DOUBLE TREE DECOMPOSITION OF LUNG SOUNDS *

E. Ademovic♣♠, J.-C. Pesquet♣ , G. Charbonneau♠
♣ Laboratoire des Signaux et Systèmes, CNRS/Univ. Paris-Sud and GdR-PRC ISIS,
ESE, Plateau de Moulon, 91192 Gif sur Yvette, France.
♠ Institut d'Électronique Fondamentale
Université de Paris-Sud, 91405 Orsay, France.
e-mail: ademovic@lss.supelec.fr

ABSTRACT

The analysis of respiratory sounds highlights the limits of commonly used techniques as a huge variety of sounds can be observed (stationary or nonstationary and of different durations) which can have themselves a great variability. New approaches have been developed in order to associate the acoustic phenomena to the respiratory flow and volume. We present here another approach only based on the wavelet packet decomposition to segment respiratory sounds.

1 Introduction

Respiratory sounds are of interest since they provide information on the respiratory system (either the upper airways or the lungs). They are therefore reliable for the physician and may help to diagnose or survey many respiratory diseases. Such pathologies induce abnormal sounds (also called adventitious sounds) caused by an obstruction in the airway path. In Lung Sound Analysis (LSA) two major adventitious sounds are considered: the *wheezes* and the *crackles.* A wheeze is defined as a continuous "musical" sound, and a crackle is a discontinuous "explosive" sound. In [1], a FFT is used to detect wheezes phenomena. For the crackling sound analysis, a time domain detection based on the first derivative is proposed in [2]. In [3], the signal is analysed with an autoregressive model, and the AR parameters are used by a classification algorithm to identify the sounds.
As adventitious sounds are generally superimposed on the respiratory sound, a simple and efficient way to analyse the signal is to perform a time segmentation. However, as the sounds exhibit a great dependence on the flow and the volume, their analysis is difficult. Moreover, because of the nonstationarity of these signals, it appears judicious to make use of a wavelet based analysis technique. We have already proposed to analyse the respiratory sounds with a time-frequency method known as the Adaptive Local Trigonometric Decomposition (ALTD or also named Malvar's wavelet transform)[6]. In this work, we show the application of a new segmentation method: the Double Tree Decomposition (DTD) [4]. This method performs a time-scale analysis which corresponds to an adaptive segmentation both in time and scale domains.
A brief description of the DTD algorithm is given in section 2 and, its application to LSA is described in section 3. Results are discussed in section 4 and a comparison is made with the ALTD.

2 Adaptive Segmentation Algorithm

The DTD is a generalization of the Wavelet Packet Decomposition (WPD). It consists of the construction of a binary tree structure describing the optimal time segmentation of a signal. Each node of this temporal tree corresponds to a segment of the signal and, is itself associated with a binary tree corresponding to its optimal WPD (see Fig. 1).

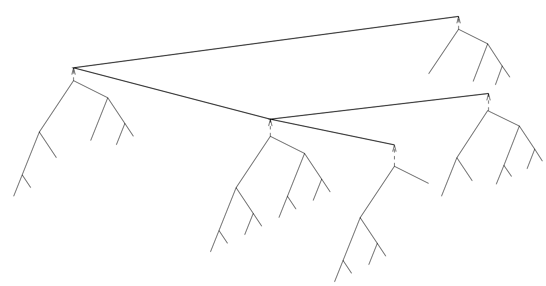

Figure 1: Double Tree segmentation.

Alike the WPD, the DTD relies on the application of the best basis search algorithm [5]. The best representation depends on the minimization of a given criterion, *e.g.* the unnormalized entropy:

$$\mathcal{H}(x) = -\sum_{n=0}^{N-1} |x_n|^2 \log |x_n|^2,$$

where N denotes the data length The elementary procedure compares the costs of representation of a "father" node C_f to its two "sons" $\{C_{sl}, C_{sr}\}$, and is iterated for

* This work was partly supported by the CORSA/BIOMED I project of the EC.

each node of each tree following a bottom-up strategy. The local minimization of the criterion leads to a solution which is also globally optimal due to the additivity of the criterion. The application of this adaptive segmentation algorithm requires some settings to be carefully chosen. The choices which must be made are the following:

- length and selectivity of the analyzing filters
- depths of analysis in time and scale domains
- cost functions.

3 Experiments

The algorithm has been tested with samples of lung sound records sampled at 5512 Hz pre-preprocessed with a high-pass filter (4^{th} order Butterworth, with cut-off frequency of 100 Hz) to remove heart components.
We use two techniques to represent data. The first one depicts the final segmentation in time domain (Fig. 2) with vertical lines superimposed on the sequence. The second one operates in the time-scale plane with a representation called *energy map* (Fig. 6). This figure requires to reorganize the wavelet packet coefficients according to Paley's order so that they correspond to an increasing frequency range. This plot shows the tiling provided by the analysis where each tile is associated with the energy of the corresponding subband. These two representations allow to appreciate the accuracy and the quality of the segmentation. Each analyzed sequence was compared to its spectrogram computed from a Short Time Fourier Transform (256 point Hanning window) (see Fig. 4).

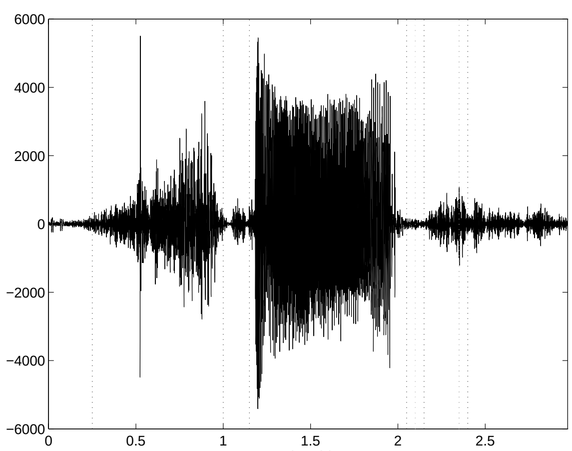

Figure 2: Example of a lung sound sequence with an arbitrary segmentation.

Analyzing filter

Some of the wavelet packet basis functions appeared to be multimodal in the frequency domain if the analyzing QMFs are not selective enough. This may create some artefacts in the tiling of the time-scale plane. To alleviate this problem, the filters have been optimized according to Rioul's method [7]. These filters present a good selectivity for our purpose and are more efficient than Daubechies' ones. Using filters without side lobes tremendously reduces these phenomena of delocation at the expense of a loss in time segmentation. We run the decompositions with filters of length 8 and 16. The flatness factor controls the selectivity of the frequency response as shown in Fig. 3. For fixed length and flatness, many filters can be synthetized which are similar in frequency response but, different in the phase domain.

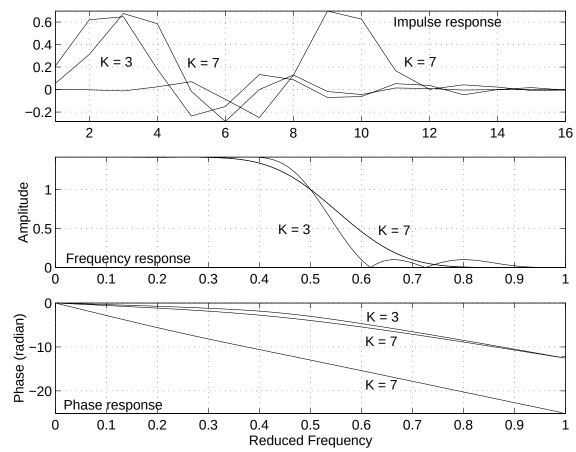

Figure 3: Rioul's filters characteristics for various flatness factors ($K = 3$ and $K = 7$).

Usually, the phase response of the filter is not considered to be of great influence in WPD. However, as the DTD algorithm segments the signal in the time domain according to the wavelet packet analysis, this feature becomes now more relevant. For instance, one can observe the differences in the energy map between two anlayses performed with the same parameters and the same filter selectivity but different phases (Fig. 5). A valuable partition is obtained when the filter is ruled by a phase which is almost linear (lower plot). Hence, the distortion of the group delays induced by the filtering stage is highlightened: the filter could not be only chosen according to selectivity perfomances but phase should also be handle with caution when using with DTD.
Practice has shown that filters of length 8 provide a picture of the time-scale plane acute enough and leading to a low complexity analysis.

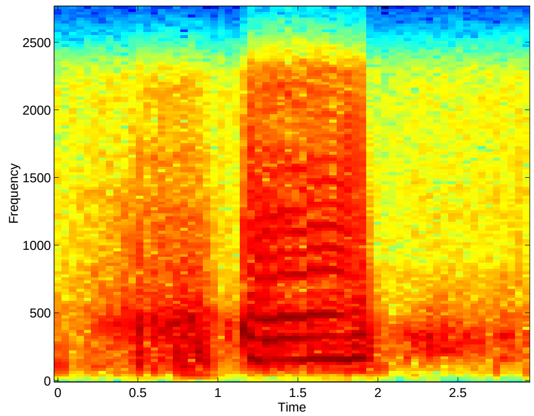

Figure 4: Spectrogram of the respiratory sound sequence in Fig 2.

Levels of analyses

The adventitious sounds we want to detect are of various duration. The shortest events are the fine crackles (> 30 ms) and the longest ones are the wheezes ($>$ 250ms). Thus, the segmentation algorithm must take into account these specifications to work properly. We impose the initial segmentation to be of 256 points (*i.e.* 46.4 ms) at the finest time resolution to allow the detection of the transient signals. The coarse time resolution was fixed to 4096 points (*i.e.* 743 ms). It is useless to proceed larger sequences as the algorithm cannot merge longer lung sound segments. Indeed, even if wheezes can be of great duration, they are not stationary on the whole interval. So, the event will be described with a sequence of various sized segments and all of them can be considered stationary. Moreover, limitating the size of analyzed sequences reduces the computational cost of the DTD algorithm. Therefore, the number of levels in time analysis is 5.
Because of decimation effects, the minimum resolution level that can be reached in the WPD is lower for long than for short segments. For this reason, localized events are sometimes not properly segmented. It appears therefore more meaningful to fix the same lowest resolution level of analysis for each segment. In the scale domain, we decide to impose 7 levels of decomposition. At the bottom of the time tree it provides a resolution of 86 Hz per subband (32 subbands), and at the top, the resolution is 43 Hz (64 subbands).

Cost functions

The cost function is another important item of the algorithm. It evaluates the representation of the decomposition. Several additive criteria have been used and compared to define the best representation. One of this

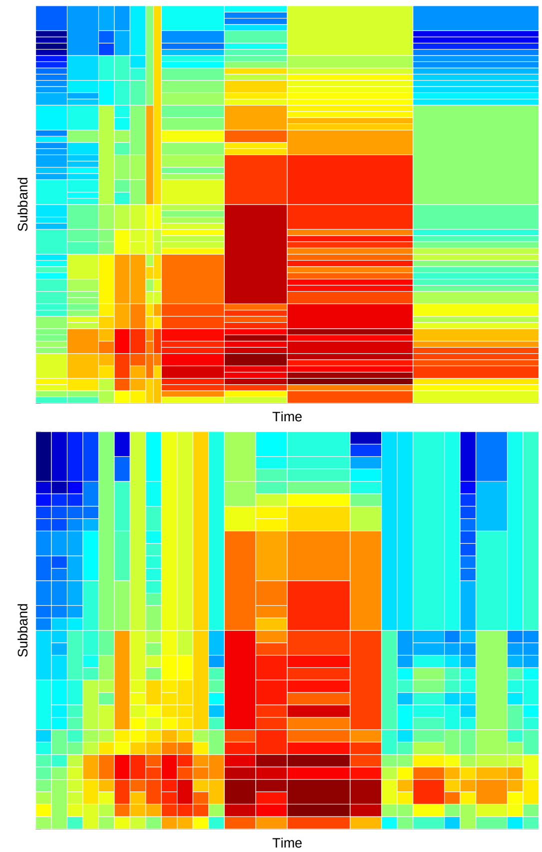

Figure 5: Influence of the phase response on the segmentation.

criterion is derived from the expression of the L^p norm:

$$\mathcal{L}^p(x) = \sum_i |x_i|^p .$$

It is also a measure of energy concentration like the entropy but, even if the final segmentation can be comparable or better, the tuning of the parameter p is sensitive.

4 Discussion

The DTD algorithm provides an optimal segmentation of a sequence. Running this method on signal is sometimes difficult because of the interdependencies of several parameters. Filters can be chosen to be selective to obtain satisfactory time-scale maps without forgetting that, their phase response can influence the result in the time domain. Levels of analysis depend on the features of the phenomena we want to detect. When DTD is run on lung signals, the resulting partition provides a precise analysis. The different events are well described in time and scale domain. Figure 6 presents the energy map computed with the promoted settings (filter of length 8, flatness of 3, sequence size 4096 pts, 5 time levels, 7 scale levels, unnormalized entropy cost

function). An example of a similar segmentation algorithm (ALTD) computed on the same lung sound signal is depicted in Fig. 7. The ALTD is a real time-frequency analysis whose mapping is very close to a spectrogram even if it is also a wavelet based technique. The two partitions are very close in the time domain.
Some improvements could be achieved concerning the cost function and especially the selection of two different criteria for the time-tree pruning and the WPD algorithm.

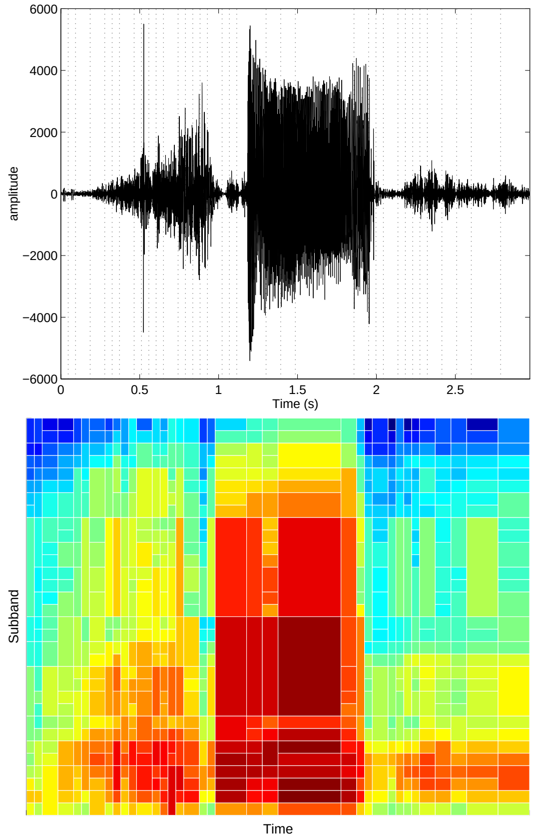

Figure 6: DTD time segmentation and energy map of a respiratory sound sequence.

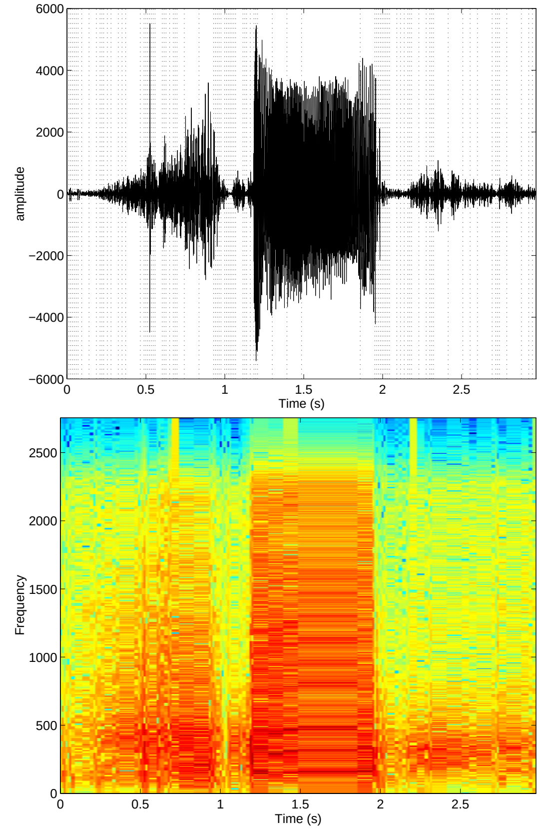

Figure 7: Time-frequency segmentation and mapping of an ALTD analysis.

References

[1] D. Li, S. Haltsonen, K. Kallio, P. Helistö, T. Katila, "*Filtered spectrum in the detection of the wheezing sound*", First CORSA WPIII Symposium on signal Processing in Lung Sound Analysis, report TKK-F-C170, Helsinki University of Technology, June 1995. ISSN 0358-0741.

[2] L. Vanuccini, M. Rossi, G. Pasquali, R. Palatresi, "*A new method to detect cracles in respiratory sounds*", First CORSA WPIII Symposium on signal Processing in Lung Sound Analysis, report TKK-F-C170, Helsinki University of Technology, June 1995. ISSN 0358-0741.

[3] T. Engin, E,Ç. Güler, B, Sankur and Y.P. Kahya, "*Comparison of AR-based classifiers for respiratory sounds*", in Proc. EUSIPCO-92 (Brussels, Belgium), pp. 1745-1748, 1992.

[4] C. Herley, J. Kovaćevič, K. Ramchandran and M. Vetterli, "*Tilings of the Time-Frequency plane: Construction of arbitrary orthogonal bases and fast tiling algorithms*", IEEE Trans. SP, pp. 3341-3359, vol. 41, Dec. 1993.

[5] R. Coifman and M. V. Wickerhauser "*Entropy-based Algorithms for Best Basis Selection*", IEEE Trans. Informat. Theory, pp. 713-718, vol. 38, Mar. 1992.

[6] E. Ademovic, G. Charbonneau and J.-C. Pesquet, "*Segmentation of infant respiratory sounds with Malvar's wavelets*", IEEE IC EMBS Workshop on Wavelets in Medicine and Biology, pp. 20a-21a, Baltimore (USA), november 1-2, 1994.

[7] O. Rioul and P. Duhamel, "*A remez exchange algorithm for orthonormal wavelets*", IEEE Trans. Circuits and Systems II, pp. 550-560, vol. 41, Aug. 1994.