International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 21 (2017) pp. 11493-11499

Hybrid Speech Compression method Based on DWT and SVD

Maytham Salman azeez[1] and Aqeel Salman azez[2]

[1]*Univ. Al- Kafeel College, Province of Najaf, Republic of Iraq.*
[2]*Univ. Al- Kafeel College, Province of Najaf, Republic of Iraq.*

Abstract

Speech compression attractive topic related to human communication so it is necessary to update methods to find efficient and low bit rate method. The main issue is keep speech quality. Methodology present compression method consist of sequences of steps; reemphasis speech using high pass filter as preprocessing followed by voice activity detection to remove silence interval. Framing a result to specific number of samples, each frame will reordered into two dimension matrix. It transformed with discreet wavelet transform (four component). Three of them analyzed using single value decomposition and get parts of them with the remaining part to reconstruct frame. Huffman codding applied to selected part to reduced bitrate of data. The result shows that compression factor reach about 6 to 9 rate with acceptable quality.

Keywords: Discreet Wavelet Transform DWT, Single Value Decomposition SVD

INTRODUCTION

Data compression is general is attractive topic in information technology. Compression allows to share the bandwidth as possible, also compression allows storing longer massages. The samples of Digital speech signals were at rate of rate of 8000 Hz, each one is represented by 8 or 16 bits. Speech coding is represent using many algorithms. Majority of low-rate speech coding is relied on the principle of linear predictive coding (LPC). Transformation always used in loss compression such as Fourier transform, discreet cosine transform, discreet wavelet transform etc. these transform will present the signal that present in time domain in another domain [1]. Research and business communities are highly interested with the increasing demand for speech communication, speech compression or coding technology. This lead to investigate structures for speech coding [2]. Speech coding is performed using operations specified as a procedure that takes a value, or several values, as an input and results a value, or several values, as an output. Thus, a sequence of computational steps that transform the input into the output is called coding algorithm.

RELATED WORK

Speech compression is a part of signal compression, there are many work deal with this topic some of these work illustrate as follow:

- ***Rohit Thanki, Komal Borisagar and Vedvyas Dwivedi in 2013*** [3]: suggest compressive sensing (CS) theory and various signal processing transforms such

  Discrete wavelet transform (DWT), singular value decomposition (SVD) and discrete cosine transform (DCT), then generate hybrid measurements by application of CS theory on these coefficients. The method is analyzed and tested for various signals such as ECG signal and the human speech signal. The performance results that an equal for electrocardiogram ECG signal and Speech signal.

- ***Yuhui Li, Wei Gou and Bo Li in 2011*** [4]: suggested DWT and SVD in the digital watermark system in order to guard the digital content of the owners, which is a technique for digital watermarking using the attributes of SVD and DWT. He showed that the technique was robust against adding noise, cropping, zoom, rotating, compression and filtering.

- ***Ranjeet Kumar, A. Kumar and G.K. Singh in 2016*** [5]: proposed a hybrid compression method with SVD and Embedded Zero-trees of Wavelet transforms (EZW) for ECG signals. And in order to obtain intra and inter beat correlation, he used three various techniques as zero padding, interpolation exploited and period length. The compression ratio achieved by him is of 24.25:1 with a high quality of signal reconstruction in percentage-root-mean square difference (PRD) as 1.89% for ECG signal Rec.

DISCRETE WAVELET TRANSFORM

There is a large number of applications in mathematics computer science, engineering and science for discrete wavelet transform. Most notably using it to code, which is representing the discrete signal in an excessive form as a preconditioning for data compression. Also, there are practical applications can be found in signal processing of accelerations for gait analysis, [6] image processing, [7] in digital communications and many others. To design low-power pacemakers, it was found that discrete wavelet is implemented as an analog filter bank in biomedical signal processing. Images are two dimensional signals, and wavelets are used to de-noise such signals. It is necessary to note that if other levels, thresholding and wavelet strategies were used, they will result in different types of filtering. In the current example, the white Gaussian noise to be removed. Although, it could just as easily have been amplified with different thresholding.

SVD ANALYZING

Is a factorization of a real matrix, It is the generalization of the Eigen decomposition of a positive semi definite normal matrix via an extension of the polar decomposition. In order to obtain the decomposition of singular value, it needs a rectangular matrix of data in that *n* rows will represent the genes and the *p* columns will represent the experimental conditions. The SVD theorem states [8]:

$$A_{nxp} = U_{nxn} S_{nxp} V_{pxp}^T$$

Where $U^T U = I_{nxn}$ and $V^T V = I_{pxp}$ (i.e. U and V are orthogonal).

Where the left singular vectors are referred as the columns of U (vectors of gene coefficient); S (the same dimensions as A) has singular values and is diagonal (mode amplitudes); and VT has rows that are the right singular vectors (expression level vectors).

The SVD referrers to an expansion of the original data in a coordinate system where the matrix of covariance is diagonal [9]. The calculation of SVD involves finding the eigenvectors and eigenvalues of both AAT and ATA. The columns of V are made up of the eigenvectors of ATA, while the columns of U are made up of eigenvectors of AAT. Also, the square root of eigenvalues of ATA or AAT are the singular values in S. The diagonal entries of S matrix are the singular values, where they are ordered in a descending matter. The singular values are always real numbers. U and B are real if the matrix A is a real matrix. The SVD are employed in several applications, these application include determining the range, rank and null space of a matrix, approximation of matrix, multivariable control, least squares fitting of data and computing the pseudoinverse [10].

METHODOLOGY

The methodology of proposal depends on treating the speech signal as image signal by re-allocation of the speech signal data into two dimensional data after partition it into previous specification fixed length that will be sub block. Each sub block will be silence (that will replaced with label) or speech (not silence) that will convert to two dimension array and treat it as sub block. Figure 1 explains general frame work of proposed method. The proposed method consists of several steps as follow:

Speech Pre-Processing

The pre-processing step include speech pre emphasis using low pass filter. The segmentation step is partition speech signal into fix length signal 64, 256, or 1024, these will convert to 8×8, 16×16 or 32×32 as a two dimensional array respectively. Each segment will test if in silence interval or not. The silence interval will replaced all samples with on flag as indicator. The other will convert into two dimensional array. The preprocessing steps shown in figure 2.

Voice Activity Detection

The speech signal may have silence before, after and between words which can effect on recognition process. Therefore, we can separate silence (non- speech data) parts and speech data from speech signal using Voice Activity Detection (VAD). The VAD is implemented by calculating the energy of each frame first and then these energy values are compared with threshold that has certain value. If the energy of a frame is below threshold value, a frame is discarded (silence frame), otherwise, a frame is labeled as speech frame.

Recombination

The recombination step consist of select the size of recombination array (8×8, 16×16 or 32×32) depend on segment size as shown in figure 3.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 21 (2017) pp. 11493-11499

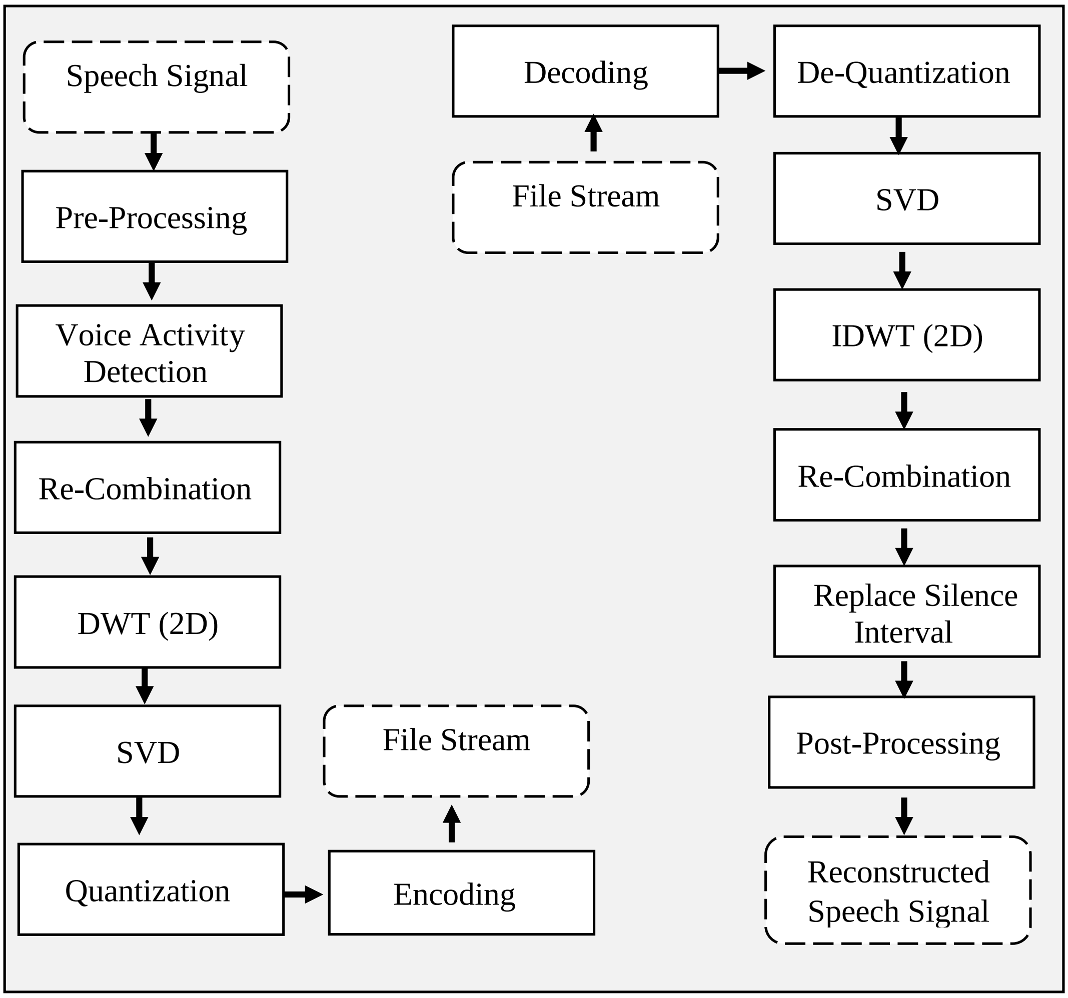

Figure 1: General frame work of proposed method

The total speech signal will sequence of sub blocks and these are not silence. The conversion two another shape array will satisfy some characteristic.

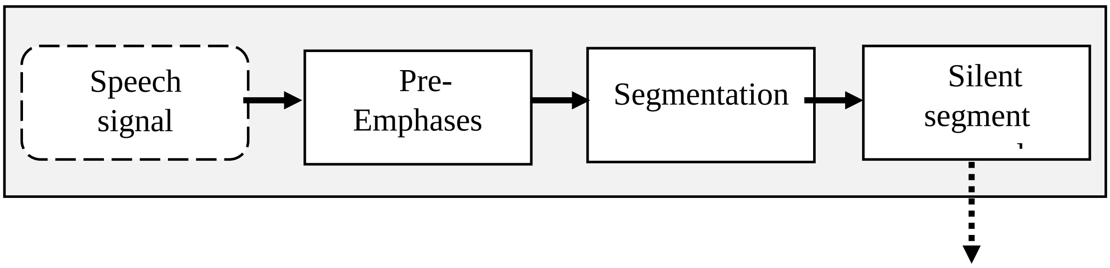

Figure 2: Speech signal preprocessing

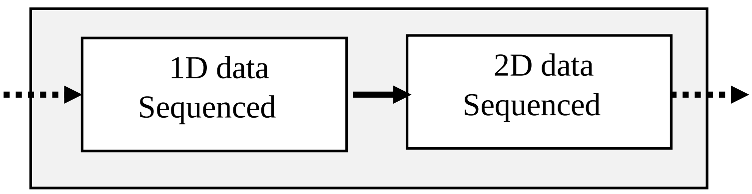

Figure 3: frame recombination

Discrete Wavelet Transform

In general, the desirable properties of wavelets are illustrated by Haar DWT. First, it is possible to preformed it in operations; second, it not only captures a notion of the input frequency content by testing it at various scales, but also the temporal content, the time at which these frequencies occur. By combining these two properties, the Fast wavelet transform (FWT) is made, instead of the conventional fast Fourier transform (FFT). No change in Low-Low block, but the other blocks will be processed as we will explain later. The two dimension discreet wavelet transform applied on each frame to produced four component Low-Low, High-Low, Low-High, High-High as shown in figure 4. The relevant features concentrated in Low-Low component, while irrelevant features concentrated in High-High component. The compression will be vary four component.

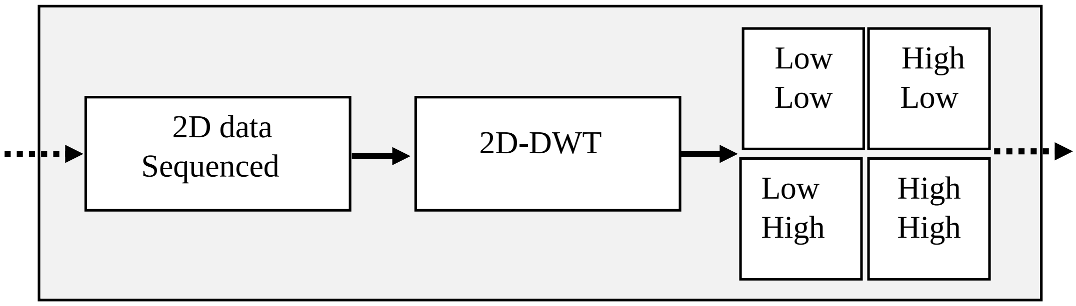

Figure 4: Discreet wavelet transform of frame (2D)

Single value Decomposition

SVD applied on the three block High-Low, Low-High and High-High with different k values. For each block three arrays will found U, S, and V, these will used in quantization step. Single value decomposition control the compression method by finding suitable k values that reduce the Mean Squared as possible as shown in figure 5.

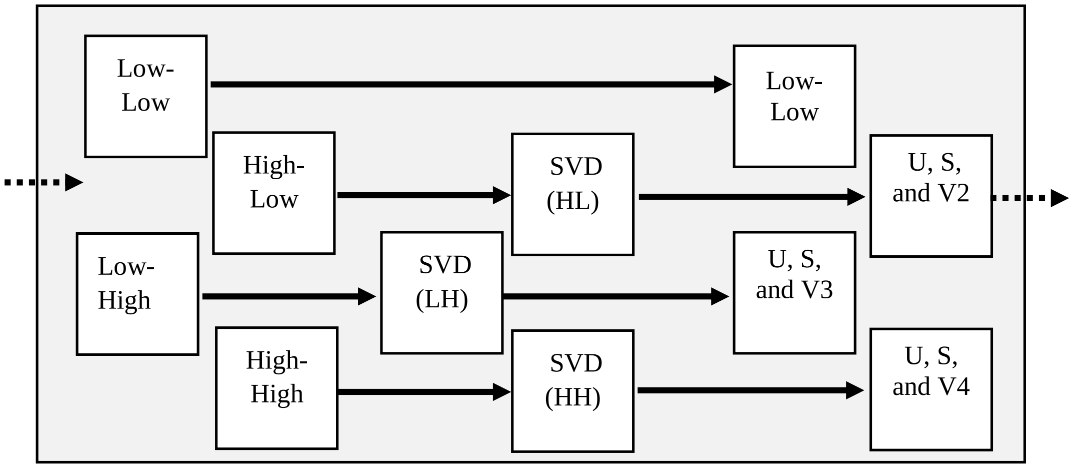

Figure 5: Single Value Decomposition of frame

Quantization

In this step select k for High-Low block, Low-high block, and High-High block with predefined MSE. PSNR (depend on MSE) value will produce value of k. the total result will selected are quantized to reduced bit rate. The value of k used as quantization parameter that will control the compression factor. As shown in figure 6. The result will reconstructed for verification step as shown in figure 7.

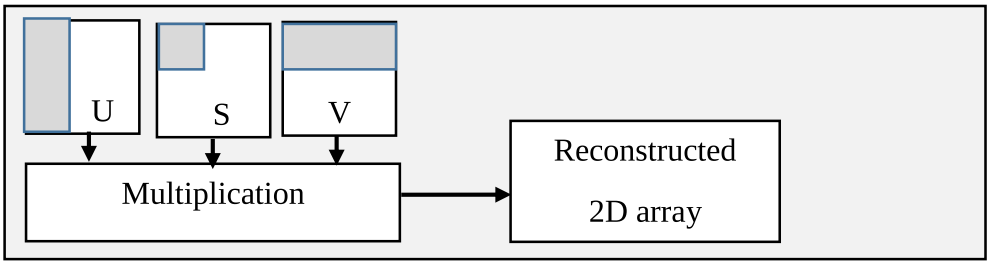

Figure 6: Quantization of frame

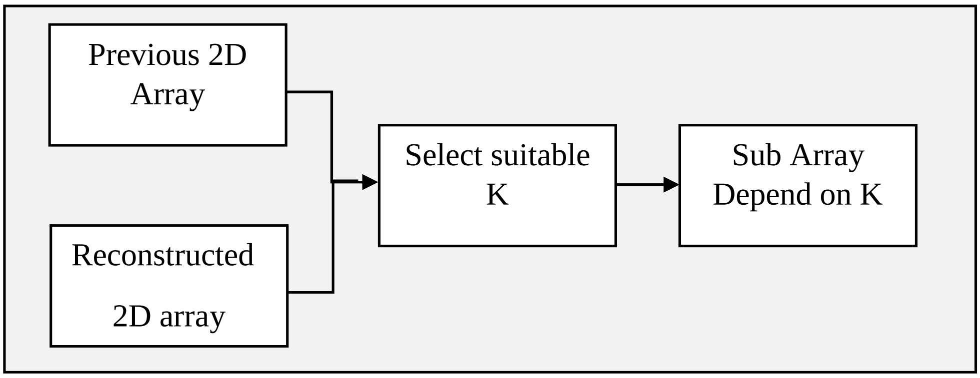

Figure 7: Quantization of frame verification

Huffman Coding

The samples results from previous step coded using lose less coding (Huffman coding) and produced binary stream that could transfer via communication channels. The Huffman coding provide extra compression that provided in previous steps.

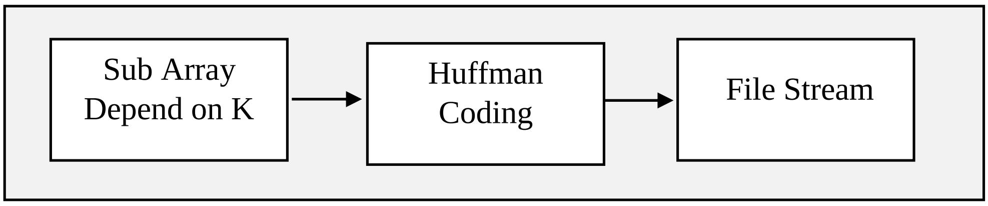

Figure 8: Huffman Coding of frame

IMPLEMENTATION TEST RESULTS

the experement of proposed method applied on fife different speech file size (100 to 500 kb) shown that there are first reduced the file size in silance removal and keep voice activity for next step as shown in table 1. the result will be input for framing and transformation and feature extraction and the result shown table 2.

Table1: silance removal

File#	File Size	Silance removal	Voice Activity
Wave 1	100 kb	27552	72448
Wave 2	200 kb	62016	137984
Wave 3	300 kb	34784	265216
Wave 4	400 kb	95360	304640
Wave 5	500 kb	124448	375552

Experement on fife speech files shown in table 2 with two selected k value 2 and 4. The minimum compression factors 6.765 and maximum is 10.822.

Table2: Compression Factor when K=2

File#	Voice Activity	File Size byte	Huffman Coding	Compression factor	Time consuming
Wave 1	72448	25753	14594	6.8521	0.227799
Wave 2	137984	49049	18481	10.822	0.316467
Wave 3	265216	94276	44406	6.756	0.534573
Wave 4	304640	108290	46476	8.606	0.549655
Wave 5	375552	133497	55036	9,085	0.652244

Experement on fife speech files shown in table 3 with two selected k value 4. The minimum compression factors 5.503 and maximum is 6.412.

Table 3: Compression Factor when K=4

File#	Voice Activity	File Size byte	Huffman Coding	Compression factor	Time consuming
Wave 1	72448	35092	15708	6.366	0.224498
Wave 2	137984	66836	24957	8.013	0.298217
Wave 3	265216	128464	54503	5.503	0.453953
Wave 4	304640	147560	62551	6.394	0.505406
Wave 5	375552	181908	77971	6.412	0.591665

Speech signals are reconstucted and ploted to find the simlarity and differences in hole of it. figure 9. Explian the speech signal and reconstucted signal.

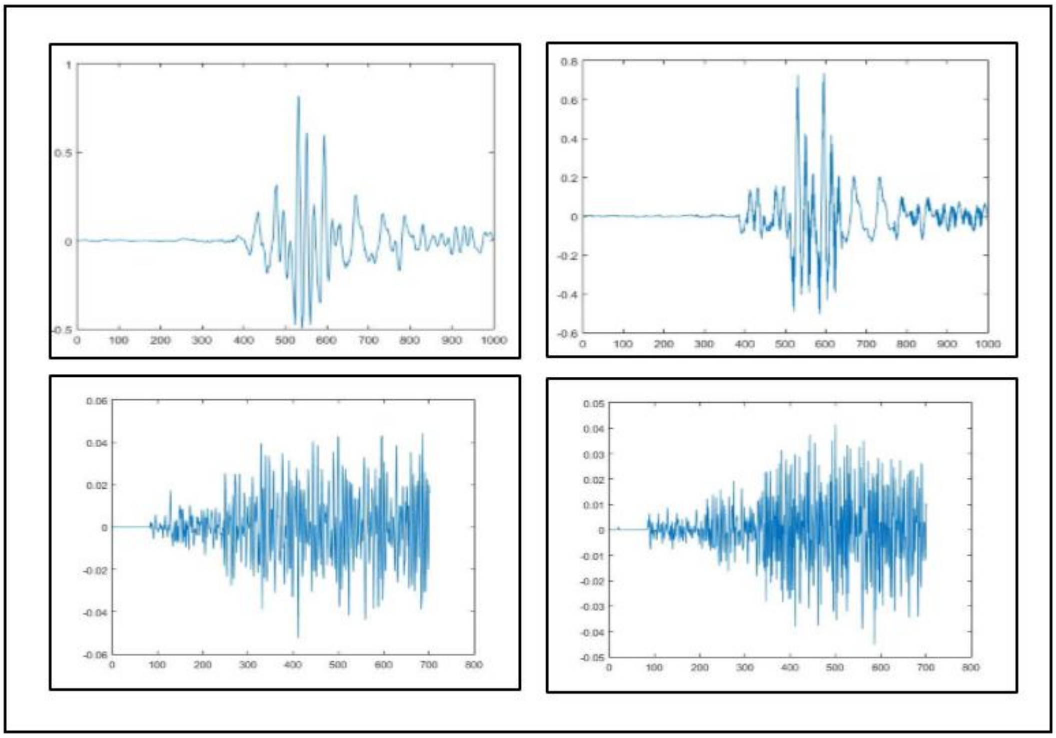

Figure 9: speech signal and reconstucted signal

CONCLUSION

Hybrid speech signal compression method based on transform are tested and investigated. The proposed method explored different compression factor, in all try to keep quality of speech in acceptable range. The experimental results show the compression method is equally performed for various types of signals. The experimental results show compression factor up to 10.822. In future work, other transforms may apply such discreet cosine transform (DCT) or Hilbert transform to generate data in other domain and can quantization to reduce the bit rate of stream of compressed data.

REFERENCES

[1] Brigham, E. Oran "*The fast Fourier transform and its applications*" Englewood Cliffs, N.J. Prentice Hall, 1998.

[2] Wai C. Chu,"*Speech Coding Algorithm Foundation and Evolution of Standardized Coders*", WILEY, 2003.

San Jose, California

[3] Rohit Thanki, Komal Borisagar and Vedvyas Dwivedi, "*A Hybrid Signal Compression Technique using CS Theory and Various Transform for Speech and ECG", Atmiya Institute of Technology and Science Rajkot, India 2013.*

[4] Yuhui Li, Wei Gou and Bo Li, *"A New Digital Watermark Algorithm Based on the DWT and SVD",* IEEE, 2011.

[5] Ranjeet Kumar, A. Kumar and G.K. Singh, *"Hybrid method based on singular value decomposition and embedded zero tree wavelet technique for ECG signal compression"* IEEE, 2016.

[6] Mallat, S. G. "*An Image coding using wavelet transform Sign In or Purchase*" IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, No. 7, 1989.

[7] Wai C. Chu *"Speech Coding Algorithms Foundation and Evolution of Standardized Coders"* Wiley Interscience, 2003.

[8] Amol R. Madane, Zalak Shah, Raina Shah, and Sanket Thakur, "*Speech Compression Using Linear Predictive Coding"* Machine Intelligent Research, 2009.

[9] Wissam Abdulhadi "*Design and Implementation of High Speech Quality Vocoder System at 2400 bps Using MELP Model*" MSc. Thesis, Electrical Engineering University of Baghdad, 2002.

[10] Steven W. Smith *"The scientist and Engineers Guides to Digital Signal Processing"* 2nd edition, California Technical publishing, USA, 1999.