A Novel Approach to Synthesize Sounds of Some Indian Musical Instruments using DWT

Raghavendra Sharma
Department of Electrical Engineering
Dayalbagh Educational Institute
Dayalbagh, Agra, UP, India

V Prem Pyara
Department of Electrical Engineering
Dayalbagh Educational Institute
Dayalbagh, Agra, UP, India

ABSTRACT
In this paper a novel algorithm based on Discrete Wavelet Transform (DWT) approach has been applied to synthesize the sounds produced by a few traditional Indian musical instruments, viz. flute, shehnai and sitar. In this algorithm, the level of decomposition of wavelets is varied till the error norm between the original signal and that generated through DWT is below a desired level. It is observed that when the wavelet decomposition level is varied, the energy retained in wavelet coefficients varies with the type of the wavelet and its decomposition level. It is further observed that the maximum level of decomposition for the three sound signals is different and the signals are also reconstructed with the wavelet coefficients only up-to the maximum level of decomposition with lesser number of samples. The quality of sound as obtained through this algorithm is perceptually close to original sound signal.

General Terms
Wavelet, Quadrature mirror filters, Discrete Wavelet Transform.

Keywords
Approximate Coefficients, Detail Coefficients. Filter bank, Energy in coefficients.

1. INTRODUCTION
Theory of Discrete Wavelet Transform (DWT) and its application to various signal processing problems has been thoroughly developed and documented over last few decades [1]. It has been successfully employed in the area of signal processing, in general, and, speech, music and image processing in particular [2][3]. Further, wavelets have also been used for synthesizing musical signals [4]. In DWT, the signal is analyzed by decomposing it into high and low frequency bands; then these bands are recombined to reconstruct the original signal. Such examples can be found in Sub Band Coding (SBC), de noising signals etc [5]-[7]. The fundamental property of such analysis is the perfect recovery of the original signal from the sub sampled signals. Perfect reconstruction filter bank decomposes the signal into low pass and high pass frequency bands through quadrature mirror filters (QMF) and then, down sampling, which further reconstructs the original signal by up sampling, filtering and summation [8].

An algorithm for designing QMF for perfect reconstruction of signals is proposed in [9]. Algorithms which are essentially linear in nature and easy to implement, are classified on the basis of superiority in terms of peak reconstruction error, and computation time [10]. In this paper, a simple and fast algorithm providing a good compression ratio is proposed. Basically the algorithm makes use of the fact that if we determine the appropriate wavelet on the basis of the maximum energy concentrated in the wavelet coefficients, then, nearly perfect reconstruction of the musical signals is possible by considering these coefficients up to certain specified levels. The rest of the paper is organized as follows: In section 2, the concept of DWT for filter bank identification is discussed briefly. Algorithm used to determine the appropriate wavelet based on the energy concentrated in wavelet coefficients is discussed in section 3. In section 4, experimental results based on the algorithm in section 3 are discussed. Finally, concluding remarks are given in section 5.

2. DISCRETE WAVELET TRANSFORM
It is well known that DWT of a signal is equivalent to passing it through an analysis filter bank followed by decimation operation [11]. This analysis filter bank consists of a low pass and a high pass filter at each decomposition stage. When a signal is passed through these filters, it splits into two bands. The low pass filter which corresponds to averaging operation, extracts the coarse information of the signal. The high pass filter, which corresponds to a differencing operation, extracts the detail information of the signal. The output of the above filters is further decimated by two [12]. The DWT is computed by successive low pass and high pass filtering of the discrete time domain signals as shown in the figure 1. In this figure the signal, denoted as s(j) is in the form of sequence, where j is the scale of resolution, passed through a low pass filter and the high pass filter with unit sample response h(-n) and g(-n) respectively. At each level the high pass filter produces the detail information, represented as wavelet function and low pass filter produces the approximate information, represented as scaling function. At each decomposition level, the filters h(-n) and g(-n) produce signal spanning only half the frequency band. This doubles the frequency resolution as the uncertainty in frequency is reduced to half.

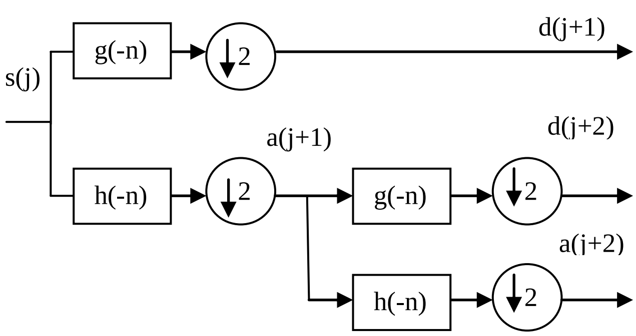


Figure1: Two stage DWT as Analysis Filter bank

As per the well known Nyquist's sampling theorem [13], a signal band limited to $f/2$ Hz can be recovered back uniquely from its samples taken at f sample/sec. So the output of the two filters is sampled at a frequency f, discarding half the samples without any loss of information. Thus this decimation by two halves the time resolution as the entire signal is now represented by half the samples only, while the low pass filtering removes the half of the frequencies, thus halves the resolution, then decimation by two doubles the scale. The filtering and decimation process continues up to the desired level. The maximum level of decomposition is dependent upon the length of the signal and the attributes of the analysis done. The reconstruction of the original signal can be done with the help of the figure 2.

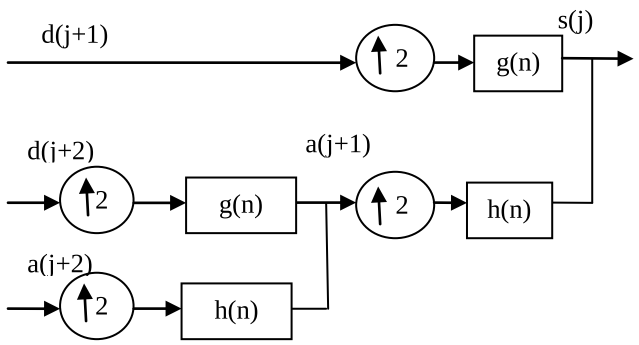


Figure2: Two stage DWT as synthesis Filter bank

3. WAVELET MATCHING ALGORITHM

By using the concept of DWT, an algorithm has been developed for the synthesis of the musical sound signals, which are represented by a specific wavelet without any loss of information. The following steps can be used for the analysis:

1) Choose a specific wavelet.
2) Apply the wavelet on the signal of study and calculate the wavelet (detail) coefficients and the approximate coefficients at level 1.
3) Evaluate the energy concentrated in the wavelet coefficients and approximation coefficients.
4) Calculate the percentage of energy concentrated in the wavelet coefficients and approximation coefficients at level 1 to that of energy of the original signal.

The energy mentioned above is based on Parseval's Theorem:

"The energy that a time domain function contains is equal to the sum of all energy concentrated in the different resolution levels of the corresponding wavelet transformed signals". This is mathematically expressed as [14]:

$$\sum_{n=1}^{N} |f(n)|^2 = \sum_{n=1}^{N} |a_j(n)|^2 + \sum_{j=1}^{J}\sum_{n=1}^{N} |d_j(n)|^2 \qquad (1)$$

Where:

$f(n)$: Signal of study in time domain

N : Total No. of samples of the signal

$\sum_{n=1}^{N} |f(n)|^2$: Energy concentrated in signal

$\sum_{n=1}^{N} |a_j(n)|^2$: Energy concentrated in the jth level of the approximation coefficients

$\sum_{j=1}^{J}\sum_{n=1}^{N} |d_j(n)|^2$: Total energy concentrated in the wavelet coefficints of the signal from level 1 to J

$$eng\_per\_wav = \frac{eng\_wav}{eng\_sig} * 100 \qquad (2)$$

Where

eng_per_wav: Percentage of energy concentrated in the wavelet coefficients only.

eng _wav : Energy concentrated in wavelet coefficients till that level.

eng _ sig : Energy concentrated in the original signal of study.

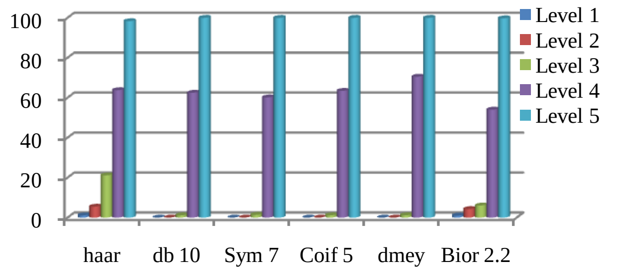
5) Steps 1, 2, 3 and 4 are repeated for different wavelets and at different level of resolution.
6) Choose the wavelet which has maximum energy concentrated in the wavelet coefficients up to this level and the relative difference of energy to the previous level is negligible, which is the maximum level of decomposition.
7) Reconstruct the signal by taking the wavelet coefficients from level 1 to the maximum level of decomposition.
8) Compare the hearing perception of the original signal and the reconstructed signal obtained by taking wavelet coefficients from level 1 to the maximum level.

4. EXPERIMENTAL RESULTS

The algorithm developed in the previous section is applied to the sound samples of the various musical instruments. For experimental purpose the sounds of three musical instruments flute, shehnai and sitar are taken. Results are shown in the figure 3, 4 and 5.

Wavelet	Level 1	Level 2	Level 3	Level 4	Level 5
haar	1.13	5.48	21.13	63.78	98.31
db 10	0.01	0.03	1.16	62.48	**99.97**
Sym 7	0.01	0.04	1.39	60.15	99.75
Coif 5	0.01	0.02	1.16	63.43	99.76
dmey	0.01	0.01	1.19	70.50	99.70
Bior 2.2	1.01	4.29	6.05	54.03	99.76

(a)

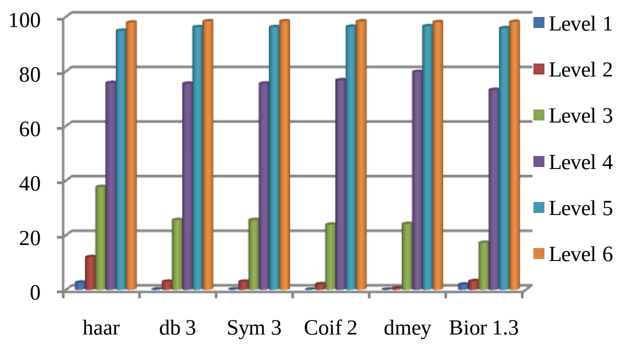


(b)

Figure 3(a, b): Percentage of energy present in wavelet coefficients only at different levels in flute sound

Wavelet	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
haar	2.60	11.94	37.63	75.72	94.85	97.87
db 3	0.10	2.94	25.53	75.48	96.19	**98.32**
Sym 3	0.10	2.94	25.53	75.48	96.19	98.22
Coif 2	0.03	2.04	23.85	76.73	96.35	98.20
dmey	0.01	0.52	24.12	79.76	96.52	98.02
Bior 1.3	1.9	3.15	17.19	73.20	95.81	98.13

(a)

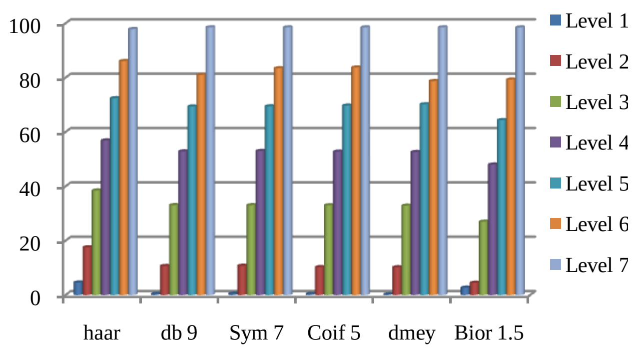


(b)

Figure 4(a, b): Percentage of energy present in wavelet coefficients only at different levels in shehnai sound

Wavelet	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7
haar	4.60	17.55	38.46	56.90	72.44	86.11	97.88
db 9	0.41	10.67	33.07	52.87	69.41	81.02	98.41
Sym 7	0.49	10.78	33.07	53.02	69.45	83.45	98.40
Coif 5	0.37	10.25	33.03	52.78	69.72	83.74	98.41
dmey	0.23	10.22	32.89	52.62	70.20	78.73	**98.52**
Bior 1.5	2.70	4.50	26.99	48.02	64.37	79.31	98.42

(a)



(b)

Figure 5(a, b): Percentage of energy present in wavelet coefficients only at different levels in sitar sound

It is observed from the figure3 that flute sound can be reconstructed with the help of the wavelet db10 taking wavelet coefficients only till the level 5, due to the presence of maximum energy concentrated in the wavelet coefficients with this wavelet. Similarly figure4 shows that shehnai sound is decomposed till the level 6 with maximum energy concentrated in the wavelet coefficients of wavelet db3. But from figure5 it is observed that sitar sound is decomposed till level 7 and the maximum energy is concentrated in the wavelet dmey. The original waveforms along with the synthesized waveform of flute shehnai and sitar are shown in figures 6-11.

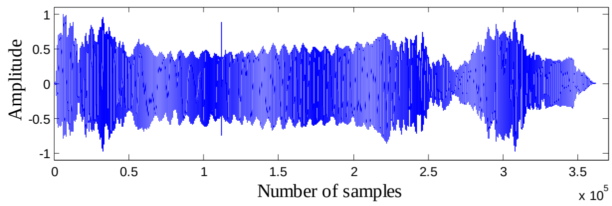


Figure 6:Original flute waveform

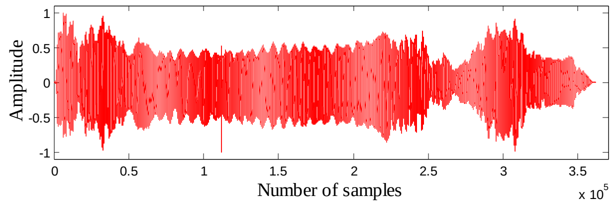


Figure 7: Flute waveform after synthesis with db10 wavelet at level 5 taking wavelet coefficients only

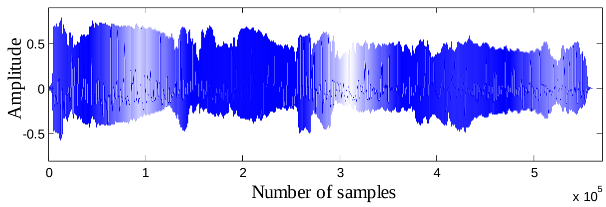


Figure8: Original shehnai waveform

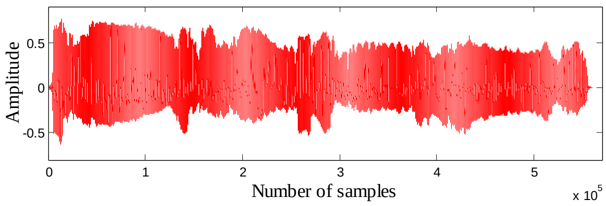


Figure 9: Shehnai waveform after synthesis with db3 wavelet at level 6 taking wavelet coefficients only

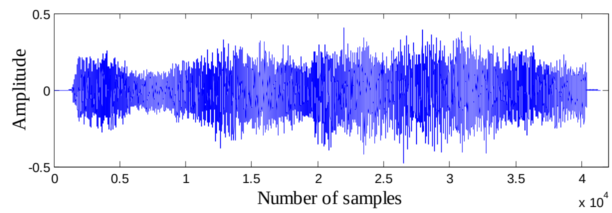


Figure 10: Original sitar waveform

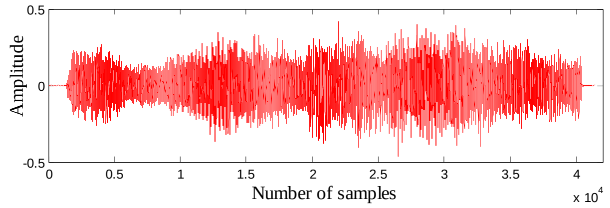


Figure11: Sitar waveform after synthesis with dmey wavelet at level 7 taking wavelet coefficients only

5. CONCLUSION

In this paper discrete wavelet transform based analysis of the sounds produced by some Indian musical instruments is presented. Here an algorithm is proposed which analyses the musical sound signals with the help of standard wavelets and the amount of energy concentrated in wavelet coefficients at each level of decomposition is calculated. It is shown in figures 3, 4 and 5 that the flute sound is having maximum energy concentrated in wavelet coefficients when analyzed with db10 till the fifth level, shehnai sound with wavelet db3 till sixth level and sitar sound with wavelet dmey till seventh level. Hence it is concluded that each sound signal is better represented with the different wavelet and at different level of decomposition. The wavelet db10 is most suited for flute sound, db3 for shehnai and dmey for the sitar waveform. This analysis also shows that these waveforms can be reconstructed with the help of the respective wavelets and at the respective level of decomposition with wavelet coefficients only. So lesser numbers of samples are required for the reconstruction of the sound signal with a little amount of error, hence a good amount of compression is achieved. This analysis can be extended to our future work of identification of signature wavelet of each musical instrument sound, which may be useful in the synthesis of sound signals.

6. REFERENCES

[1] Mallat, S.G. 2009. A Wavelet Tour of Signal Processing. The Sparse Way, 3rd Edition, Academic Press.

[2] Tobias, J. S. 1970. Foundations of Modern Auditory Theory, Vol. 1, Academic Press, New York.

[3] Bultheel, A. 2003. Wavelets with applications in signal and Image processing. Text pages 1-181.

[4] Panos E. Kudumakis and Mark B. Sandler. 1993. Synthesis Of Audio Signals using the Wavelet Transform. IEEE Colloquium on Audio DSP – Circuits and Systems (Digest No. 1993/219), pp. 1- 4, vol. 4.

[5] Vetterli, M., and Kovacevic, J. 2007. Wavelets and Sub band Coding. Prentice Hall PTR, Englewood Cliffs, New Jersey.

[6] Ramchandran, K., Vetterli, M. and Herley, C. 1996. Wavelets, Sub band Coding and Best Bases, Proceedings of IEEE, Vol. 84, No. 4, pp. 541-560.

[7] Vetterli, M., and Kovacevic, J. 1995. Wavelet and Sub band Coding. *Prentice Hall.*

[8] Vetterli, M. and Harley, C. 1992. Wavelets and Filter Banks: Theory and Design, IEEE Transaction on Signal Processing. Vol.40, No. 9.

[9] Galand, C. R. and Nussbaumber, H. J. 1984. New Quadrature Mirror Filter Structures, IEEE Transaction on Acoustics Speech and Signal Processing, Vol. 32, No. 3, pp. 531-552.

[10] Kumar, A., Singh, G. K, and Anand, R. S. 2008. Near Perfect Reconstruction Quadrature Mirror Filter, World Academy of Science, Engineering and Technology, Vol.37, pp. 204-207.

[11] K.P Soman, K.I Ramchandran, N.G Resmi, 2010. Insight into Wavelets from theory to practice, PHI Learning Private Limited New Delhi.

[12] Szu, H., Sheng, Y. and Chen, J. 1992. Wavelet Transform as a Bank of Matched Filters, Appl. Optic Vol. 31, pp. 3267-3277.

[13] Mallat, S.G. 1989. A Theory for Multi-resolution Signal Decomposition: The Wavelet Representation, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 11, pp. 674 – 693.

[14] Burrus, C.S., Gopinath, A.R, Wickerhauser, M.V. 1996. Wavelets and Time frequency Analysis, Proceedings of IEEE, Vol. 84, No.4, pp.523-540.