International Journal of Future Generation Communication and Networking
Vol.7, No.5 (2014), pp.229-238
http://dx.doi.org/10.14257/ijfgcn.2014.7.5.19

Speech Coding

P. Srinivasa Rao#, Ch. ChandraLekha*, B. SrinivasaReddy*, G. LakshmiDeepthi* and K. Bhargav*

#Asst.Professior, Department of Electronics & Communication Engineering,
St.Anns college of engineering &Technology, India
* Student, Department of Electronics & Communication Engineering,
St.Anns college of engineering &Technology, India
psraoece@gmail.com#, chandralekha.chimakurthy@gmail.com*

Abstract

In modern world, communication service providers are continuously met with the challenge of accommodating more users with in a limited allocated bandwidth. Due to this motivation service providers and manufactures are continuously in search of low bit rate speech coders that deliver high quality speech. The main objective of speech compression is to reduce the bits rate of speech for communication or storage without significant loss of its intelligibility. Wavelet based speech coding is a new technique used for speech compression. The main issue regarding wavelet based speech coding is to choose optimal wavelet to transform the speech signal. In this experiment we use two types of wavelets orthogonal and bi-orthogonal wavelets to transform the speech signal. Run length encoding algorithm is used to encode the threshold coefficient. Performance of this experiment is evaluated in terms of MSE, SNR, PSNR, RSE, and in terms of compression percentage

Keywords*: orthogonal wavelet, bi-orthogonal wavelet, compression*

1. Introduction

The goal of compressing a signal is to minimize the storage or medium capacity needed to hold or to convey the information contained in the signal. The main approaches towards data compression are removing redundancy, thresholding and coding the remaining signal efficiently. Removing redundancy means removing away the correlated data from the original signal and it can be done through a transforms. The thresholding is to set small data to zero when their absolute values are lower than threshold. The coding methods are to code the data with special strategies which are mainly for reducing the transferring volume of data through a communication link.

Speech synthesizing systems generally carry out synthesis via time frequency representations such as Short Time Fourier Transforms (STFT) or Linear Predictive Coding (LPC) techniques. The FT assumes that signals are stationary within a given time frame and may therefore lack the ability to analyze localized events accurately. The main disadvantage of a Fourier expansion however, is that it has only frequency resolution and no time resolution. So, on looking at the signal we cannot say when a particular event occurs. That means in frequency domain we lost the time information.

Speech coding is a vital area of research because of its economical importance. Various algorithms are already applied in different levels of speech coding. If compression is applied on speech data more users can be accommodated at the same time because less bandwidth is required. So it cuts down the cost of communication. Speech compression plays an important

ISSN: 2233-7857 IJFGCN

role in teleconferencing, wireless communication, and audio for videophones or video teleconferencing systems. Other applications include the storage of speech synthesis and play back, or for the transmission of voice at a later time applications. Wavelet is a new tool for analyzing and compressing non stationary signals like speech and audio. Wavelet is a small wave with limited duration and that has an average value of zero. More over wavelets have a beginning and an end. In this paper we are evaluating the performance of two subclasses of discrete wavelet such as orthogonal and bi-orthogonal wavelets in speech coding. Haar, Daubechies, Symlets, Coiflets and Discrete Meyer wavelets are the examples of orthogonal wavelets and Bior is an example of the second category. These are the two classes of wavelet which are commonly used for 1-D and 2-D signal analysis and compression.

Much of the later work in speech compression was motivated by military research s digital communications for secure military radios, where very low data rates were required to allow effective operation in a hostile radio environment. At the same time, far more processing power was available, in the form of VLSI integrated circuits, than was available for earlier compression techniques. As a result, modern speech compression algorithms could use far more complex techniques than were available in the 1960s to achieve far higher compression ratios.

These techniques were available through the open research literature to be used for civilian applications, allowing the creation of digital mobile phone networks with substantially higher channel capacities than the analog systems that preceded them.Speech encoding is an important category of audio data compression. The perceptual models used to estimate what a human ear can hear are generally somewhat different from those used for music. The range of frequencies needed to convey the sounds of a human voice are normally far narrower than that needed for music, and the sound is normally less complex. As a result, speech can be encoded at high quality using a relatively low bit rate.

If the data to be compressed is analog (such as a voltage that varies with time), quantization is employed to digitize it into numbers (normally integers). This is referred to as analog-to-digital (A/D) conversion. If the integers generated by quantization are 8 bits each, then the entire range of the analog signal is divided into 256 intervals and all the signal values within an interval are quantized to the same number. If 16-bit integers are generated, then the range of the analog signal is divided into 65,536 intervals.

Wavelet provides an alternative approach to traditional signal processing techniques such as Fourier analysis for breaking a signal in to its constituent parts. In wavelet transform the basic functions are compact in time. This feature allows the wavelet transform to obtain time information about a signal in addition to frequency information. In DWT the original signal is successively decomposed in to low frequency and high frequency components. The high frequency components are not analysed any further. The approximation signal is subsequently divided into new approximation and detailed signals. The process of down sampling by 2 is that of keeping every second sample of x[n] and removing them in between samples thus generating an output sequence $d_2[n]$ .The successive high pass and low pass filtering of the signal can be depicted by the following equations:

$$Y_{high}=\sum_n x[n]g[2k-n] \text{------}(1)$$

$$Y_{low}=\sum_n x[n]h[2k-n] \text{-------}(2)$$

A **wavelet** is a wave-like oscillationwith an amplitude that begins at zero, increases, and then decreases back to zero.

There are two types of discrete wavelets (a) Orthogonal wavelets (b) Bi-Orthogonal wavelets

Orthogonal Wavelets

An orthogonal wavelet is a wavelet whose associated wavelet transform is orthogonal. That is, the inverse wavelet transform is the adjoint of the wavelet transform. If this condition is weakened you may end up with bi-orthogonal wavelets.

The scaling function is a refinable function. That is, it is a fractal functional equation, called the refinement equation.

$$\phi(x) = \sum_{k=0}^{N-1} a_k \phi(2x-k),$$

where the sequence $(a_0, \ldots, a_{N-1})$of is called a scaling sequence or scaling mask. The wavelet proper is obtained by a similar linear combination,

$$\psi(x) = \sum_{k=0}^{M-1} b_k \phi(2x-k),$$

where the sequence $(b_0, \ldots, b_{M-1})$of real numbers is called a wavelet sequence or wavelet mask.

A necessary condition for the orthogonality of the wavelets is that the scaling sequence is orthogonal to any shifts of it by an even number of coefficients:

$$\sum_{n\in\mathbb{Z}} a_n a_{n+2m} = 2\delta_{m,0}$$

In this case there is the same number *M*=*N* of coefficients in the scaling as in the wavelet sequence, the wavelet sequence can be determined as$b_n = (-1)^n a_{N-1-n}$.

Properties of Orthogonal Wavelets

Vanishing moments

A wavelet has m vanishing moments if and only if its scaling function can generate polynomials of degree smaller than or equal to m. While this property is used to describe the approximating power of scaling functions, in the wavelet case it has a "dual" usage, *e.g.*, the possibility to characterize the order of isolated singularities.The number of vanishing moments is entirely determined by the coefficients h[n] of the filter h which is featured in the scaling equation.

If the Fourier transform of the wavelet is p continuously differentiable, then the three following conditions are equivalent:

- The wavelet y has p vanishing moments
- The scaling function j can generate polynomials of degree smaller than or equal to p
- The transfer function of the filter h and its p-1 first derivatives vanish at w=p.

Compact Support

Compactly supported wavelets and scaling functions exist. The scaling function is compactly supported if and only if the filter h has a finite support, and their supports are the same. If the support of the scaling function is $[N_1,N_2]$, then the wavelet support is $[(N_1-N_2+1)/2,(N_2-N_1+1)2]$.Atoms are thus compactly supported if and only if the filter h is

Daubechies has proved that, to generate an orthogonal wavelet with p vanishing moment, a filter h with minimum length 2p had to be used. Daubechies filters, which generate Daubechies wavelets, have a length of 2p. The Daubechies filter coefficients are available as ASCII text files which can be used in a spreadsheet, for instance.

Regularity

Wavelet regularity is much less important than their vanishing moments. Il is studied in a theorem by TchamitchianThe following two properties are important:

- There is no compactly supported orthogonal wavelet which indefinitely differentiable
- For Daubechies wavelets with a large p, the scaling function and wavelet are l-Lipschitz, where l is of the order of 0.2 p. For large classes of orthogonal wavelets, more regularity implies more vanishing moments.

Meyer wavelets are indefinitely differentiable orthogonal wavelets, with an infinite support. They are generally implemented in the Fourier domain.

Symmetry

Symmetric scaling functions and wavelets are important because they are used to build bases of regular wavelets over an interval, rather than the real axis. Daubechies has proved that, for a wavelet to be symmetric or anti-symmetric, its filter must have a linear complex phase, and the only symmetric compactly supported conjugate mirror filter is the Haar filter, which corresponds to a discontinuous wavelet with one vanishing moment. Besides the Haar wavelet, there is no symmetric compactly supported orthogonal wavelet.

Bi-Orthogonal Wavelets

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

Scaling Equation

As in the orthogonal case, y(t) and j(t/2) are related by a scaling equation which is a consequence of the inclusions of the resolution spaces from coarse to fine:

$$\frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{+\infty} g[n]\,\phi(t-n),$$

Similar equations exist for the dual functions which determine the filters h_2 and g_2.

Properties of Bi-Orthogonal Wavelets

Vanishing Moments

A bi-orthogonal wavelet has m vanishing moments if and only if its dual scaling function generates polynomials up to degree m. This can be verified by looking at the bi-orthogonal decomposition formulas. Hence there is an equivalence theorem between vanishing moments and the number of zeroes of the filter's transfer, provided that duality has to be taken into account. Thus the following three properties are equivalent:

- The wavelet y has p vanishing moments
- The dual scaling function j_2 generates polynomials up to degree p
- The transfer function of the dual filter h_2 and it p-1 first derivatives vanish at w=p

and the dual result is also valid. Duality appears naturally, because the filters determine the degree of the polynomials which can be generated by the scaling function, and this degree is equal to the number of vanishing moments of the dual wavelet.

Compact Support

If the filters h et h_2 have a finite support, then the scaling functions have the same support, and the wavelets are compactly supported. If the supports of the scaling functions are respectively $[N_1,N_2]$ and $[M_1,M_2]$, then the corresponding wavelets have support $[(N_1-M_2+1)/2,(N_2-M_1+1)/2]$ and $[(M_1-N_2+1)/2,(M_2-N_1+1)]$ respectively. The atoms are thus compactly supported if and only if the filters h et h_2 are.

Regularity

Tchamitchian's theorem provides again a sufficient regularity condition. Remember that this condition bears on the filter h which determines the scaling equation. Hence the regularity of the primal atoms are related to the primal filters.

Wavelet Balancing

Consider the following decomposition of f:

$$f = \sum_{n,j=-\infty}^{+\infty} \langle f, \psi^{*}_{j,n} \rangle \, \psi_{j,n}$$

The number of vanishing moments of a wavelet is determined by its dual filter. It corresponds to the approximating power of the dual multiresolution sequence. This is why it is preferred to synthesize a decomposition filter h with many vanishing moments, and possibly with a small support.

On the other hand, this same filter h determines the regularity of j, and hence of y. This regularity increases with the number of vanishing moments, that is, with the number of zeroes of h.

Symmetry

Unlike the orthogonal case, it is possible to synthesize bi-orthogonal wavelets and scaling functions which are symmetric or antisymmetric and compactly supported. This makes it possible to use the folding technique to build wavelets on an interval.

If the filters h and h_2 have and odd length and are symmetric with respect to 0, then the scaling functions have an even length and are symmetric, and the wavelets are also symmetric. If the filters have an even length and are symmetric with respect to n=1/2, then the scaling functions are symmetric with respect to n=1/2, while the wavelets are anti-symmetric.

Table 1. Comparison of Orthogonal Wavelets

Parameters	Orthogonal wavelets							
	db2	**db4**	**db6**	**db8**	**db10**	**db12**	**db16**	**db20**
MSE	0.0021	0.0018	0.0017	0.0017	0.0017	0.0016	0.0016	0.0016
PSNR	73.5096	74.1765	74.3443	74.3309	74.4223	74.5539	74.5368	74.6164
RSE	83.4198	83.5799	83.5433	83.5755	83.5671	83.4865	83.5427	83.4387
COMP %	85.6167	87.6639	88.1314	88.0949	88.3429	88.6908	88.6460	88.8523
SNR	28.0998	29.1726	29.2835	29.3714	29.3791	29.7240	29.5096	29.7104

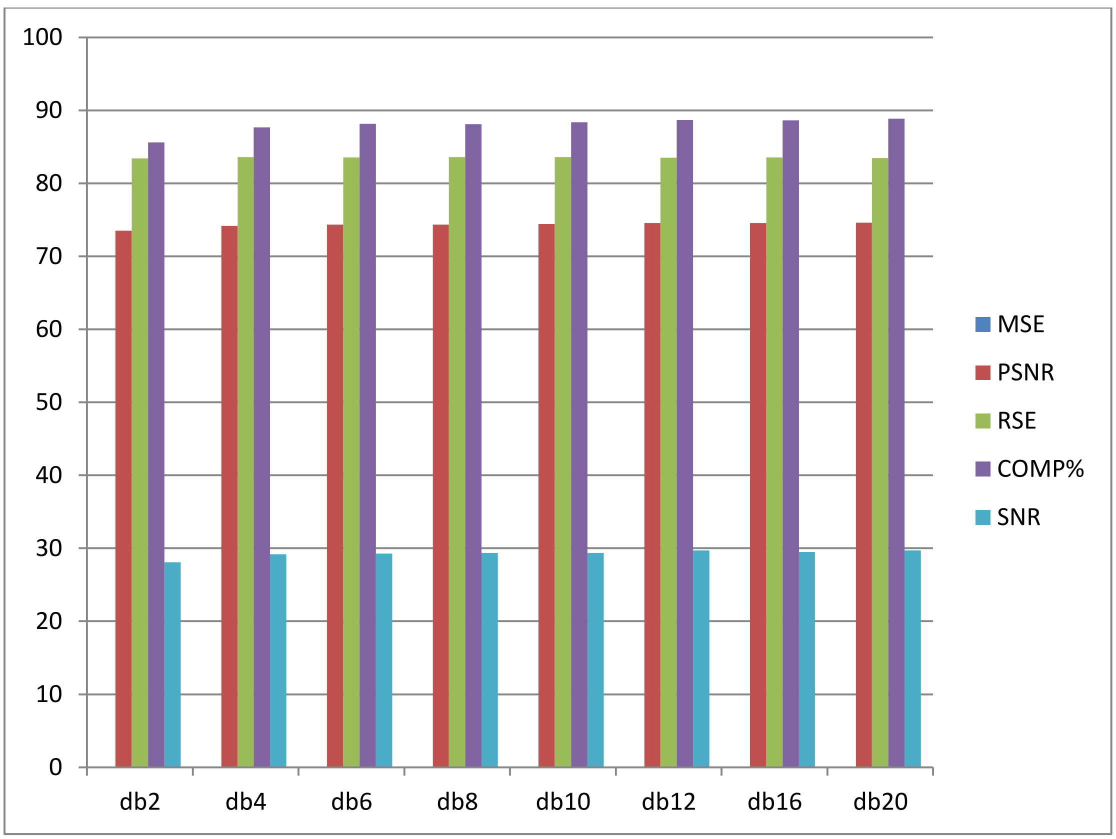

Figure 1. Analysis of Orthogonal Wavelet based Speech Coding

Table 2. Comparison of Bi-orthogonal Wavelets

parameters	Bi-orthogonal wavelets						
	Bior1.3	Bior2.2	Bior2.6	Bior3.1	Bior3.5	Bior3.9	**Bior5.5**
MSE	0.0024	0.0019	0.0017	0.0023	0.0017	0.0016	0.0020
PSNR	72.9149	73.8476	74.2456	72.9827	74.3968	74.6115	73.7335
RSE	82.3288	83.3138	83.0536	82.5323	82.4590	82.3445	84.1007
COMP%	86.1010	91.6526	92.5422	96.2514	95.7644	95.8277	81.1037
SNR	27.1069	27.8556	28.8233	26.1467	28.9790	29.4704	28.2812

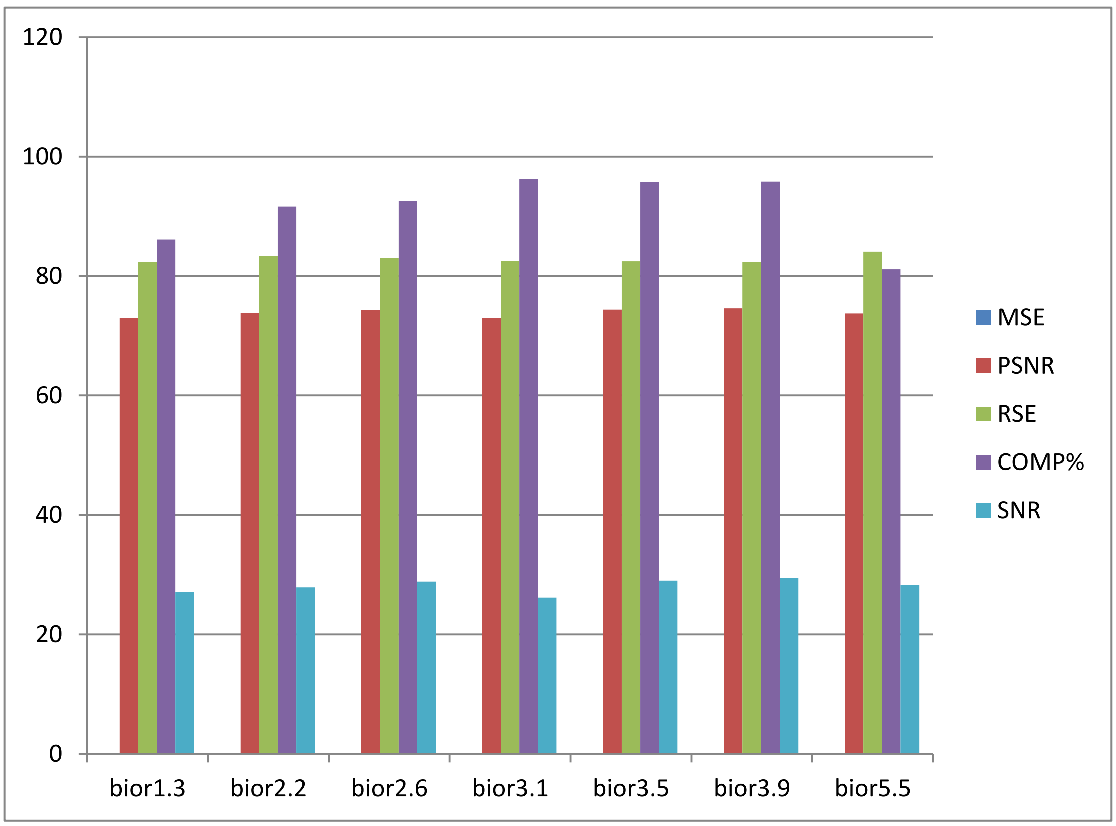

Figure 2. Analysis of Bi-orthogonal Wavelet based Speech Coding

Table 3. Comparison of Orthogonal Wavelets

Parameters	Decomposition Levels				Constant			
	3	4	5	6	18	19	20	21
Db2	85.6167	84.3031	83.6538	83.2518	85.6167	84.5290	83.5815	82.4704
Db4	87.6639	86.4964	85.8077	85.4005	87.6639	86.7998	85.7909	84.9962
Db6	88.1314	86.9898	86.3850	85.9619	88.1314	87.1355	86.2612	85.3964
Db8	88.0949	87.0472	86.3594	85.9369	88.0949	87.2858	86.4258	85.3435
Db10	88.3429	87.1977	86.4993	86.0685	88.3429	87.4485	86.6885	85.6160
Db12	88.6908	87.5208	86.8431	86.4080	88.6908	87.6123	86.6752	85.6991
Db16	88.6460	87.5002	86.7956	86.3556	88.6460	87.8167	86.9400	86.1746
Db20	88.8523	87.8360	87.1137	86.6244	88.8523	87.8669	86.9581	86.9581

Table 4. Comparison of Bi-orthogonal Wavelets

Parameters	Decomposition Levels				Constant			
	3	4	5	6	18	19	20	21
Bior1.3	86.1010	84.7951	84.0922	83.7017	86.1010	82.0250	83.8049	82.6952
Bior2.2	91.6526	90.9988	90.5052	90.2035	91.6526	91.0031	90.3315	89.6310
Bior2.6	92.5422	91.9170	91.3426	90.9695	92.5422	91.9087	91.3410	90.7768
Bior3.1	96.2514	97.2052	97.7624	98.3075	96.2514	95.8968	95.5735	95.2326
Bior3.5	95.7644	95.7430	95.3992	95.2319	95.7644	95.3475	94.9869	94.6482
Bior3.9	95.8277	95.4333	95.0946	94.8569	95.8277	95.4487	95.0532	94.6961
Bior5.5	81.1037	80.1149	79.5003	78.9972	81.1037	79.6118	78.2225	76.9045

This paper discusses the merit of orthogonal and bi-orthogonal wavelet based speech encoding and decoding. We also evaluate the performance in terms of qualitative and quantitative parameters. The difference between the original and the reconstructed speech signal is insignificant. This proves that wavelet is a good tool for speech coding. It may be observed from out study that there is no significant difference between orthogonal and bi-orthogonal wavelets.

Our future work will be to encode continuous speech signal with frames. We shall also examine the performance of different wavelets in speech synthesis and also try to improve the compression ratio by removing silence and noises from the speech signal.

References

[1] L. R. Rabimner , "Digital processing of speech signals", Pearson Education , **(2005)**.
[2] T. F. Quayeri , "Discrete Time Speech Signal Processing Principles and Practices'', Person Publication, **(2001)**.
[3] M. A. Osman, N. Ai, H. M. Magbough and S. A. Alfandi, "Speech compression using LPC and wavelet", 2nd International conference on Computer Engineering and Technology, vol. 7, **(2010)**.
[4] S. Sharma, "Digital Communications", (5th Ed) S.K Kataria&Sons, **(2010)**.
[5] Mallet SA, "Theory for Multiresolution Signal Decomposition: The Wavelet Representation", IEEE Transactions on Pattern Analysis Machine Intelligence. vol. 31, **(1989)**, pp. 674-693.
[6] S. Mallet, "A Wavelet Tour of Signal Processing", San Dieago: Academic Press, **(1999)**.
[7] L. R. Litwin, "Speech Coding with Wavelets", IEEE Potentials, vol. 17, no. 2, **(1998)**, pp. 38-41.
[8] B. Lin, B. Nguyen and E. T. Olsen, , "Orthogonal Wavelets and Signal Processing, Signal Processing Methods for Audio, Images and Telecommunications", P.M. Clarkson and H. Starked, Academic Press London, **(1995)**, pp 1-70.
[9] Daubechies, "Orthogonal Bases of compactly Supported Wavelets", Communication on pure and Applied Mathematics, vol. 41, **(1998)**, pp. 909-906.
[10] S. M Joseph, R. Sebastian and P. BabuAnto, "The effect of different wavelets on speech compression", ACM Proceedings of the 2011 International Conference on Communication, Computing & Security.

Authors

P. Srinivasa Rao, he is currently working as Asst. professor in Electronics & Communication Engineering Department St.anns College of Engineering &Technology. His areas of interest are Signal processing and Communications.

Ch.Chandra Lekha pursuing the B.TECH. degree in Electronics & Communication Engineering in SACET, JNTUK. Her areas of interest include Speech Communication, Wireless Communication. She is a member of ISTE.

B. SrinivasaReddy pursuing the B.TECH. degree in Electronics & Communication Engineering in SACET, JNTUK. His areas of interest include speech communication, wireless communication .He is a member of ISTE.

G. Lakshmi Deepthi pursuing the B.TECH. degree in Electronics & Communication Engineering in SACET, JNTUK. Her areas of interest include Speech Communication, Wireless Communication. She is a member of ISTE.

K. Bhargav pursuing the B.TECH. degree in Electronics & Communication Engineering in SACET, JNTUK. His areas of interest include speech communication , wireless communication .He is a member of ISTE.

