESS

The need for an efficient representation of the parameters is evident when dealing with low rate coding. At the same time, we need to reduce the parameters, often in a compromise between quality and bit rate, as well as represent the reduced parameters in a robust way with regard to quantization. This work deals with both aspects, a new way of modeling the parameters and a method of representing the model parameters in an efficient manner. This section involves a new technique in the process of modeling and encoding of the amplitudes and locations of the residual pulses as well as LPC parameters.

7.1 Encoding the Locations of the Residual Pulses

We developed a procedure to represent the reduced parameters in a robust way with regard to quantization.

This procedure can be summarized as follows:

Suppose we have locations: $x_n = [loc_1, loc_2, \ldots, loc_N]$; where N is the number of the considered pulses, and it is known that, the locations are integer number, then:

1. Sort the locations in ascending order, keeping the corresponding amplitude index assigned; this step is justified because we note that there is a small difference between successive locations in the same frame.
2. Floor the first location in the frame and convert it to binary using (7) bits.
3. Subtract all other locations in the same frame by the value found in the step 2 above to obtain the residue of the location.
4. If the value of the residue of the location is less than 32, then set the first bit to (0), and Convert the value into 5 bits binary number.
5. If the value of the residue of the location is grater than 32, then set the first bit to (1), and the encoding procedure consists of the following steps:
   a. Subtract 32 from the location.
   b. To reduce the dynamic range, take $Log_{10}(.)$ of the result.
   c. Multiply the result by β and we choose β =13 to scale the range [0-31].
   d. Floor the value found in the previous step.
   e. Convert the value into binary number using (5 bits).
6. If the first bit of the previous residue is (0) then subtract from all remaining locations in the same frame the value found in the step 4 above, otherwise, subtract from all remaining locations (x_n's) the output of the step 5(d) divided by β and converted to $10^{(.)}$.
7. Repeat steps (4-6) until you finish all xn's.

7.2 Encoding the Amplitudes of the Residual Pulses

In this section, we developed a method to encode the reduced parameters in a robust way with regard to quantization.

The key of the algorithm used to encode the sample amplitudes is to find the proper scale to make the length of the amplitudes within the range [1-10] as shown in Table 1, and then representing the amplitudes using 5 bits (the last bit is the sign bit) after scaling to explore the full resolution [0-15]. This method can be summarized as follows:

Suppose we have amplitudes: $a_n = [amp_1, amp_2, \ldots, amp_N]$ where N is the number of the considered pulses, then:

1. Find the maximum absolute value of the amplitudes in the frame (a_{max}).
2. If a_{max} multiplied by 10^2 is greater than 1, as shown in Fig. 6, then the first 2 bits in the frame are [00].
3. Else if a_{max} multiplied by 10^3 is greater than 1 and a_{max} multiplied by 10^2 is less than 1 then the first 2 bits in the frame are [01].
4. Else if a_{max} multiplied by 10^4 is greater than 1 and a_{max} multiplied by 10^3 is less than 1, as shown in Fig. 7, then the first 2 bits in the frame are [10].
5. Otherwise, the first 2 bits in the frame are [11].
6. If the first two bits are [00], then multiply all the amplitudes in the frame (a_n) by 100, else if the first two bits are [01], then multiply all the amplitudes in the frame (a_n) by 1000, else if the first two bits are [10], then multiply all the amplitudes in the frame (a_n) by 10^4, otherwise, multiply all the amplitudes in the frame (a_n) by 10^5 as shown in Table (1).
7. Round the amplitudes in the frame and convert into 4 bits binary number.

Table (1): Scale Code.

Scale Code	Range before scaling (Linear)	Range before scaling (Log_{10})	Scale Number	Range after scaling
--	$0 – 10^{-5}$	[-∞ , -5]	--	ignored
11	10^{-5} -10^{-4}	[-5 , -4]	10^5	1-10
10	10^{-4} -10^{-3}	[-4 , -3]	10^4	1-10
01	10^{-3} -10^{-2}	[-3 , -2]	10^3	1-10
00	10^{-2} -10^{-1}	[-2 , -1]	10^2	1-10
--	10^{-1} - ∞	[-1 , ∞]	--	rare

7.3 Modeling and Encoding LPC Parameters

A set of linear prediction coefficients $\{a_k\}$, for k =1, 2,… p describes the vocal tract filter. To be useful in low bit rate speech coding it is necessary to quantize and transmit the LPC parameters using a small number of bits. Direct quantization of these LPC coefficients is inappropriate due to their large dynamic range (8-10 bits/coefficient). Thus for transmission purposes, especially at low bit rates, other forms are used to represent the LPC parameters (Spanias, 1994: 1541-1582), (Papamichalis, 1987), (Kim and Lee, 1999), (Kim et al., 2000), (Rothweiler, 1999), (Kang et al., 2004). Therefore, direct quantization of these LPC coefficients is avoided, since quantization error can lead to instability of the synthesis filter. Consequently, one of the major issues in LPC is the quantization of the LP parameters. Linear Prediction Coefficients (LP coefficients) can be described by several representations such as Line Spectral Frequencies (LSF), Reflection Coefficients (RC), Autocorrelations Coefficients (AC), Log Area Ratios (LAR), Arcsine of Reflection Coefficients (ASRC)… etc. They effectively have a one-to-one relationship with the LP coefficients, and they preserve all the information from the LP coefficients. Among them, some are computationally efficient. Some of them have special features which make them attractive for certain purposes. That is why a good understanding of those representations and their features is needed prior to further processing.

The line spectral frequencies (LSF) have a well behaved dynamic range. On the other hand, The LSP coefficients represent the LPC model in the frequency domain and lend themselves to a robust and efficient quantization of the LPC parameters. Therefore, the first LSP represents low frequency components and the last LSP represents high frequency components. Hence, if the LP coefficients are encoded as LSFs, we do not need to spend the same number of bits for each LSF. This is because higher LSFs correspond to the high frequency components and high frequency components have less effect in speech perception. Therefore, higher LSFs can be quantized using fewer bits than lower LSFs. This reduces the bit rate while keeping the speech quality almost the same. The LSPs are related to the poles of the LPC filter H(z) (the zeros of the inverse filter A(z)). Usually, the LSFs are more concentrated around formants. Moreover, spectral sensitivity of each LSF is localized. Hence, in order to allow for efficient and robust quantization the LPC parameters are encoded as LSPs. In addition, the LSP parameters are an ordered set of values between 0 and π (ascending order). In this section, we developed a procedure to encode the LSPs coefficients in a robust way with regard to quantization. This procedure can be summarized as follows:

Multiply all LSPs by ($0.5/\pi$), to make the range of LSP between 0-0.5.

1. Scale the first LSP to fit into a 6 bits binary number.
2. Convert the result to binary number using (6 bits).
3. Subtract from all other LSPs in the same frame the value found in step 2 above.
4. Round the result found in the previous step.
5. Convert the value found in the previous step into binary number.
6. Subtract from all other LSPs in the same frame the value found in step 5 above.
7. Repeat steps (5-7) until finish all LSPs.

The number of bits used from LSP (2) to LSP(10) is {6/6/6/5/5/4/3/3/3}.

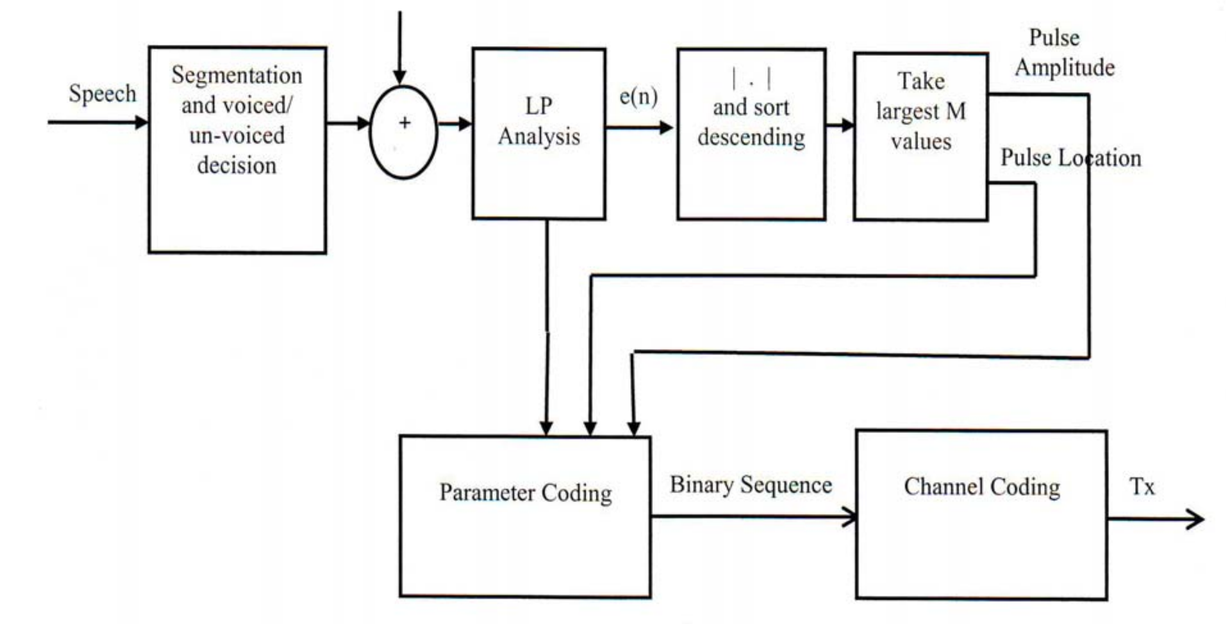

Figure 3. Encoder.

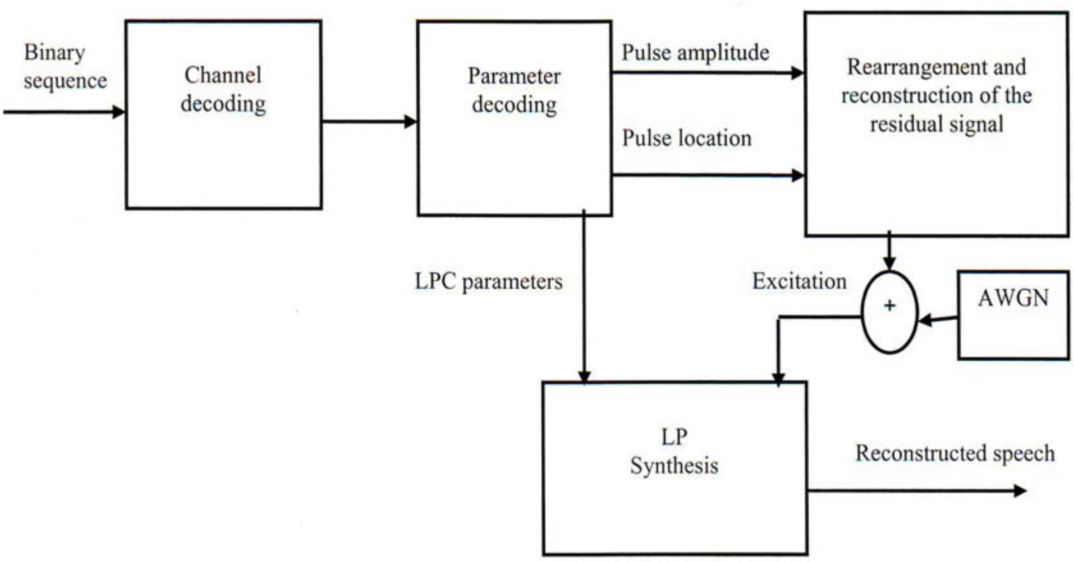

Figure 4. Decoder.

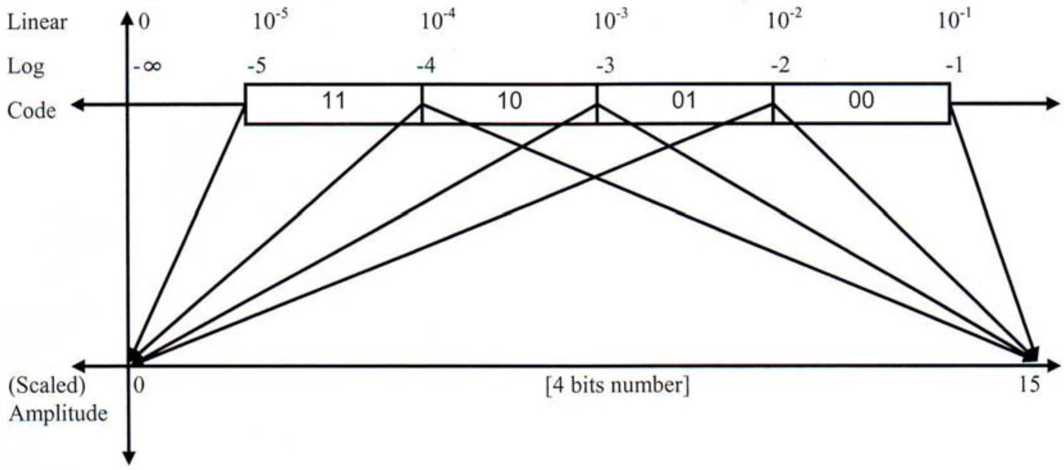

Figure 5. Scale Code.

8. ADVANTAGES OF THE PROPOSED SCHEME

Efficient techniques and algorithms were presented, which have resulted in reconstructed speech of good quality and intelligibility at bit rate from 8 to 3.6 kbps.

The advantages of our scheme:

1. It gives a reconstructed signal with high quality
2. It reduces the data rate to (3.6-8) kbps.
3. This work doesn't depend on pitch.
4. It reduces the transmitted power because it reduces the bit rate.
5. It allows the error detection and correction procedure. It has a very efficient coding procedure.
6. It is a simple method (low complexity and small delay) and easy to implement.

9. CONCLUSIONS

In this paper, we developed a new model for linear predictive coding of speech that can be used to produce high quality speech at low data rate. In this model, a new technique to reduce the number of the pulses in the residual signal is introduced. Therefore, this vocoder forms an excitation sequence which consists of multiple non-uniformly spaced pulses. In an analysis part, both the amplitude and location of the pulses are determined. In addition, new techniques to model and encode the amplitude and location of each pulse as well as linear prediction parameters are developed.

Hence, we presented an efficient vocoder operating between (3.6 to 8 k bit/s) and depending on the number of pulses used to encode the residual signal. The speech quality of this vocoder at rate above 4.8 is high, because of the emphasis on coding of the perceptually important residual components. Moreover, the speech quality of this vocoder is also limited by the information lost in the residual signal. However, the residual information can be modeled using other methods like Multi-Pulse Linear Prediction, on the expense of increasing the bit rate of the vocoder and its complexity.

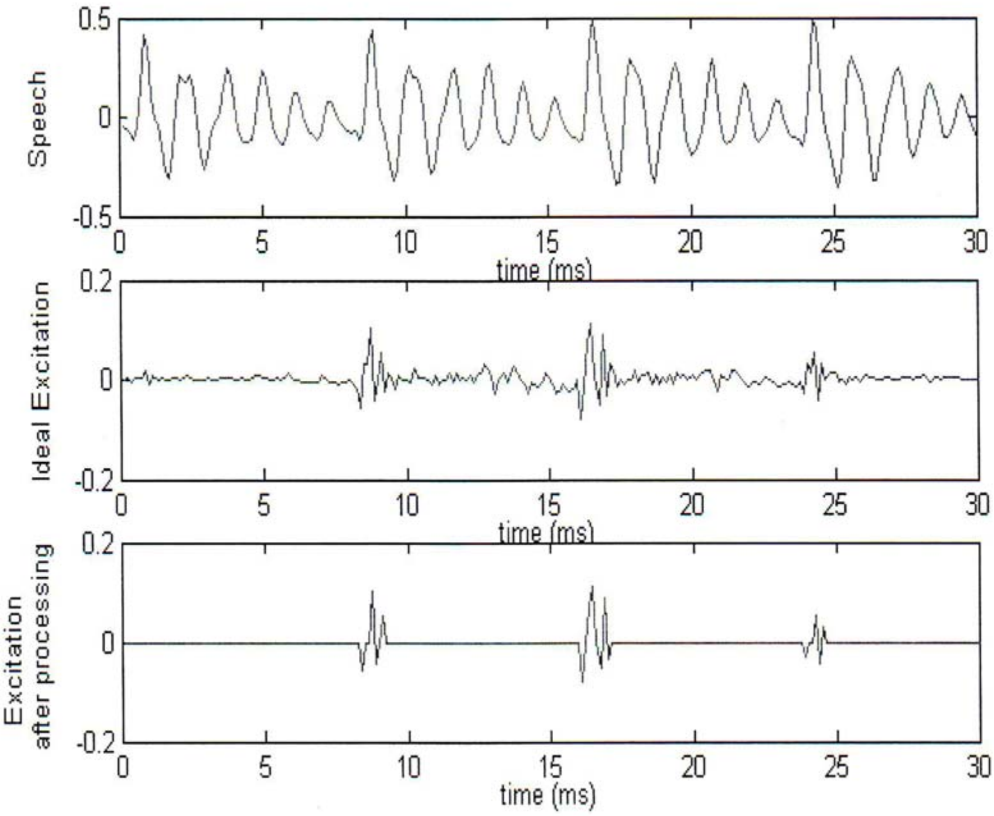

Figure 6. **The Amplitudes of an Unvoiced Frame and the Corresponding Residual Signal Before and After Coding.**

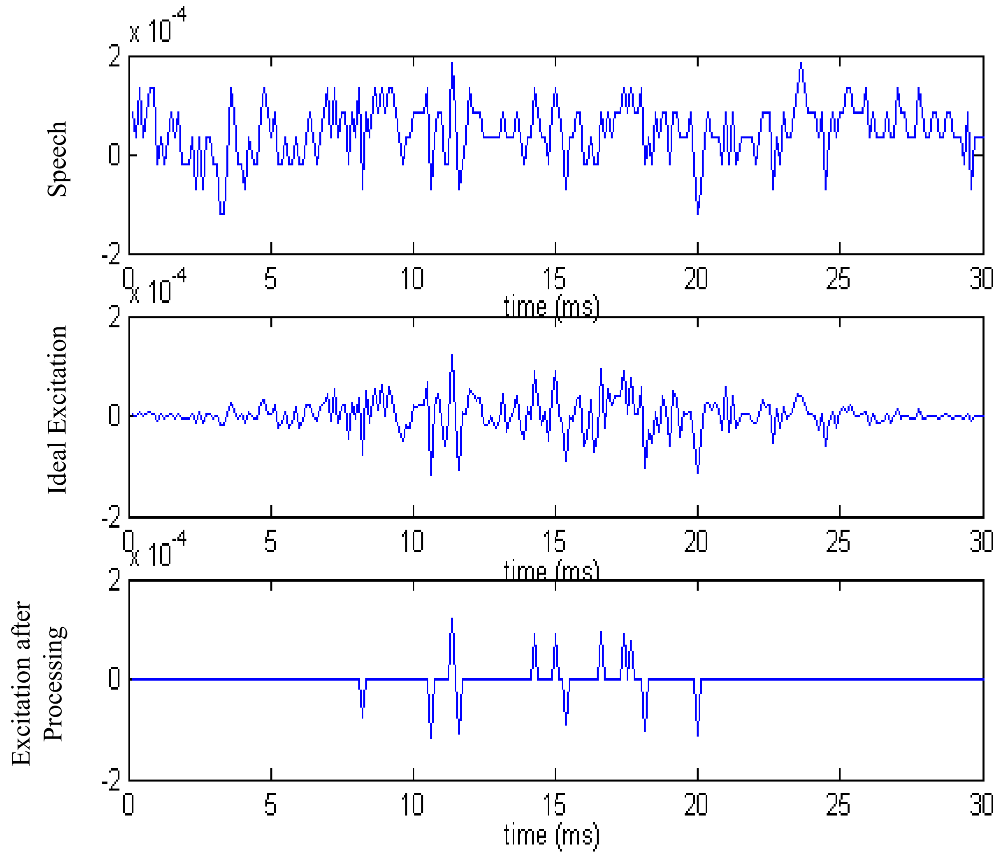

Figure 7. The Amplitudes of a Voiced Frame and the Corresponding Residual Signal Before and After Coding.

REFERENCES

Alku, Paavo and Bäckström, Tom. 2004. Linear Predictive Method for Improved Spectral Modeling of Lower Frequencies of Speech with Small Prediction Orders. *IEEE Transactions on* Speech and Audio *Processing*, 12 (2).

Atal, Bishnu. 1982. Predictive Coding of Speech at Low Bit Rates. *IEEE Transactions on Communications*, COM-30 (4):600-614.

Barnwell, Thomas. 1980. Windowless Techniques for LPC Analysis. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP- 28(4): 421.

Brinker, Albertus C. den, Voitishchuk, V. and Eijndhoven, Stephanus J. L. van. 2004. IIR-Based Pure Linear Prediction. *IEEE Transactions* on *Speech and Audio Processing*, 12(1).

CCITT Recommendation G.721. 1984. 32kb/s Adaptive Differential Pulse Code Modulation ADPCM. CCITT.

Chong, Nicola and Cox, Richard. 2003. An Intelligibility Enhancement for the Mixed Excitation Linear Prediction Speech Coder. *IEEE Signal Processing Letters*, 10(9).

Galand, C., Lancon, E., Rosso, M. and Menez, J. 1986. *A New Architecture of Multi-pulse Excited Linear Predictive Coder*. Elsevier Science Publishers.

Gray, A. and Wong, D. 1980. The Burg Algorithm for LPC Speech Analysis/ Synthesis. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP- 28 (6): 609.

Härmä, Aki. 2001. Linear Predictive Coding With Modified Filter Structures. *IEEE Transactions* on *Speech and Audio* Processing, 9 (8).

Hu Hwai-Tsu and Wu Hsi-Tsung. 2000. A Glottal-Excited Linear Prediction (GELP) Model for Low-Bit-Rate Speech Coding, *Proc. Natl. Sci, Counc. ROC (A*), 24: 134-142.

Hu, Jianling, Xu, Sheng, and Chen, Jian. 2001. A Modified Pitch Detection Algorithm. *IEEE Communications Letters*, 5 (2).

Ingle, Vinay and Proakis, John. 2000. *Digital Signal*

Processing Using MATLAB. BookWare Companion Series, Northeastern University.

Johansson, Anders. 1986. *Linear Prediction of Speech Using Recursive Analysis*, Elsevier Science Publishers.

Kang, Sangwon, Shin, Yongwon and Fischer, Thomas. 2004. Low-Complexity Predictive Trellis-Coded Quantization of Speech Line Spectral Frequencies. *IEEE Transactions on Signal Processing*, 52 (7).

Kim, Hong Kook and Lee, Hwang Soo. 1999. Interlacing Properties of Line Spectrum Pair Frequencies. *IEEE Transactions* on *Speech and Audio Processing*, 7 (1).

Kim, Hong Kook, Choi, Seung Ho and Lee, Hwang Soo. 2000. On Approximating Line Spectral Frequencies to LPC Cepstral Coefficients. *IEEE Transactions on Speech and Audio Processing*, 8 (2).

Papamichalis, Panos. 1987. Practical Approaches to Speech Coding, Prentice Hall, Inc. Texas Instruments, Inc. Rice University.

Proakis, John and Manolakis, D. 1996. *Digital Signal Processing: Principles, Algorithms and Applications*, 3rd edition, Macmillan, New York.

Proakis, John. 2001. *Digital Communications*. 4th Edition. McGraw-Hill, New York.

Rothweiler, Joseph. 1999. On Polynomial Reduction in the Computation of LSP Frequencies. *IEEE Transactions on Speech and Audio Processing*, 7 (5).

Spanias, Andreas. 1994. Speech Coding: a Tutorial Review. *Proc. IEEE*, 82: 1541-1582.

Viswanathan, V., Higgins, A. and Russell, W. 1982. Design of a Robust Baseband LPC Coder of Speech Transmission Over 9.6 Kbit/s Noisy Channels. *IEEE Transactions on Communications*, COM- 30(4): 663.

نموذج جديد لترميز إشارة الصوت باستخدام التنبؤ الخطي يعمل بمعدل معلومات منخفض

*ابراهيم منصور وسامر العبد**

ملخـص

يعد ترميز الصوت من المجالات المهمة جدا وله تطبيقات في المجال المدني والعسكري؛ فهو يعالج مشكلة تقليل معدل المعلومات اللازم لتمثيل إشارة الصوت مع القدرة على استرجاعها بجودة عالية. إن ترميز إشارة الصوت باستخدام التنبؤ الخطي من أهم طرق ترميز الصوت لإنتاج صوت بمعدل معلومات منخفض وجودة عالية. قمنا في هذا البحث بتطوير تقنيات وطرق جديدة لإيجاد مرمز لإشارة الصوت يعمل بمعدل معلومات منخفض وبجودة عالية.

يقدم هذا العمل نموذجاً جديداً لترميز إشارة الصوت بالتنبؤ الخطي الذي يمكن استخدامه لإنتاج صوت ذي جودة عالية وبمعدل معلومات منخفض. فقد اقترحنا في هذا النموذج ترميز الإشارة المتبقية بالاعتماد على اختيار النبضات المهمة من الإشارة المتبقية عوضا عن ترميز جميع النبضات. وبالتالي فإن هذا المرمز يمثل الإشارة المتبقية بحيث تتكون من عدة نبضات موزعة على فترات غير متماثلة. ويحدد خلال مرحلة التحليل قيمة كل نبضة ومكانها. بالإضافة إلى ذلك، يقترح البحث تقنية جديدة لترميز قيمة كل نبضة ومكانها والمكونات اللازمة للتنبؤ الخطي.

* قسم الهندسة الكهربائية، كلية الهندسة والتكنولوجيا، الجامعة الأردنية. تاريخ استلام البحث 2005/6/15، وتاريخ قبوله 2006/10/2.