Hybrid Quantization on Discrete Wavelet Coefficients for Efficient Speech Coding

Shijo M Joseph
Department of Computer Science
Mahatma Gandhi College Iritty, Kannur, Kerala, India 670 703
: shijomjose71@gmail.com

Abstract—Among many compression techniques discrete wavelet transform- based speech coding is the most recent and powerful one. The main advantage of wavelet-based speech or image compression is that it provides better compression percentage and maintains good speech intelligibility or image quality. If speech compression is performed using wavelet transform and quantization, there is an enhancement in compression performance without much distortion of the reconstructed signal. This paper proposes a new technique for speech coding by applying hybrid quantization on wavelet coefficients

Keywords— *Speech coding, wavelet transform, hybrid quantization*

I INTRODUCTION

Data compression is the process of converting data files into smaller files for efficiency of storage and transmission. As one of the enabling technologies of multimedia revolution, data compression is a key to the rapid progress being made in information technology. Data compression algorithms are used to reduce the number of bits required to represent a speech signal or an image [1]. There have been many types of compression algorithms developed.

Quantization is one of the simplest and earliest methods of lossy compression. Quantization is a process of mapping an infinite set of scalar or vector quantities by a finite set of scalar or vector quantities. Quantization has applications in the area of signal processing, speech processing and image processing [2]. In speech coding, quantization is required to reduce the number of bits used to represent a sample of speech signal. When less number of bits is used to represent a sample the bit-rate, complexity and memory requirement get reduced [3]. Quantization results in the loss in the quality of a speech signal, which is undesirable. So a compromise must be made between the reduction in bit-rate and the quality of speech signal [4].

Speech coding techniques are mostly based on lossy algorithms. Lossy algorithms are considered acceptable when encoding speech quality is undetectable by the human ear [5].

Usually uncompressed speech signals are transmitted at 64 kb/s, using 8 KHz for sampling. Any bit rate below 64 kb/s is considered as compression [6]. There are many lossy speech compression techniques that are in use. They are Linear Predictive Coding (LPC-10), Code Excited Linear Prediction (CELP), Pulse Code Modulation (PCM), Adaptive Differential Pulse Code Modulation (ADPCM) etc. [7].

Aim of this paper is to analyze the effectiveness of wavelet-based hybrid quantization in Malayalam speech coding. This paper is organized as follows: A brief introduction of speech compression is presented in section I. Section II gives overview about wavelet decomposition and reconstruction. Sections III furnish the details of speech data base created for this experiment. Section IV discuss the methodology used in this experiment. Finally Section V presents the simulation and result and concludes this paper with remarks in section VI.

II DISCRETE WAVELET TRANSFORMS

Discrete Wavelet Transform (DWT) has emerged as a powerful mathematical tool in many areas of science and engineering, especially in the field of speech and image compression which uses multi resolution filter banks for the signal analysis. A wavelet is a waveform of effectively limited duration that has an average value zero [8].The basic idea of the wavelet transform is to represent an arbitrary signal, 'S' as a super position of a set of such wavelets or basis functions. These basis functions are obtained from a single prototype wavelet called the mother wavelet by dilation and translation. The Discrete Wavelet Transform for one dimensional signal is defined in the following equation.

$$W(j, K) = \Sigma_j \Sigma_k X(k)\, 2^{-j/2} \Psi(2^{-j}n-k) \tag{1}$$

Where $\Psi(t)$ is the basic analyzing function called the mother wavelet

In DWT a time-scale representation of a signal is obtained by digital filtering techniques. A low frequency component of a signal is more significant than high frequency components since the low frequency components have maximum information content. The DWT is computed by successive low pass filtering and high pass filtering of the discrete time domain signal. This algorithm is called the Mallat algorithm

[9]. At each level the decomposition of the input signal has two kinds of outputs. The low frequency components are known as the approximations a[10] and high frequency components are known as the detailed d[n]. At each decomposition level, the half band filters produce signals spanning only half the frequency band. This doubles the frequency resolution as the uncertainty in frequency is reduced by half. With this approach, the time resolution becomes good at high frequencies, while the frequency resolution becomes approximately good at low frequencies [11]. The filtering and decimation process is continued until the desired level is reached. Fig.1 shows the decomposition by using Discrete Wavelets Transforms.

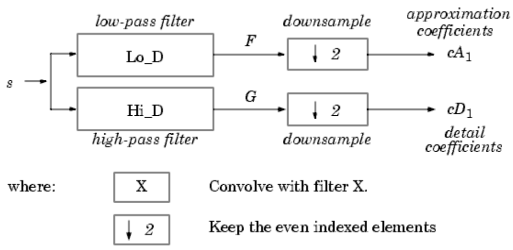

Figure 1 Signal Decompositions (DWT)

The DWT of the original signal is then obtained by concatenating all the coefficients, a[n] and d[n], starting from the last level of decomposition. The successive high pass and low pass filtering of the signal can be depicted by the following equations:

$$Y_{high}[k]= \Sigma_n x[n]g[2k-n] \quad (2)$$

$$Y_{low}[k]= \Sigma_n x[n]h[2k-n] \quad (3)$$

Where Y_{high} and Y_{low} are the outputs of the high pass and low pass filters obtained by sub sampling by 2 [12].

A. *Signal Reconstruction*

The original signal can be reconstructed or synthesized using the Inverse Discrete Wavelets Transforms (IDWT)

The synthesis starts with the approximation and detail coefficient cAj and cDj, and then reconstructs cAj-1 by up sampling and filtering with the reconstruction filters. Fig 2 shows the signal reconstruction process by using Inverse Discrete Wavelet Transforms.

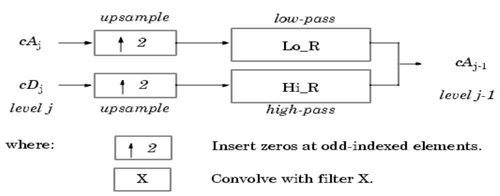

Figure.2. Signal Reconstruction (IDWT)

The reconstruction filters are designed in such a way as to cancel out the effect of aliasing introduced in the wavelet decomposition phase. The reconstructed filters (Lo_R and Hi_R) together with the low and high pass decomposition filters, form a system known as Quadrature Mirror Filters (QMF) for multilevel analysis. The reconstructed process can itself be iterated producing successive approximation at finer resolution and finally the original signal can be synthesized [13].

The discrete wavelet transforms based decomposition and inverse discrete wavelets transform based reconstruction of speech signal is depicted in fig.3.

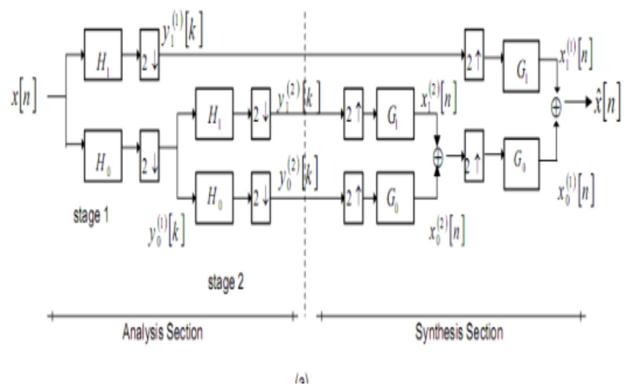

Figure. 3 Multi level DWT based signal analysis and synthesis

III. SPEECH DATA BASE

Aim of this section is to give details of the speech data base created for this experiment. For this part of the experiment continuous speech samples were collected from six hundred individuals of different age and gender. First sentence of the Indian National pledge was selected for the creation of database. This sentence was uttered in Malayalam (one of the regional languages). The first sentence consisted of three words and had an average duration of 1.8 seconds. Each word contained vowels, consonants, fricatives and diphthongs. The Malayalam sentences and its corresponding International Phonetic Alphabet (IPA) are given in table 1. Speech samples were recorded using unidirectional USB microphone with following specifications: frequency response 100-10,000 Hz and sensitivity: 62dBV/μBar -42 dBV/Pa +/-d 3dB. "Goldwave" the software at a sampling rate of 8KHz (4KHz) band limited was used for capturing and processing the speech signal. Speech samples were recorded at different environments like office rooms, class rooms, research labs and normal noise environments etc. The speakers were native speakers of Malayalam and free from speech disabilities. The recorded speech samples are trimmed by removing prefix and suffix silence from the utterances and it is labelled and stored in the speech data base for further processing.

.

TABLE 1.
SPEECH SENTENCE AND ITS IPA FORMAT

Malayalaam	§Lc /®æa/øÞ¼cÎÞÃí
English	India/ ente/ rajamanu
IPA	'indiaː'entə rəːdzrːæm aːnʊ

IV HYBRID QUANTIZATION

Hybrid quantization on wavelet coefficient is a new algorithm is proposed for low complexity and efficient speech coding. The proposed algorithm has the following steps.

1. Divide the speech signal in to 'n' number of frames of size 1024 bytes (1KB).
2. Speech signal transformed from time domain to time scale domain using discrete wavelet transform technique. This transform unreel the redundancy within the input signal. This process creates the wavelet feature vector.
3. Find the mean value of the wavelet coefficients by using the formulae:

$$m = \frac{1}{k}\sum_{i=1}^{k} X_i \tag{1.1}$$

where i is ranging from 1 to the number of the wavelet coefficient.

4. Subtract mean value from each wavelet coefficients. Resultant coefficient is known as mean removed residual (MRR) or simply residual coefficient.
5. The mean removed (MRR) signal has so many insignificant coefficients. Remove such insignificant coefficients by applying uniform dead zone quantization (Scalar quantization). Uniform dead zone quantization will not make significant degradation in the reconstructed speech signal.
6. Encode the remaining wavelet coefficients using any entropy coding algorithm.
7. Transmit the encoded signal along with mean value.
8. On receiver side decode the encoded signal.
9. Add mean value 'm' to the entire wavelet coefficient. Where i is ranging from 1 to number of wavelet coefficient.
10. Apply inverse transform to reconstruct the original signal.
11. Repeat the step 2 to 10 for n times. Where n is number of frames.

Various steps involved in the proposed algorithm is depicted in figure 1.

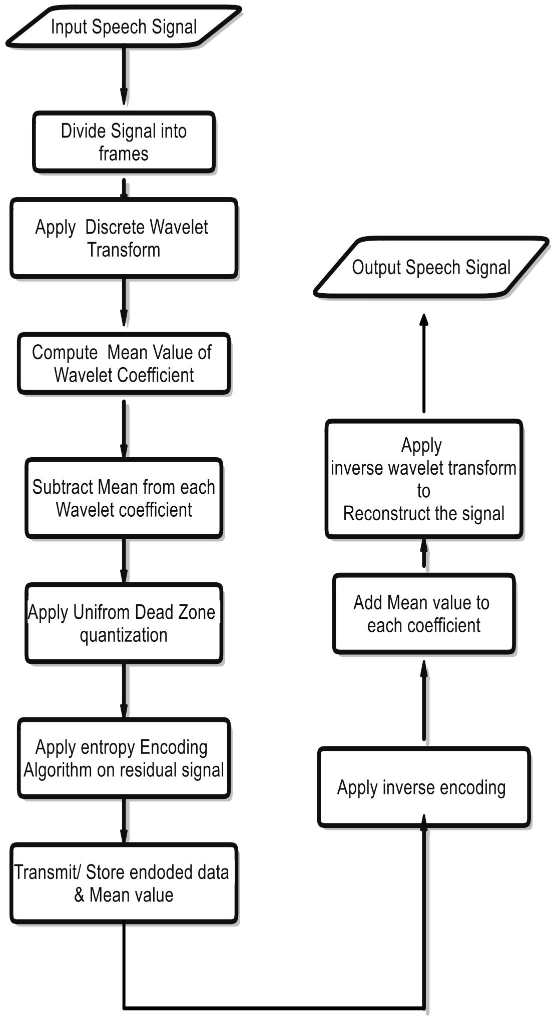

fig 4 Various steps involved in the proposed algorithm

V EXPERIMENTATION

In this section, we present the details of experiments conducted to demonstrate our proposed algorithm for coding the speech signal. We use one continuous speech signal from KUSDBase (Mlalayalam) speech databases for the experiment. The results of the proposed speech compression algorithm are presented.

We chose bior 3.1 as the optimal wavelet for decomposing the speech signal and level 3 is the best decomposition level beyond that it is observed that there is degradation in the qualitative and quantitative parameters[14]. That means we may get more compression ratio but it decrease quality of reconstructed signal. Decompose the signal using bior 3.1 wavelet up to the level 3.

Apply non uniform dead zone quantization on mean removed Wavelet coefficient. In this stage we are removing only insignificant wavelet coefficients, which do not make much distortion in the reconstructed signals at the same time it gives better compression in the speech signal.

Encode the remaining coefficients with a lossless entropy encoding algorithm. Here we are using run length entropy encoding scheme to code the coefficients.

A. *Experimental Setup &Results*

To examine the efficiency of the candidate system we conducted hybrid vector quantization experiment on wavelet feature vector. Mother wavelet used for this experiment is bior 3.1 and decomposed speech signal at level 3.

Different sampling frequency signals were used for the research.

B. *Experiment on Malayalam Database*

The following parameters were fixed for the experiment.

Bandwidth of speech signal	=4 KHz
Sampling rate	=8000Hz (Samples/Sec)
Window type	= Rectangular
Window size	=1024 samples per frame

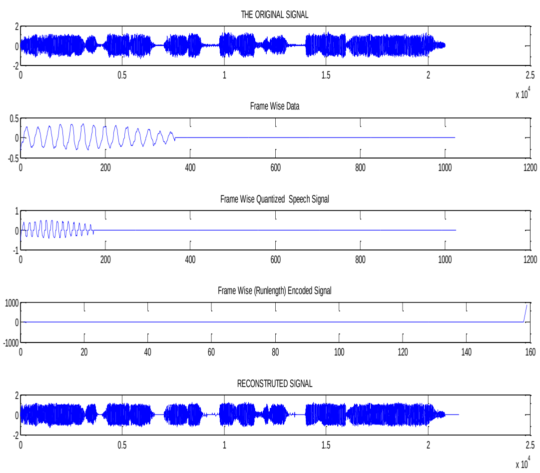

Figure 5. Hybrid quantization(WT-HQ) based speech coding a) the original speech signal b)original speech signal within the frame, c) quantized speech signal. d) Compressed speech signal. E) the reconstructed speech signal.

TABLE 2 COMPARATIVE STUDY BETWEEN CONVENTIONAL SPEECH CODING AND THE PROPOSED SYSTEM

Method	Sentence 1				
	SNR	PSNR	NRMSE	RSE	C-PR%
WT based Speech Coding	12.11	21.95	0.66	96.47	82.99
WT based HQ Speech Coding	4.257	20.95	0.735	88.92	89.52

Table 2 gives the result obtained by a comparative study between simple wavelet transform-based speech coding and wavelet transform-based hybrid quantization (WT-HQ) technique. Figure 3 graphically illustrates this study.

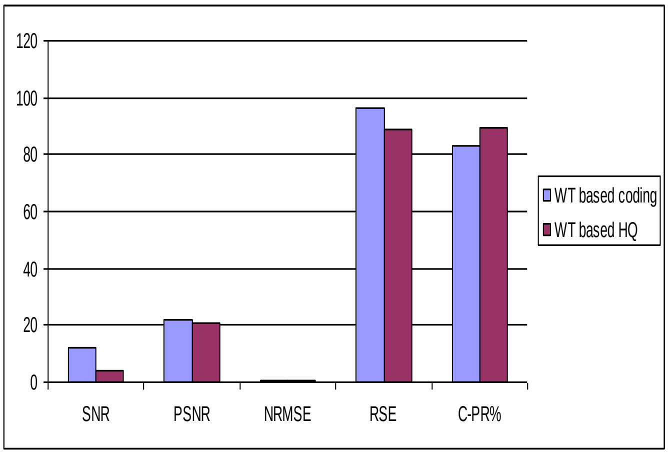

Figure 6 Performance comparison between classical wavelet and WT-HQ based speech coding

Experiment conducted on Malayalam speech database with different sampling frequency to compare the effect of different frequency variation in speech coding.

Experiment conducted on Malayalam speech database with different wavelet families to compare the effect of different wavelet features in speech coding Table 3 gives performance evaluation of different wavelet families on the WT-HQ based speech coding algorithm.

TABLE 3 PERFORMANCE EVALUATION OF DIFFERENT WAVELET FAMILY ON WAVELET ENERGY-BASED AND WT-HQ BASED SPEECH CODING

wavel	SNR	PSNR	NRMSE	RSE	C-PR%
bior.1	4.257	20.95	0.735	88.927	89.529
coif5	3.261	19.77	0.624	82.82	84.141
db8	4.187	19.71	0.627	82.687	88.798
dmey	4.429	19.96	0.612	85.365	82.642
rbio3.	4.892	17.63	0.79	84.364	89.229
sym5	3.097	19.62	0.642	82.562	87.711

Figure 7 illustrates the performance of the selected wavelets from different wavelet families on the proposed speech coding algorithm.

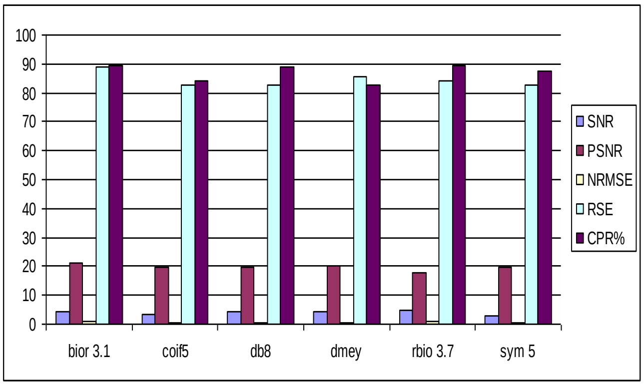

Figure.7 Performance of selected wavelets from different wavelet families on WT-HQ based speech coding

VI SUMMARY

In this paper we introduced the new technique, i.e., the wavelet transform-based hybrid quantization (which is a combination of mean removed vector quantization and dead zone scalar quantization) on discrete wavelet transform feature vector. The candidate system has the good qualities of both scalar quantization and vector quantization. The proposed technique is simple to implement and hybrid quantization requires less computation. The system is evaluated quantitatively and qualitatively and found superior in terms of compression percentage and mean opinion score. The reconstructed signal has structural similarity with the original one; it is a clear indication that there is further possibility of processing the reconstructed signal.

VII. REFERENCE

[1]. Rao.R.M&A.S, Wavelet Transforms: Introduction to Theory and Applications, Pearson Education Pvt.Ltd, 2004.
[2]. Taubman&Marcellin, Jpeg2000: Image Compression Fundamentals, Standards, and Practice, Springer-Verlag GmbH, 2002.
[3]. Rao, Wavelet Analysis and Applications, New Age International Publishers, India, 2004.
[4]. Ram, Hybrid Vector Quantizers for Low Bit Rate Speech Coding Application, in: Electronics and Communication Engineering, Jawaharlal Nehru Technological University, India, 2010.
[5]. Benesty, Sondhi&Huang, Springer Handbook of Speech Processing, Springer, Germany, 2007.