A ROBUST RESPIRATORY PHASE IDENTIFICATION SCHEME BASED ON A NEW MIXING INDEX

F. Jin[1], *F. Sattar*[1], *DYT Goh*[2], *and I.M. Louis*[3]

[1]School of Electrical & Electronic Engineering, Nanyang Technological University
email:{jinf0001,efsattar}@ntu.edu.sg

[2]Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore
email: paegohyt@nus.edu.sg

[3]University Children's Medical Institute, National University Hospital, Singapore
email: irenemelinda@yahoo.com

ABSTRACT

This paper introduces a novel method to identify inspiratory and expiratory phases from single channel tracheal breath sound (TBS) of different types, by proposing a new annotating index name as "mixing index" (MI). An alignment scheme based on phase shift difference information has been firstly introduced to align the consecutive respiratory phase segments. MI is then proposed based on similarity measurements to annotate the respective inspiration/expiration in each aligned respiratory phase segment pair. By incorporating the novel alignment scheme, the presented index overcomes the problem of phase cancellation which affects the cross-coherence of the input segment pairs. As MI is invariant to spectral content and amplitude dynamics, the proposed method maintains a good performance even in the presence of adventitious sounds. A high averaged accuracy of 97.4% for adventitious sounds and 100% for normal TBS have been thereby achieved. The proposed method has been a successful attempt to solve the clinical application challenge faced by the existing phase identification methods in terms of respiratory dysfunctions.

1. INTRODUCTION

Respiratory phase information is essential for the automation of respiratory signal analysis. In the studies of flow in heart [1] and adventitious sounds quantification [2], it is crucial to correctly differentiate individual respiratory phases as inspiration/expiration. Conventionally, direct airflow measurements such as chest movement measurements and mouthpiece pneumotachograph, are widely used for respiration monitoring. To perform the TBS segmentation into respective inspiration and expiration, the measured Forced Expiratory Volume (FEV) readings [3] have been used. However, direct flow measurements would be difficult when studying patients with high obstruction in tracheal [4] or children with neurological impairments [5].

Indirect respiration monitoring methods by acoustical analysis of TBS have therefore been recently proposed. Respiratory phase identification methods based on spectral and temporal analysis of the transformed TBS have been suggested in [6][7][8]. As presented in [6] and [7], lung sound information instead of TBS has been used to distinguish inspiration from expiration. The phase identification accuracies depend solely on the distinct amplitude and spectral differences between inspiration and expiration of lung sounds. However, compare to lung sound, TBS as a more appropriate source for adventitious sound analysis, is having less distinctive respiratory phases with inconsistent spectral differences between inspiration and expiration. In [8], the respective respiratory phases are distinguished by the combined features of sound envelope, frequency content as well as disturbance characteristics of TBS, but an extra microphone is used to capture the ambient noise. In order to solve the above mentioned problem, a triplet Markov chain in wavelet packet domain is adopted by [9] to improve accuracy of identifying inspiration and expiration by removing its dependency on signal amplitude. However, *priors* on the respiratory cycle structure for normal respiratory sounds have been exploited for chain adaptation that makes this method not robust for adventitious TBS.

As the presence of adventitious sounds affects the amplitude and energy distribution of the TBS signals to a large extend, the aim of this paper is to propose a method which is able to identify inspiration and expiration from single channel TBS recordings. The proposed method therefore should be independent of amplitude variation between these two respiratory phases and able to perform accurate annotation without knowing the structures of the respiratory cycles. The presented method is applied on the consecutive inspiration/expiration segments as obtained by using the segmentation method presented in [10]. Each pair of consecutive respiratory phase segments are first aligned using phase shift information. A frequency domain MI based on cross-segment similarity measurements is lastly proposed for inspiratory/expiratory phase identification within each segments pair. The performances of the proposed MI are compared for different types of real TBS recordings to show its robustness.

2. METHODOLOGY

The overall structure of our respiratory phase identification scheme is shown in Fig. 1. The input noisy TBS sequence is indicated by $y(n)$ and $y_{1,2}(n)$ denote the consecutive respiratory phase segments pair. The proposed phase identification scheme consists of a respiratory phase segment alignment sub-scheme and a phase identification sub-scheme based on the newly proposed MI calculation. The sub-schemes together with the notations used in the block diagram are described in Section 2.2 and 2.3 respectively. Here, phase segmentation method which has been proposed in our previous paper [10] is used for the correct localization of the respiratory phase segments.

2.1 Data Acquisition

The recording environment and equipments are chosen based on the standard given by [11]. Recordings were done in audio laboratory with the subjects in sitting position. Single

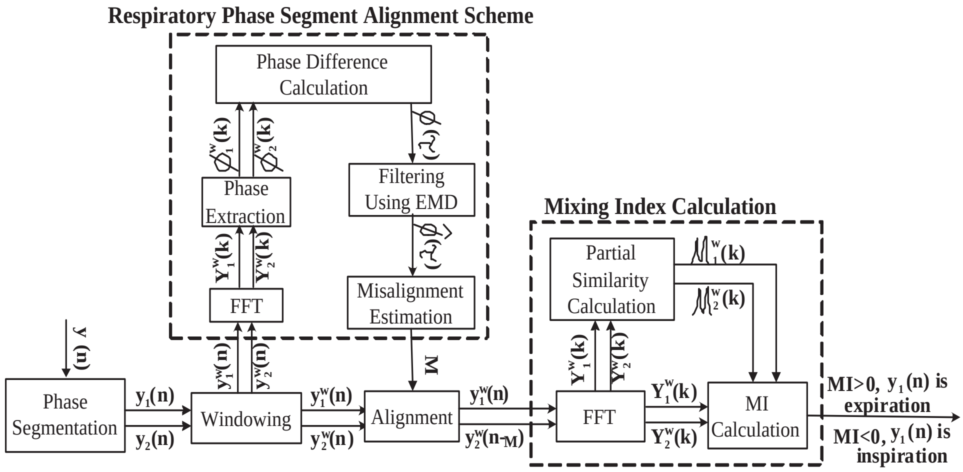

Figure 1: A overall block diagram of the proposed respiratory phase identification scheme.

electret condenser microphone (ECM-77B, Sony Inc., Japan) was inserted into a hemispherical rubber chamber of 2 cm in diameter, and placed over suprasternal notch. Recording software WAVEPAD (V3.05, NCH Swift Sound Software) was used and the TBS recordings have been saved as mono-channel '*.wav' files with sampling frequency $F_s = 11025$ Hz. Test subjects were asked to breathe normally and 600 seconds recording was saved for each subjects.

TBS is chosen due to its distinct respiratory phases and close relationship with respiratory flow. The origin of TBS is the vibrations in tissues caused by the turbulence occurred during the airflow into or out of the lungs. A small time delay is present related to the distance between sound source and microphone (typically 0.03 ms) [12]. TBS can be segmented into four successive phases: inspiratory phase, end-inspiratory pause, expiratory phase, and end-expiratory pause. In this paper, only respiratory phases (inspiratory/expiratory phases) are involved in phase identification.

2.2 Respiratory Phase Segment Alignment Scheme

As consecutive inspiration/expiration segments pairs are used as the input of the presented method, the effectiveness of the proposed annotating index depends on the correct boundary locations of segmented respiratory phases (inspiratory/expiratory phases). Although segmentation method with high accuracy in [10] is adopted here to obtain the inspiration/expiration segments pairs, any other suitable methods can also be used. Therefore, respiratory phase segment alignment scheme is introduced as signal conditioning in order to make the annotation index independent of respiratory phase segmentation accuracy.

As the proposed MI in Section 2.3 is calculated based on Fourier transform of the windowed signal, we define the alignment in terms of the phase shift of the signal in frequency domain. The unwrapped phase $\phi^w(k)$ of the windowed signal is calculated in frequency domain with k as the frequency index and w being the windowed frame index.

Fast Fourier transform (FFT) is first applied to every windowed frame $y_1^w(n)$ and $y_2^w(n)$ of the adjacent respiratory phase segments $y_1(n)$ and $y_2(n)$ to obtain $Y_1^w(k)$ and $Y_2^w(k)$. A rectangular window with minimum length of $N = 1250$ is used to ensure uniform energy distribution within each windowed frame for accurate phase approximation based on averaging. $1 \leq w \leq W$ and W indicates the total number of frames of 1250 samples. Also, the choice of window length is required to be large enough to maintain a low time resolution for reducing the inter-segment variance due to the presence of possible discontinuous adventitious sounds. $Y_1^w(k)$ or $Y_2^w(k)$ consists of W complex numbers that are represented as magnitude and phase pairs. Factors of 2π are added or subtracted to obtain the unwrapped phases of $\phi_1^w(k)$ and $\phi_2^w(k)$.

The phase shift difference $\Phi(\tau)$ between two consecutive respiratory phase segments with a varying time delay $\tau = [1, N]$ is then calculated as $\Phi(\tau) = \frac{1}{W}\sum_{w=1}^{W}\Phi^w(\tau)$, where

$$\Phi^w(\tau) = \frac{1}{N}\sum_{k=0}^{N-1} ||\triangle\phi_1^w(k) - \triangle\phi_2^w(k,\tau)||^2 \qquad (1)$$

with $\triangle\phi^w = \phi^w(k+1) - \phi^w(k)$ being the phase shift and $||\cdot||^2$ indicating the L_2 norm. $\triangle\phi_2^w(k,\tau)$ represents the phase shift of $y_2^w(n-\tau)$.

Since the resulting phase shift difference $\Phi(\tau)$ is very spiky, the idea of empirical mode decomposition (EMD) has been adopted here to extract the underlying phase shift difference information. EMD being a general nonlinear, nonstationary signal processing technique, is suitable for biomedical signal analysis. As the basis functions (intrinsic mode functions, $s_{j|j=0...J}$ with J being the total number of functions) are directly derived from the signal, the analysis is adaptive instead of being combinations of fixed sinusoids. The s_j functions are calculated according to [13] and the filtered phase shift difference $\hat{\Phi}(\tau) = \sum_{j=l}^{J} s_j(\tau)$. $l = 6$ is used in this paper as higher l is not able to approximate the signal for maintaining the trend, while lower l values capture more irregularity information.

Fig. 2 shows the plots of the filtered phase shift difference $\hat{\Phi}(\tau)$ extracted by EMD and that by conventional low-pass filtering together with the raw $\Phi(\tau)$ between two consecutive respiratory phase segments of a normal TBS recording. It is shown that low-pass filtering makes the analysis of $\Phi(\tau)$ inaccurate, while EMD enables the tracking of the underlying information through $\hat{\Phi}(\tau)$ for misalignment estimation.

The relative time shift $\tau = M$ between the misaligned signals $y_1^w(n)$ and $y_2^w(n)$ is estimated as the first local minima of the extracted phase shift difference $\hat{\Phi}(\tau)$ as illustrated in Fig. 2. And the aligned consecutive segments $y_1^w(n)$ and $y_2^w(n-M)$ are then used in MI calculation.

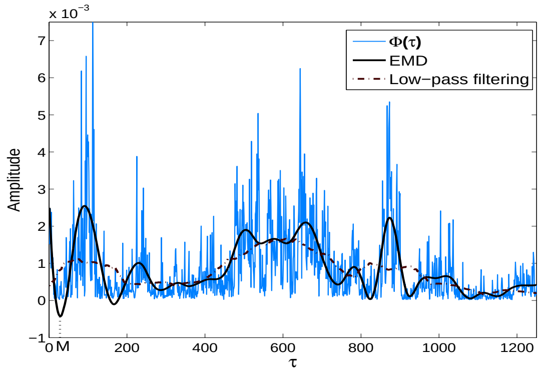

Figure 2: Illustrative plots of the filtered $\hat{\Phi}(\tau)$ extracted by EMD with $l = 6$ and that by conventional low-pass filtering displayed together with $\Phi(\tau)$.

2.3 Mixing Index (MI)

This section introduces a frequency domain annotating index as named by the mixing index (MI) for respiratory phase identification, based on inter-segment similarity measurements. For each frequency bin, MI is having values [−1, 1] centered at 0. The basic idea behind MI is to compare the consecutive phase segment signals in the time-frequency plane to derive an time-averaged mixing gain associated with each frequency bin. The final MI values for each pair of signals are then obtained by taking the mean over the selected frequency bins.

Similarity between $Y_1^w(k)$ and $Y_2^w(k)$, which are the FFT of $y_1^w(n)$ and $y_2^w(n-M)$, can therefore be defined as

$$\mu(k) = \frac{2}{W}\sum_{w=1}^{W}\frac{|Y_1^w(k)Y_2^w(k)^*|}{|Y_1^w(k)|^2+|Y_2^w(k)|^2} \tag{2}$$

with '$*$' indicating complex conjugate. Also, we defined another set of partial similarities based on cross-correlation function of the phase aligned signals to avoid ambiguity:

$$\begin{cases} \mu_1^w(k) & = \frac{|Y_1^w(k)Y_2^w(k)^*|}{|Y_1^w(k)|^2} \\ \mu_2^w(k) & = \frac{|Y_2^w(k)Y_1^w(k)^*|}{|Y_2^w(k)|^2} \end{cases} \tag{3}$$

The annotating index MI(k) at frequency bin k can be then obtained as

$$\mathrm{MI}(k) = [1-\mu(k)] \times \hat{D}(k) \tag{4}$$

where $\hat{D}(k)$ is a mapped version of the difference $D(k) = \frac{1}{W}\sum_{w=1}^{W}\mu_1^w(k) - \mu_2^w(k)$ between the partial similarities:

$$\hat{D}(k) = \begin{cases} +1 & \text{for } D(k) > 0 \\ 0 & \text{for } D(k) = 0 \\ -1 & \text{for } D(k) < 0 \end{cases} \tag{5}$$

Therefore, the averaged MI for each pair of consecutive phase segments is:

$$\mathrm{MI} = \frac{1}{K2-K1}\sum_{k=K1}^{K2}\mathrm{MI}(k) \tag{6}$$

with $K1$ and $K2$ being the frequency bin indices. As an inherent property of respiratory sounds, the inspiration segments have stronger frequency components than expiration segments. Therefore, the partial similarity $\mu_1^w(k) < \mu_2^w(k)$ and thus MI<0 for $y_1^w(n)$ being annotated as inspiration and $y_2^w(n-M)$ being annotated as expiration. It is vice versa for MI>0.

The proposed annotating index is useful since MI being the relative index, is not affected by any spectral content or amplitude dynamics. Thereby, any amplitude changes or even complete modification of spectra to both respiratory phases (as long as the changes are proportional to the similarity coefficient), do not affect the MI values.

3. RESULTS

3.1 Experimental Dataset

In this study, the real recording dataset consists of TBS from 7 healthy subjects and 14 subjects with different degrees of airway obstructions (8 males/13 females, 15 ± 9 years old). The characteristics due to sex, age, weight were not taken into consideration. Also, preclassified and preprocessed normal/wheezing respiratory sound recordings were drawn from 2 databases [14][15]. A total of 60 normal TBS inspiration/expiration (IE) pairs, 40 wheezing IE pairs, 32 stridor IE pairs, and 45 IE pairs from mixture of wheeze and pleural sounds (WNP) recordings were available.

3.2 Experimental Results

The performance of the proposed method is depicted in terms of the quantitative measure of accuracy (%). The accuracy of the proposed method relies on the choice of frequency bins ($K1$ and $K2$) and the incorporation of the alignment scheme.

3.2.1 The Role of Respiratory Phase Segment Alignment Scheme

Fig. 3 indicates the performance of the proposed phase segment identification method on different types of TBS with and without the incorporation of respiratory phase segment alignment. The significance of the alignment scheme can be realized in terms of the improved results by temporally shifting the consecutive respiratory phase segments for achieving minimum phase shift difference. Since the similarity measurements $\mu(k)$ and $\mu_{1,2}^w(k)$ in (2) and (3) are affected by the presence of undesired phase shift due to the segmentation method adopted, misalignment would result in inaccurate MI calculation and thus inaccurate respiratory phase identification. Therefore,the incorporation of this phase alignment scheme minimizes the effect of phase cancellation to ensure an accurate annotation as well as identification.

Since time shifting $y_2^w(n-M)$ which implies $e^{j\Phi(M)}Y_2^w(k)$ is applied, the presented scheme improves the performance for TBS signals interfered by the intermittent discontinuous adventitious sounds of broad band nature. However, continuous adventitious sounds interfere with normal TBS by having additional frequency components. These frequency components enlarge the phase misalignments for the interfered frequency bins. Therefore, being frequency invariant, the phase shift $\Phi(M)$ might not be sufficient to compensate these frequency varying misalignments caused by the presence of continuous adventitious sounds. In general, the presence of the alignment scheme makes the proposed MI less sensitive to the accuracy of segmentation method.

3.2.2 Choice of the Frequency Bins

The effectiveness of the proposed annotating index MI is shown by Fig. 4 with $K1$ varies in (6). As the presence of heart sound is mainly dominated within the frequency band of 20-200 Hz for $F_s = 11025$ Hz [11], and would be even higher when being captured over suprasternal notch, the analysis of TBS signals is confined within [300, 1000] Hz to minimize the effects of heart sounds and high-frequency noise. Therefore, with $K2 = 113$ (1000 Hz) at $F_s = 11025$ Hz ($K2 = 1250$) have been selected for MI calculation. At the same time, $K1$ varies from $K1 = 34$ (300 Hz) to 68 (600 Hz) to show the effect of frequency bins selection.

As average phase shift $\Phi(M)$ is used for respiratory phase alignment scheme, it can be a good approximation of instantaneous phase shift for narrow band signals. Also, since the MI measurement is computed as an average value over all selected frequency bins of $[K1, K2]$, the choice of $K1$ should be small enough to include all desired frequency components for an unbiased average MI. Therefore, it is noticed that the performance has been improved when $K2$ reduces from 45 (400 Hz) to 34 and from 68 to 56.

On the other hand, considerable improvements as observed for wheeze, stridor and WNP at $K1 = 56$ (500 Hz) indicate the significance of the narrow bandwidth used. As the energy distribution for normal TBS is more uniform within each windowed frame $Y^w(k)$, average phase shift can be a good approximation of its instantaneous value for relatively larger frequency bandwidth. Thus $K1 = 34$, 45, 56 are all producing accurate results. Since both stridor and wheeze are characterized by the periodic waveforms having dominant frequency over 100 Hz and pleural sounds are characterized by transient waveforms, their presence can affect the uniform energy distribution within the windowed frames. In such cases, narrow bandwidth should be used to ensure accurate average phase shift approximation and thereby good performances have been observed with $K1 = 56$.

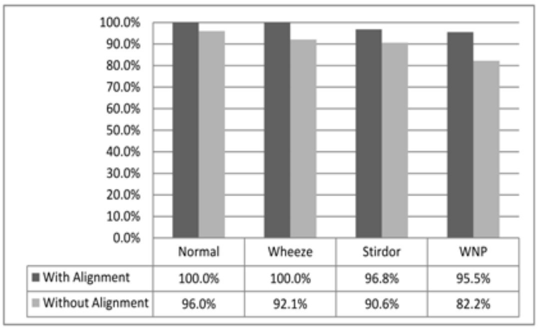

Figure 3: Performance of the proposed phase identification method with and without alignment scheme for MI calculation.

Figs. 5 and 6 illustrate the performance of the proposed phase identification method on a preprocessed normal respiratory sound and a real wheeze recording. As shown by Fig. 5, normal TBS is having stronger inspiration than expiration. While on the other hand, wheeze signal is having opposite amplitude dynamics as depicted by Fig. 6(b). Despite this large diversity in amplitude dynamics due to the presence of adventitious sounds, the proposed MI is able to identify the respective respiratory phases correctly.

3.2.3 Comparison with Other Method

A frequency index has been calculated in [8] as the summation of spectral components within frequency range of [300, 1100] Hz. The respiratory phase identification is therefore relied on the signal nature of normal TBS such that the calculated index at the beginning of the respiratory phases is higher for expiration than that for inspiration. Also, the information including respiratory pause length and the identification results of previous phase have been incorporated to improve the reliability of this method. The comparison results are presented in Table. 1 where the proposed method outperforms the method in [8]. As the presence of adventitious sounds causes variations in signal nature, the identification criteria based on nature of normal TBS signal no longer sustains. Especially in the presence of stridor which is usually inspiratory, the frequency index at the beginning of the inspiratory phases is even higher. Therefore, the method in [8] is having a much lower identification accuracy for inspiratory adventitious sounds but a high accuracy for expiratory adventitious sounds (wheeze). Furthermore, due to the absence of distinct pauses, the existing method does not show a good performance for signals with high respiratory rate. Thus, only 86.8% accuracy has been achieved for normal TBS when infant TBS signals have been included.

4. CONCLUSION AND FUTURE WORK

In this paper, we propose a robust method to identify respiratory phases into inspiration/expiration for different types of TBS. The incorporation of respiratory phase segment alignment scheme enhances the robustness of the proposed method by providing consistent high respiratory phase identification accuracies irrespective of the types of TBS recordings. The proposed MI is insensitive to the amplitude vari-

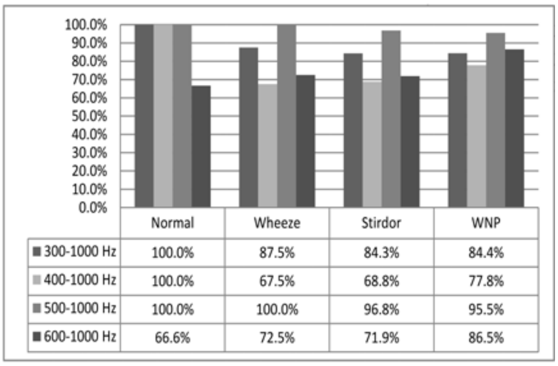

Figure 4: Performance of the proposed phase identification method with $K1 = 34$, 45, 56, 68 and $K2 = 113$ for MI calculation.

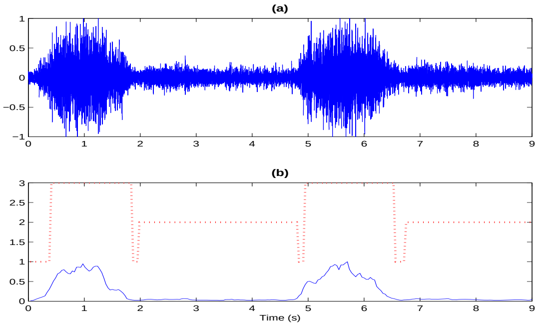

Figure 5: Performance of the proposed phase identification scheme on a preprocessed normal TBS. (a) Waveform of the signal; (b) The phase identification result (dotted line, inspiration=3, expiration=2, pause=1) along with the signal envelope.

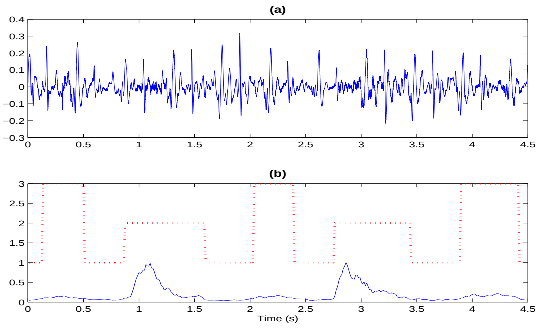

Figure 6: Performance of the proposed phase identification scheme on a real wheeze recording. (a) Waveform of the signal; (b) The phase identification result (dotted line, inspiration=3, expiration=2, pause=1) along with the signal envelope.

ation, and it makes the algorithm automatic even in the absence of any *a priori* information of the input signal types. As the performance of the proposed method still slightly varies among different types of adventitious sounds, a source extraction scheme may be incorporated in the future to extract the underlying normal TBS prior to MI calculation, for further improving the robustness of the method.

Table 1: Performance comparison between the proposed MI and the identification method in [8] on different types of TBS signals

Signal Type	Accuracy (%)	
	The proposed method	Method in [8]
Normal TBS	100	86.8
Stridor	100	20.2
Wheeze	96.8	95.2
WNP	95.5	74.4

REFERENCES

[1] P. Hult, B. Wranne, and P. Ask, "A Bioacoustic Method for Timing of the Different Phases of the Breathing Cycle and Monitoring of Breathing Frequency", *Medical Eng. and Physics*, vol. 22, pp. 425–433, 2000.

[2] N. Meslier, G. Charbonneau, and J. L. Racineux, "Wheezes", *Eur. Respir. J.*, vol. 8, pp. 1942–1948, 2006.

[3] S. A. Taplidou and L. J. Hadjileontiadis, "Nonlinear Analysis of Wheezes Using Wavelet Bicoherence", *Computers in Biology and Medicine*, vol. 37, pp. 563–570, 2007.

[4] S. Cortés, R. Jané, J. A. Fiz, and J. Morera, "Monitoring of Wheeze Duration During Spontaneous Respiration in Asthmatic Patients", *Proc. 27th IEEE EMBS Conf.*, 2005.

[5] A. Yadollahi and Z. Moussavi, "A Robust Method for Estimating Respiratory Flow Using Tracheal Sounds Entropy", *IEEE Trans. Biomed. Eng.*, no. 4, vol. 53, pp. 662–668, 2006.

[6] J. Chuah and Z. Moussavi, "Automated Respiratory Phase Detection By Acoustical Means", *University of Manitoba, Dept. of Electrical and Computer Engineering*, 2004.

[7] I. Yildirim, R. Ansari, and Z. Moussavi, "Automated Respiratory Phase and Onset Detection Using Only Chest Sound Signal", *Proc. 30th IEEE EMBS Conf.*, pp. 2578–2581, 2008.

[8] P. Hult, T. Fjällbrant, B. Wranne, O. Engdahl, and P. Ask, "An Improved Bioacoustic Method for Monitoring of Respiration", *Tech. Health Care*, vol. 12, pp. 323–333, 2004.

[9] S. L. Cam, Ch. Collet, and F. Salzenstein, "Acoustical Respiratory Signal Analysis and Phase Detection", *Proc. 33rd IEEE ICASSP Conf.*, pp. 3629–3632, 2008.

[10] F. Jin, F. Sattar, and M. Pwint, "Phase Segmentation of Noisy Respiratory Sound Signals Using Genetic Approach", *Proc. 1st BIOSIGNALS Conf.*, vol. 2, pp. 122–127, 2008.

[11] A. R. A. Sovijärvi, J. Vanderschoot, and J. R. Eavis, "Standardization of Computerized Respiratory Sound Analysis", *European Respiratory Review*, no. 77, vol. 10, pp. 585–649, 2000.

[12] N. Gavriely and D. W. Cugell, "Airflow Effects on Amplitude and Spectral Content of Normal Breath Sounds", *Journal of Applied Physiology*, no. 1, vol. 80, pp. 5–13, 1996.

[13] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis", *Proc. R. Soc. London A*, pp. 903–1005, 1998.

[14] S. Lehrer, "Understanding Lung Sounds", Philadelphia, Pa.: W.B. Saunders, 2002, Audio CD.

[15] R. L. Wilkins, J. E. Hodgkin, and B. Lopez, "Fundamentals of Lung and Heart Sounds", Mosby, 2004, Audio CD.