Extended Wavelet Packet Decomposition and Best Basis Search Algorithm, Evaluation of entropic gain for audio signals

Gilles Gonon[a], Silvio Montresor[b] and Marc Baudry[a]

[a]*Laboratoire d'Informatique, Université du Maine, Av. F. Laennec, 72000 Le Mans, France*
[b]*Laboratoire d'Acoustique,UMR CNRS 6613, Université du Maine, 72000 Le Mans, France*

The Wavelet Packet Decomposition (WPD) is a efficient tool in audio coding because of its frequency adaptation skills through the best basis search algorithm. The entropic minimization of this algorithm is bounded by the dyadic structure of the decomposition. In order to decrease the entropy of the best basis, a low cost extended tree in the WPD is used. It is still compatible with the classical WPD and insures perfect reconstruction. The entropic test is updated to take into account the new basis. We present an example of the resulting best basis on a simulation signal and evaluate the average entropic gain obtained on various audio signals.

INTRODUCTION

The Wavelet Packet Decomposition (WPD) is a powerful tool in signal processing because it offers the opportunity to give an adaptive representation of the signal trough the Best Basis Search Algorithm (BBSA). Such an algorithm finds a partition of the frequency axis adapted to the signal in the sense that single tones are isolated in narrow bands, and that grouped wide bands are low energy noisylike or empty. Thus, the BBSA offers to expand the signal on an optimal wavelet basis according to a cost function, such as entropy or energy.

A problem subsists due to the dyadic structure of the decomposition. The optimality of the basis remains bounded by artificial segmentation induced by the BBSA because the entropic test does not operate on adjacent bands not coming from the same father in the decomposition (Figure 1).

In [1], we proposed an extension of the WPD and BBSA to take into account a new packet in the best basis search. The main steps of this extension are recalled and the results in terms of entropic diminution are presented.

The proposed decomposition is noted Sigma-WPD (SWPD) and the new fathers constructed are indexed $(d, 2b + \frac{1}{2})$, where d is the depth of the decomposition and b the packet number, $b \in [0, 2^d - 1]$.

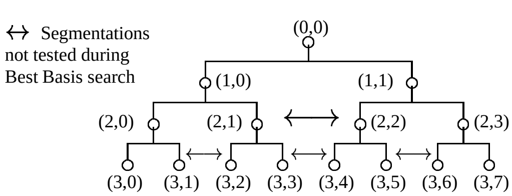

FIGURE 1. Artificial segmentation induced by the dyadic structure of the WPD

WPD AND BBSA EXTENSION

The extension proposed in [1] constists in constructing the wavelet packet corresponding to the father of nodes (2,1) and (2,2), and then to include it in the entropic test of the BBSA, also called "Split and Merge" algorithm.

The derivation of the new packet is based on the equivalence between two wavelet decompositions that lead to depth 2 packets in different ways (Figure 2). One of the packets in the alternative decomposition is the father of packets (2,1) and (2,2), so that it can be used to perform the entropic test among these packets.

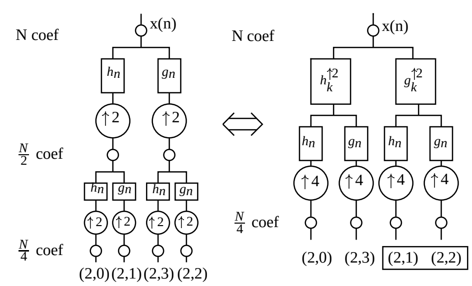

FIGURE 2. Left : classical WPD filterbank, Right : alternative decomposition leading to equivalent depth 2 WPD filters.

Letting h_n be the coefficients of the lowpass QMF $h(n)$ and $g_n = (-1)^n h_{-n+1}$ be the highpass QMF, we denote $h_k^{\uparrow 2}$ and $g_k^{\uparrow 2}$ their Impulse Responses upsampled by a factor 2. The filters equivalent to father of nodes (2,1) and (2,2) is $h_{(1,\frac{1}{2})}(n) = g_k^{\uparrow 2}$. The wavelet coefficients for packet $(1, \frac{1}{2})$ are obtained with a $\frac{\pi}{4}$ Single Side Band modulation, as described in [4].

The entropic test now deals with tree fathers for four children. The band $(1, \frac{1}{2})$ is kept as final only if the entropy of the basis $[(2,0) \cup (1, \frac{1}{2}) \cup (2,2)]$ is lower than any other possible dyadic basis.

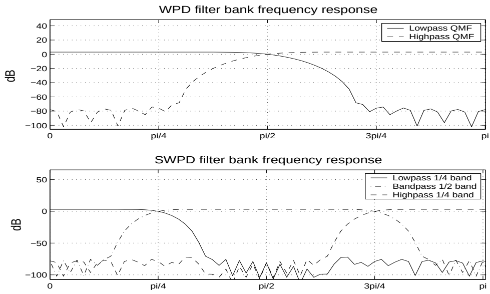

FIGURE 3. Frequency responses of Vaidyanathan filter banks. Top : WPD [(1,0),(1,1)], bottom : SWPD [(2,0),$(1,\frac{1}{2})$,(2,1)].

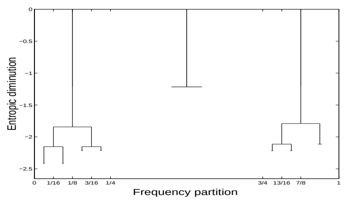

FIGURE 4. S-dyadic best tree that includes node $(1,\frac{1}{2})$.

RESULTS

We generate a simple simulation signal where the s-dyadic band $(1,\frac{1}{2})$ should be kept in the best basis. The signal is composed of two sinusoidal components, one in subband $(2,0)$ and the other in $(2,3)$. They both have different temporal supports. To insure that band $(1,\frac{1}{2})$ has minimum entropy, it is left empty. The filters used to perform the decomposition and the best basis search are Vaidyanathan's QMF because of their fair frequencial selectivity. Frequency responses of QMF bank and S-QMF bank are shown in figure 3 and figure 4 shows s-dyadic the depth 4 best basis tree. As expected the band $(1,\frac{1}{2})$ is a final node of the SWPD best basis search and the entropic gain of SWPD versus WPD is about 2%. In order to test the general interest of the method and particularly to apply it to audio coding [3], we tested it on different types of audio signals : metronome, harpsichord, castanets and speech. For the two decomposition, the entropic gain was evaluated for two types of wavelets : one with a good temporal resolution, the Daubechies 4 wavelet (D4) and one with a fair time-frequency resolution compromise, the Vaidyanathan wavelet (V24) (Table 1).

First, if we compare the results for both wavelets, we can see that the V24 wavelets give better result, due to the nature of audio signal. This just confirms that a fair frequencial selectivity is needed to obtain a better frequencial segmentation of signals. In the case we compare both decompositions, each time that s-dyadic nodes are kept as final, the entropic gain is measurable. The average entropic gain lies between 5% and 10%. While no s-dyadic node is kept, the classical best basis gain is not affected. Better gains have been obtained with percussive signals because the signals tested were not temporally segmented before the analysis and so the attacks overcomes tonal components in an energetic or entropic cost sense.

Table 1. Entropic gain (in %) comparison for both methods WPD and SWPD BBSA. D4 and V24 respectively stand for Daubechies 4 and Vaidyanathan wavelets.

Signal Entropy	Depth	BB Gain		S-BB Gain	
		D4	V24	D4	V24
Metronome	5	25.1	25.9	50.1	50.4
9.07	9	36.5	42.3	56.2	58.6
Harpsichord	5	23.9	24.4	23.9	24.4
9.65	9	36.1	42.3	47.6	42.6
Speech	5	16.1	17.1	16.1	17.2
7.39	9	20.2	23.8	23.4	24.0
Castanets	5	9.0	11.1	9.0	11.4
8.35	9	13.5	15.4	21.5	22.9

CONCLUSION

The proposed decomposition SWPD extends the dyadic bases library of the WPD and allows perfect reconstruction. The 7% average entropic gain obtained on various audio signals proves that the extended library of bases leads to a better adaptation to the signal, and so a better knowledge of its frequencial content.

It would be interesting to use this decomposition in a frequency adaptive audio coder [3, 1] and evaluate the coding gain rather than entropy diminution.

REFERENCES

1. G. Gonon, S. Montresor and M. Baudry, *Improved entropic gain for speech signals analysis/synthesis based on an adaptive time-frequency segmentation scheme*, Eurospeech 2001. **45**, 815-885 (1982).
2. R.R. Coifman and M.V. Wickerhauser,*Entropy-based algorithms for best basis selection*, IEEE Trans. on Information Theory, **38**, 713-718 (March 1992).
3. M. Erne, G.S. Moschytz and C. Faller, *Best Wavelet-Packet Bases for Audio Coding using Perceptual and Rate-distortion Criteria*, ICASSP 99, Phoenix (May 1999).
4. R. Crochiere and L. Rabiner, *Multirate Digital Signal Processing*, Prentice-Hall Signal processing series(1983).