IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)
Volume 4, Issue 1, Ver. II (Jan. 2014), PP 34-38
e-ISSN: 2319 – 4200, p-ISSN No. : 2319 – 4197
www.iosrjournals.org

Analysis of Different Discrete Wavelet Transform Basis Functions in Speech Signal Compression

Sai Lakshmi Bhamidipati[1], Sai Sudha Mindagudla[2],
Harsha Vardhan Devalla[3], Hima Sagar Goodi[4], Hemanth Nag[5]
Student(IV), Dept of ECE, LIET, Vizianagaram, Andhra Pradesh, India.
Student(IV), Dept of ECE, LIET, Vizianagaram, Andhra Pradesh, India.
Student(IV), Dept of ECE, LIET, Vizianagaram, Andhra Pradesh, India.
Student(IV), Dept of ECE, LIET, Vizianagaram, Andhra Pradesh, India.
Assistant Professor, Dept of ECE, LIET, Vizianagaram, Andhra Pradesh, India.

Abstract: *In this paper we attempt to evaluate the challenge of compression of speech signals. Compression of speech signals have pre-dominantly occupied a considerable position in the present era of multimedia. Speech compression is one area of digital signal processing that focuses on reducing the bit rate of the speech signal for transmission or storage without significant loss of quality. In recent years a new technique called wavelet transform has been proposed for signal analysis. It has been successfully used in image compression application. So far, less attention has been paid to the research in the speech compression using wavelet. Here we assess the compression of speech signal using different discrete wavelet transform basis functions. There are different wavelet basis families like haar wavelet, daubechies wavelet, biorthagonal spline wavelet, coiflet wavelet, meyer wavelet, reverse biorthogonal wavelet, Shannon wavelet, symlet wavelet. The auditory masking method and psycho acoustic methods are used to compress the speech. At first the speech is divided into number of frames and upon each frame wavelet transformation is used to minimize number of bits required to represent frame while keeping any distortion inaudible. MATLAB code is implemented to perform the compression. This paper performs evaluation of different wavelet families on the basis of compression scores each contributing itself in the compression of speech.*

Key words*: wavelet transforms, wavelet families, psychoacoustic method.*

I. Introduction:

The growth of the computer industry has invariably led to the demand for quality audio data. Compared to most digital data types, the data rates associated with uncompressed digital audio are substantial. For example, if we want send high-quality uncompressed audio data over a modem, it would take each second's worth of audio about 30 seconds to transmit. This means that the data would be gradually received, stored away and the resulting file played at the correct rate to hear the sound. However, if real-time audio is to be sent over a modem link, data compression must be used. In a digital system, the bit rate is the product of the sampling rate and the number of bits in each sample. The difference between the information rate of a signal and its bit rate is known as redundancy. Compression systems are designed to eliminate this redundancy. So, compression of speech is done in order to achieve reduced storage requirements and over all execution time. With respect to transmission of data, the data rate is reduced at the source by the compressor (coder) ,it is then passed through the communication channel and returned to the original rate by the expander(decoder) at the receiving end. The compression algorithms help to reduce the bandwidth requirements and also provide a level of security for the data being transmitted. In this paper wavelet analysis is being used to achieve speech compression.

1.1Wavelets:

The advent of wavelets has been started with the failure of fourier series in speech analysis. Wavelet theory is applicable to several subjects. All wavelet transforms may be considered forms of time frequency representation for continuous-time (analog) signals and so are related to harmonic analysis. Almost all practically useful discrete wavelet transforms use discrete-time filterbanks. These filter banks are called the wavelet and scaling coefficients in wavelets nomenclature. These filterbanks may contain either finite impulse response (FIR) or infinite impulse response(IIR) filters. The wavelets forming a continuous wavelet transform (CWT) are subject to the uncertainty principle of Fourier analysis respective sampling theory: Given a signal with some event in it, one cannot assign simultaneously an exact time and frequency response scale to that event. The product of the uncertainties of time and frequency response scale has a lower bound. Thus, in the scaleogram of a continuous wavelet transform of this signal, such an event marks an entire region in the time-scale plane, instead of just one point. Also, discrete wavelet bases may be considered in the context of other forms of the uncertainty principle.

1.2Discrete Wavelet Transform:

It is computationally impossible to analyze a signal using all wavelet coefficients, so one may wonder if it is sufficient to pick a discrete subset of the upper halfplane to be able to reconstruct a signal from the corresponding wavelet coefficients. One such system is the affine system for some real parameters $a > 1$, $b > 0$. The corresponding discrete subset of the halfplane consists of all the points (a^m, na^mb) with m, n in **Z**. The corresponding *baby wavelets* are now given as $\psi_{m,n}(t) = a^{-m/2}\psi(a^{-m}t - nb)$. A sufficient condition for the reconstruction of any signal x of finite energy by the formula $x(t) = \sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}}\langle x,\ \psi_{m,n}\rangle \cdot \psi_{m,n}(t)$ is that the functions $\{\psi_{m,n} : m, n \in \mathbb{Z}\}$ form a thight frame of $L^2(\mathbf{R})$.The function $\varphi_{mn}(t)$provides sampling points on the scaletime plane.linear sampling in the time (y-axis) direction but logarithmic in the scale (x-axis) direction.The most common situation is that a_0 is chosen as: a_0=$2^{\frac{1}{v}}$, where v is an integer value,and that v pieces of $\varphi_{mn}(t)$ are processed as one group,which is called a voice.the integer v is the number of voices per octave;it defines a well-tempered scale in the sense of music.This is analogous to the use of a set narrowband filters in conventional fourier analysis. Wavelet analysis is not limited to dyadic scale analysis.By using an appropriate number of voices per octave,wavelet analysis can effectively perform the $1/_3$- octave,$1/_6$-octave,$1/_{12}$octave analyses that are used in acoustics.The wavelet basis function can be implemented as an FIR OR an IIR filter depending on the particular properties needed.

1.3Basis Wavelet Functions:

HAAR WAVELET: In mathematics, the Haar wavelet is a sequence of rescaled "square-shaped" functions which together form a wavelet family or basis. Wavelet analysis is similar to Fourier analysis in that it allows a target function over an interval to be represented in terms of an orthonormal function basis. The Haar sequence is now recognised as the first known wavelet basis and extensively used as a teaching example. The Haar wavelet's mother wavelet function $\psi(t)$can be described as

$$\psi(t) = \begin{cases} 1 & 0 \le t < 1/2, \\ -1 & 1/2 \le t < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Its scaling function $\phi(t)$can be described as

$$\phi(t) = \begin{cases} 1 & 0 \le t < 1, \\ 0 & \text{otherwise.} \end{cases}$$

BIORTHOGONALWAVELET: A biorthogonal wavelet is a wavelet where the associated wavelet transform is invertible but not necessarily orthogonal . Designing biorthogonal wavelets allows more degrees of freedom than orthogonal wavelets. One additional degree of freedom is the possibility to construct symmetric wavelet functions.

$$\sum_{n\in\mathbb{Z}} a_n \tilde{a}_{n+2m} = 2 \cdot \delta_{m,0}$$

.

Then the wavelet sequences can be determined as $b_n = (-1)^n \tilde{a}_{M-1-n} \qquad (n = 0, \ldots, N-1)$ $\tilde{b}_n = (-1)^n a_{M-1-n} \qquad (n = 0, \ldots, N-1)$.

DAUBECHIES (dB): The Daubechies wavelets, based on the work of Ingrid Daubechies, are a family of orthogonal wavelets defining a discrete wavelet transform and characterized by a maximal number of vanishing moments for some given support. With each wavelet type of this class, there is a scaling function (called the *father wavelet*) which generates an orthogonal multi resolution analysis.

SYMLET WAVELET (SYM): Symlet Wavelet also known as "least asymmetric" wavelet, defines a family of orthogonal wavelets. Symlet Wavelet is defined for any positive integer n. The scaling function (ϕ) and wavelet function (ψ) have compact support length of 2n. The scaling function has n vanishing moments. Symlet Wavelet can be used with such functions as discreet wavelet transform and Wavelet Phi, etc.

COIFLET WAVELET: Coiflet scaling functions also exhibit vanishing moments. In `coifN`, `N` is the number of vanishing moments for both the wavelet and scaling functions. These filters are also referred to in the literature by the number of filter taps, which is `2N`.

MEYERWAVELET: Dmey wavelet is the discrete format of meyerwavelet function. Mayer's wavelet as shown in equation is fundamentally a solvent method for solving the two-scaleequation. Given a basis '□' for the approximation space, Meyer employed Fourier techniques to derive the DTFT of the two-scale education coefficients.

$$G_0\left(e^{jw}\right) = \sqrt{2}\ \ \varepsilon_k\ \emptyset(2\omega + 4k\pi)$$

SHANNONWAVELET: In functional analysis Shannon wavelet may be either of real or complex type. signal analysis by ideal band pass filters defines a decomposition known as Shannon wavelets(sinc wavelets).The haar and sync wavelets are fourier duals of each other. The Shannon transform is given by.

$$\Psi^{(\mathrm{Sha})}(w) = \prod\left(\frac{w-3\pi/2}{\pi}\right) + \prod\left(\frac{w+3\pi/2}{\pi}\right).$$

II. Proposed Algorithms:

ENCODING: Signal compression is achieved by first truncating small valued coefficients and then efficiently encoding them. Another approach to compression is to encode consecutive zero valued coefficients, with two bytes. One byte to indicate a sequence of zeros in the wavelet transforms vector and the second byte representing the number of consecutive zeros.

PSYCHOACOUSTIC MODEL**:**

The human auditory system has a dynamic frequency range from about 20Hz - 20 kHz, and the intensity of the sound as perceived by us varies. However, we are not able to perceive sounds equally well at all frequencies. In fact, hearing a tone becomes more difficult close to the extreme frequencies (i.e. close to 20 Hz and 20 kHz). Further study exhibits the concept of critical bands which is the basis of audio perception. A critical band is a bandwidth around a center frequency, within which sounds with different frequencies are blurred as perceived by us. Critical bands are important in perceptual coding because they show that the ear discriminates between the energy in the band and the energy outside the band. It is this phenomenon that promotes masking.

In this implementation, the following were determined:

_ Tone maskers

Tone Masker determining whether a frequency component is a tone requires knowing whether it has been held constant for a period of time, as well as whether it is a sharp peak in the frequency spectrum, which indicates that it is above the ambient noise of the signal. A frequency f (with FFT index k) is a tone if its power P[k] is:1. greater than P[k-1] and P[k+1], i.e., it is a local maxima 2. 7 dB greater than the other frequencies in its neighborhood, where the neighborhood is
dependent on f:

o If 0.17 Hz < f <5.5 kHz, the neighborhood is [k-2…k+2].

o If 5.5 kHz _ f <11 kHz, the neighborhood is [k-3…k+3].

o If 11 kHz _ f <20 kHz, the neighborhood is [k-6…k+6].

COMPANDING: In telecommunication and signal processing companding (occasionally called **compansion**) is a method of mitigating the detrimental effects of a channel with limited dynamic range. The name is a portmanteau of **com**pressing and ex**panding**. The use of companding allows signals with a large dynamic range to be transmitted over facilities that have a smaller dynamic range capability. Companding is employed in telephony and other audio applications such as professional wireless microphones and analog recording.

QUANTISATION:

Quantization, in mathematics and digital signal processing, is the process of mapping a large set of input values to a (countable) smaller set – such as rounding values to some unit of precision. A device or algorithmic function that performs quantization is called a **quantizer**. The round-off error introduced by quantization is referred to as quantization error.

III. Implementation:

The implementation in mat lab takes place with the execution of code, which is made based on the proposed algorithms. The steps used are reading a sound file, performing wavelet decomposition, selecting a threshold to truncate the coefficients, using suitable encoding scheme to get rid of truncated coefficients, decoding the received signal, reconstructing the samples of speech signal and construction of samples. Some of the instructions used are [y, fs, bfs]= wavread (‘path of the file’); to read sound file. To perform wavelet decomposition [C L]=wavdec (y, N, ’wname’) is used. Where N=number of wavelets and ‘wname’ is wavelet name. to perform wavelet compression [X, CXC, LXC, PERF0, PERFL2]= wdencmp (‘gbl’, C, L, wlet, decomplevel, the,sorh, keepapp); is used. To decode the wavelet used command is Rx= decode1 (y, posnum, N); to reconstruct the signal Y= waverec(C, L, ‘wavelet’); is used

IV. Results:

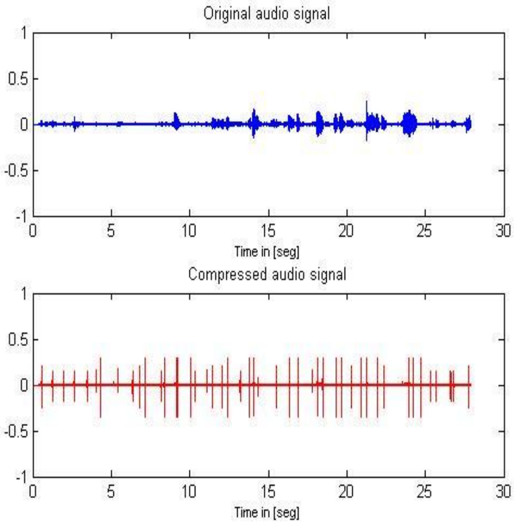

Fig1**:** Output of dB10

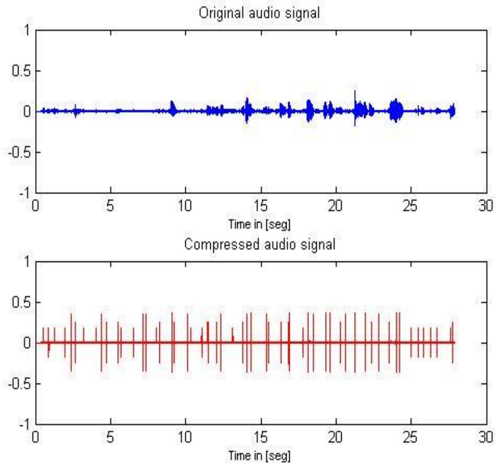

Fig2: Output of haar

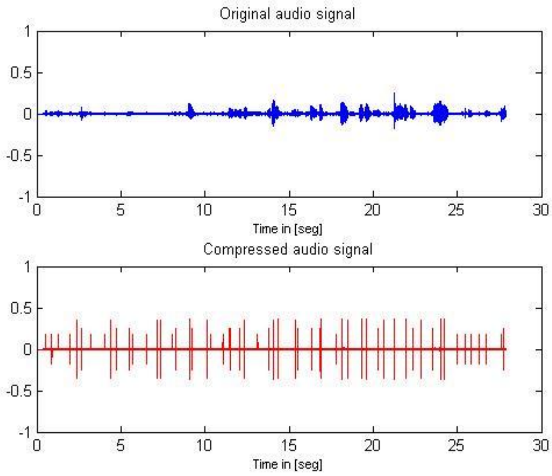

Fig3: output of bior1.3

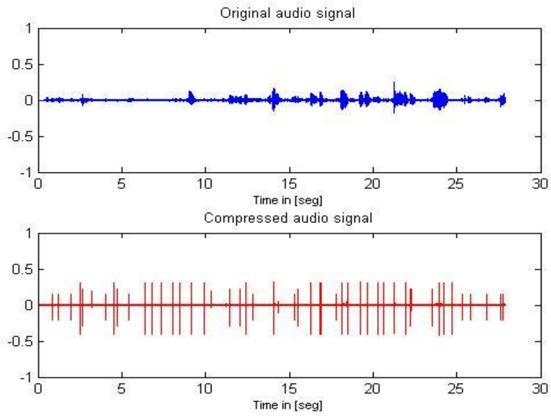

Fig4: output of demy:

S.NO	WAVELET FAMILY	ZEROS (%)	RETAINED ENERGY	DISTORTION	SNR
1	db10	50.2140	99.3756	4.395e-004	8.6577
2	haar	35.6274	99.439	4.8547e-004	12.9152
3	symlet	43.1929	99.3735	4.7812e-004	20.8795
4	coiflet	43.5101	99.3709	4.7234e-004	9.2899
5	biorthogonal	35.304	99.54	4.7234e-004	19.8053
6	Reverse biorthogonal	47.3401	99.3030	4.9166e-004	13.4852
7	dmey	49.0894	99.4450	4.5393e-004	10.0332

V. Conclusion Remarks:

In this paper we have made an attempt to analyse various basis functions of discrete wavelet transform. some of the wavelet families contributed less for the purpose of compression of speech, while db10 was robust in the application. zeros(%),retained energy, signal to noise ration and compression ratio was considerable in case of certain wavelet families. The noticed point upon the performed comparisions between different wavelet families is that ”symlet” is having high snr ratio of order20.8795.

VI. Future Scope:

The output audio file which is obtained after compression has included certain amount of noise. Though the noise is negligible, our future work is concentrated on eliminating the noise using wiener filters using techniques of Two-Step Noise Reduction(TSNR), Harmonic Regeneration Noise Reduction(HRNR) and for further better performance we can also use “kalman filter”.

References:

[1] S. M. Joseph, “Spoken digit compression using wavelet packet,”, International conference on signal and image processing, PP: 255-259, 2010.

[2] M. A. Osman and N. Al, “Speech compression using LPC and wavelet,” 2nd International conference on computer engineering and technology, PP: V7-92 – V7-99, 2010.

[3] J. Y. Zhang and L. X. Gao, “Research on the selection of wavelet function for the feature extraction of shock fult in the bearing diagnosis,” International conference on wavelet analysis and pattern recognition, PP: 1630-1634, 2007.

[4] R.S.Cai and Y.M.Guo, “Wavelet based multi feature V/UV speech classification algorithm,” International conference wireless mobile and sensor networks, PP: 897 – 900, 2007.

[5] S. Ranjan. “Exploring the DWT as a tool for Hindi speech recognition,” International journal of computer theory and engineering, PP: 642-646, 2010.

[6] Z. Dan and M. S. Qian, “Sppeech compression with best wavelet package transform and SPIHT algorithm,” 2nd International conference on computer modeling and simulation, PP. 360-363, 2010.

[7] Kinsner and A.Langi, “Speech and Image Signal Compression with wavelets”, IEEE Wescanex conference proceedings, IEEE, New York, NY-1993, PP. 368-375.

[8] Kornsing.S.; Srinonchat.J., "Enhancement Speech Compression Technique Using Modern Wavelet Transforms," Computer, Consumer and Control (IS3C), 2012 International Symposium on , vol., no., pp.393,396, 4-6 June 2012.

[9] E.B.Fgee, W.J.Philips, W.Robertson, “Comparing Audio Compression using Wavelets with other Audio Compression Schemes”, IEEE Canadian Conference on Electrical and Computer Engineering, IEEE