

AN ENHANCED RESPIRATORY RATE MONITORING METHOD FOR REAL TRACHEAL SOUND RECORDINGS

F. Jin¹, F. Sattar¹, DYT Goh², and I.M. Louis³

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

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

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

ABSTRACT

This paper addresses the problem of non-invasive respiratory rate (RR) monitoring using single channel tracheal sound (TS) recordings. We have recently developed a robust respiratory phase segmentation method based on genetic algorithm (GA) which works well only for preprocessed clean TS. Therefore, an enhanced respiratory phase monitoring method is proposed in this paper by exploiting the signal redundancy to our existing method. In this context, appropriate overlapping windows have been applied to ensure sufficient redundancy of TS signals. The performance of the enhanced method is analyzed for different types of real TS and standard preprocessed TS. The average accuracy of respiratory phase segmentation found for real TS is comparable to that of the standard preprocessed data by our proposed method.

1. INTRODUCTION

Accurate estimation of respiratory rate (RR) plays an important role in many clinical situations. Many adventitious sounds as indications of infectious and respiratory diseases, can be clinically characterized by their duration in respiratory cycle and relationship to the phase of respiration [1]. Thus it is necessary in adventitious sound quantification to segment the respiratory sound into individual respiratory cycles their respective inspiratory and expiratory phases. Direct airflow measurements such as spirometry, are widely used for RR monitoring, but they have suffered from the limitation which the accurate flow measurement depends on mechanisms that affect the natural breathing pattern [2]. This makes the direct flow measurement fails in many situation, especially for patients with high obstruction in tracheal [3].

Indirect flow measurement and RR monitoring methods by acoustical analysis of TS has therefore been recently proposed as an alternative solution for RR monitoring. Respiratory phase segmentation methods based on spectral and temporal analysis of transformed TS have been proposed in [2][4][5]. Since the existing methods depend on either spectral content or short-term energy of the signal, the detection results are strongly affected by the amplitudes of the input signals. In addition, attempts have also been made to relate flow with respiratory sounds. Airflow has been estimated using respiratory sounds by applying different models with training dataset required [6][7]. A relatively high estimation accuracy has been achieved in [7] but the predefined linear model applied does not support flow estimation for different types of TS other than normal TS.

In this paper, an enhanced RR monitoring method for real TS recordings has been proposed by extending a re-

cently developed GA based respiratory phase segmentation method [8]. The relationships between signal redundancy, disjointness and wide-sense stationarity have been investigated. The redundancy of the signals using different overlapping windows are measured in terms of spectrum conformity with 1st order auto-regressive (AR(1)) spectrum. This ensures the appropriately windowed signal with sufficient redundancy is used for *SampEn* calculation. As the estimated number of respective respiratory phases present in TS is initially obtained, a new evaluation function is calculated based on *SampEn* of the appropriately windowed signal and a heterogeneity measure. A multi-population GA is then employed using the evaluation function to determine the locations of phase boundaries. The proposed method is thus effectively monitoring RR by improving the signal redundancy.

2. METHOD

2.1 Data

TS is chosen due to its distinct respiratory phases and close relationship with respiratory flow. The origin of TS is the vibrations in tissues caused by the turbulence occurred during the airflow into or out of the lungs. It is captured by a microphone placed over the suprasternal notch. A small time delay is present related to the distance between sound source and microphone (typically 0.03ms) [9]. TS can be segmented into four successive phases: inspiratory phase, end-inspiratory pause, expiratory phase, and end-expiratory pause.

The preprocessed TS recordings from [10][11] and different types of real TS recordings corrupted with heart sounds and ambient noise are used for testing and analysis. TS recordings are acquired by using a single electret condenser microphone (ECM-77B, Sony Inc., Japan). Test subjects were asked to breathe normally with no targeted flow when a 600 seconds recording was done for each subjects with sampling frequency of 44.1kHz. The segmentation algorithm has been tested using a total of 17 TS recordings as well as 10 preprocessed TS signals. Each TS recording consists of 10 breathing cycles whereas each preprocessed data has duration of 20 seconds.

2.2 Signal Redundancy Enhancement

As a measure of information richness of the signal, a high redundancy implies high ability that past sections of the signal to provide information about the present sections. Disjointness of a signal can be defined as the degree of non-

overlapping of the signal. As a measure of information richness of the signal, signal with high redundancy results in low disjointness. Since the proposed method is based on GA which works effectively only under high data sufficiency condition, information redundancy has to be ensured or the signal has to be modified for highly disjoint conditions. In this way, redundancy measurement is required as the initial step for the proposed algorithm ensuring the signal to have sufficient redundancy.

In this paper, overlapping technique has been adopted to enhance information redundancy of the signal. It observes that only correct overlapping can produce AR WSS sequence that is highly redundant with low disjointness. The signal is first enframed using a gliding window with different overlapping ratios ranging from 0% to 80%. The spectra of the ensemble average of the frames obtained using different overlapping ratios are calculated. For an input TS signal, the most appropriate overlapping ratio is found by measuring the relative conformity between the obtained spectra and an AR(1) smooth spectrum. The conformity is simply measured in terms of spectral distortion as

$$D_{ii} = \sum_{k=1}^{k_o} |\Gamma(k) - E_{ii}(k)|^2 \quad (1)$$

where k is the frequency spacing index $k = 1, 2, \dots, K$, with $K = 64$ in this context. k_o is the band of interest for conformity calculation corresponding to the first spectral peak since AR(1) is applied here. E_{ii} refers to the spectrum of the ensemble average of the windowed frames with i^{th} overlapping ratio. In this paper, $ii = [1 \ 17]$ refers to overlapping ratio [0% 80%] with step size 5%. Furthermore, Γ representing an AR(1) spectrum, is calculated as a frequency correlation function (FCF) using the expression

$$\Gamma(k) = \frac{1 - \alpha}{1 - \alpha e^{-j2\pi k/K}} \quad (2)$$

where $0 < \alpha < 1$ which is set as $\alpha = 0.9$ in this paper. The FCF is WSS and depends only on frequency spacing k . The conformity test is done iteratively by selecting the overlapping index ii with the minimum D_{ii} . The corresponding overlapping ratio is used lastly to modify the signal for redundancy improvement.

2.3 Sample Entropy

Sample entropy (*SampEn*) is applied on the appropriately windowed signal to measure the complexity and regularity of time series. $SampEn(m, r, N)$ is chosen as it does not count self-matches of the time series. This ensures the consistency of the measurement and reduces the dependency on the signal length. It is defined as the negative natural logarithm of the conditional probability that a data set of length N , having repeated itself within a tolerance r for m points, will also repeat itself for $m + 1$ points, without allowing self-matches.

For an input signal u of length N , $\{u(j) : 1 \leq j \leq N\}$ forms the $N - m + 1$ vectors $x_m(i)$ for $\{i | 1 \leq i \leq N - m + 1\}$, where $x_m(i) = \{u(i+k) : 0 \leq k \leq m - 1\}$ is the vector of m data points from $u(i)$ to $u(i+m-1)$. In this context, only the first $N - m$ vectors of length m are considered to ensure that, $x_m(i)$ and $x_{m+1}(i)$ are defined for $1 \leq i \leq N - m$. Let $B^m(r)$ is the probability that two sequences will match for m points and $A^m(r)$ is the probability that two sequences will

match for $m + 1$ points. $B_i^m(r)$ is defined as $(N - m - 1)^{-1}$ times the numbers of vectors $x_m(j)$ within r of $x_m(i)$, where $1 \leq j \leq N - m$, and $j \neq i$ to exclude self-matches. Then $B^m(r)$ is defined as

$$B^m(r) = (N - m)^{-1} \sum_{i=1}^{N-1} B_i^m(r) \quad (3)$$

Similarly, $A_i^m(r)$ is defined as $(N - m - 1)^{-1}$ times the numbers of vectors $x_{m+1}(j)$ within r of $x_{m+1}(i)$, where $1 \leq j \leq N - m$ and $j \neq i$. Then set $A^m(r)$ as

$$A^m(r) = (N - m)^{-1} \sum_{i=1}^{N-1} A_i^m(r) \quad (4)$$

Finally, sample entropy (*SampEn*) is calculated by

$$SampEn(m, r, N) = -\ln \frac{A^m(r)}{B^m(r)} \quad (5)$$

A low value of *SampEn* reflects a high degree of self-similarity in time series. With increasing irregularity, a larger value of *SampEn* is obtained.

2.4 Multi-population Genetic Algorithm

GAs are numerical optimization algorithms operate on a population of strings as a group of potential solutions of a problem [12]. Fitness of each string is calculated in decoded form by applying an evaluation function to measure how good or bad the solutions within the population. At each generation, a new set of solutions are produced by selecting the fittest strings in the problem domain and through the application of the genetic operators such as crossover and mutation.

2.4.1 Initial Population

In order to detect both start and end locations of each segment, a population of GA is generated with the strings whose length is two times the total number of segments as obtained earlier. A string is real-valued which represents the locations of the candidate segment boundaries in ascending order. Although the binary-coded GAs are the most commonly used representation, a more natural real-valued representation is used in this system to increase the efficiency of GA. Using the real-valued strings, there is no need to convert strings to solution vectors to evaluate their fitness. Thus it would be faster in computation.

2.4.2 Evaluation Function

In GAs, an evaluation function or fitness function is usually used to evaluate the performance of the strings in the problem domain. In order to obtain accurate boundaries of each segment, an evaluation function is designed using the heterogeneity measure and *SampEn*. This function simultaneously maximize the homogeneity within the segments and heterogeneity among different segments using sample entropy.

In this way, *SampEn* of the original segmenting signal is calculated first to investigate the dynamics. To achieve feasible computational time and to make the proposed algorithm tractable, *SampEn* is calculated on each data set of length 100 (i.e. $N=100$) within a tolerance r of $0.15 \times SD$ for 1 point (i.e. $m=1$). Here, SD is the standard deviation of the

data set. Let H_w be the total within-segment homogeneity and H_b denotes the total between-segment heterogeneity, a segmentation evaluation function is defined as

$$H = \frac{H_b + 1}{H_b + H_w + 1} \quad (6)$$

where total within-heterogeneity H_w is defined as

$$H_w = \frac{\sum_{i=1}^S L_i \sigma_i^2}{L} \quad (7)$$

where L is the total length of the segmented signal, L_i is the length of i^{th} segment, σ_i^2 is the variance of the *SampEn* of the i^{th} segment and S is the number of segments in the segmented signal. The between-segment heterogeneity, H_b , is defined as the average Euclidean distance between the mean value of the *SampEn* of any two adjacent segments.

$$H_b = \frac{\sum_{(i,j) \in adjacent, i \neq j} \|\mu_i - \mu_j\|^2}{ns} \quad (8)$$

where ns is the total number of the adjacent segments in the segmented signal, μ_i and μ_j are the mean value of the *SampEn* of the i^{th} and j^{th} segments. H becomes one when the internals of all segmented respiratory signals are completely homogeneous.

2.4.3 Evolution Procedure

The proposed algorithm applies the multiple subpopulations approach for the evolutionary process. The initial population is created using 8 subpopulations containing 20 individuals each where each of them can evolve parallelly using *crossover* and *mutation* over generations. At each generation, 90% of the individuals with higher fitness values within each subpopulation are selected for breeding using a *stochastic universal sampling* function which has minimum spread and zero bias. By applying *discrete recombination crossover*, the new offsprings within each subpopulation are produced. In this segmentation method, offsprings are inserted into the appropriate subpopulations depending on *fitness-based reinsertion* with a rate of 0.9. In this multi-population GAs, migration of individuals between subpopulations is performed at every 20 generations with a migration rate of 0.2. After GA iterates for *maxgen* times (here *maxgen*=80), the evolution of this GA stops. The best individual with the maximum fitness value presents the optimized solution for the segments boundaries of the segmented signal.

3. RESULTS AND DISCUSSION

3.1 Redundancy Improvement

Fig. 1 shows the plots of the spectra E_{ii} of a windowed wheeze recording with 0%, 50%, and 80% overlapping ratios together with the FCF Γ of the AR(1) model. These plots illustrate the differences in conformity between E_{ii} and Γ . As seen in Fig. 1, the windowed signal with 50% overlapping window is the one with the closest conformity within the band of interest and thereby is chosen for *SampEn* calculation as well as the following GA-based phase segmentation.

With 0% overlapping, the energy spreading in spectral domain is not sufficient implying disjointness and information insufficiency in temporal domain. While with 80% overlapping ratio, the high decimation effect causes severe aliasing for which extra spectral peaks are warped back to distort the spectrum heavily.

Therefore, high redundancy in time has been introduced by applying highly overlapping window to the signal. This implies energy spreading towards higher frequency so that spectral overlapping is less likely to occur. Then the information redundancy in spectral domain becomes insufficient. However, increasing spectral redundancy suffers from the reduced temporal resolution. Thereby a compromise needs to be taken in order to balance between temporal and spectral redundancy by choosing appropriate overlapping ratio in windowing. Furthermore, compared to non-tonal signal (such as normal TS), a highly redundant tonal type signal like wheeze/stridor is less likely to induce significant frequency spreading. Therefore, time resolution should be favored by increasing the overlapping ratio for wheeze/stridor.

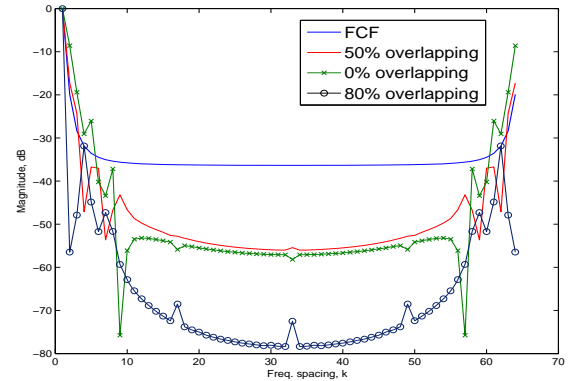


Figure 1: Frequency correlation function Γ for AR(1) model and spectra of the ensemble average of the frames E_{ii} of a wheeze recording with 0%, 80%, and 50% overlapping windows.

3.2 Performance Analysis

Table 1 presents the performances of the proposed segmentation method on both the preprocessed data from [10][11] and the real TS recordings. Mean and standard deviation ($\mu \pm \sigma$) for estimation error between actual respiratory phase locations identified by highly experienced doctors and the estimated phase locations by the proposed method are calculated. While the estimation error ε is defined as

$$\begin{cases} \varepsilon_{start} &= \frac{1}{JS} \sum_{js=1}^{JS} \left| \frac{P_{est}^{js} - P_{std}^{js}}{D_{std}^{js}} \right| \\ \varepsilon_{end} &= \frac{1}{JE} \sum_{je=1}^{JE} \left| \frac{P_{est}^{je} - P_{std}^{je}}{D_{std}^{je}} \right| \\ \varepsilon &= \frac{\varepsilon_{start} + \varepsilon_{end}}{2} \end{cases} \quad (9)$$

where ε is the percentage errors for starting/end positions of each respiratory phases respectively, with JS/JE representing the total number of starting/end points where $J = \frac{JS+JE}{2}$. Also, $P_{est}^{js/je}$ are the starting/end positions of the j^{th} respiratory phase identified by the proposed method and $P_{std}^{js/je}$ are those identified by the doctors. D_{std}^{js} is the same as D_{std}^{je} which is defined as the duration of the j^{th} respiratory phase.

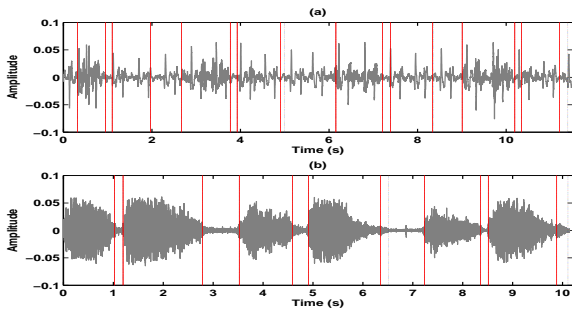


Figure 2: Illustrative plots of (a) real stridor recording and (b) preprocessed normal TS signal. Along with their respective actual respiratory phase locations identified by highly experienced doctors (dotted line), and the estimated phase locations by the proposed method (solid line).

Table 1: The estimation errors ($\mu \pm \sigma$)(%) of the proposed method on preprocessed signals and real recordings of different types of TS

Signal Type	Segmentation Error ($\mu \pm \sigma$)(%)	
	Preprocessed Signals	Real Recordings
Normal TS	1.78 ± 1.19	1.30 ± 1.21
Stridor	2.71 ± 3.33	3.29 ± 4.83
Wheeze	1.22 ± 1.90	1.76 ± 1.74
WNS	1.35 ± 1.18	3.46 ± 3.97

For performance evaluation on different types of TS signals, the error is calculated for each subject using (9) and then averaged over the subjects. The types of TS signals used include 9 normal TS (5 real recordings), 7 pure wheezes (5 real recordings), 6 pure stridor (4 real recordings), and 5 mixture of wheeze and stridor (WNS) (3 real recordings).

Fig. 2 illustrates the performance of the proposed segmentation method on the samples of the data used. The actual respiratory phase locations identified by highly experienced doctors are displayed together with the estimated phase locations by the proposed method. Despite the large diversity in their waveforms, the performance of the proposed method is comparable for the preprocessed data and noisy TS recordings, as indicated in Table 1. This shows that presence of the heart sound interference as well as other ambient noises do not affect the wide-sense stationarity of the appropriately windowed signals. The robustness of the proposed method has thus been verified under noisy conditions.

4. CONCLUSION AND FUTURE WORK

In this paper, an enhanced respiratory phase segmentation method for different types of TS is introduced based on genetic (GA) approach. The variability of the signals in terms of disjointness and information redundancy has been investigated thoroughly. The incorporation of disjointness measurement enhances the robustness of our GA method by modifying signals to be sufficiently redundant. By using *SampEn* and heterogeneity measure, the evaluation function of GA is designed. The segmentation results for various types of TS corrupted with heartbeats and ambient noise are found quite accurate, especially when the existing method only perform

well on the processed signals without these noise.

Furthermore, for initial estimation of total segment number, many approaches (e.g. using onset or other detection techniques) can be suitable for the proposed segmentation method. As the performance of the proposed method does not depend heavily on the accuracy of the total segment number estimated, only a rough estimation by using any detection technique is required. Moreover, the independency on threshold values makes the method robust and suitable for segmentation of real TS recorded under noisy conditions.

To our best knowledge, GA based segmentation method for real tracheal sound has not been published earlier. Thus the comparisons with a few existing ANN-GA or GA methods applied in other applications including flow rate measurement with the trained phase segments and lung sound identification, will be included in the future full-version of our paper.

REFERENCES

- [1] N. Meslier, G. Charbonneau, and J. L. Racineux, "Wheezes", *Eur. Respir. J.*, vol. 8, pp. 1942–1948, 2006.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] G. Sierra, V. Telfort, B. Popov, L. G. Durand, R. Agarwal, and V. Lanzo, "Monitoring Respiratory Rate Based on Tracheal Sounds. First Experiences", *Proc. 26th IEEE EMBS Conf.*, 2004.
- [6] M. Golabbakhsh, "Tracheal Breath Sound Relationship With Respiratory Flow: Modeling, the Effect of Age and Airflow Estimation", *M. Sc. Thesis, University of Manitoba*, 2004.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] S. Lehrer, "Understanding Lung Sounds", Philadelphia, Pa.: W.B. Saunders, 2002, Audio CD.
- [11] R. L. Wilkins, J. E. Hodgkin, and B. Lopez, "Fundamentals of Lung and Heart Sounds", Mosby, 2004, Audio CD.
- [12] D. A. Coley, "An Introduction to Genetic Algorithms for Scientists and Engineers", World Scientific, 2001.