

Speech Enhancement Based Compression Method

Zinah S. Abduljabbar*
Computer Science, College of Science, Al-Mustansiriya University

Noor Khaled
Computer Science, College of Science, Al-Mustansiriya University

The research is financed by Asian Development Bank. No. 2006-A171(Sponsoring information)

Abstract
This paper presents enhanced signal based compression method depending on the Normalized Least Mean Square (NLMS) filter and Linear Predictive coding (LPC) technique. The aim of this proposed system is to enhance speech signal then reducing the size of speech file while maintaining the quality. NLMS is used to enhance the additive noisy signal , while the LPC technique is depended in compression operation for analyzing and synthesizing speech signals ,this represented by coding signal with low bit rate and its affected on reconstructed quality of speech signal. Discrete Cosine Transform (DCT) is adopted to obtain presentation of information in the frequency domain for compressed and improved speech quality.
Keywords: LPC; NLMS; DCT

1. Introduction

Speech signal is based in many speech signal processing. The purpose of enhance algorithm is to reduce the noise in a noisy speech signal. Commonly, noise can be additive, convolution, or multiplication. Algorithms for noise reduction are classified into: spectral restoration filters techniques and model based methods [1].adaptive filters considered variable filters that means its coefficients can be adapted or modified to enhance its performance in related to some criterion, this allow filters to adjust the variations in the input signal characteristics . The more effective algorithm is Least Mean Square (LMS) because its simplicity and less computation are needed [2].

Every day there are a huge amount of speech data transmitted such as education, e-learning and phone banking, due to the recent evolution of speech data dense in multimedia based web applications so the need to compression application have become an urgent necessity to overcome the problem of considerable storage capacity and transmission bandwidth. Compressing data is the process of changing the representation of the input source data to compressed form that has a smaller size [3]. In recent years many algorithms have been emerged for this purpose and most of these algorithms are based on using sub-band coding or transform coding. Band-pass filter used in sub-band coding to splits the signals into a number of sub-bands, mathematical transformation such as fast Fourier transform and discrete cosine transform are used in transform coding [4, 5]. In terms of methodologies previous this work adopted based on LPC technique for example in [6, 7, 8].

In this paper speech enhancement based compression operation is introduced and once of the challenges facing the compression process is to find a balance between the conflicting requirements demands (i.e. small size in representation while the quality should not be affected). This work depends on NLMS filter and LPC technique, and another tool like DCT transform. LPC is adopted for coding speech signal with low bit rate. The rest paper has been coordinated as follows: section 2 explains tools that used in this work such as NLMS, DCT transform and LPC technique. Section 3 illustrated the methodology of this introduced system in details. Finally the obtained results with figures are presented in section 4, followed by conclusion in section 5.

2. Algorithms Used in This System

2.1 Adaptive Filter

The normalized Least Mean Square (NLMS) is updated from the standard LMS algorithm in [9], where the adaptive filter coefficients are adjusted as in equation below [10,11]:

$$\overline{W}(n+1) = \overline{W}(\mathrm{n}) + \mu(\mathrm{n}).\mathrm{e}(\mathrm{n}).\bar{\mathrm{u}}(\mathrm{n}) \quad (1)$$

Where

$$\mu(\mathrm{n}) = {}^{\mu}/_{\|\bar{\mathrm{u}}(\mathrm{n})\|^2} \quad (2)$$

This equation illustrated that the difference between the standard LMS algorithm and NLMS algorithm lies in time-varying step size μ (n) of NLMS algorithm. This step size has benefit to enhance the convergence speed of an adaptive filter.

2.2 DCT

The purpose of using DCT in speech compression belongs to the high correlation in adjacent coefficient. The sequence is reconstructed most precisely from less coefficients of DCT. This possibility of DCT leads to effective reduction of data [4, 12, 13].

$$x(m) = \left[\frac{2}{N}\right]^{1/2} c_m \sum_{m=0}^{N-1} x(n) \cos\left[\frac{(2n+1)m\pi}{2N}\right] \quad (3)$$

Where m=0, 1, …… , N-1
The inverse discrete cosine transform is

$$x(n) = \left[\frac{2}{N}\right]^{1/2} \sum_{m=0}^{N-1} c_m \; x(m) \cos\left[\frac{(2n+1)m\pi}{2N}\right] \quad (4)$$

In the two equations cm is:
C_m= (1/2)1/2 for m=0

2.3 LPC

LPC is a methodology adopted for analyzing and synthesizing speech signals, it used for estimating parameters related to pitch, formants and spectra. Basically, LPC is used for reducing the summation of squared differences in finite interval between the original and estimated speech signals to produce predictors' coefficients, transfer function of the time varying of digital filter is presented in equation below [7]:

$$H(Z) = \frac{G}{1-\sum_{k=1}^{p} a_k Z^{-k}} \quad (5)$$

3. Propose System

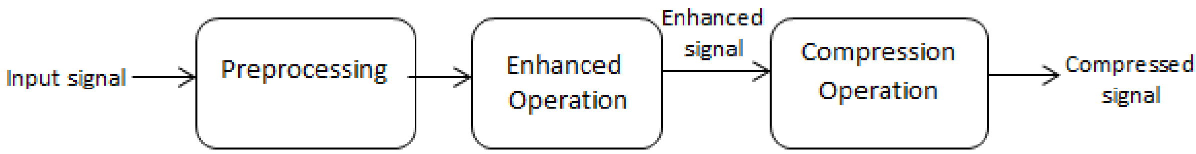

Figure 1. Layout of the proposed system

As shown in Fig1, the proposed system consists of the following operations: preprocessing, enhanced and compression operations.

1. preprocessing consists of two steps:
   1.1. Read speech file (mono) and load the data part in the vector S of dimension 1*n, where n is the number of samples.
   1.2. The data in S is normalized to the range [-1, 1] as represented in the following equation.

$$NS(i) = \frac{S(i)-2^{(n\,bits-1)}}{2^{(nbits-1)}} \quad (6)$$

   Where:
   NS(i): the ith element of speech file after normalization process.
   S(i): the ith element of speech file.
   nbits: number of bits per samples.
2. Enhanced Operation
   When the noisy signal is as input the NLMS adaptive filter is used to enhance the additive noisy signals as explain in section (2.1). Enhanced signal is passed through the second process where the compression operation is made.
3. Compression operation
   Like any other compression operations, the proposed compression in Fig1 involves two essential Phases namely encoding and decoding phases, each phase has its own stages which illustrated as follows:
   a. Encoding Phase

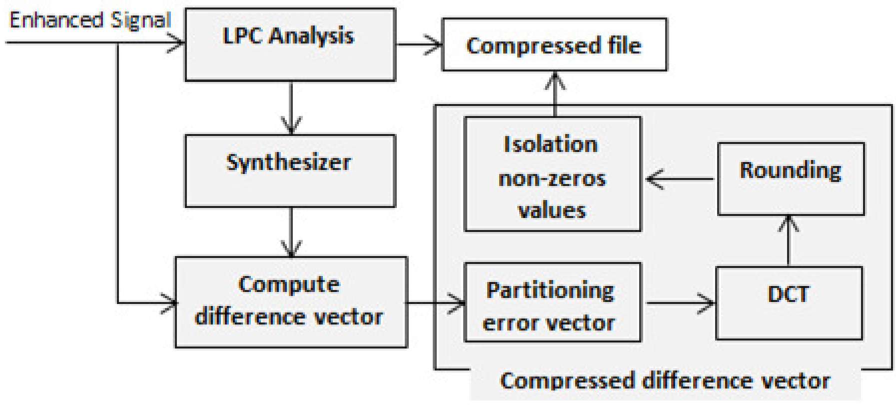

Figure 2. Encoding Phase

As seen in Fig2, encoding phase consists of the following stages:

1. LPC analysis:

LPC is used for coding signal with low bit rate signal representation and affected on its reconstructed quality resulted twelve parameters that represent pitch, gain voiced\unvoiced bits.

2. Synthesizer

In synthesizer the speech signal is reconstructed by reversing the operation, the quality of the reconstructed speech signal is affected, to overcome this situation difference vector is computed by finding the difference between the original speech signal and the synthesized signal which has length as the same of original signal.

3. Compressed difference vector

Number of steps have been applied to the difference vector to reduce its size, these steps are:

3.1 Partition the difference vector to number of non-overlapping frames, the number of frames depending on the length of difference vector and the predefined length of the frame.
3.2 DCT is used to obtain presentation of information in the frequency domain for each frame by using eq(2) for increasing the small values in each frame while the energy of the frame is concentrated in few coefficients at the beginning of frame.
3.3 Importantly, in rounding process converted DCT coefficients to integer numbers by mapping these coefficients to nearest integer number for reducing bits count and rising zeros values.
3.4 Neglecting zeros value that obtained in the previous step produce compression factor with rising value and maintained PSNR value.

b. Decoding phase

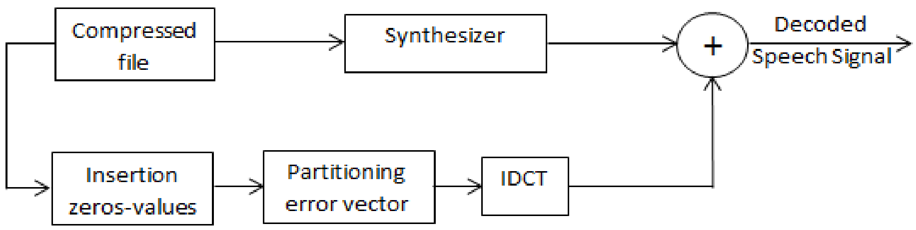

Figure 3. Decoding Phase

As shown in Fig3, in decoding phase the twelve parameters are extracted from compressed file to decode and reconstruct speech signal by synthesizer**.** The non-zeros values is arranged in vector has the same length of the difference vector then it must be partitioned, each frame is inversely transformed using IDCT. Finally decoded signal is obtained by adding each value of difference vector to its corresponding in reconstructed vector.

To evaluate the above system Psnr and compression factor (CF) are used as measurements and calculated as below [14, 15]:

$$\mathrm{Psnr(dB)} = 10\log_{10}\frac{(2^{nb}-1)^2}{\frac{1}{N}\sum_{n=1}^{N}(y1_n-y2_n)^2} \tag{7}$$

nb number of bits per sample
$y1_n$ is the original signal.
$y2_n$ is the reconstructed signal.
N is the total number of samples.

$$\mathrm{CF} = \frac{\text{size of the original file}}{\text{size of the compressed file}} \tag{8}$$

4. Experimental Results

This section is devoted to show and study results of tests to evaluate the performance of introduced system. Table (1) shows the wave file symbols have been considered in this introduced system.

Table 1. Wave file symbols

Speech file	Attributes		
	Sampling Rate (Hz)	Sample Resolution (bps)	Size (KB)
S1.wav	22050	8	107
S2.wav	16000	16	74.4
S3.wav	8000	16	46.9

Control parameters that investigated are: speech signal enhanced effect, effect of rounding process, neglecting zeros values and frame length.

4.1 Enhanced Effect

Fig4 shows enhanced operation effect on examples of symbols (S1 and S2) that used in this system. In which White Gaussian Noise is added to original speech signal, then NLMS is adopted for noise reduction.

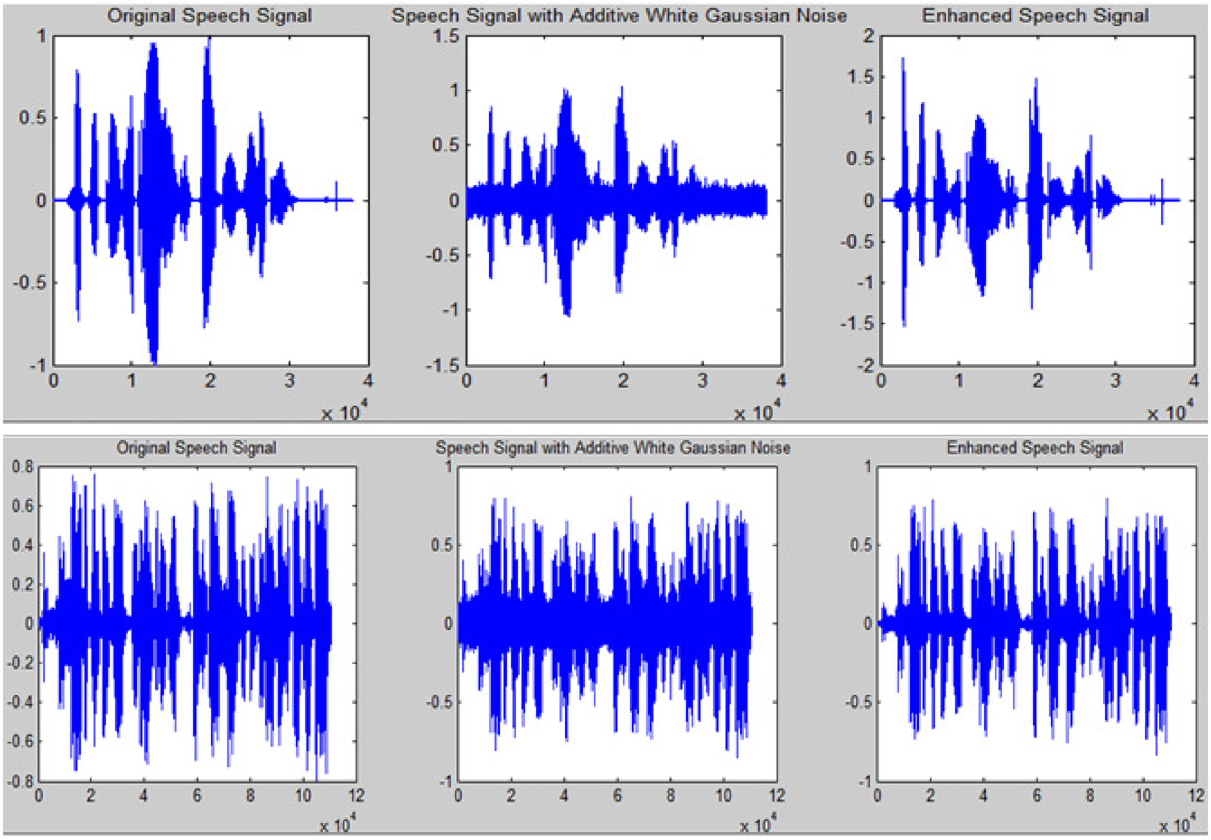

Figure 4. Enhanced operation

4.2. Rounding Process Effect

Fig5 illustrates rounding process effect on CF and Psnr. in this process DCT coefficients values are mapped to the nearest integer numbers causing increased in zeroes values.

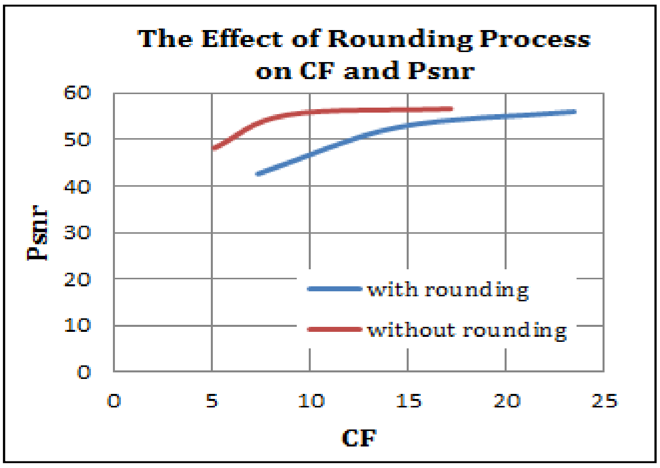

Figure 5. Rounding process effect on CF and Psnr

4.3. Neglecting Zeros Values

Fig6 describes zero values neglected. These zeros values increased by rounding process, so their effect reflected significantly on the value of compression factor when neglecting them, and in the same time Psnr value is maintained.

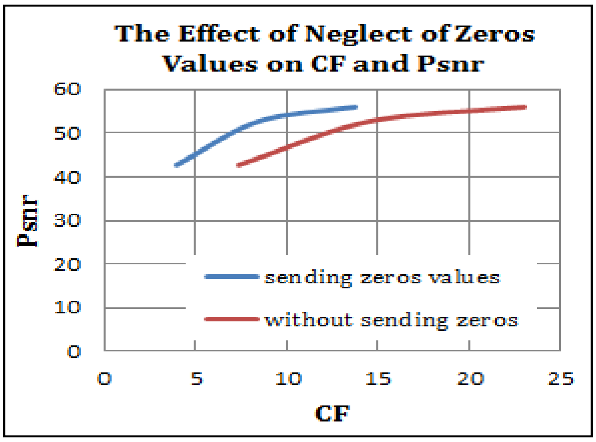

Figure 6. Zeros value Neglecting Effect

4.4. Frame Length Effect

Fig 7 clarifies frame length effect on compression factor and Psnr values, as explain the rising in CF resulted from the increasing in frame length while Psnr value is reduced. The length of frame which adopted by trail in this system equals to 60.

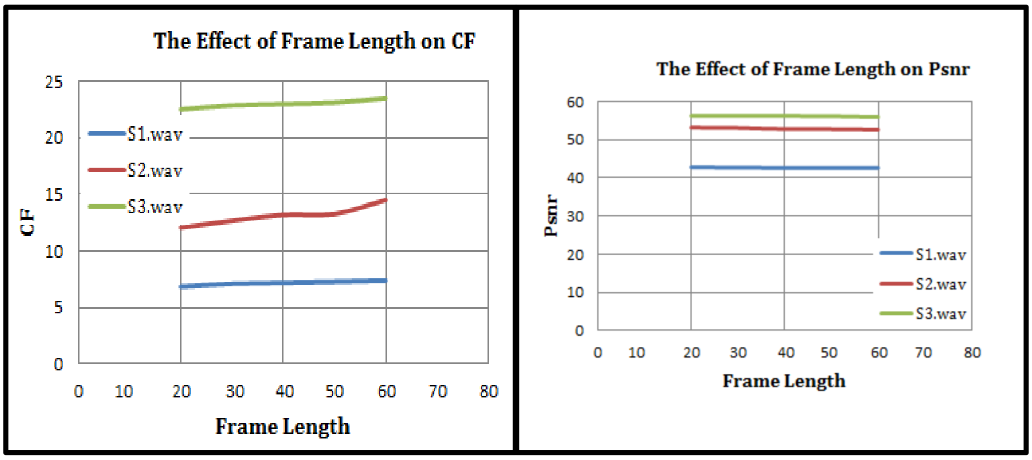

Figure 7. Frame length effect on compression factor and Psnr

Fig8 appears samples (S1 and S2) that considered in the proposed system, as illustrated waveforms show the great effect that the signal is subjected to using LPC, whereas the retrieved signal is very close to the original one.

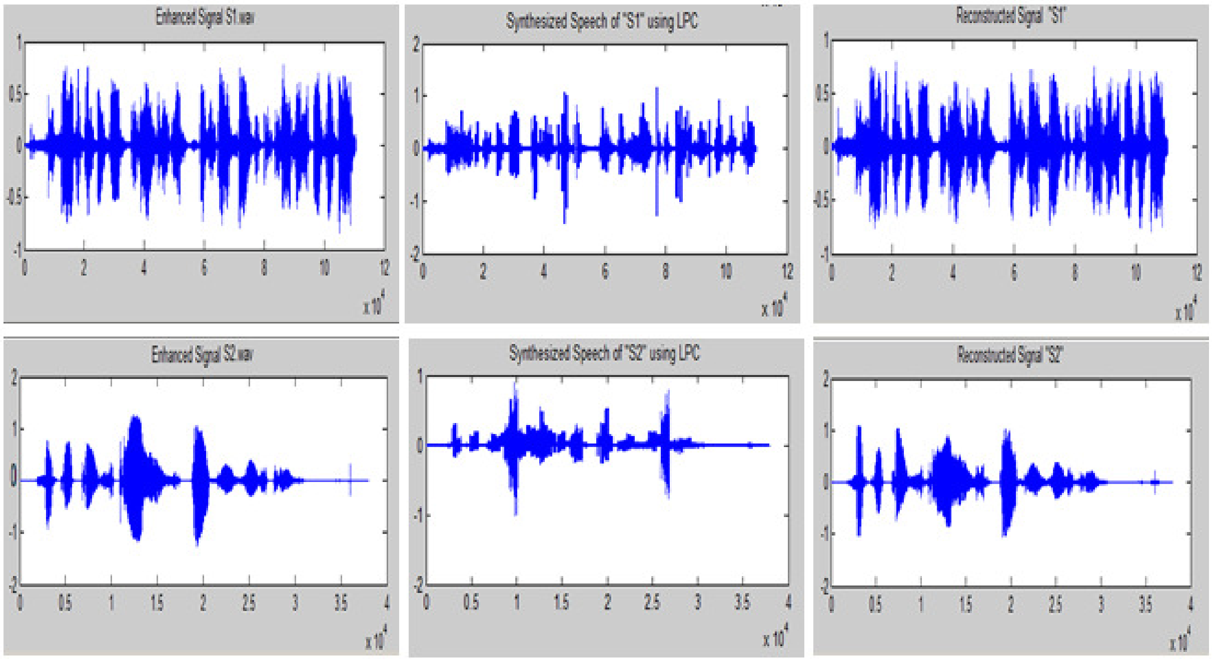

Figure 8. Samples used in proposed system

5. Conclusions

In this work the quality of the speech signal is satisfied by using NLMS to enhance it from additive noisy. The speech quality is significant issue in any speech processing method like compression operation that introduced in this work. LPC has its effect on reconstruction operation of speech signal and DCT is used for preserving recovered speech signal quality by making compression to difference vector. (The result of tests conducted on this system reveals that using some parameters as neglecting zeros values have significant effect on compression operation where frame length has less effect.

References

[1] Alugonda Rajani ,Soundarya. S.V.S. "A Review on Various Speech Enhancement Techniques" *International Journal of Advanced Research in Computer and Communication Engineering,* vol. 5, no. 8 (2016), pp. 296-301.

[2] G.V.P.Chandra Sekhar Yadav, B. Ananda Krishna. "Study of different adaptive filter algorithms for noise cancellation in real-Time environment" *International Journal of Computer Applications,* vol. 96, no. 10 (2014), pp. 20-25.

[3] David Salmon. "Data Compression The Complete Reference". London: Springer-Verlag, (2007).

[4] Zainab T. DRWEESH, Loay E.GEORGE. "Audio Compression Based on Discrete Cosine Transform, Run Length and High Order Shift Encoding" *International Journal of Engineering and Innovative Technology ,* vol. 4, no. 1(2014), pp. 45-51.

[5] David J. Katz, Rick Gentile. "Embedded Media Processing". UK: Elsevier Science, (2012).

[6] Amol R. Madane, Zalak Shah, Raina Shah, Sanket Thakur. "Speech compression using linear predictive coding" Proceedings of the International Workshop on Machine Intelligence Research MIR, (2009), pp. 119-122.

[7] Nikhil Sharma. "Speech Compression Using Linear Predictive Coding (LPC)" *International Journal of Advanced Research in Engineering and Applied Sciences ,* vol. 1,no. 5(2012):, pp. 16-28.

[8] Uday Mithapelli, Suraj P. Patil. "Implementation of Speech Compression Using Linear Predictive Coding (LPC) with Tms320c6713dsk and Comparison with Other Platforms" *International Journal of Emerging Engineering Research and Technology,* vol. 2, no.2 (2014), pp. 91-95.

[9] Prajna Kunche, K.V.V.S. Reddy. "Metaheuristic Applications to Speech Enhancement". Switzerland: Springer Briefs in Speech Technology, (2016).

[10] Allam Mousa, Marwa Qados, Sherin Bader. "Speech signal enhancement using adaptive noise cancellation techniques" *Canadian Journal on Electrical and Electronics Engineering ,* vol. 3, no. 2 (2012), pp. 375-383

[11] Jashvir Chhikara. "Noise cancellation using adaptive algorithms" *International Journal of Modern Engineering Research ,* vol. 2, no.3 (2012), pp. 792-795.

[12] M. V. Patil, Apoorva Gupta, Ankita Varma, Shikhar Salil. "Audio and speech compression using dct and dwt techniques" *International Journal of Innovative Research in Science, Engineering and Technology,* vol. 2,

no. 5 (2013), pp. 1712-1719.
[13] P. Sandhya Rani, D.Nanaji, V.Ramesh, K.V.S. Kiran. "Audio Signal Compression using DCT and LPC Techniques" *International Journal of Engineering Trends and Technology ,* vol. 21, no. 5 (2015), pp. 261-266.
[14] Nigel Chapman, Jenny Chapman. "Digital Multimedia". John Wiley & Sons, Ltd, (2009).
[15] Jeroen Breebaart, Christof Faller. "Spatial Audio Processing: MPEG Surround and Other Applications". John Wiley & Sons Ltd, (2007).