Wavelet-based bowel sounds denoising, segmentation and characterization

R. Ranta[1], C. Heinrich[1], Valérie Louis-Dorr[1], D. Wolf[1], and F. Guillemin[2]
[1] Centre de Recherche en Automatique de Nancy – CRAN, 2 av. de la Forêt de Haye, 54500 Vandoeuvre, France
[2] Centre Alexis Vautrin – CAV, av. de Bourgogne – Brabois, 54500 Vandoeuvre, France

Abstract— **The general framework of this communication is phonoenterography. The ultimate goal is the development of a clinical diagnostic tool based on abdominal sound monitoring. Bowel sounds are recorded using several microphones. Unsupervised data processing should lead to diagnosis assessment.**

We address here the early stages of data processing, *i.e.*, denoising, segmentation and characterization of detected events. The denoising algorithm is based on former work by Coifman and Wickerhauser [1] and Hadjileontiadis *et al.* [2], [3]. Their wavelet-based algorithm is revisited, allowing to significantly reduce the computational burden. Sound segmentation and event characterization are based on the wavelet representation of the phonoenterogram. Real data processing examples are given.

Keywords— **Bowel sounds, auscultation, phonoenterogram, wavelet transform, noise cancelling, sound segmentation, feature extraction.**

I. Introduction

Bowel sounds have been attracting attention for a long time, since they carry diagnostic information about the abdominal tract. Initial studies were dedicated to the physiological interpretation of bowel sounds (BS) [4], [5], [6], [7]. The field of phonoenterography is now regaining attention, thanks to modern signal processing techniques and computerized data processing enabled by advances in computer capacities [8], [9], [10], [3]. The goal is the development of an unsupervised research and clinical diagnostic tool.

Main uses of this tool for research purposes are the study of the effect of various drugs on the digestive system and the elaboration of a map of abdominal activity. Elaboration of such a map is expected to improve knowledge of the abdominal tract functioning. The use of this tool as a clinical diagnostic mean may be oriented towards post-operative diagnosis, intestinal obstruction detection, acute appendicitis, irritable bowel syndrome and more generally motility disorders diagnosis [11].

Such a system is supposed to complement classical auscultation, which takes precious time, is repetitive and subjective. It would also allow monitoring over longer periods of time and quicker alarm activation. Furthermore, it is a non invasive technique. It doesn't interfere with the physiological phenomenon under investigation.

Nevertheless, two problems will have to be settled:

- recorded data will have to be interpreted accurately. This point is far from being settled since, as many papers point it, abdominal auscultation seems to be lacking of support in scientific fact [9], [10], [3];
- the ambient noise contamination of BS will have to be addressed, at least for a clinical use of the system.

To the authors' best knowledge, those problems are still open. Moreover, the difficulty of interpreting data is worsened by the variability of sounds from one patient to another and by the fluctuation of sounds with time.

According to previous studies [9], [10], the following characteristics of individual sounds and complete signals will have to be scrutinized: frequential sound content, sound intensity, sound duration, silence duration, sound localization. All those features are embedded in the wavelet representation, sound localization excepted. Localization is allowed by the use of several microphones.

In particular, Hadjileontiadis *et al.*'s clinical results are used and their wavelet-based algorithm is revisited, allowing significant reduction of the computational burden. We extend the data processing step so as to include segmentation and feature extraction. Features are extracted for all detected events (an event may also be named a sound in this communication) which should lead to event classification in further work: bowel sounds usually described subjectively (as "staccato pops", "gurgling sounds" or "clicks" for example) should be given an objective and quantitative correspondance in the space of characteristics.

The communication is organized as follows. First, we focus on Hadjileontiadis *et al.*'s denoising algorithm. Then we deal with segmentation and parameter extraction. Finally, real data processing results are given.

II. Wavelet denoising

In this section, we briefly remind Hadjileontiadis *et al.*'s denoising algorithm [2], [3]. This algorithm will be reinterpreted as a fixed-point algorithm.

The mentioned algorithm is derived from the classical denoising method of Coifman and Wickerhauser [1]. The basic principle of the method is an "information viewpoint" definition of noise: a signal which is not well correlated with a waveform basis. The algorithm consists in thresholding the coefficient vector obtained by a wavelet transform and reconstructing a "denoised" signal from the resulting vector. The threshold may be different for each scale (for a non-white noise) and the thresholding can be "soft" or "hard" [1]. Furthermore, it can be done iteratively, possibly decomposing on different basis, until a certain stop criterion is validated.

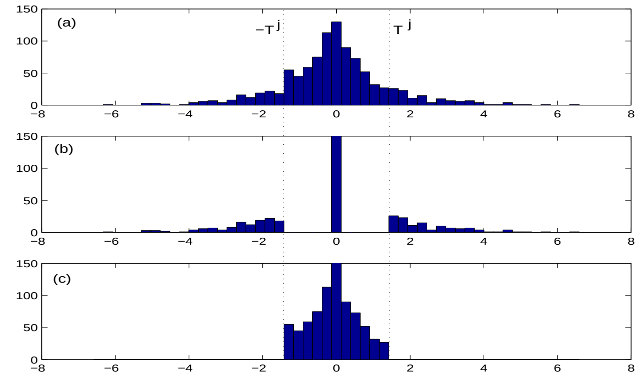

Fig. 1. Wavelet coefficients histograms at scale j. (a) original signal $w^{j,p}$; (b) denoised signal $w_c^{j,p}$; (c) residual noise $w_r^{j,p}$. For this example the wavelet coefficients of the supposed original signal have been drawn according to an exponential law. The y axis is not entirely represented on (b) and (c).

Basically, Hadjileontiadis's algorithm works as follows. The measured signal is decomposed on a wavelet basis, yielding a wavelet coefficient vector $\mathbf{w} = \{w^{j,p}\}$, where j is the scale and p is the index. Estimated denoised signal and noise wavelet coefficients will be computed from $\{w^{j,p}\}$, yielding $\{w_c^{j,p}\}$ and $\{w_r^{j,p}\}$. For "large" $w^{j,p}$, we have: $w_r^{j,p} = 0$, $w_c^{j,p} = w^{j,p}$ whereas for "small" $w^{j,p}$ we have: $w_r^{j,p} = w^{j,p}$, $w_c^{j,p} = 0$. The "large" or "small" decision results from comparison to a threshold T^j (see Fig. 1).

In Hadjileontiadis's approach, this threshold is calculated by an iterative (k) decomposition-reconstruction algorithm:

$$T_k^j = F_a \cdot \sigma_k^j, \qquad (1)$$

where $F_a = 3$ is a multiplicative empirical constant (justified by medical expertise) and σ_k^j is the standard deviation of the estimated noise coefficients $\{w_{r,k}^{j,p}\}$ at iteration k (see Fig. 1(c)). The wavelet coefficients $\{w_{r,k}^{j,p}\}$ are obtained by a thresholding of $\{w_{r,k-1}^{j,p}\}$ with a threshold T_{k-1}^j. The threshold T_k^j may be seen as a function of T_{k-1}^j: T_{k-1}^j being given, we obtain $\mathbf{w_{r,k}^j} = \{w_{r,k}^{j,p}\}$ and its standard deviation σ_k^j, which permits the computation of T_k^j (eq. 1). This allows to interpret Hadjileontiadis's approach as an iteration:

$$T_k^j = f(T_{k-1}^j) = 3\sigma_{\mathbf{w_{r,k}^j}}. \qquad (2)$$

We have to deal with the initialization and the convergence of these iterations. The $\mathbf{w_{r,k}^j} = \{w_{r,k}^{j,p}\}$ vector is initialized as $\mathbf{w_{r,0}^j} = \mathbf{w^j}$, the vector containing the wavelet coefficients of the original signal at scale j. In Hadjileontiadis's method, the iterations end when the following stop criterion, depending on a user chosen ε, is verified:

$$STC_k = |E\{\mathbf{r_k^2}\} - E\{\mathbf{r_{k-1}^2}\}| < \varepsilon, \qquad (3)$$

that is, when no significant energy is lost by the "noise" between two iterations.

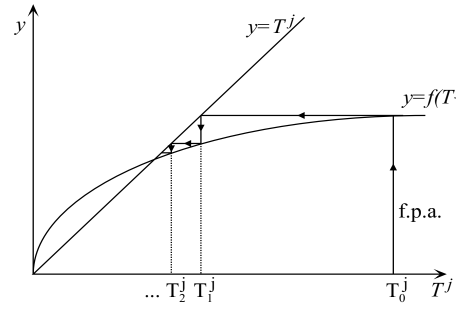

Fig. 2. Fixed point algorithm

A. Fixed point interpretation

In the sequel, we redefine the iteration (2) as a fixed-point algorithm and we give a stop-criterion, based on this interpretation.

One can prove that, under mild and realistic conditions on the probability distribution of the wavelet coefficients at scale j, $f(T^j) = 3\sigma^j$ is non-decreasing, that when $T^j \to \infty, f(T^j) < T^j$ and when T^j is close to 0, $f(T^j) > T^j$ (see Fig. 2). This means that we have at least one intersection point between $f(T^j)$ and the function $y = T^j$. From this set of points, we select the largest one, noted T_{fin}^j. This point will be a convergence point for a fixed point descent algorithm starting from $T^j = \infty$. In a fixed-point descent algorithm $f(T_1^j) = T_2^j < T_1^j$ and the convergence is achieved when the stop criterion is verified:

$$f(T_{fin}^j) = T_{fin}^j. \qquad (4)$$

We still have to prove that Hadjileontiadis's algorithm and the fixed-point descent converge to the same point. Indeed, the stop criteria (3) and (4) are equivalent, because of the orthogonality of the wavelet transform. In fact, (3) can be rewritten as:

$$STC_k = |E\{\mathbf{w_{r,k}^2}\} - E\{\mathbf{w_{r,k-1}^2}\}| < \varepsilon. \qquad (5)$$

This means, first of all, that we don't need to perform a reconstruction of the residual noise at every iteration in order to calculate its variance, as required by eq. (3), so an important reduction of the computational burden is obtained. Furthermore, as the wavelet transform is also linear, $\mathbf{w_{r,k}}$ and $\mathbf{w_{r,k-1}}$ will differ only by the few thresholded coefficients. For $\varepsilon \to 0$, the stop criterion (5) is equivalent to the fact that no coefficient is "lost" between $\mathbf{w_{r,k-1}}$ and $\mathbf{w_{r,k}}$. This means that $T_{fin}^j = 3\sigma_{fin}^j > max(|\mathbf{w_{r,fin}}|)$.

This work reveals the fixed-point nature of Hadjileontiadis's algorithm and exploit this observation in order to determine the threshold T^j by a fast descent directly on the histogram of the wavelet coefficients.

B. Maximum scale decomposition limit

Concerning the wavelet decomposition, let us say that we have used the same wavelet basis (Daubechies 4) as in [2], [3]. This choice was justified by Hadjileontiadis *et al.* by medical expertise.

Still, as we have said previously, we have to process long signals in order to extract useful data. A complete wavelet decomposition will be in this case very time consuming. Besides, we know from our experimental setup that the microphone we have used is band-limited at 50 Hz in the low frequencies, so all large scales wavelets coefficients will be zero: a complete decomposition is therefore useless. In a frequential and algorithmical interpretation, the output of the low-pass filter corresponding to the scale function in Mallat's algorithm [12] will have a zero output. Therefore, another modification of the denoising method is a limited wavelet decomposition. Anyway, as one can easily see, in the experimental setup of Hadjileontiadis *et al.*, the combination between the length of the signal and the choice of the threshold leads to the elimination of the large scale coefficients after the first iteration, because at least 9 coefficients are needed in order to have one of them greater than or equal to 3σ.

III. Segmentation and parameter extraction

As we have said earlier, our goal is to characterize the bowel sounds. Naturally, after denoising, the next step is the identification of these sounds *i.e.*, the segmentation of the signal. As we have considered that a wavelet decomposition offers also the possibility to characterize a sound by its wavelet coefficients, we have chosen a segmentation method applied directly on the wavelet coefficients vector $\mathbf{w}$.

This segmentation must be done with care, considering the upsampling-filtering reconstruction algorithm of Mallat [12] for 'db4' wavelets. A discrete wavelet transform (DWT) represents a signal as a discrete sum of continuous wavelets:

$$s(t) = \sum_{j,p} w^{j,p} \cdot \psi_{j,p}(t), \tag{6}$$

where $\psi_{j,p}(t)$ is the wavelet function (in our case 'db4') and $w^{j,p}$ is the wavelet coefficient obtained by the DWT. As shown in [12], such a wavelet is calculated by a filtering-upsampling algorithm starting from the corresponding coefficient – the inverse discrete wavelet transform (IDWT). This means that we can obtain the scale, hence the temporal dimension of a wavelet, knowing the reconstruction filter and the position of the corresponding coefficient in the transformed vector $\mathbf{w}$. Next, as the signal is a sum of wavelets, we can obtain the temporal position of a sound, its starting and ending instants. More precisely, the contribution of the $w^{j,p}$ coefficient in the reconstructed signal ends at $d = p \cdot 2^j$ and starts at $u = d - (L-1) \cdot (2^j - 1)$, where p is the temporal position, j is the scale of the corresponding wavelet and L is the length of the reconstruction filter ($L = 8$ for a Daubechies 4 filter).

As the segmentation is performed on the wavelet coefficients vector, we obtain directly the wavelet decomposition of an event (a sound) and use this decomposition to characterize it. By its construction, the wavelet denoising algorithm outputs all the events that are correlated with the basis waveform, including those of very low energy. As a processing option, we propose a second thresholding after the segmentation, based on the power of an event. More precisely, we have calculated the power for each detected sound (e_i) and we have fixed a threshold equal to the standard deviation of these powers $E = \sigma(\mathbf{e})$, where $\mathbf{e} = \{e_i\}$.

Concerning the characterization of the bowel sounds, in the actual stage of our research we propose few rather empirical features, easy to extract and physically significant: the duration of the sound and the power distributed upon each frequency band (scale). We present all the results in the sequel.

IV. Processing results

A. Experimental setup

In order to facilitate an auditive medical validation, we have attached three electret pressure microphones to mechanical stethoscope heads, placed on the abdominal area. After band-pass [50-2250] Hz anti-aliasing filtering, the signals were digitized by a 12 bit Analog to Digital Converter, at a sampling rate of 5 kHz. The signals were recorded on several healthy subjects, for 10 to 15 minutes, before and after lunch.

B. Processing results

As seen before, our limited wavelet decomposition depth approach is justified both by the physical nature of the signal and by the experimental setup. This observation reduces the computational burden from $O(K \cdot N \cdot \log N)$, for Hadjileontiadis's original algorithm, to $O(K \cdot N \cdot M)$, where K is the number of iterations and M is the depth of the decomposition (in our case $M = 7$, which implies 8 frequency bands). Our optimisation, based on the orthogonality of the wavelet transform (5) reduces it to $O(N \cdot M)$. The fixed-point interpretation speeds up the computation by eliminating the iterative testing and thresholding of the wavelet coefficient vector. In fact, it performs an initial step dedicated to the threshold finding, followed by a unique test on the coefficient vector. For a 2^{18} points signal, on a Pentium III/500 MHz platform, our Matlab implementation showed a 4 times faster execution of our optimized algorithm compared to the original.

The results of Hadjileontiadis's and our denoising algorithms are slightly different (the difference between the variances of the resulting signals is of magnitude 10^{-8}), because of our $\varepsilon = 0$ in the stop criteria (3) and (5). We consider the difference insignificant, compared to the difference between the original signal and the denoised one

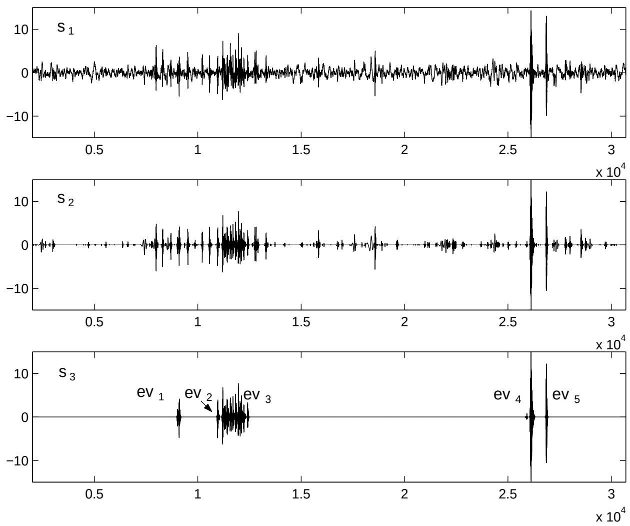

Fig. 3. Denoising and segmentation results. s_1: a section of 6.5 seconds of the normalized original signal - var(s_1)=1; s_2: the denoised segmented signal after the fixed-point algorithm - var(s_2)=0.4477 (116 events); s_3: the denoised segmented signal after the second thresholding - var(s_3)=0.3034 (5 events).

(the variance differs by 10^{-1}). Furthermore, no difference between our reconstruction and the one performed by Hadjileontiadis's algorithm can be heard listening the denoised signals.

On the other hand, there is an audible difference between the denoised signal and the final segmentation (Fig. 3 (b) and (c)), due to the second optional thresholding E. In fact, this step eliminates all the low energy events, even if important diagnostic information might be extracted from the discarded sounds. The difference between the variance of the output signal after this last processing step and the variance after the first denoising is of magnitude 10^{-1} (see Fig. 3 for the results).

Concerning the characterization of the sounds, we present the extracted features for the five events on the s_3 signal from Fig. 3 in table I (e^j is the variance on the scale j and τ is the duration of the event, in seconds). We have discarded the energy on the first scale, as it represents the frequencies lower than 20 Hz.

TABLE I

Feature extraction for the 5 events from Fig. 3.

	e^2	e^3	e^4	e^5	e^6	e^7	e^8	τ
ev_1	0	0	0.00	0.93	0.49	0.02	0.00	0.05
ev_2	0	0	0.00	0.53	0.85	0.04	0.00	0.04
ev_3	0	0	0.10	1.33	0.78	0.08	0.01	0.28
ev_4	0	0	0.11	1.13	6.58	0.35	0.15	0.10
ev_5	0	0	0.00	1.94	6.78	0.41	0.78	0.04

V. Conclusion

The aim of this paper is to facilitate the extraction of some useful diagnostic information from recorded bowel sounds. In order to achieve this, we have optimized by a fixed-point approach the denoising algorithms of Coifman and Wickerhauser [1] and Hadjileontiadis[2], [3].

We have proposed a segmentation method performed directly upon the wavelet coefficients vector and a feature extraction method, in view of a future classification.

In perspective, the use of several microphones simultaneously, as proposed by [13], will allow us to include another feature in the characterization of an event: its localization. This will permit to elaborate a map of abdominal activity, which is also of much help for diagnosis, since sound localization is an important factor.

References

[1] R. Coifman and M. Wickerhauser, "Experiments with adapted wavelet de-noising for medical signals and images," in *Time-Frequency and Wavelets in Biomedical Engineering*, pp. 323–346, IEEE Press, 1998.

[2] L. Hadjileontiadis and S. Panas, "Separation of discontinuous adventious sounds from vesicular sounds using a wavelet-based filter," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 12, pp. 1269–1281, 1997.

[3] L. Hadjileontiadis, L. Liatsos, C. Mavrogiannis, T. Rokkas, and S. Panas, "Enhancement of bowel sounds by wavelet-based filtering," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 7, pp. 876–886, 2000.

[4] W. Cannon, "Auscultation of the rythmic sounds produced by the stomach and intestines," *Am. J. Physiol.*, vol. 14, pp. 339–353, 1905.

[5] M. Gaussen, "L'auscultation de l'abdomen," *Journal de Médecine de Bordeaux*, vol. 1, pp. 25–31, 1952.

[6] D. Du Plessis, "Clinical observation on intestinal motility," *S. Afr. Med. J.*, vol. 28, pp. 27–33, 1954.

[7] G. Milton, "Normal bowel sounds," *Med. J. Aust.*, vol. 2, pp. 490–493, 1958.

[8] H. Yoshino, Y. Abe, T. Yoshino, and K. Ohsato, "Clinical application of spectral analysis of bowel sounds in intestinal obstruction,"*Dis. Col. Rectum*, vol. 33, no. 9, pp. 753–757, 1990.

[9] M. Sugrue and M. Redfern, "Computerized phonoenterography: the clinical evaluation of a new system," *J. Clin. Gastroenterol.*, vol. 18, no. 2, pp. 139–144, 1994.

[10] T. Tomomasa, A. Morikawa, R. Sandler, H. Mansy, H. Koneko, T. Masahiko, P. Hyman, and Z. Itoh, "Gastrointestinal sounds and migrating motor complex in fasted humans," *Am. J. Roentgenol.*, vol. 94, no. 2, pp. 374–381, 1999.

[11] D. Bray, R. Reilly, L. Haskin, and B. McCormack, "Assessing motility through abdominal sound monitoring," in *Proc. IEEE EMBS*, pp. 2398–2400, 1997.

[12] S. Mallat, *A wavelet tour of signal processing*, Academic Press, 1999.

[13] C. Garner and H. Ehrenreich, "Non invasive topographic analysis of intestinal activity in man on the basis of acoustic phenomena," *Res. Exp. Med. (Berl.)*, vol. 189, no. 2, pp. 129–140, 1989.