APPLICATION OF WALSH TRANSFORM BASED METHOD ON TRACHEAL BREATH SOUND SIGNAL SEGEMENTATION

Jin Feng, Farook Sattar
School of Electrical& Electronic Engineering, Nanyang Technological University, Singapore

Moe Pwint
Dept. of Information Science, The University of Computer Studies, Yangon, Myanmar

Keywords: Segmentation, Walsh Transform, Refinement Scheme, Inspiratory/Expiratory Phase, Various Types of Tracheal Breath Sounds, End-Inspiratory/Expiratory Pause.

Abstract: This paper proposes a robust segmentation method for differentiating consecutive inspiratory/expiratory episodes of different types of tracheal breath sounds. This has been done by applying minimal Walsh basis functions to transform the original input respiratory sound signals. Decision module is then applied to differentiate transformed signal into respiration segments and gap segments. The segmentation results are improved through a refinement scheme by new evaluation algorithm which is based on the duration of the segment. The results of the experiments, which have been carried out on various types of tracheal breath sounds, show the robustness and effectiveness of the proposed segmentation method.

1 INTRODUCTION

For early detection of diverse illnesses, accurate estimation of respiratory rate is very important (Sierra et al., 2005). Many adventitious lung sounds, which are indications of infectious and respiratory diseases, can be clinically characterized by their duration in respiratory cycle and relationship to the phase of respiration (Meslier et al., 1995). Therefore, segmentation of respiratory sound into individual respiratory cycles and further subdividing into its inspiratory and expiratory phases is necessary in quantifying adventitious sounds.

Generally, phonopneumography or spirometer together with sound recording devices are always used in respiratory sound analysis, in which amplitude of the sound signal is displayed simultaneously with the airflow as a function of time. Signals can be segmented into consecutive inspiratory phase, end-inspiratory pause, expiratory phase, and end-expiratory phase according to the provided Forced Expiratory Volume (FEV) readings (Taplidou and Hadjileontiadis, 2007)(Cort*é*s et al., 2005). However, it could be difficult to carry out a spirometric test for patients with high obstruction in tracheal (Cort*é*s et al., 2005).

Acoustical flow estimation is one of the first attempts to relate respiratory sounds and flow. In (Hossain and Moussavi, 2002) and (Golabbakhsh, 2004), airflow has been estimated using the respiratory sounds by applying different models, while exponential model between flow and averaged sound power has been found with the highest estimation accuracy. The model coefficients calculation in the above mentioned methods require samples of breath sound with known flow. However, the calibration process is not always possible. Therefore, a modified entropy-based linear model describing relationship between flow and tracheal sound has been derived in (Yadollahi and Moussavi, 2006) without prior acoustical flow knowledge. Also, other segmentation methods using spectral and temporal analysis of transformed respiratory sounds have been developed in (Hult et al., 2000)(Sierra et al., 2004). As these researches are still in preliminary stage, the segmentation is restricted to normal tracheal breath and the accuracy depends mainly on signal-to-noise ratio (SNR) for various types of tracheal breath sounds.

In this paper, an automatic and robust respiratory sound signal segmentation method is developed. The proposed method is based on the modification of input sound signal using a modified analysis and synthesis

Feng J., Sattar F. and Pwint M. (2008).
APPLICATION OF WALSH TRANSFORM BASED METHOD ON TRACHEAL BREATH SOUND SIGNAL SEGEMENTATION.
In *Proceedings of the First International Conference on Bio-inspired Systems and Signal Processing*, pages 116-121
DOI: 10.5220/0001057501160121

scheme based on Walsh basis functions. Without the aid of any other features, a decision module is then applied on the modified signal by adaptive thresholding for segmentation. The preliminary segmentation result is optimized lastly by the refinement scheme based on the segment duration. This scheme ensures the segmentation process to perform equally accurate irrespective of flow and types of tracheal breath sounds. The proposed method is tested to be effective for both normal tracheal breath sounds as well as adventitious respiratory sounds such as, wheeze and stridor.

2 BACKGROUND

The Walsh transform is a matrix consisting of a complete orthogonal function set having only two values +1 and -1 over their definition intervals (Beauchamp, 1984). The motivation for using Walsh transform rather than other transforms is its computational simplicity giving a realistic processing time. The Walsh function of order N can be represented as

$$g(x,u) = \frac{1}{N}\prod_{i=0}^{q-1}(-1)^{b_i(x)b_{q-1-i}(u)} \tag{1}$$

where $u = 0,1,...,N-1$, $N = 2^q$ and $b_i(x)$ is the i--th bit value of x. In this context, the Walsh functions are arranged into sequential order, the number of zero crossings of Walsh function per definition interval, to obtain a set of basis functions. The number of zero crossings increases with the order of basis functions $W = [\phi_0, \ \phi_1, \ \cdots \ ,\phi_{N-1}]$.

3 PROPOSED SEGMENTATION METHOD FOR RESPIRATORY SOUND SIGNAL

The proposed respiratory sound signal segmentation approach is based on segmentation of the respiratory sounds using Walsh functions. The segmentation method is based on the reconstruction/modification of the analyzed signals by efficient linearly combined Walsh functions. A simple decision scheme is then followed for segmentation of our recorded respiratory sound signals based on the statistics of the modified/reconstructed signal. The details of our minimal Walsh functions based segmentation method is presented here.

3.1 Modification of Signal

The modification of the input signal consists of two stages - sinusoidal signal analysis (Arfib et al., 2002) followed by our signal reconstruction scheme using minimal Walsh functions.

3.1.1 Signal Analysis

The input signal $x(n)$ is multiplied by a Hann window to yield successive windowed segments of $x_s(n)$. These window segments are mapped into the spectral domain by using FFTs. In this way, a time varying spectrum $X_s(n,k) = |X_s(n,k)|e^{j\varphi(n,k)}$ with $n = 0,1,...,N-1$ and $k = 0,1,...,N-1$ for each windowed segment is obtained. Here, $X_s(n,k)$ denotes the spectral component of the input signal at frequency index k and time index n, while $|X_s(n,k)|$ and $\varphi(n,k)$ denote the time-varying magnitude and phase responses, respectively.

3.1.2 Modified Signal Synthesis

The recorded input respiratory signal is reconstructed as a modified sequence based on our modified analysis/synthesis approach. Prior to synthesis, each s--th windowed segment is modified as the weighted sum of the magnitude $|X_s(n,k)|$ using binary Walsh basis functions. Using basis functions, the number of parameters required to track along the variations of the inspiration and expiration phases of the noisy signal can be reduced. For this reason, SVD (Singular-Value Decomposition) is used to determine the minimal number of Walsh basis functions to be applied. The detailed procedure for the identification of the minimal number of Walsh basis functions and the new modified basis function used based on the selected basis functions, are described in the following section. Applying the i-th basis function ϕ_i, a modified sequence, $y_s(n)$, for each windowed segment is then obtained as

$$y_s(n) = \sum_{k=0}^{N-1} |X_s(n,k)|.\phi_i(k) \tag{2}$$

All the modified segments are finally concatenated to generate an output signal $y(n)$ having the time-varying magnitude responses.

$$y(n) = \sum_{s=0}^{S-1} y_s(n - sN) \tag{3}$$

3.1.3 Selection of Minimal Walsh Functions for Modified Synthesis

It is very important to select appropriate basis functions so that variations between the dynamics of the two phases can be captured more precisely. A method used to select the global natural scale in discrete wavelet domain (Quddus and Gabbouj, 2002) is adopted to determine the minimal number of basis functions. This method adaptively selects the optimal scale using SVD, while decomposition is being carried out. Consider an input noisy respiratory signal x of length $\mathcal{V}$, and $y_d(\nu)$ be its modified sequence obtained by applying the basis functions of order d into Eq(2) and Eq(3). Modified sequences $\{y_d(\nu)\}_{d=0}^{D-1}$ can be represented as a matrix of size $D \times \mathcal{V}$. To determine the order of basis functions with dominant eigenvalues, the SVD of the $D \times \mathcal{V}$ matrix is calculated adaptively begin with the first two orders (i.e. ϕ_0 and ϕ_1) while adding the Walsh functions of higher orders.

Here, the proposed algorithm defines the minimal order of basis functions N_{min} as 3 throughout the simulations and found very robust against various situations. In the original algorithm (Quddus and Gabbouj, 2002), optimal scale is defined as the average of the details from the first level to the natural scale, the level associated with the dominant eigenvalue. However, this averaging may introduce clipping effect for the signals at low signal level. To avoid this effect, a shifting operator which swaps the right and left halves of the basis function coefficients is applied first. Then a good estimate of a modified binary Walsh basis function within dominant eigenvalues is defined as

$$\phi_m = \frac{\phi_0 - \sum_{i=1}^{N_{min}} CS(\phi_i)}{\max\{|\phi_0 - \sum_{i=1}^{N_{min}} CS(\phi_i)|\}} \tag{4}$$

where $N_{min} = 3$ is the largest order referring to the most prominent eigenvalues and $CS(\cdot)$ is the shifting operator. This new basis function ϕ_m provides sharper representation and higher discriminating features.

3.2 Decision Strategy

3.2.1 Preliminary Decision Module

First, 0-order basis function, ϕ_0 is used to produce a modified sequence, $y_0(\nu)$, to get the global information of the original sample signals. This modified sequence is used as a reference or pilot sequence as used in the areas of telecommunication. Containing the local characteristics, another modified signal, $y_m(\nu)$, is formed using the new basis function ϕ_m. From this new sequence, locations and durations of inspiration and expiration phases can be located more precisely even for adventitious respiratory sounds such as wheeze and stridor. In this way, approximate locations of inspiration and expiration segments are first determined from the modified signal, $y_0(\nu)$. Then, the results to determine respiratory phases can be improved by using the second modified signal, $y_m(\nu)$, which contains the detailed information. Applying the reconstructed signals y_0 and y_m, the procedure of detection scheme can be described as below:

- Extract two sequences of local minima, $\{\alpha_{0i}\}_{i=1}^{L}$ and $\{\alpha_{mi}\}_{i=1}^{L}$, where L is the number of frames, from every 4 ms frame of $y_0(\nu)$ and $y_m(\nu)$.
- Set thresholds, τ_0 and τ_m, for each minima sequence which are obtained using a simple statistics: $\tau_0 = \mu_0 - \kappa\delta_0$ and $\tau_m = \mu_m - \kappa\delta_m$, where μ_0 and δ_0 are the mean and the standard deviation of the first set of local minima, and μ_m and δ_m are those of the second set of local minima while κ is a positive value which depends on the dynamic range of modified sequence $y_0(v)$.
- Set threshold coefficient, κ, which is the same for τ_0 and τ_m. As shown by Eq(5), κ is proportional to global average of $y_0(v)$, and a is a constant value. After experimenting with 10 reconstructed waveforms of different respiration types (stridor and wheeze, normal tracheal breath for adult and infant), a is found to be 3.4, and universal for all types of tracheal breath sounds.

$$\kappa = a \times \frac{1}{N} \sum_{v=0}^{N-1} y_0(v) \tag{5}$$

- Declare a frame as an respiration frame if either $\alpha_{0i} < \tau_0$ or $\alpha_{mi} < \tau_m$. As it is mentioned earlier, respiratory cycle is divided into four consecutive phases: inspiratory phase, end-inspiratory pause, expiratory phase, and end-expiratory pause. Respiration frames is defined in this context as the frames belong to either inspiratory or expiratory phases. In this way, the respiration frame indices are obtained from $y_0(\nu)$ and $y_m(\nu)$ as $\mathcal{R}$ and $\mathcal{T}$:

$$\mathcal{R} = \{r_1, r_2, ..., r_P\} \tag{6}$$
$$\mathcal{T} = \{t_1, t_2, ..., t_Q\} \tag{7}$$

- Combine the two initial boundary decisions as follows:

$$\mathcal{C} = \mathcal{R} \cap \mathcal{T} \tag{8}$$

where $\mathcal{C}$=$\{c_1, c_2, ..., c_J\}$ is the set of elements common to $\mathcal{R}$ and $\mathcal{T}$. Considering that the members of $\mathcal{C}$ are the indices of either inspiration or expiration frames, the final decision for detecting respiration frames are obtained.

In the above, we decide that there exist respiration frames whenever some or all of the prominent local minima obtained from the first modified signal $y_0(\nu)$ would coincide with the local minima found from the second modified signal $y_m(\nu)$. For those detected frames when their corresponding local minima are not obtained from both modified sequences of $y_0(\nu)$ and $y_m(\nu)$, are discarded as outliers.

3.2.2 Refinement Scheme

Due to the quasi-stationary nature of the adventitious respiratory sounds and their relatively small dynamic range due to shallow breath, there are chances where frames are wrongly identified because of the inflexibility of the global threshold value used: small spikes happen during end-inspiratory/expiratory pauses being wrongly identified as respiration segments which are denoted by peaks; and small fluctuations during inspiration/expiration might be wrongly identified as pause segments which are denoted by troughs as indicated in Fig.1(c). In order to ensure the accuracy of the segmentation, the results obtained from the preliminary decision module will be fine-tuned by the refinement scheme to avoid wrong identification of the respiratory frames. The scheme consists of two stages:

- Identify error segments with durations shorter than threshold σ_t, where σ_t varies for patients with different respiratory rate. Since the duration of end-inspiratory/expiratory pauses range from 0% to 30% and inspiration time range from 10% to 80% of a complete breath cycle (Li, 2004), we defined error segment to be with duration less than 5% of individual's averaged breath cycle. Therefore σ_t is defined as:

$$\sigma_t = 5\% \times \frac{60}{RR} \times F_s \tag{9}$$

  where RR as Respiration Rate, is the number of breath cycle per minute and F_s is the sampling rate of the signal. Since the averaged RR is the highest for infant which is 44 breaths/min (Keszler and Abubakar, 2004), the scheme adopts this value to minimize the wrong identification. The selected parameter values are listed in Table 1. The error segments are then divided into error respiration segments and error pause segments, where the number of segments for each error segment type is counted.

Table 1: Values of parameters for refinement scheme.

Parameter	Value
F_s	8000 Hz
RR	44 breaths/min
σ_t	545 samples

- Evaluate the error segments based on segment duration. This process is applied for evaluating error respiration segments first. The procedure can be described using our following pseudo code, where respiration segment is denoted by R(s) and pause segment by P(s) and s is the positional index of the segment along time line.

```
Begin
    T = threshold;
    Pd(s) = duration of P(s);
    Rd(s) = duration of R(s);
    I = number of error R(s);
    for i=1:I,
        locate first error R(s);
        if duration of Pd(s-1) & Pd(s) < T
            if Pd(s) > Pd(s-1)
                R(s) combine with R(s-1);
            else
                R(s) combine with R(s+1);
        else if Pd(s-1) < T or Pd(s) < T
           R(s) combine with R(s-1) or R(s+1);
        else
           R(s) is considered as pause segment;
        end
    end
End.
```

This procedure is then applied for the second time to evaluate error pause segments by interchanging R(s) with P(s) in the pseudo code.

4 EXPERIMENTAL RESULTS

4.1 Data and Parameter Selection

Five different types of tracheal sound signals are chosen from (Lehrer, 1993) and (Wilkins et al., 2004). Tracheal breath sound is chosen due to its relatively larger amplitude compared with the sounds recorded over chest. Also, it has distinct inspiratory/expiratory phases and is related closely to respiratory flow.

The segmentation algorithm has been tested on total 10 sound signals, each consists of 8 breathing cycles. Four phases are distinct in every breathing cycle for all signals chosen. Since the segmentation method is working based on the overall trend instead of the detail fluctuations of the signals, the order m for reconstructed signal $y_m(v)$ should be kept low. Therefore, $m = 3$ is used in the experiments.

4.2 Illustrative Results and Analysis

Fig.1 illustrates the outputs of individual segmentation steps on a signal of inspiratory stridor and expiratory moderate wheeze. Fig.1(a) shows the original signal containing wheeze and stridor whereas Fig.1(b) shows its transformed version, the reference modified sequence $y_0(v)$, together with the reference threshold τ_0. In Fig.1(c), output of preliminary decision module is depicted. As indicated by arrows A, B, C, D, there are 4 locations of preliminary results containing error segments. Being optimized by the refinement scheme, the final segmentation result is displayed in Fig.1(d).

Also, the results for infant normal tracheal breath are shown by Fig.2. By comparing these two figures, no error segments are detected in Fig.2(c). This is due to the different nature of the signals: The quasi-stationary nature of wheeze and stridor signals gives them more prominent components at low frequency, while the fast transient nature of the normal tracheal breath makes it emphasize more on the high frequency components. Since $y_0(v)$ focusses on the signal trend which is represented by the low frequency components, it captures more spikes (low frequency details) for wheeze and stridor, but provides smoother waveforms for normal breath sound signal. Therefore, after thresholding by τ_0, segments with short duration are detected for abnormal breath sound signals. However, due to the optimization by refinement scheme, the final segmentation results are equally accurate for both normal tracheal breath sounds and adventitious breath sounds.

Moreover, illustrative results of the segmentation algorithm for different types of respiratory sound signals are shown by Fig.3(a)-(e). These results demonstrate the robustness of our proposed method on different types of tracheal breath.

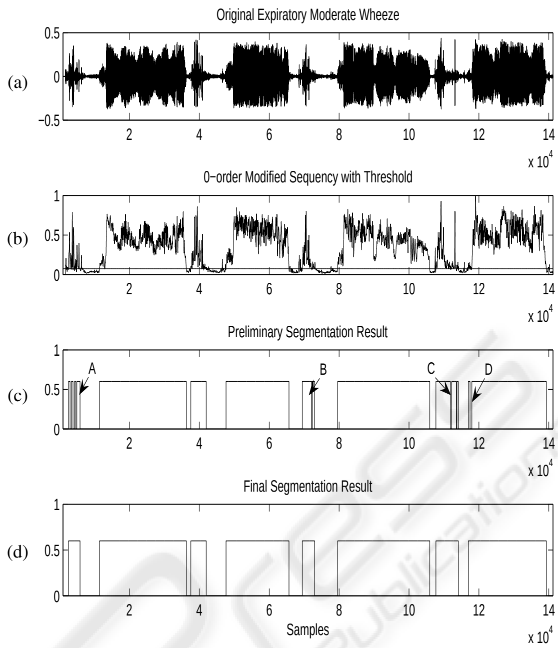

Figure 1: (a) Original signal waveform; (b)0-order modified sequence $y_0(v)$ with threshold τ_0; (c) preliminary segmentation result; (d) final segmentation result for inspiratory stridor and expiratory moderate wheeze.

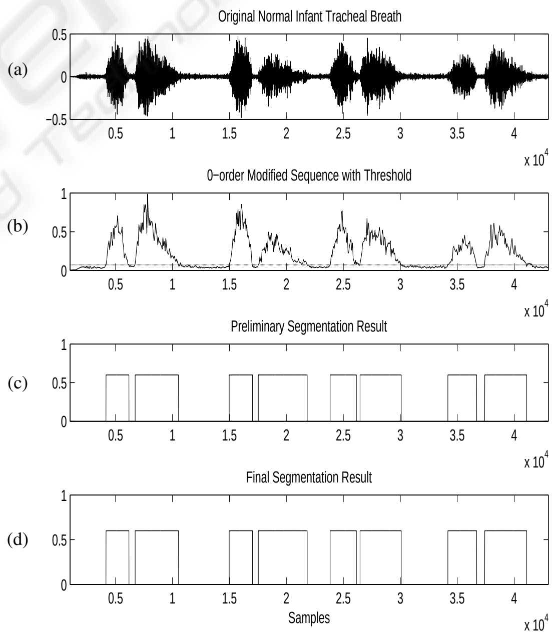

Figure 2: (a) Original signal waveform; (b)0-order modified sequence $y_0(v)$ with threshold τ_0; (c) preliminary segmentation result; (d) final segmentation result for infant normal tracheal breath.

5 DISCUSSION

In this paper, we have presented an algorithm to locate and differentiate inspiratory/expiratory phases with end-inspiratory/expiratory pauses for different types of tracheal breath sounds. The use of binary Walsh transform simplifies the proposed algorithm to a large extend and left only few parameters for adjustment. This makes the algorithm fast and automatic even in the absence of any *a priori* information of the input signal types. It performs equally accurate for both normal as well as adventitious sounds due to the incorporation of refined decision module. Thus it is more robust compared to existing methods as by using these conventional methods, accurate segmentation is still restricted within normal breath sounds.

As the only limitation, the proposed method does not perform well on raw recorded tracheal breath sound signals. This is due to the presence of the prominent heartbeat. Since the frequency range of heartbeat is below 300Hz, it interferes with the normal breath sounds and contaminates the signal with large amount of low frequency components. This can be solved by taking recording at positions with low heart sound to respiratory sound amplitude ratio, or preprocessing using a notch filter to suppress the effect of heartbeat. However, the algorithm is immune to other ambient noises due to the wide spectrum occupied by the noises.

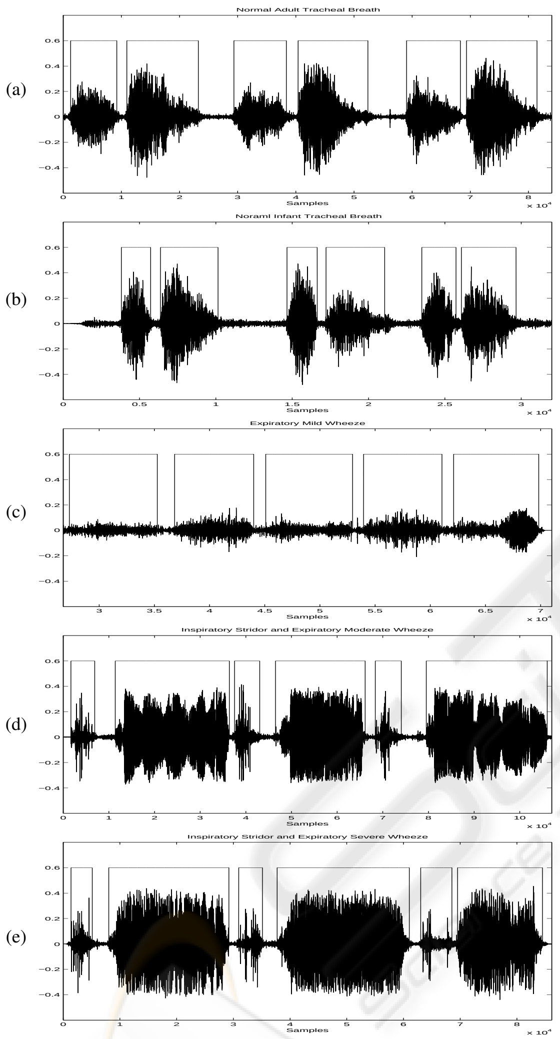

Figure 3: The segmentation results displayed with original signal waveform for (a)-(b) normal tracheal breath of adult/ infant; (c) expiratory mild wheeze; (d)-(e) inspiratory stridor and expiratory moderate/ severe wheeze.

REFERENCES

Arfib, D., Keiler, F., and Zöler, U. (2002). *DAFX - Digital Audio Effects*. John Wiley Publisher.

Beauchamp, K. G. (1984). *Applications of Walsh and Related Functions*. Academic Press.

Cort*é*s, S., Jan*é*, R., Fiz, J. A., and Morera, J. (Sept, 2005). Monitoring of wheeze duration during spontaneous respiration in asthmatic patients. *IEEE Proceedings of Engineering in Medicine and Biology*, pages 6141–6144.

Golabbakhsh, M. (2004). Tracheal breath sound relationship with respiratory flow: Modeling, the effect of age and airflow estimation. *M. Sc. theses, Electrical and Computer Engineering Department, University of Manitoba*.

Hossain, I. and Moussavi, Z. (2002). Respiratory airflow estimation by acoustical means. *[Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] EMBS/BMES Conference, 2002. Proceedings of the Second Joint*, 2.

Hult, P., Wranne, B., and Ask, P. (2000). A bioacoustic method for timing of the different phases of the breathing cycle and monitoring of breathing frequency. *Medical Engineering and Physics*, 22:425–433.

Keszler, M. and Abubakar, K. (2004). Volume guarantee: Stability of tidal volume and incidence of hypocarbia. *Pediatric Pulmonology*, 38(3):240–245.

Lehrer, S. (1993). *Understanding lung sounds, Audio CD*. Saunders.

Li, T. (2004). Invasive mechanical ventilation. *Respiratory problems: Invasive mechanical trainee manual*.

Meslier, N., Charbonneau, G., and Racineux, J. L. (1995). Wheezes. *European Respiratory Journal*, 8(11):1942–1948.

Quddus, A. and Gabbouj, M. (2002). Wavelet-based corner detection technique using optimal scale. *Pattern Recognition Letters*, 23:215–220.

Sierra, G., Telfort, V., Popov, B., Durand, L. G., Agarwal, R., and Lanzo, V. (2004). Monitoring respiratory rate based on tracheal sounds. first experiences. *Annual International Conference of the IEEE Engineering in Medicine and Biology*.

Sierra, G., Telfort, V., Popov, B., Pelletier, M., Despault, P., Agarwal, R., and Lanzo, V. (2005). Comparison of respiratory rate estimation based on tracheal sounds versus a capnograph. *Annual International Conference of the IEEE Engineering in Medicine and Biology*.

Taplidou, S. A. and Hadjileontiadis, L. J. (2007). Nonlinear analysis of wheezes using wavelet bicoherence. *Computers in Biology and Medicine*, 37:563–570.

Wilkins, R. L., Hodgkin, J. E., and Lopez, B. (2004). *Fundamentals of lung and heart sounds, Audio CD*. Mosby.

Yadollahi, A. and Moussavi, Z. (2006). A robust method for estimating respiratory flow using tracheal sounds entropy. *IEEE Transactions on Biomedical Engineering*, 53(4):662–668.