HEART SOUND REDUCTION FROM LUNG SOUND RECORDINGS APPLYING SIGNAL AND IMAGE PROCESSING TECHNIQUES IN TIME-FREQUENCY DOMAIN

BY

MAHSA TALEBPOURAZAD

A Thesis
Submitted to the Faculty of Graduate Studies
In Partial Fulfillment of the Requirements for the
Degree of

MASTER OF SCIENCE

Department of Electrical and Computer Engineering
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

HEART SOUND REDUCTION FROM LUNG SOUND RECORDINGS APPLYING SIGNAL AND IMAGE PROCESSING TECHNIQUES IN TIME-FREQUENCY DOMAIN

BY

MAHSA TALEBPOURAZAD

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

MASTER OF SCIENCE

ABSTRACT

During lung sound recordings, an incessant noise source exists due to the heart beating that masks the clinical interpretation of lung sounds. The Heart Sound (HS) interference on lung sounds is significant especially at low flow rates. Several methods such as adaptive filtering, higher order statistics and wavelet denoising have been investigated to reduce HS interference on lung sound recordings. The performance of all the recent proposed methods is not satisfactory or relies on additional signal recordings and/or suffers from costly computational loads. In this study, two novel HS cancellation methods, using spectrogram Independent Component Analysis (ICA)-based technique and spectrogram filtering-based method in Time-Frequency (TF) domain along with three new techniques for HS localization in respiratory sound recordings are presented. To separate HS from lung sound, the spectrogram ICA-based method applies the ICA algorithm independently to every frequency on the spectrogram of two simultaneously lung sound recordings from two different locations on the chest and yields the independent components at that frequency. Then the proper independent components from each frequency are chosen and combined with each other to produce the spectrogram of separated signals. By implementing Inverse Short Time Fourier Transform (ISTFT), the separated signals are reconstructed in time domain. On the other hand, the spectrogram filtering-based method detects the HS-included segments in the spectrogram of a recorded lung sound signal using one of the proposed HS localization techniques. Afterwards, the algorithm removes those segments and estimates the missing data via a 2D interpolation in the TF domain. Finally, the signal is reconstructed into the time domain. The efficiency of the proposed methods for HS localization and HS

cancellation from lung sound recordings was examined quantitatively and qualitatively by visual and auditory means. The results show that the spectrogram ICA-based method is promising in term of HS reduction from lung sound recordings and the spectrogram filtering-based method successfully removes HS from lung sound signals, while preserving the original fundamental components of the lung sounds. The computational cost and the speed of both proposed methods were found to be much more efficient than other HS reduction methods.

ACKNOWLEDGMENT

I would like to thank my advisor, Dr. Zahra Moussavi, for her consistent guidance, stimulating suggestions, encouragement, commitment, enthusiasm and invaluable and faithful support throughout the last two years living and studying in Canada. Also, sincere thanks to Dr. Gabriel Thomas, for his helpful inputs and assistance; Dr. Hans Pasterkamp, for his valuable comments and his collaborations in data collection; University of Manitoba Biomedical Engineering group, for their time and participation in data recordings and Dr. Andreas Ziehe, for his time and cooperation. In addition, I am grateful to the Department of Electrical and Computer Engineering staff for their friendly supports. This study was sponsored by the Natural Sciences and Engineering Research Council (NSERC) of Canada; I want to thank them as well for their support. Finally, I wish to thank my parents, my brother and specially my husband, for their precious support, patient love and understanding in difficult times.

To the Dearest People in My Life

My Parents, My Husband and My Brother

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviations

ANC	Adaptive Noise Cancellation
AV	atrioventricular
BSS	Blind Source Separation
db4	Daubechies Mirror Filters order 4
ECG	electrocardiogram
FOS	Fourth-Order Statistics
HPF	High Pass Filter
HS	Heart Sounds
ICA	Independent Component Analysis
ISTFT	Inverse Short-Time Fourier Transform
JADE	Joint Approximate Diagonalization of Eigen-matrices
LMS	Least Mean Square
PCA	Principle Component Analysis
PSD	Power Spectral Density
RLS	Recursive Least Square
ROKF	Reduced Order Kalman Filtering
STFT	Short-Time Fourier Transform
SL	semilunar
TF	Time-Frequency

CHAPTER 1

INTRODUCTION

1.1 Motivation

Auscultation and acoustical analysis of lung sounds provide important and helpful information in the diagnosis and monitoring of the clinical course of patients with lung disorders such as pleurisy, pneumothorax, bronchial asthma, emphysema and pneumonia. Lung sound analysis has distinct advantages in patients who are unable to carry out pulmonary function tests properly, i.e., children. During auscultation of chest, lung sounds are not the only audible sounds and Heart Sounds (HS) are also perceptible. Thus, respiratory sounds encompass an intrusive quasi-periodic interference sound due to heart beating. This interference is more distinguishable at low flow rate (7.5 ml/s/kg). At higher flow rates, HS may be masked due to the increase of the lung sounds' average power. The main components of the HS are in the range of 20-100 Hz that overlaps with the low-frequency components of the breath sound [Yang-Sheng et al., 1988], which obscure the clinical interpretation of lung sound over those frequencies. Removing HS from respiratory sound recordings is a challenging task due to their overlap in the time and frequency domain. High Pass Filtering (HPF) with an arbitrary cut-off frequency between 70-100 Hz [Vannuccinni et al., 2000] cannot be efficient in this case because lung sound has major components in that region. Therefore, HS removal from lung sounds without altering the main characteristic features of the lung sound has been of interest for many researchers.

Adherence to the advances in computer technology and latest acoustical pattern recognition techniques, computer based acoustical analysis of respiratory sounds has become of great interest in the last three decades. During the last few years, several computerized methods for HS cancellation from lung sound recordings have been developed by researchers, but the performance of all those methods is not completely satisfactory or relies on additional signal recordings and/or suffers from costly computational loads. Hence, new methods need to be developed in this field of study.

1.2 Thesis Objective and Scope

The objective of this study was to develop new efficient methods in terms of computational speed and successful HS cancellation from lung sound without hampering the main components of the original signal using a single recording or two simultaneous recordings at low (7.5 ml/s/kg) and medium (15 ml/s/kg) flow rates. In order to evaluate the effectiveness of the proposed methods, the results were assessed using the same quantitative evaluation method that was employed in [Gnitecki et al., 2003; Hossain et al., 2003; Pourazad et al., 2003], i.e. comparing Power Spectral Density (PSD) values of the lung sound recordings before and after filtering with the HS-free segments of the original lung sounds. The hypothesis was that the average PSD of the filtered signal would fall between the average PSD of HS-free segments and the average PSD of the original signal including HS and that it would be closer to the former than to the latter.

1.3 Organization of the Thesis

Chapter 1 provided the motivation for this study and briefly discussed about the objective and scope of this research.

Chapter 2 overviews the respiratory system anatomy and the characteristics of the breath sounds and heart sound. Afterward this chapter gives an overlook to the previous studies that have been done on HS cancellation from respiratory sound recordings.

Chapter 3 presents the methodology used in this study. First, the spectrogram-ICA based method is described, including statement of the problem in frequency domain, the applied algorithm to solve the problem, overall data acquisition and signal pre-processing and the application of this method on the acquired data. Second, the spectrogram filtering-based method and its application are fully discussed along with proposing three techniques for HS localization. Third, the routines for evaluating the effectiveness of the proposed methods are explained.

Chapter 4 elaborates on the results obtained from performing the proposed algorithms on the acquired data. Furthermore the performance of these algorithms is evaluated quantitatively and qualitatively.

Chapter 5 concludes on this study and presents recommendations for future research in this area.

CHAPTER 2

BACKGROUND

2.1 Respiratory System Anatomy

In breathing procedure, the respiratory system supplies oxygen to the blood and expels waste gases (CO2) from the body. The upper structures of the respiratory system are combined with the sensory organs of smell and taste (in the nasal cavity and the mouth) and the digestive system (from the oral cavity to the pharynx). The larynx, or voicebox, is located at the head of the trachea. The trachea or the windpipe is pipe shape with rings of cartilage. The trachea divides into two branches called bronchi, which carry air into each side of the lung. In humans, the lung is not symmetric because heart is located at the center of the chest (thorax) and leaned slightly to the left. Thus the right side of the lung has three lobes (sections) and the left side has two lobes (Fig. 2.1).

The bronchi branch off at the tracheal bifurcation to enter the hilus of the left or right lung. Inside the lung, each bronchi divides into narrower branches called bronchioles, which carry air to the functional unit of the lung, i.e., the alveoli. Capillaries are small blood vessels that are wrapped around the alveolar chambers. There, in the thousands of tiny alveolar chambers, oxygen is transferred through the membrane of the alveolar walls to the blood cells within the capillaries. The walls are so thin and close to each other that the air easily seeps through. Likewise, waste gases diffuse out of the blood cells into the air in the alveoli, to be expelled upon. Fig. 2.1 demonstrates the anatomy of the respiratory system.

2.2 Auscultation over the Chest

Lung sound auscultation using stethoscopes is a primary diagnostic tool for physicians. Regardless of the high cost of many modern stethoscopes, these instruments provide an imperfect and subjective perception of the respiratory sounds [Pasterkamp et al., 1997]. They selectively amplify or attenuate sounds within the spectrum of clinical interest. The amplification tends to take place below 112 Hz and the attenuation occurs at higher frequencies [Abella et al., 1992]. This feature is inherent in the design of the stethoscopes. During auscultation of chest, lung sound is not the only audible sound, while heart sound is also perceptible. Since the main components of the heart sound are in the range of 20-100 Hz, the amplification of lung sounds over low frequencies will also result in amplifying the HS. Therefore the clinical interpretation of lung sound over low frequencies is obscured even if modern and expensive stethoscopes are employed for auscultation.

Recording respiratory sounds by sound transducers provide almost an ideal way for analyzing the lung sounds. The electret microphones with coupling chamber and the contact accelerometers are the two common types of transducers for lung sound recordings and research applications [Pasterkamp et al., 1993]. Electret microphones with a sealed chamber coupled to the skin is a sensitive lung sound transducer. It has been found that the size and the shape of coupling chambers affect the overall frequency response of this type of transducers. For example the smaller, conically shaped chambers are more sensitive to higher lung sound frequencies [Kraman et al., 1995; Wodicka et al., 1994], and also highly susceptible to the ambient noise. Contact accelerometers are popular in lung sound studies; however they are typically more expensive and fragile

compared to the electret microphones, but have a much wider frequency response than microphones.

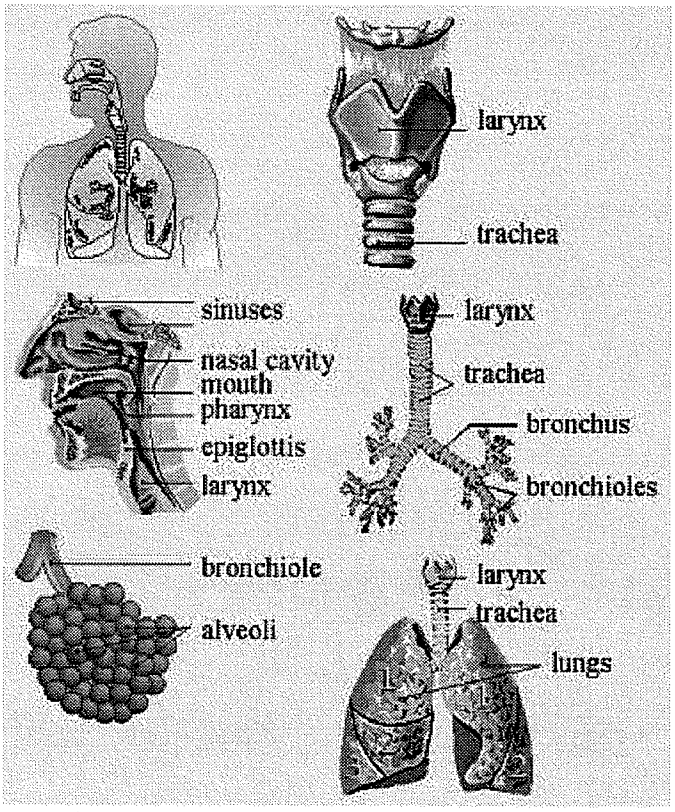


Figure 2.1 The anatomy of respiratory system

Computer based acoustical analysis of respiratory sounds has become of great interest in the last three decades. Advances in computer technology and latest acoustical pattern recognition techniques are likely to yield a clearer understanding of the lung sound acoustic basis. During the last two decades, numerous computerized techniques for detection and quantification of lung sounds with several clinical applications have been developed by researchers. The applications of lung sound technology are proven to be clinically useful since lung sounds provide helpful information in the diagnosis and monitoring the clinical course of patients with respiratory diseases. Lung sound analysis

has also distinct advantages in patients who are unable to carry out pulmonary function tests properly, i.e., children.

2.2.1 Breath Sounds Characteristics

Breath sounds are generated by the turbulent of airflow through the respiratory tree and characterized by their frequency, intensity, qualities and duration (of the inspiratory and expiratory phases). The quality of breath sounds is modified as they are filtered through the respiratory tree. Breath Sounds can be heard and recorded over the neck (tracheal sound) and over the chest wall either in front or back (lung sound). Lung sounds are of more diagnostic value than tracheal sounds; hence most of respiratory researches are focused on lung sounds analysis.

The breathing sound that one hears over the chest of a healthy person is called the normal lung sound [Pasterkamp et al., 1997]. Over the chest, breathing sounds appear more muffled, since they lose much of the higher frequency components on the passage through lung and chest wall. It is likely that these tissues behave like a low-pass filter. The peak of lung sound is in frequencies below 100 Hz. The lung sound energy decreases rapidly between 100-200 Hz [Gavriely et al., 1995] but it is still detectable at or above 800 Hz with sensitive sensors. Lung sounds are much louder during inspiration than expiration. Lung sound's amplitude differs between persons and locations on the chest surface. Body size also affects the respiratory sounds. In children, the transmission of respiratory sound through smaller lungs and thinner chest walls imposes a distinct quality to the lung sound [Pasterkamp et al., 1997]. Acoustical measurements have shown that

normal lung sounds in infants include higher median frequencies of that in older children and adults [Hidalgo et al., 1991; Kanga et al., 1986].

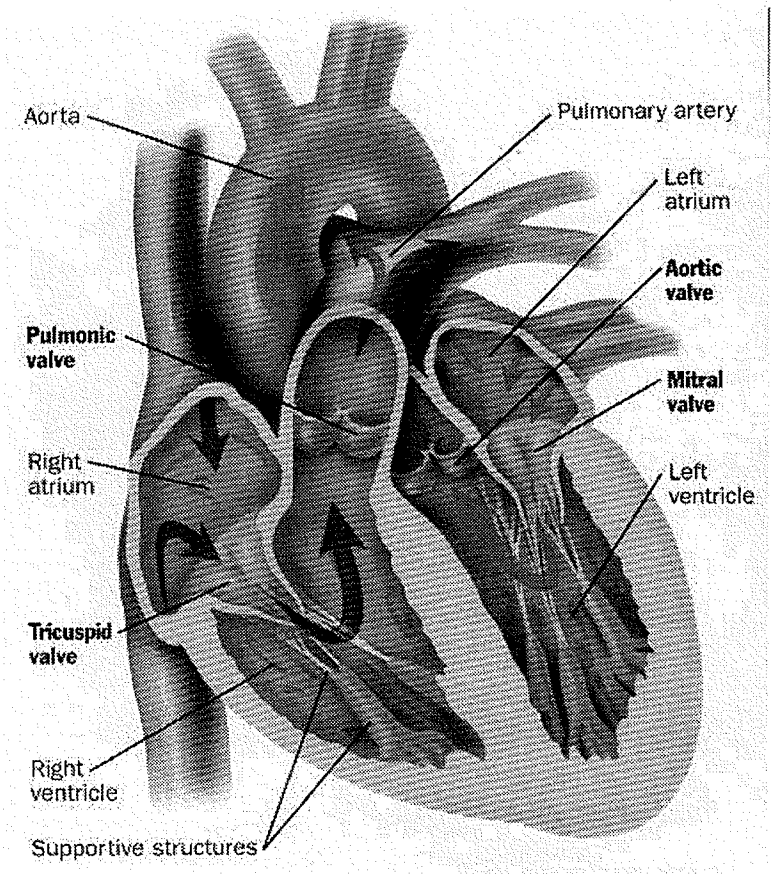


Figure 2.2 The chambers and valves of heart organ
(From http://www.ramhospital.com/valvesoftheheart.jpg)

2.2.2 Heart Sounds Characteristics

During auscultation of chest, lung is not the only organ that generates an audible sound since the heart organ is beating as well. Heart sounds are vibrations or sounds that are caused by the acceleration or deceleration of blood [Webster, 1998]. In fact, the closure of heart valves causes the blood to rebound against the ventricular walls or blood vessels, and due to the produced vibration, the HS is generated. Fig. 2.2 depicts the interior valves and chambers of heart organ.

Heart sounds are extremely small amplitude signals with the frequency range of 0.1 to 2000 Hz. There are four major HS, but only the first two can be heard without the use of special amplification:

- *First heart sound*: The first HS is produced at the beginning of ventricular systole [Rushmer, 1970], when the atrioventricular (AV; the mitral and tricuspid) valves are closed and the semilunar (SL; the aortic and pulmonary) valves are opened. This sound has a low-frequency tone commonly termed the "lub" sound of the heartbeat.
- *Second heart sound*: The second heart sound is a low-frequency vibration which occurs during the end of the systole and is produced by the closure of the SL valves, the opening of the AV valves, and the resulting deceleration and reversal of flow in the arteries and ventricles. Due to the higher blood pressures in the arteries, the second HS has higher frequencies relative to the first heart sound. It is commonly referred to as the "dub" sound. The second HS is associated with the completion of the T wave of the ECG [Webster, 1998].
- *Third heart sound*: The third sound occurs during the rapid filling of the ventricles after the AV valves are opened and is probably produced by vibrations of the ventricular walls. This low-amplitude, low frequency vibration is perceptible in children and some adults [Webster, 1998].
- *Fourth heart sound*: The forth or atrial heart sound -which is not audible- occurs at the time of atrial contraction and is probably due to the accelerated rush of blood into the ventricles [Webster, 1998].

Fig. 2.3 illustrates relationship among the four heart sounds and the electrical and mechanical events of cardiac cycle. The last 30% of the time span of the R-R intervals may be considered void of HS [Pasterkamp et al., 1985]. Note that the percent of R-R interval is chosen conservatively at 30%. Although figure 2.3 shows that the last two HS may occur within the last 30% of the R-R intervals, the third HS does not appear in all subjects as mentioned before and the forth HS is not audible [Webster, 1998].

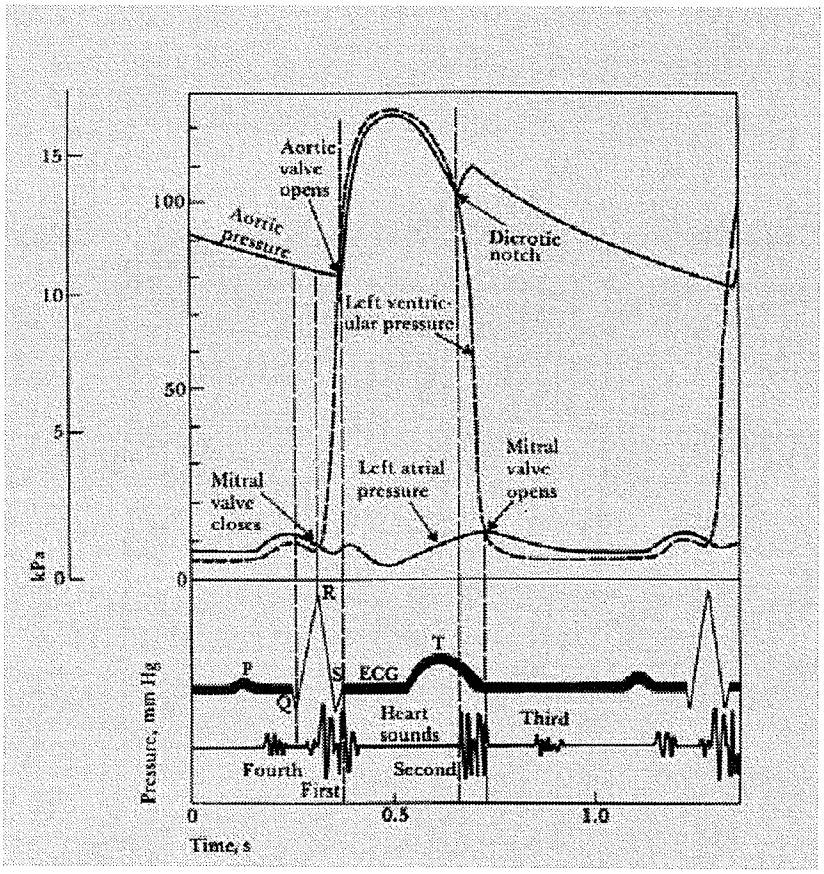


Figure 2.3 Relation among four heart sounds and the electrical and mechanical events of cardiac cycle (Adopted from Webster, J.G., *Medical Instrumentation: Application and Design.* New York, NY: John Wiley & Sons, Inc., ch. 7, 1998).

Heart sounds travel through the body. The acoustical characteristics of their transmission path from the heart and major blood vessels to the skin surface do not let

them to be reflected but attenuated. HS has largest attenuation in the most compressible tissues like the lungs and fat layers [Webster, 1998]. Fig. 2.4 shows the four main locations over the chest, at which the intensity of sound from the four valves is highest because the sound is being transmitted through solid tissues or inflated lung tissues with a minimal thickness.

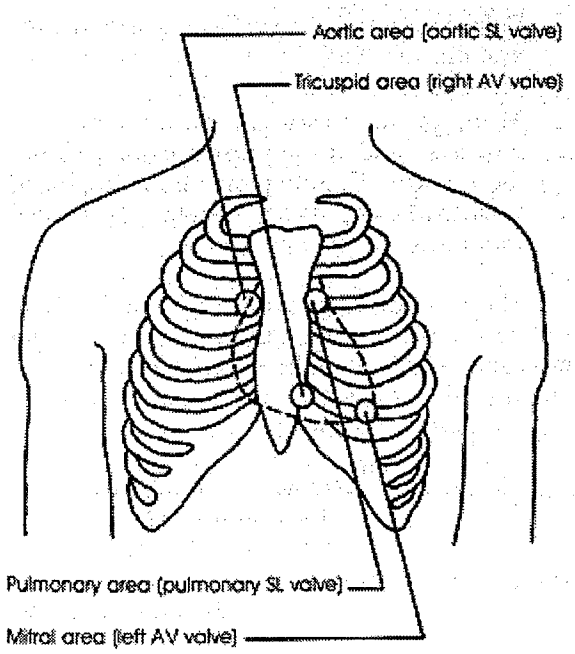


Figure 2.4 Heart auscultatory areas on the chest

2.3 Heart Sound Reduction Methods

During breathing, lung sound propagates through the lung tissue and can be heard over the chest wall. The tissue acts as a frequency filter with special characteristics based on pathological changes [Schuttler et al., 1996]. Therefore, auscultation and acoustical analysis of lung sounds are primary diagnostic assessments for respiratory diseases. However, heart beating produces an intrusive quasi-periodic interference sound that masks the clinical interpretation of lung sounds over the low frequency components. The

main components of the HS are in the range of 20-100 Hz, in which lung sound has major components as well [Yang-Sheng et al., 1988]. Therefore, HS reduction from lung sounds without altering the main characteristic features of the lung sound has been of interest for many researchers.

High Pass Filtering (HPF) with an arbitrary cut-off frequency between 70-100Hz [Vannuccinni et al., 2000] cannot be efficient in this case because lung sounds' major components are in that region as well. One group of researchers designed a signal pre-processing system with varying amplifier gain using an adaptive filter [Yang-Sheng et al., 1988]. Another group developed a portable breath sound analysis system that used an adaptive filter based on the Least Mean Square (LMS) algorithm for removing HS interferences [Guangbin et al., 1992]. Both of the aforementioned groups used HS signals recorded on the patient's heart location as the reference input for the adaptive filtering, which inevitably included lung sounds as well. Two other groups [Iyer et al., 1986; Yip et al., 2001] used electrocardiogram (ECG) signal information as the reference signal for adaptive filtering to reduce heart sounds. In both of these studies, adaptive filtering based on the LMS algorithm with hundreds of taps (i.e. 1000 and 300, respectively) was applied which resulted in a high adaptation time. Another group of researchers proposed the use of Reduced Order Kalman Filtering (ROKF) for HS reduction [Charleston et al., 1996]. In order to facilitate the estimation of respiratory sounds using ROKF, an autoregressive model was fitted to the HS segments free of respiratory sounds (i.e., breath hold segments including HS) based on three assumptions: *i*) heart and respiratory sounds are mutually uncorrelated, *ii*) these sounds have additive interaction and *iii*) the prior and subsequent heart sounds are linearly related to the heart sounds corrupted by the respiratory signal.

These prior assumptions make the method inefficient in practical implementations. The performance of their proposed method was tested only with synthesized data and not actual lung sounds. This ROKF approach was also computationally costly.

In order to eliminate the need for a reference signal when using adaptive filtering, some researchers proposed a single recording technique based on the modified version of the adaptive LMS algorithm by adding a low pass filter with a cut-off frequency of 250 Hz in the error signal path to the filter [Kompis et al., 1992]. However, their results showed that heart sounds were still clearly audible due to the improper identification of the HS segments within a lung sound recording. Later, another single recording Adaptive Noise Cancellation (ANC) technique based on Fourth-Order Statistics (FOS) was examined [Hadjileontiadis et al., 1997]. In the proposed technique a specialized algorithm for HS localization, based on band pass filtering and thresholding was developed. The performance of ANC-FOS in HS reduction was evaluated respect to the HPF technique (as mentioned before, an inappropriate method for HS reduction in lung sound recordings) and two other reported techniques in [Kompis et al., 1992; Iyer et al., 1986]. In fact, only HPF was employed and the basis of the rest of the comparisons was the reported quantitative analysis in [Kompis et al., 1992; Iyer et al., 1986] using different sensors and processing methods. Furthermore the method was computationally costly.

Recently, a group of researchers applied a band pass filtered version of the original signal as the reference input for the Recursive Least Squares (RLS) adaptive noise cancellation filtering technique [Gnitecki et al., 2003]. To gauge the effectiveness of this RLS-ANC filtering scheme, an absolutely heart-sound-free lung sound reference was extracted from lung sound recordings based on the recorded ECG signal (using the last

30% of R-R time interval). Then the PSD values of pre-filtered lung sound segments in areas including and void of HS were compared with the PSD values of RLS-ANC filtered recordings over different frequency bands. Although the performance of their proposed method was promising qualitatively and quantitatively, its computational load was high.

Wavelet based filtering techniques have been also applied for HS reduction from lung sounds. In one of the reported studies the performance of the proposed method was evaluated with respect to the performance of HPF technique (an inappropriate standard) [Hadjileontiadis et al., 1998]. Moreover the resulted signals were only compared to the HS-included signals and not to the ones free of HS. Another group of researchers applied the same procedure as [Gnitecki et al., 2003] to gauge the effectiveness of the proposed wavelet based filtering technique quantitatively [Hossain et al., 2003]. According to the reported results, this technique was not successful in removing HS from lung sound recordings completely.

The performance of all the mentioned studies relies on additional signal recordings and/or suffers from costly computational loads. Also in most of these studies, the reported results have not been compared with HS-free sections of lung sounds extracted in a strict quantitative manner. Hence, the objective of this study was to develop and investigate new efficient methods in terms of computational speed and successful HS cancellation from lung sounds without hampering the main components of the lung sound signal. In order to evaluate the effectiveness of the proposed methods in this study, the results were assessed using the same quantitative evaluation technique employed in [Gnitecki et al., 2003; Hossain et al., 2003; Pourazad et al., 2003], i.e. comparing average

PSD values of lung sound signals before and after filtering with those of original lung sound free of HS.

2.4 Chapter Summary

Although several techniques have been developed for reducing HS from respiratory sound recordings, new methods still need to be developed in this field of research, since the performance of all the current techniques is not satisfactory. This Chapter has provided a solid background on respiratory system and the characteristics of heart and lung sounds along with a brief report on the previous studies in the HS removal from lung sound recordings. The following Chapter elaborates on the methodology of the two new methods for HS cancellation from lung sound recordings.

CHAPTER 3

METHODOLOGY

3.1 Introduction

In Chapter 2, the fundamentals of respiratory system and respiratory sound analysis were discussed. Furthermore the characteristics of HS and lung sounds were investigated. The following sections elaborate on the data acquisition details and two novel methods for HS removal from lung sound recordings using spectrogram Independent Component Analysis (ICA)-based technique and spectrogram filtering-based method in Time-Frequency (TF) domain along with three new techniques for HS localization in respiratory sound recordings.

3.2 Spectrogram ICA-Based Method

3.2.1 Introduction

Recovering independent sources only from the sensor outputs, which are linear mixtures of independent source signals, i.e., Blind Source Separation (BSS), has been of great interest for researchers in the last two decades. The term "blind" refers to the fact that both the source signals and the way that they have been mixed, are unknown. ICA is a technique used for solving BSS problems. This technique finds a linear coordinate system (the unmixing system) such that the recovered signals are statistically independent [Lee, 2000].

Multichannel blind deconvolution refers to the BSS problem of reconstructing unknown source signals, which are not simultaneously mixed but are convolved due to

transmission through several systems with unknown transfer functions. In other words, only mixed and distorted (convolutive or filtered) versions of the source signals are available at the sensor outputs. HS cancellation from lung sound recordings can be considered as a multichannel blind deconvolution problem. In this case lung sound and HS signals are considered as independent source signals. However, due to the delays and reflections of the lung tissues, the recorded mixed signals on the skin are correlated and are not instantaneous mixtures but convolutive mixtures.

One alternative solution to multichannel blind deconvolution problems is performing ICA on the spectrogram of the signal [Lee, 2000]. This technique has been previously applied for separating speech signals [Murata et al., 2001]. The following subsections describe the application of this technique for HS separation from lung sounds using two simultaneous breath sound recordings from two different locations over the chest.

3.2.2 Formulating the Problem

As mentioned before, HS cancellation from lung sound recordings can be treated as a multichannel blind deconvolution problem. Assuming n sources that produce the breath and HS, a respiratory sound record over the chest can be denoted as a source matrix $S(t)$ consists of n source signals as:

$$S(t)=[s_1(t),...,s_n(t)]^T, \tag{3.1}$$

where $s_i(t)$, $i=1,...,n$ are assumed to be independent of each other, i.e., the joint probability density function of source signals is factorized by their probability density function:

$$p\{s_1(t),...,s_n(n)\}=p\{s_1(t)\}\times...\times p\{s_n(t)\}, \tag{3.2}$$

Without loss of generality, it is assumed that $S(t)$ is zero mean (a full description on centering non-zero mean signals has been provided in Section A.1 of the Appendix). The n source signals are assumed to be delayed, filtered and mixed, while they are transmitted through the medium (i.e., lung organ) before being picked up by an array of n sensors (i.e., accelerometers) on the skin. These simultaneously recorded signals can be represented by a matrix called observation matrix:

$$X(t) = [x_1(t), ..., x_n(t)]^T, \tag{3.3}$$

where $x_i(t)$, is the recorded signal at sensor i, which can be expressed as:

$$x_i(t) = \sum_{j=1}^{n} \left(\sum_{\tau} a_{ij}(\tau) s_j(t-\tau) \right) = \sum_{j=1}^{n} a_{ij}(t) * s_j(t), \tag{3.4}$$

where $*$ denotes convolution and a_{ij} represents the transfer function of the transmission path from the source j to the sensor i. Eq. (3.4) can be written in matrix format as:

$$X(t) = A(t) * S(t), \tag{3.5}$$

where, $A(t) = \begin{bmatrix} a_{11}(t) & . & . & . & a_{1n}(t) \\ . & . & & & . \\ . & & . & & . \\ . & & & . & . \\ a_{n1}(t) & . & . & . & a_{nn}(t) \end{bmatrix}$.

$A(t)$ can be considered as a filter matrix that acts upon the sensors. Assuming that the number of sensors is equal to the number of sources, the filter $A(t)$ will be a full rank matrix and respectively invertible. The case in which the number of sources is more than the number of sensors, is still of theoretical and practical interest for researchers. As the preliminary results in [Lewicki et al., 1998] suggest, an overcomplete representation of data can to some extent, extract the independent components using a prior knowledge of

the source distribution. On the other hand, in the cases which the number of sources is less than the number of sensors, Principle Component Analysis (PCA) should be applied as a preprocessing stage to reduce dimensions and prevent overlearning [Hyvarinen et al., 2001].

The objective of BSS is to determine a filter, $B(t)$, from observation matrix $X(t)$ without knowing the filter $A(t)$ and the probability density function of the source signals, such that the estimated source signals are being mutually independent. This can be formulated as:

$$Y(t) = B(t) * X(t), \tag{3.6}$$

where $Y(t)$ is the estimated source matrix. Comparing Eq. 3.6 with Eq. 3.5 shows that if $B(t)$ is a perfect inverse filter of $A(t)$, then $Y(t) = S(t)$. However, due to the lack of information about the amplitude and the order of source signals, there will be ambiguities about scaling factors and permutation [Hyvarinen et al., 2001]. In fact, the optimum $B(t)$ satisfies the following relationship:

$$Y(t) = B(t) * (A(t) * S(t)) = PDS(t), \tag{3.7}$$

where P is a permutation matrix, and D is a diagonal matrix representing scaling factors. In the permutation matrix all the elements of each row are equal to zero except for one element with value 1, where the column number corresponds to the source order number.

In order to avoid the scaling indefiniteness, instead of estimating source signals, the problem could be moderated to find a decomposition as [Murata et al., 2001]:

$$X(t) = V_1(t) + V_2(t) + ... + V_n(t), \tag{3.8}$$

such that $V_i(t)$s are mutually independent. In fact, each $V_i(t)$ is a signal originated from the i^{th} independent source. By the use of filter $B(t)$ and its inverse, the observation matrix can be then decomposed as:

$$\begin{aligned} X(t) &= B(t)^{-1} * B(t) * X(t) \\ &= B(t)^{-1} * IB(t) * X(t) \\ &= B(t)^{-1} * (E_1 + ... + E_n) B(t) * X(t) \\ &= B(t)^{-1} * E_1 B(t) * X(t) + ... + B(t)^{-1} * E_n B(t) * X(t). \end{aligned} \tag{3.9}$$

where E_i is a diagonal matrix with 1 for i^{th} diagonal element and zero for all the other elements [Murata et al., 2001]. Also the sum of E_i matrices, $i = 1,...,n$, should be equal to the identity matrix I. Consequently comparing the results of Eq. 3.9 with Eq. 3.8 yields the desired decomposition $V_i(t)$ as:

$$V_i(t) = B(t)^{-1} * (E_i B(t)) * X(t), \tag{3.10}$$

As can be observed from Eq. (3.10), rescaling of $B(t)$ does not affect the representation of $V_i(t)$ [Murata et al., 2001]. This can be proven by rescaling the observation matrix $B(t)$ by an arbitrary non-singular diagonal matrix D and substitution in Eq. (3.10) as:

$$\begin{aligned} V_i(t)|_{B(t) \to D.B(t)} &= (DB(t))^{-1} * (E_i (DB(t))) * X(t) \\ &= B(t)^{-1} * (D^{-1} E_i DB(t)) * X(t) \\ &= B(t)^{-1} * (E_i B(t)) * X(t) \\ &= V_i(t). \end{aligned}$$

In order to solve the multichannel blind deconvolution problem stated by Eq. (3.5) the spectrogram ICA-based approach proposed in [Murata et al., 2001] was implemented. This technique performs ICA in TF domain on the spectrogram of the recorded signals.

3.2.3 Spectrogram Estimation

The recorded signals can be represented in TF domain by using the Short-Time Fourier Transform (STFT). The discrete STFT can be expressed as:

$$\hat{f}(\omega,t_s)=\sum_t e^{-j\omega t} f(t)w(t-t_s),$$
$$\omega = 0,\frac{1}{N}2\pi,...,\frac{N-1}{N}2\pi, \qquad t_s = 0,\Delta T,2\Delta T,... \tag{3.11}$$

where ω represents the frequency, N denotes the number of uniformly spaced samples in the discrete-time Fourier transform, $w(t)$ is a window function such as Hanning, Hamming, etc., t_s indicates the window position and ΔT specifies the shifting time [Nawab et al., 1988]. This type of representation is commonly referred as spectrogram.

The Inverse STFT (ISTFT) is defined by:

$$f(t)=\frac{1}{2\pi}\frac{1}{W(t)}\sum_{t_s}\sum_{\omega} e^{j\omega(t-t_s)}\hat{f}(\omega,t_s), \tag{3.12}$$

where

$$W(t)=\sum_{t_s} w(t-t_s). \tag{3.13}$$

Performing ISTFT on the spectrogram of a signal can reconstruct the signal in time domain.

3.2.4 Time-Frequency Representation

The relationship between observation and source matrices can be represented in TF domain by employing the Fourier transform properties. In this regard Eq. (3.5) can be rewritten as:

$$\hat{X}(\omega,t_s)=\hat{A}(\omega)\hat{S}(\omega,t_s), \tag{3.14}$$

where $\hat{A}(\omega)$ is the Fourier transform of $A(t)$ and $\hat{S}(\omega,t_s)$ is the STFT of source matrix $S(t)$. As can be observed from Eq. (3.14), for a particular frequency ω, the $\hat{X}(\omega,t_s)$ is defined as a non-convolutive or instantaneous mixture of complex-valued time series $\hat{S}(\omega,t_s)$. Consequently, at the frequency ω, the convolutive mixture problem can be represented as non-convolutive mixture problems to solve. Among the several approaches of BSS in the literature, eliminating the cross-correlation of reconstructed signals was selected for solving these non-convolutive mixture subproblems.

3.2.5 Eliminating Cross Correlation

Eliminating the cross-correlation approach is based on the second order statistics, which makes the estimation robust and can be performed through a non-iterative algorithm [Murata et al., 2001]. This algorithm consists of two stages: sphering and rotation [Ziehe et al., 2000; Ziehe, 1998]. Fig. 3.1 graphically explains the stages of this algorithm.

Recorded signals are generally distributed over a non-orthogonal direction in the coordinates. Sphering (pre-whitening) is an operation that rearranges these directions such that the recorded signals become orthogonal to each other in a new coordinate system (full explanation is provided in Section A.2 of the Appendix). However, there still remains the ambiguity of rotation. The correct rotation is determined by implementing Joint Approximate Diagonalization of Eigen-matrices (JADE) algorithm on the spectrogram of the recorded signals at each frequency, ω. JADE is referred as one of the algorithms for ICA on complex valued data [Cardoso, 1999; Cardoso et al., 1993] and is an appropriate candidate for our case since breath sound spectrograms are complex

valued. The JADE algorithm employs Jacobi rotation for finding an optimal rotation between orthogonal matrices. Therefore the resulted matrix is the closest orthogonal compromise between all cumulant solutions for one frequency, ω (for more details see [Cardoso et al., 1993]).

Implementing the algorithm of eliminating the cross-correlation independently on each frequency will yield the independent components at that frequency. For reconstructing the independent source signals in time domain, proper independent components from each frequency should be chosen and combined with each other, i.e. the un-permutation procedure should be performed before reconstructing the independent source signals.

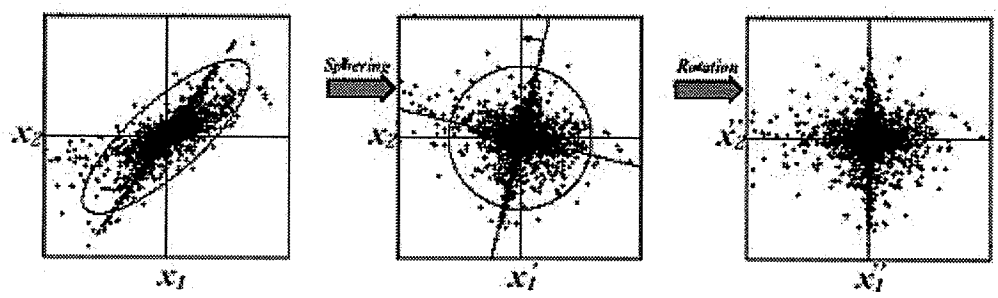


Figure 3.1 The stages of applied blind source separation technique (Adapted from [Murata et al., 2001])

3.2.6 Un-Permutation

As mentioned in the previous section, proper independent components at each frequency should be chosen and combined with each other to construct the spectrogram of each separated signal. Finding such an appropriate combination is only possible for non-stationary signals [Murata et al., 2001], which is the case in this study. This is due to the correlation that exists between the components of a non-stationary signal at different

frequencies [Murata et al., 2001]. Let's write the i^{th} source signal in exponential format, i.e., with magnitude and phase:

$$\hat{s}_i(\omega,t_s)=a_i(\omega,t_s)e^{j\phi_i(\omega,t_s)}. \qquad (3.15)$$

The non-stationarity of the signal causes its amplitude $a_i(\omega,t_s)$ to change in time corresponding to the $\hat{s}_i(\omega,t_s)$ envelope. Since at frequency ω, the i^{th} and j^{th} source signals are independent, their magnitudes are uncorrelated:

$$corr\left(a_i(\omega,t_s),a_j(\omega,t_s)\right)=\frac{1}{T}\sum_{s=1}^{T}a_i(\omega,t_s)a_j(\omega,t_s)-\frac{1}{T}\sum_{s=1}^{T}a_i(\omega,t_s)\frac{1}{T}\sum_{s=1}^{T}a_j(\omega,t_s)=0, \qquad i\neq j$$

if T is sufficiently large.

Similarly, different frequency components of different source signals are also uncorrelated. Therefore, measuring the correlation coefficient of different source signal envelopes at different frequencies,

$$r\left(a_i(\omega,t_s),a_j(\omega',t_s)\right)=\frac{corr\left(a_i(\omega,t_s),a_j(\omega',t_s)\right)}{\sqrt{corr\left(a_i(\omega,t_s),a_i(\omega,t_s)\right)corr\left(a_j(\omega',t_s),a_j(\omega',t_s)\right)}}, \qquad (3.16)$$

would be helpful in finding proper combination of frequency components [Murata et al., 2001].

3.2.7 Summary of the Algorithm

The following is a summary of spectrogram ICA-based algorithm:

Spectrogram Estimation: The STFT of each signal is calculated, considering a suitable window size. The size of the window should not exceed the duration of the signal's stationarity. It was considered as 1024 (100 ms) in this study.

BSS at each frequency, ω: The explained algorithm in Section 3.2.5 is applied to the estimated spectrogram $\hat{X}(\omega,t_s)$ at each frequency, and consequently the demixing matrices $\hat{B}(\omega)$ are estimated, which give the independent components:

$$\hat{U}(\omega,t_s)=\hat{B}(\omega)\hat{X}(\omega,t_s), \tag{3.17}$$

Decomposition of Spectrograms: The spectrograms are decomposed by:

$$\hat{V}(\omega,t_s;i)=\hat{B}(\omega)^{-1}E_i\hat{B}(\omega)\hat{X}(\omega,t_s)=\hat{B}(\omega)^{-1}\begin{bmatrix}0\\ \vdots\\ \hat{u}_i(\omega,t_s)\\ \vdots\\ 0\end{bmatrix}, \tag{3.18}$$

where $\hat{u}_i(\omega,t_s)$ is the i^{th} independent component of $\hat{U}(\omega,t_s)$ at frequency ω.

Clustering Frequency Components: Based on the proposed algorithm in [Murata et al., 2001], first a moving average operator $\Im$ should be defined for estimating the time series envelope by:

$$\Im\left(\hat{V}(\omega,t_s;i)\right)=\frac{1}{2M+1}\sum_{t_s'=t_s-M}^{t_s+M}\sum_{j=1}^{n}\left|\hat{v}_j(\omega,t_s';i)\right|, \tag{3.19}$$

where M is a positive constant and $\hat{v}_j(\omega,t_s';i)$ is the j^{th} component of $\hat{V}(\omega,t_s;i)$. Then the permutation problem is solved by the following un-permutation procedure:

1. The similarity function is defined by:

$$sim(\omega)=\sum_{i\neq j}r\left(\Im\left(\hat{V}(\omega,t_s;i)\right),\Im\left(\hat{V}(\omega,t_s;j)\right)\right), \tag{3.20}$$

and is sorted in increasing order:

$$sim(\omega_1)\leq sim(\omega_2)\leq\ldots\leq sim(\omega_N). \tag{3.21}$$

This corresponds to sorting ωs in order of low correlation between ωs' independent components. N denotes the number of uniformly spaced samples in the discrete time Fourier transform.

2. For ω_1, $\hat{V}(\omega_1,t_s;i)$is assigned to $\hat{Y}(\omega_1,t_s;i)$:

$$\hat{Y}(\omega_1,t_s;i)=\hat{V}(\omega_1,t_s;i), \qquad i=1,...,n. \tag{3.22}$$

3. For ω_k, a permutation $\sigma(i)$is found such that correlation between the envelope of ω_k and the aggregated envelope from $\omega_1$to ω_{k-1} is maximized. This is accomplished by maximizing the summation of correlation coefficients,

$$\sum_{i=1}^{n} r\left(\Im\left(\hat{V}(\omega_k,t_s,\sigma(i))\right),\sum_{j=1}^{k-1}\Im\left(\hat{Y}(\omega_j,t_s;i)\right)\right), \tag{3.23}$$

within all possible permutations of $i=1,...,n$.

4. The appropriate permutation is assigned to the $\hat{Y}(\omega_k,t_s;i)$:

$$\hat{Y}(\omega_k,t_s;i)=\hat{V}(\omega_k,t_s;\sigma(i)), \qquad i=1,...,n. \tag{3.24}$$

The procedure is repeated from stage 3 till k becomes equal to N.

Restoration of Signals: The separated spectrograms $\hat{Y}(\omega,t_s;i)$ are reconstructed in time domain by applying ISTFT, which has been described in Section 3.2.3:

$$\begin{aligned} &Y(t;i)=\frac{1}{2\pi}\frac{1}{W(t)}\sum_{t_s}\sum_{\omega}e^{j\omega(t-t_s)}\hat{Y}(\omega,t_s;i), \\ &\sum_{i=1}^{n}Y(t;i)=X(t). \end{aligned} \tag{3.25}$$

and finally a set of separated signals are obtained. Fig. 3.2 graphically depicts the applied blind source separation algorithm.

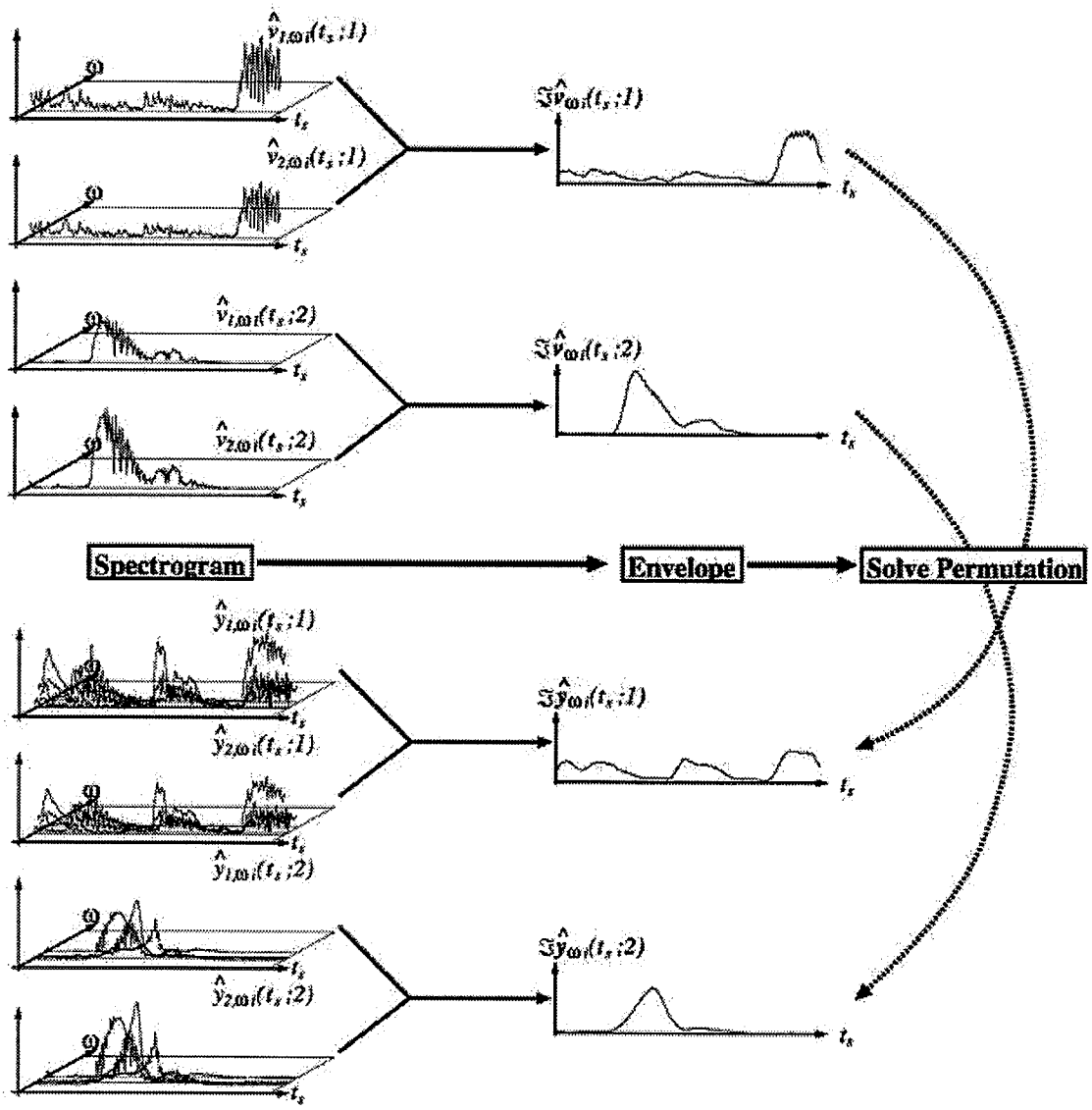


Figure 3.2 The algorithm of applied blind source separation technique (Adapted from [Murata et al., 2001])

3.2.8 Experimental Procedure for Spectrogram ICA-Based Method

3.2.8.1 Data Acquisition

The breath sound recordings were conducted at the Biomedical & Signal Processing Laboratory, Electrical & Computer Engineering Department, University of Manitoba. Two healthy volunteer subjects (one female) aged 26 and 30 years participated in this

study. Two piezoelectric contact accelerometers (Siemens EMT 25C) were used to record the respiratory sounds from the subjects in sitting position. For each recording the accelerometers were secured with double-sided adhesive tape rings on two locations of the followings over the subjects' chest:

1. left midclavicular area, 2nd intercostals space (L1)
2. left midclavicular area, 4th intercostals space (L2)
3. center of chest (C)
4. right midclavicular area, 2nd intercostals space (R1)
5. right midclavicular area, 4th intercostals space (R2)

Fig. 3.3 shows the accelerometer placement on the chest.

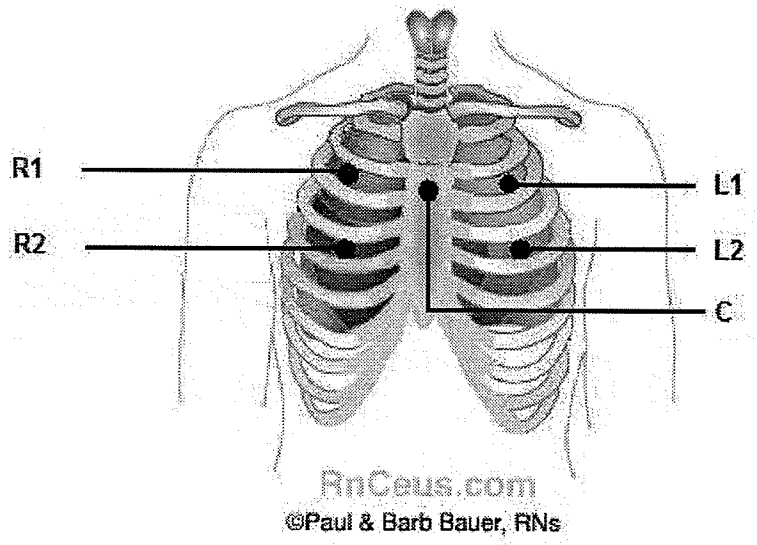


Figure 3.3 Accelerometer placements on the anterior chest. (Legend: **L1**= left midclavicular area, 2nd intercostals space; **L2**= left midclavicular area, 4th intercostals space; **C**= center of chest; **R1**= right midclavicular area, 2nd intercostals space; **R2**= right midclavicular area, 4th intercostals space)

Each recording consisted of two simultaneous signals from two different locations on the subject's chest. With a nose clip in place, airflow was also measured with a

mouthpiece attached to a calibrated pneumograph (Hans Rudolph Inc., no. 3830B) with a differential pressure transducer and was monitored on an oscilloscope. The subjects were asked to maintain their target breathing at 7.5 ml/s/kg (low) and 15 ml/s/kg (medium) flow rates by monitoring their breathing on the oscilloscope display. The subjects were instructed to breathe such that one full breath occurred per two to three seconds at every flow rate and had at least five breaths at each target flow. In order to have a reference for background noise, the subjects were instructed to hold their breath for ten seconds with a closed glottis at the end of the experiment. The recorded signals and the measured airflow were amplified and band pass filtered from 50-2500 Hz and digitized at 10240 Hz using 16 bits per sample. Each recording consisted of 50 s target breathing (five breaths at each target flow), followed by a 10 s of breath hold. Breath sounds and airflow over the anterior chest were simultaneously recorded and stored on a Pentium PC in binary format using a custom written program in Labview.

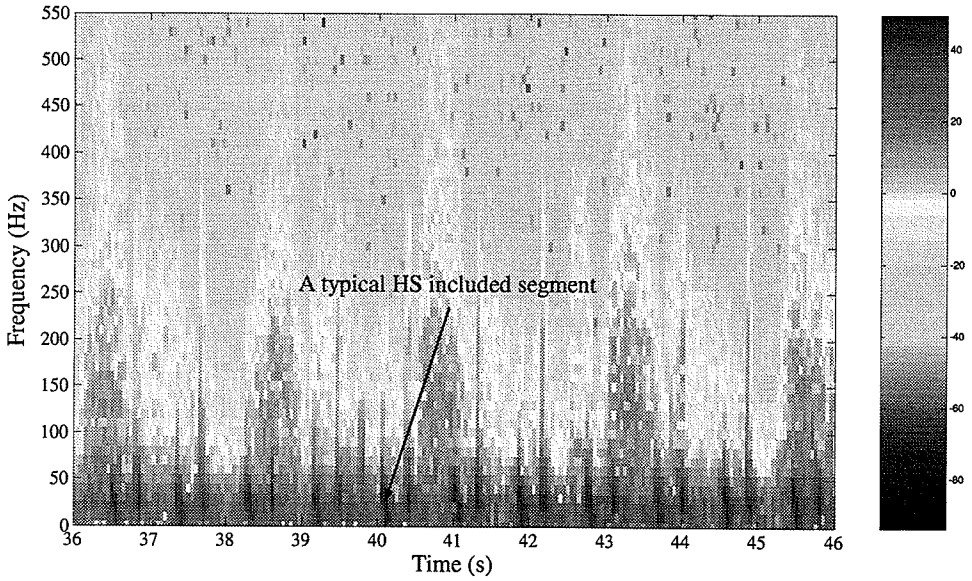


Figure 3.4 Spectrogram of a typical recorded breath sound at 7.5 ml/s/kg flow rate

3.2.8.2 Applying the Algorithm to the Acquired Data

The spectrogram of the signals at each flow rate, i.e., $\hat{X}(\omega,t_s)$ was calculated using the discrete STFT to every 100 ms segment (1024 samples) of data using a Hanning window. The overlap between adjacent segments was considered 85%, since as suggested in [Murata et al., 2001] the small overlap leads to a wrong solution (a full description on criteria for selecting an appropriate overlap has been provided in Section A.3 of the Appendix). Fig. 3.4 shows the spectrogram of a typical breath sound signal recorded at 7.5 ml/s/kg flow rate.

Then the discussed algorithm in Section 3.2.7, assuming two sources (lung and heart), was applied to the spectrograms and the separated signals were obtained.

3.3 Spectrogram Filtering-Based Method

In this section a new method based on image processing for HS removal from lung sound recordings is presented. In the proposed method, the spectrogram of the lung sound signal is analyzed as an image with M gray levels. Three different techniques are proposed for HS localization on the spectrogram of the recorded signal. The HS-included segments are then removed and 2D interpolation is employed for filling up the created gaps in the TF plane. Finally, the signal is reconstructed back in the time domain.

3.3.1 Data Acquisition

Lung sounds were recorded by placing a piezoelectric contact Siemens accelerometer (EMT 25C) on the third intercostal space anteriorly (mid-clavicular line) on the right upper lung lobe of six healthy volunteer subjects (three females) aged 10-26 years.

Fig.3.5 depicts the Respiratory Acoustics Laboratory at the University of Manitoba (Bantyne Campus), where the breath sound recordings for this part of study were conducted. While the subject was sitting inside the acoustic chamber with a nose clip in place, his/her lung sounds were recorded with the accelerometer. Airflow was also measured by a face mask pneumotacograph (Fleisch no. 3) connected to a differential pressure transducer and was recorded simultaneously with lung sounds.

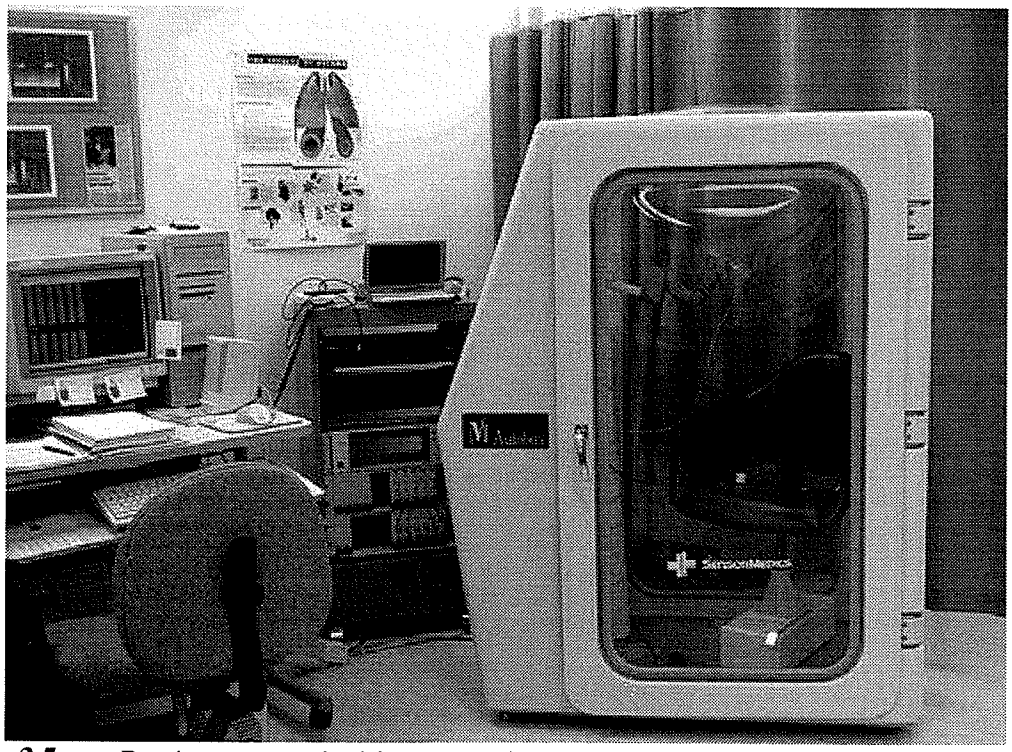


Figure 3.5 Respiratory acoustics laboratory at the University of Manitoba, Banatyne Campus

The subjects were asked to maintain their target breathing at 7.5 (low), 15 (medium), and 22.5 ml/s/kg (high) flow rates by monitoring their breathing using a custom written LabVIEW® (National Instruments) virtual instrument (VI) user interface. Again like the data acquisition described in Section 3.2.8.1, the subjects were instructed to breathe such that at every flow rate, five breaths with the duration of 2-3 s occurred and they were

asked to hold their breath for 10 s with a closed glottis at the end of each experiment. The recorded signals were amplified and band pass filtered from 7.5-2500 Hz and digitized at 10240 Hz using 12 bits per sample (National Instruments DAQ). The measured airflow was digitized simultaneously with breath sounds at the same sampling rate. However, the airflow signal was later decimated to 320 Hz.

3.3.2 Spectrogram Estimation

The spectrogram of the signals at each flow rate was calculated by applying the discrete STFT to every 100 ms segment of data with a 50% overlap between adjacent segments and using a Hanning window. Fig. 3.4 depicts the spectrogram of a typical lung sound signal recorded at low (7.5 ml/s/kg) flow rate.

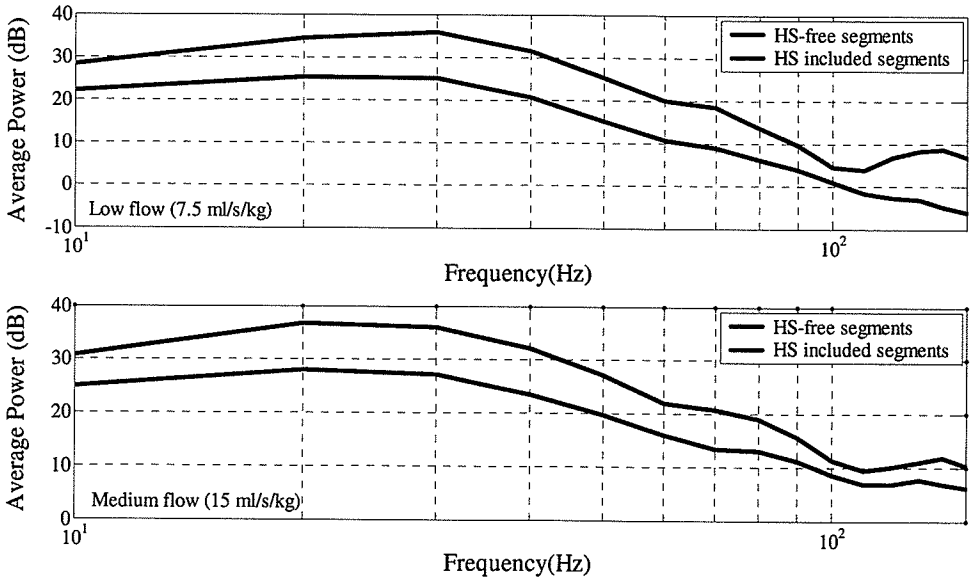


Figure 3.6 Comparing the average PSD of inspiration lung sound segments with and without heart sound

3.3.3 Detection of HS-Included Segments

3.3.3.1 Average Power-Based Technique

The spectra of lung sound segments including HS, have higher intensity in the range of 20-300 Hz as can be seen in Fig. 3.6. This fact leads to a basic approach for detecting segments including HS. If the spectrogram of the signal is analyzed as an image with M gray levels, the location of HS can be detected by using an adaptive threshold. In order to specify such a threshold, the inspiration and expiration segments of the original recorded lung sound signal were separated using the recorded airflow signal. Then, from the spectra of the inspiration and expiration segments, their average power over 20-40 Hz frequency band were calculated to define two reference thresholds (Th) as [Pourazad et al., 2003]:

$$Th = \mu_{P_{ave}} + k_{adj} \times \sigma_{P_{ave}}, \tag{3.26}$$

where $\mu_{P_{ave}}$ and $\sigma_{P_{ave}}$ are the mean value and standard deviation of the calculated average power of the inspiration or expiration segments and k_{adj} is a threshold adjustment parameter that needs to be determined in a training stage by selecting some known segments with and void of HS. Because the power spectral density of lung sound segments during inspiration is higher than that of expiration, the reference threshold for inspiration and expiration was determined separately. This threshold was then applied to the spectrogram. The power of every segment within 20-40 Hz was compared with the threshold. If a segment had one or more points above the threshold, it was considered as a segment that included HS. The detected HS-included segments were evaluated by visual inspection of the spectrogram and further by listening to the original signal.

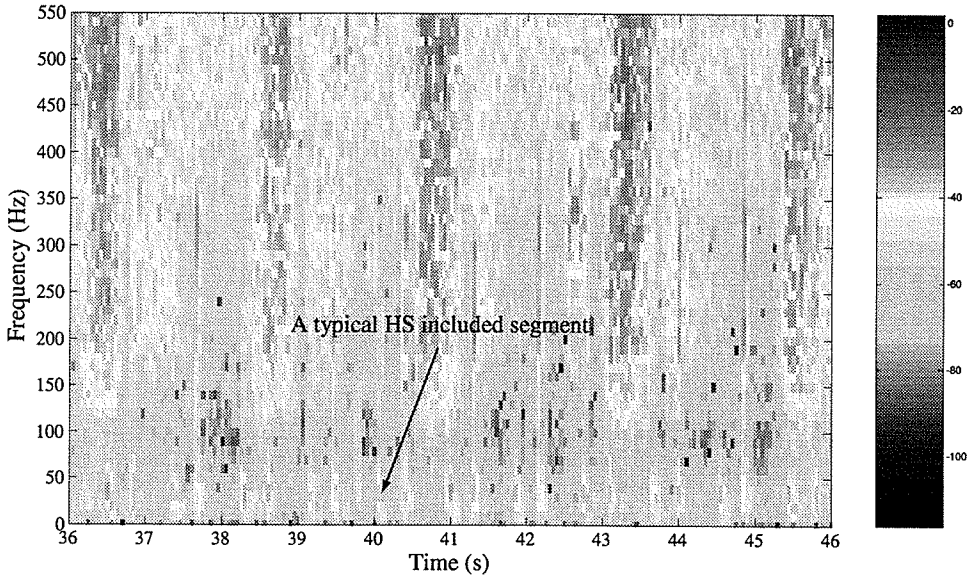


Figure 3.7 Spectrogram of a_5 coefficients at 7.5 ml/s/kg flow

3.3.3.2 Wavelet Coefficient-Based Technique

This technique is very similar, but not identical to the previous technique (Section 3.3.3.1). In the average power-based technique, training stage for adjusting the threshold parameter k_{adj} is compulsory. This parameter varies with noise considerably. In addition, it differs between subjects because the lung sound amplitude varies between the subjects and locations on the chest surface [Pasterkamp et al., 1997]. Thus, it would be crucial to apply the proposed HS removal method in a semiautomatic manner such that the program does not required to be trained or calibrated for each recording. In order to remedy this problem, a new HS localization technique using the wavelet transform was sought.

First, the wavelet transform of the original lung sound signal was calculated using Daubechies Mirror Filters order 4 (db4) with 8 levels. Then, the spectrogram of the coefficients at the fifth level (a_5) was calculated. Fig. 3.7 shows the spectrogram of the a_5 coefficient of the same signal as in Fig. 3.4. As can be observed from Fig. 3.7 the

difference between the gray levels of the segments including heart sounds and those of without HS in the range of 20-40 Hz is much more substantial compare to Fig. 3.4. A significant difference between segments' gray level corresponds to a considerable difference between the segments power spectra. This fact has been also demonstrated by Fig. 3.8.

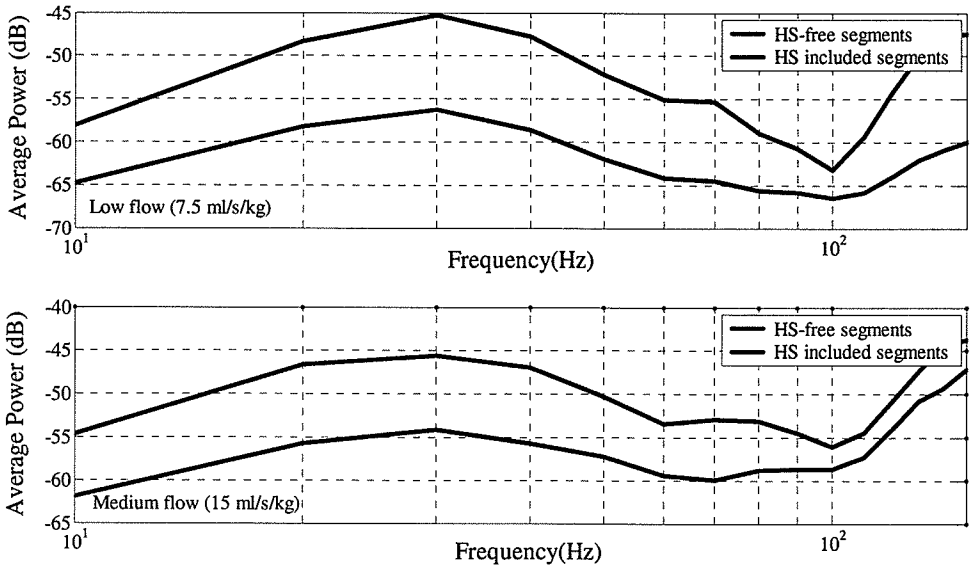


Figure 3.8 Comparing the average PSD of a_5 coefficients segments with and without heart sound, inspiratory phase

The procedure of defining an adaptive threshold in this technique is very similar to that of average power-based technique. There are only two differences: a) the whole procedure was performed on the spectrogram of the a_5 coefficients instead of the original signal's spectrogram; b) the reference thresholds were calculated based on the following expression:

$$Th = \mu_{P_{ave}}, \tag{3.27}$$

where $\mu_{P_{ave}}$ is the mean value of the calculated average power of the inspiration or expiration segments. As it can be observed from Eq. (3.27), this threshold is robust and

independent to the adjustment parameter or callibration. This threshold was applied to the spectrogram of a_5 coefficients. The power of every segment within 20-40 Hz was compared with the threshold. If a segment had one or more points above the threshold, it was considered as a HS segment. Later, the correspondent segments in the original signal spectrogram were also marked as HS-included segments.

3.3.3.3 Second Order Statistical-Based Technique

Lung sound signals free of HS are stationary within a short time-scale (100 ms), while they are intrinsically non-stationary within a long time. HS interference substantially changes the recorded signal's variance. This fact leads to another technique for HS-included segments' localization. In the third technique the HS-included segments were detected by comparing the variance of each segment by a reference threshold defined by:

$$Th = mean(Var), \tag{3.28}$$

where Var is the variances of all segments. The segments that had higher variance relative to the reference threshold were considered as HS-included segments. Similar to the previous techniques, these segments were evaluated by visual inspection of the spectrogram and further confirmed by listening to the original signal. Note that in this technique the same threshold is applicable to both respiratory and expiratory segments of the lung sound at each flow rate. Furthermore, the HS localization is performed in the time domain. Fig. 3.9 illustrates the spectrogram of a typical lung sound at low flow rate and the variance of the segments versus time with the HS localization reference threshold. As can be observed from this figure, the proposed technique was able to successfully detect all of the HS-included segments of this example.

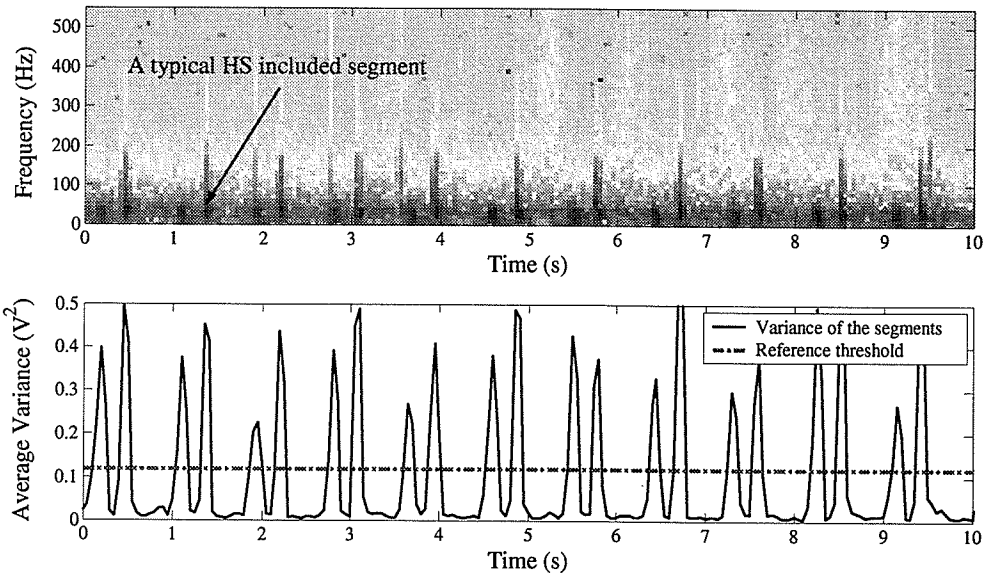


Figure 3.9 Spectrogram of a typical lung sound at low flow rate and the variance of the segments versus time with applied reference threshold for HS localization

3.3.4 Filtering Heart Sound-Included Segments

Having detected the HS-included segments in the signal spectrogram, two different filtering approaches were investigated: *a*) applying a band-stop filter on the HS-included segments in the range of 20-300 Hz; *b*) removing the entire segments including HS. Applying any of *a* or *b* approaches would create some gaps in the spectrogram of the signal. Fig. 3.10 illustrates the spectrogram of Fig. 3.4 after applying the threshold defined by the wavelet coefficient-based technique. As can be observed the segments including HS have been partially filtered. The created gaps have to be estimated by some interpolation approaches.

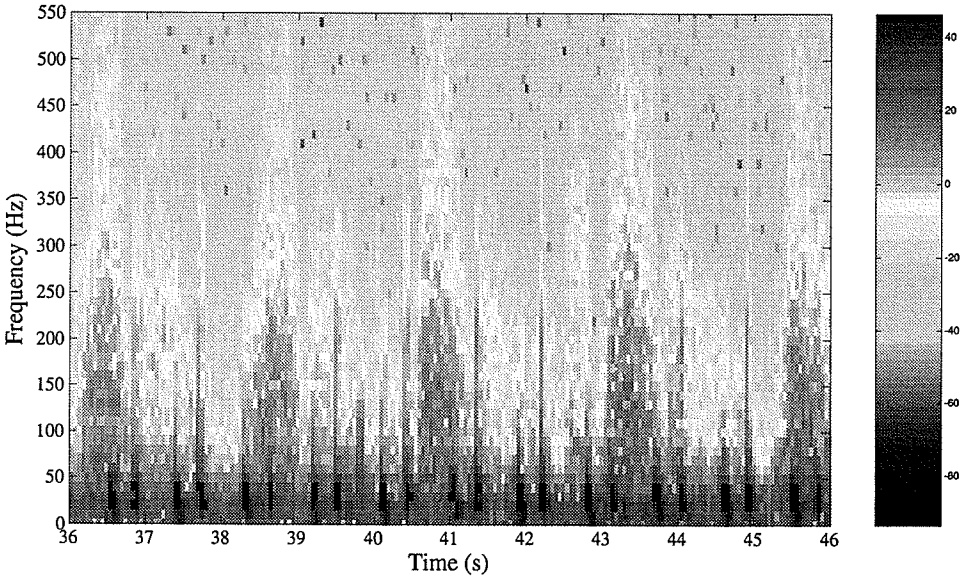


Figure 3.10 Spectrogram of a typical recorded breath sound at 7.5 ml/s/kg flow rate after thresholding

3.3.5 Interpolation Approaches

After filtering, two-dimensional interpolation was implemented to estimate the data in the created gaps. Note that the filtering and the two-dimensional interpolation were applied in the time-frequency domain. The idea of interpolating the spectrogram has been employed before for removing power line interference from electromyogram (EMG) signals [Mewett et al., 2001]. In that study the interference was assumed to be zero-phase and the type of applied interpolation approach was not clarified.

Interpolation is a numerical analysis technique, which is used to fit a function to a given set of data. The technique can be done in different ways. However, all of the interpolation approaches consider available samples to calculate the value of an unknown point within one or two-dimensional data grids. Two-dimensional functions can be interpolated through several one-dimensional interpolation steps. In this study three different interpolation techniques were investigated.

3.3.5.1 First Degree (Linear) Spline Interpolation

One of the most popular one-dimensional interpolation functions is the spline. A spline function consists of polynomial pieces joined together with certain smoothness conditions [Cheney et al., 1999]. The first-degree spline function consists of linear polynomial pieces joined together to achieve continuity. Such a function is called piecewise linear and is defined as:

$$S_1(x) = [(x_1 - x)f_0 + (x - x_0)f_1]/(x - x_0), \tag{3.29}$$

where (x_0, f_0) and (x_1, f_1) are the known positions of two adjacent data samples. The two-dimensional form of linear interpolation is called bilinear spline interpolation, which in the two dimensional space is defined by:

$$S_{ij}(x, y) = b_1 + b_2\frac{x - x_i}{\Delta x_i} + b_3\frac{y - y_j}{\Delta y_j} + b_4\frac{(x - x_i)(y - y_j)}{\Delta x_i \Delta y_j}, \tag{3.30}$$

which is solved for:

$$\begin{aligned} &b_1 = f_{ij},\ b_2 = f_{i+1,j} - f_{ij},\ b_3 = f_{i,j+1} - f_{ij},\ b_4 = f_{i+1,j+1} - f_{i,j+1} - f_{i+1,j} + f_{ij}, \\ &\Delta x_i = x_{i+1} - x_i, \\ &\Delta y_j = y_{j+1} - y_j, \end{aligned} \tag{3.31}$$

where (x_i, y_j, f_{ij}), $(x_i, y_{j+1}, f_{i,j+1})$, $(x_{i+1}, y_j, f_{i+1,j})$ and $(x_{i+1}, y_{j+1}, f_{i+1,j+1})$ are known positions of four adjacent data samples [Spath, 1995].

3.3.5.2 Third Degree (Cubic) Spline Interpolation

Higher degree splines are used whenever more smoothing is needed in approximating a function. Most frequently a 3rd degree spline function, called cubic polynomial, is used [Cheney et al., 1999]. The cubic polynomials are joined together in such a way that

adjacent splines keep continuity in their curvature and slope at the knots [Son et al., 2001]. The general format of the cubic equation is as follows:

$$S_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i, \qquad 0 \le i \le n-1 \qquad (3.32)$$

where n is the number of knots and a_i, b_i and c_i are coefficients, which bend the line such that it passes through each of the data points without any erratic behavior or break in continuity. The application of cubic spline interpolation on two dimensional grids is known as bicubic spline interpolation and can be expressed as:

$$S(x,y) = \begin{bmatrix} 1 & (x-x_i) & (x-x_i)^2 & (x-x_i)^3 \end{bmatrix} \times A \times \begin{bmatrix} 1 & (y-y_j) & (y-y_j)^2 & (y-y_j)^3 \end{bmatrix}^T, \quad (3.33)$$

where A is a 4×4 matrix of coefficients which are the functions of adjacent data samples on the rectangular grid [Spath, 1995].

3.3.5.3 Nearest Neighbor Approximation

There are also some other approaches for approximating the data in two-dimensional grids that do not attempt to fit a particular function to the data. The nearest-neighbor approximation is one of the most frequently considered approaches in this category [Mersereau et al., 1974; Jain, 1989]. In this approach, an unknown point is assigned the value of the closest data point in a two-dimensional grid [Son et al., 2001]. Fig. 3.11 depicts the results of nearest-neighbor approximation, cubic spline interpolation and piecewise linear interpolation for a set of simulated data in one-dimensional space. These three techniques were applied to interpolate the breath sound signal after filtering the segments including HS.

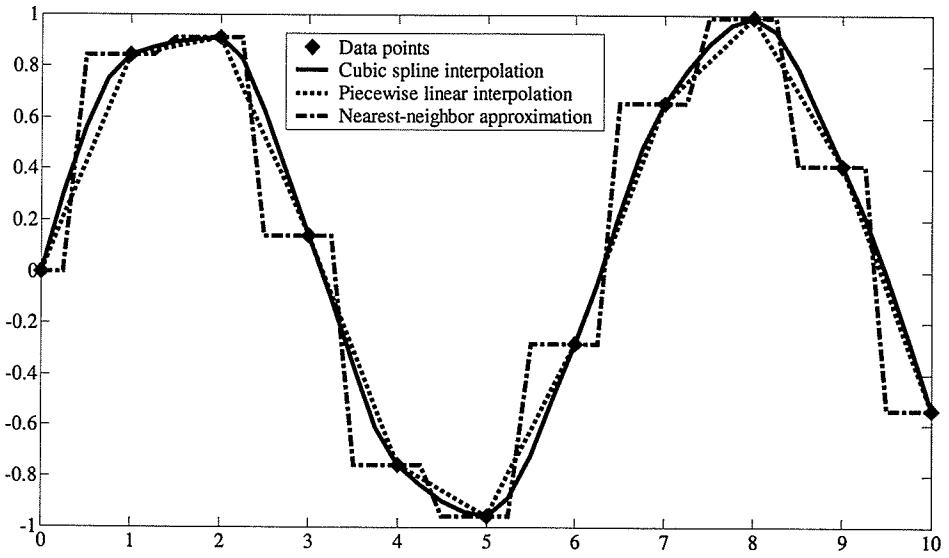


Figure 3.11 Results of nearest-neighbor approximation, cubic spline interpolation and piecewise linear interpolation for a set of data in one-dimensional space.

3.3.6 Reconstruction of the Lung Sound Signal

After approximating the data and filling up the created gaps, the TF representation of each segment must be reconstructed in time domain by applying ISTFT (a fuller explanation is available in Section 3.2.3).

3.4 Evaluating the Effectiveness of the Proposed Methods

In addition to qualitative manual validation by visual and auditory means, quantitative analysis was also performed to assess the efficiency of the proposed methods in HS localization and HS cancellation from lung sounds and to compare the results of different interpolation approaches.

3.4.1 Quantitative Analysis of HS Localization Techniques

The performance of all the presented HS localization techniques is based on selecting an appropriate threshold. Therefore the quantitative analysis of these techniques can be investigated based on linear classification error.

The recorded signals were first segmented (1024 samples per segment and 50% overlap between adjacent segments), and then the HS-included segments were manually separated from the ones without HS by visual and auditory inspections. By considering some known segments with and void of HS, a threshold adjustment parameter for average power-based technique was determined. Afterwards, the proposed HS localization techniques were separately applied to the segmented data to classify the segments in two groups: HS-free and HS-included segments. The results of each group were verified manually and the classification error of each HS localization technique was averaged between the subjects at each flow rate.

3.4.2 Quantitative Analysis of HS Cancellation Methods

The quantitative analysis of the proposed HS cancellation methods in this study was based on the standard approach of comparing average PSD of segments with and without HS introduced in [Gnitecki et al., 2003; Hossain et al., 2003; Pourazad et al., 2003] for investigating the effectiveness of any HS cancellation method.

From the spectrogram of the original and filtered signals the average PSD over four frequency bands of the segments corresponded to ±20% of the target flow was calculated. In order to provide reference spectra of HS-free segments, the average PSD of the segments in the original lung sound recording that were within ±20% of the target flow

and were also free of HS (selected manually) was computed. The four frequency bands for average PSD calculation were chosen as 20-40 Hz, 40-70 Hz, 70-150 Hz, 150-300 Hz, the same as those in [Gnitecki et al., 2003]. The power ratios (in dB) between the average PSD of reference spectrum, original sound segments including HS and filtered segments in each frequency band were statistically analyzed with paired t-tests. The hypothesis was that the average PSD of the filtered signal would fall between the average PSD of HS-free segments and the average PSD of the original signal including HS and that it would be closer to the former than to the latter. Furthermore, it would ideally be equivalent to (or not significantly different from) those of the HS-free segments. This would indicate reduction on HS without loss or alteration of lung sound information. The validation of this premise was also inspected by comparing the results of different data interpolation approaches.

3.5 Chapter Summary

This chapter elaborated on the methodology for HS removal from lung sound recordings. First, the spectrogram-ICA based method was described, including statement of the problem in frequency domain, the applied algorithm to solve the problem, overall data acquisition and signal pre-processing and the application of this method on the acquired data. Second, the proposed spectrogram filtering-based method and its application were fully discussed along with proposing three techniques for HS localization over the breath sound spectrogram. Third, the routines for evaluating the effectiveness of the proposed methods were explained. The following chapter will discuss

on the results of performing proposed algorithms on the acquired data and the effectiveness of these algorithms will be evaluated and compared with each other.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

Adherence to performing the systematic procedures described in Chapter 3 for HS localization on the recorded respiratory sounds and HS removal from lung sound recordings, both evaluative and comparative experimental results were obtained. In the following sections, the obtained results are fully described and the effectiveness of the proposed algorithms are evaluated and compared with each other.

4.2 Evaluating HS Localization Techniques

In order to evaluate the performance of the proposed HS localization techniques, the results of each technique were verified based on the manually selected HS-included segments. Then the error of each HS localization technique was averaged between the subjects at each flow rate. Table 4.1 shows the percentage of error for each technique in detecting HS-included segments at low and medium flow rates averaged between the subjects.

As can be observed from Table 4.1, the average power-based technique has the minimum total error in HS localization. However, in terms of missing HS-included segments, the wavelet coefficient-based technique showed the least error but at the cost of a higher false positive detection error. It was found that the overlap between the adjacent segments some times caused the wavelet coefficient based-technique and the second order statistical-based technique to misclassify the segments in the neighborhood

of HS-included segments. Thus the length of segments or length of the overlap between adjacent segments is crucial in the performance of these techniques. Recall that the segments' length and the overlap should be carefully chosen in a way that the estimation remains unbiased.

As described in Section 3.3, HS localization procedure is one of the preprocessing stages in spectrogram filtering-based method. Considering the fact that this technique is able to reconstruct the spectrogram of the recorded signal after filtering HS corresponding segments, it seems that applying a HS localization technique with the least error in missing HS-included segments would be more appropriate even at the cost of slightly higher false positive detection error. Hence the second order statistical-based technique and the wavelet coefficient-based technique are the recommended techniques for spectrogram filtering-based method.

Flow Rate	**HS Localization Technique**	**Type of Error**		
		Missing HS-included segment (%)	False positive detection error (%)	Total Error (%)
Low (7.5 ml/s/kg)	Average Power-Based Technique	2.85 ± 1.41	3.30 ± 1.79	6.14 ± 1.98
	Wavelet Coefficient-Based Technique	1.15 ± 1.09	6.76 ± 6.31	7.91 ± 6.51
	Second Order Statistical-Based Technique	1.95 ± 1.79	6.98 ± 1.60	8.92 ± 2.05
Medium (15 ml/s/kg)	Average Power-Based Technique	3.58 ± 2.18	3.71 ± 2.43	7.29 ± 2.00
	Wavelet Coefficient-Based Technique	0.68 ± 0.74	10.24 ± 5.65	10.92 ± 6.04
	Second Order Statistical-based Technique	2.01 ± 1.64	6.93 ± 2.02	8.94 ± 3.30

Table 4.1 The error ($\mu \pm \sigma$) of HS localization techniques at different flow rates.

Since in the average power-based technique, the threshold parameter, k_{adj}, considerably varies with noise and differs between the subjects, it is crucial to apply this technique in a semiautomatic manner such that the program is being trained or calibrated

for each recording. In contrast it was found that the performance of the two other techniques is robust for all subjects. Thus, these methods can be applied in a fully automatic manner, which is not possible for the average power-based technique.

The second order statistical-based technique uses the same threshold for both respiratory and expiratory segments, however for the average power-based and wavelet-based techniques different reference thresholds should be defined for each respiratory phase. In other words, the information about the onset of breath signal is not required for the second order statistical-based technique. Furthermore, the second order statistical-based technique is implemented in time domain and is very efficient in terms of speed and computational load, while the other two techniques should be applied in TF domain over the spectrogram of signal and in terms of speed are not as efficient as the second order statistical-based technique.

For the recordings that include crackles, average power-based and wavelet coefficient-based techniques will be much more effective than statistical-based technique since they apply the reference threshold to very low frequencies, i.e., in the range of 20-40 Hz. Hence the distortion due to crackles (at high frequencies) does not affect their performance.

As Table 4.1 shows, the rate of error, specially the false positive detection error is higher for medium flow than that for low flow. This implies at medium flow rate the respiratory sounds mask the HS to the extent that it becomes less dominant and distinguishable. The reported standard deviation of errors is very high at both flow rates and for all the proposed HS localization techniques. It was investigated that, it was due to the high rate of HS localization error in two of recordings, which were very noisy in a

way that even HS localization in manual manner was difficult for those recordings. This problem is not restricted only to this technique as the performance of the reported method in [Hadjileontiadis et al., 1997] was also distortable by ambient noise, weak heart beating and microphone's mislocation on the chest. All of the proposed techniques for HS localization suffer from low-resolution performance since all of them should be applied to the segmented recorded signals, i.e., every 1024 sample (100 ms). To remedy this problem the length of segments should be decreased but this must not go beyond having unbiased estimation.

4.3 Evaluating HS Cancellation Methods

The efficiency of the proposed methods for HS cancellation from lung sound recordings was examined quantitatively and qualitatively by visual and auditory means. The quantitative analysis was performed by comparing the average band-limited PSD of resulted signals with that of original signals including HS and without HS (selected manually) at low and medium flow rates. It was expected that the filtered signal would fall between the average PSD of HS-free segments and the average PSD of the original signal including HS and that it would be closer to the former than to the latter. Note that at each flow rate only the segments correspond to ±20% of the target flow were selected for quantitative analysis. The power ratios (in dB) between the average PSD of the original sound segments without HS, the original sound segments including HS and the filtered segments in each frequency band (20-40 Hz, 40-70 Hz, 70-150 Hz, 150-300 Hz) were statistically analyzed with paired t-tests for spectrogram filtering-based method. The

spectrogram ICA-based method did not undergo the statistical tests due to insufficient number of recordings.

The results of comparing average band-limited PSD of original lung sounds with and without HS at low and medium flow rates are shown in Table 4.2. The difference between the aforementioned average PSD values up to 70 Hz was significant ($p < 0.02$) and this difference was more pronounced at low flow than medium flow. These results confirm that the main interference of HS on a lung sound recording occurs at low frequencies. At higher flow rates heart sounds may be masked due to the increase of the lung sound average power.

In the following sections the obtained results from performing the proposed HS cancellation methods on the acquired data will be presented and fully discussed.

Flow Rate	20-40 Hz	40-70 Hz	70-150 Hz	150-300 Hz
Low (7.5 ml/s/kg)	5.61 ± 3.00 *	6.67 ± 2.79 *	3.91 ± 1.34	1.56 ± 0.95
Medium (15 ml/s/kg)	4.29 ± 2.45 *	4.33 ± 1.73 *	1.7 ± 0.52	1.00 ± 0.49

Table 4.2 Difference ($\mu \pm \sigma$, in dB) between average power spectra of the original lung sound signal with and without HS, at low and medium flow rates and over four frequency bands ($* \equiv p < 0.05$)

4.3.1 Spectrogram ICA-Based Method

Applying the spectrogram ICA-based method to the two sensor recordings (from two different locations on the chest) results in estimating two sources per each sensor, i.e., at the end of the experiment one estimation for HS and one estimation for lung sounds per each sensor are resulted. Fig. 4.1a & b show the spectrograms of the original signals recorded simultaneously from R1 (sensor 1) and L2 (sensor 2) locations on the chest at low flow rate and Fig. 4.1c & e illustrate the spectrograms of the estimated HS per

sensors 1 and 2, respectively. Fig. 4.1d & f depict the spectrograms of the estimated lung sound per sensors 1 and 2, correspondingly. By visual and auditory inspections it was found that the HS were not cancelled completely from the estimated lung sounds but were reduced. Also the estimated HS still included weakened lung sounds. Fig. 4.2 shows the same results but for the recorded signal at C and R2 locations on the anterior chest at medium flow rate.

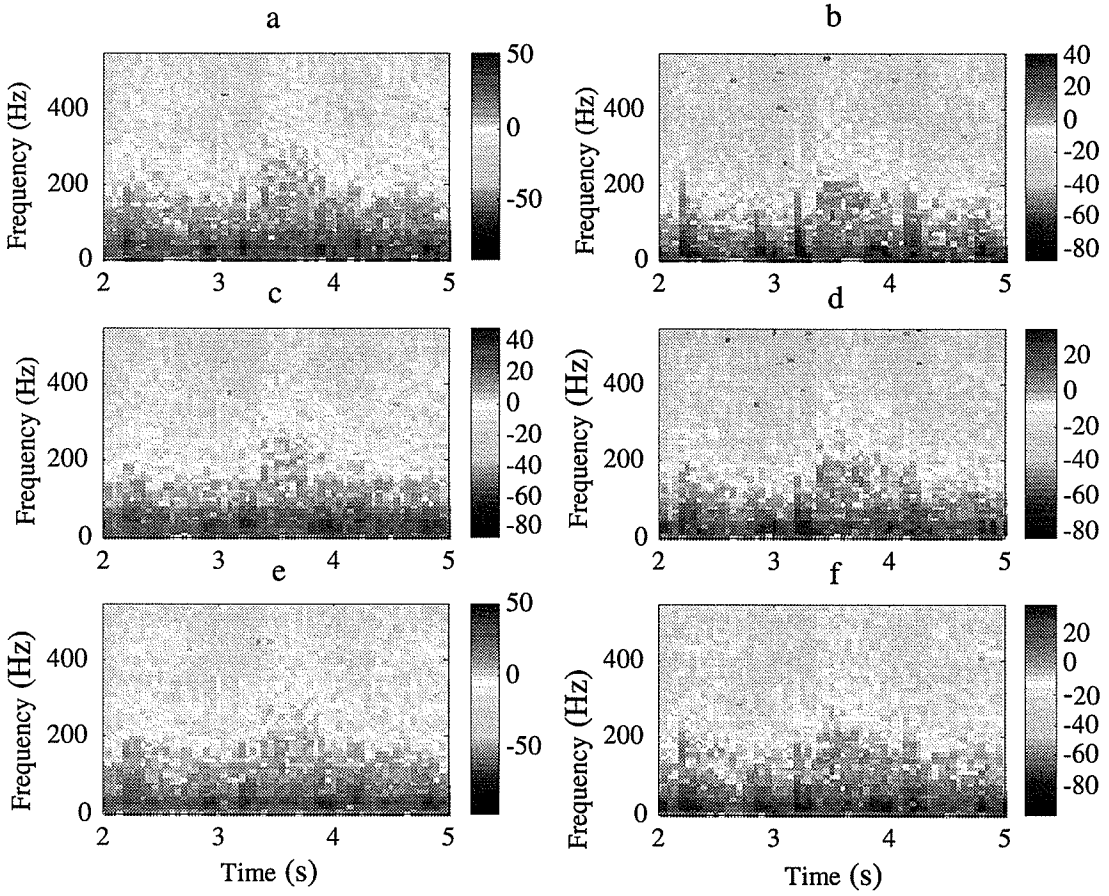


Figure 4.1 Spectrograms of the recorded signals at R1 and L2 locations and the estimated HS and lung sound signals at low flow rate (7.5 ml/s/kg) for a typical subject. a) the spectrogram of the recorded signal at the location R1 on the chest; b) the spectrogram of the recorded signal at the location L2 on the chest; c & e) the spectrograms of the estimated signals corresponding to HS at sensors R1 and L2, respectively; d & f) the spectrograms of the estimated signals corresponding to lung sounds at sensors R1 and L2, respectively.

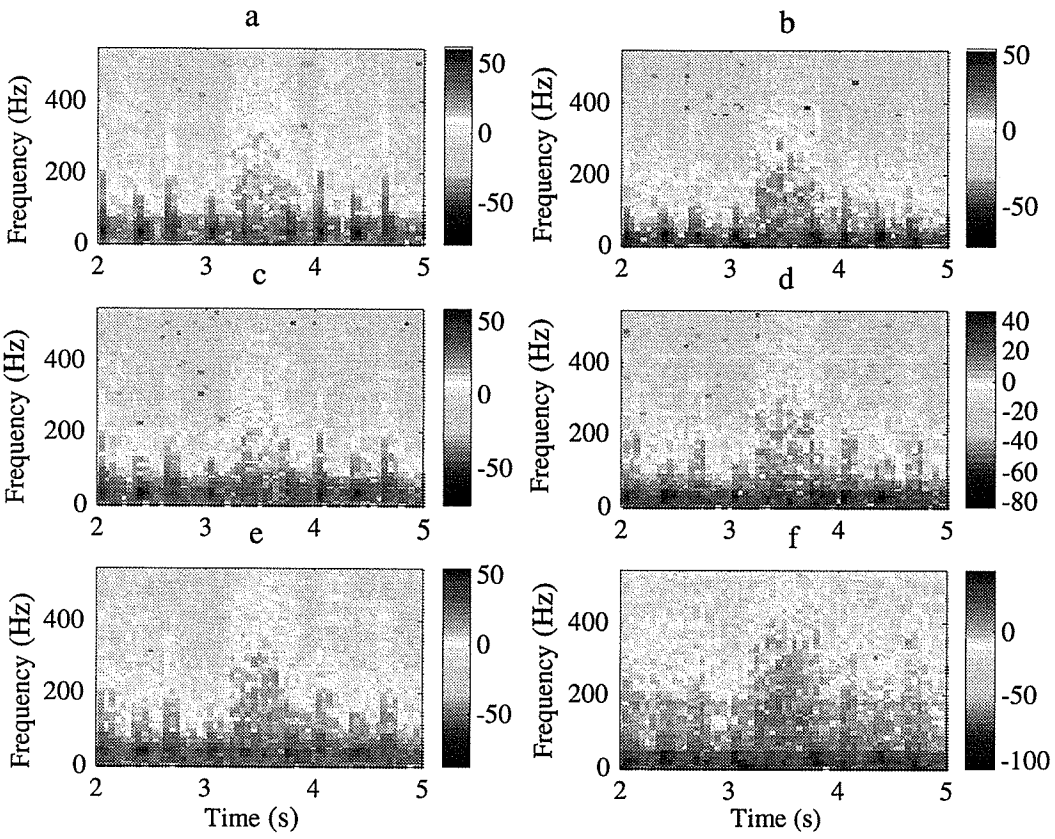


Figure 4.2 Spectrograms of the recorded signals at C and R2 locations and the estimated HS and lung sound signals at medium flow rate (15 ml/s/kg) for a typical subject. a) the spectrogram of the recorded signal at the location C on the chest; b) the spectrogram of the recorded signal at the location R2 on the chest; c & e) the spectrograms of the estimated signals corresponding to HS at sensors C and R2, respectively; d & f) the spectrograms of the estimated signals corresponding to lung sounds at sensors C and R2, respectively.

Fig. 4.3 shows the average PSD of the original signals recorded at R1 and L2 locations on the chest and the average PSD of the estimated HS and lung sound signals using spectrogram ICA-based method in comparison to the average PSD of the original signals free of HS for the ±20% of the target flow, at low flow rate. Fig. 4.4 illustrates the same but for the recorded signals at C and R2 locations on the chest at medium flow rate. The average PSD of the estimated lung sound signal was found to be lower than the

average PSD of the original lung sounds free of HS for both cases. On the contrary, the average PSD of the estimated HS signal was higher with respect to the average PSD of the HS-free original lung sounds.

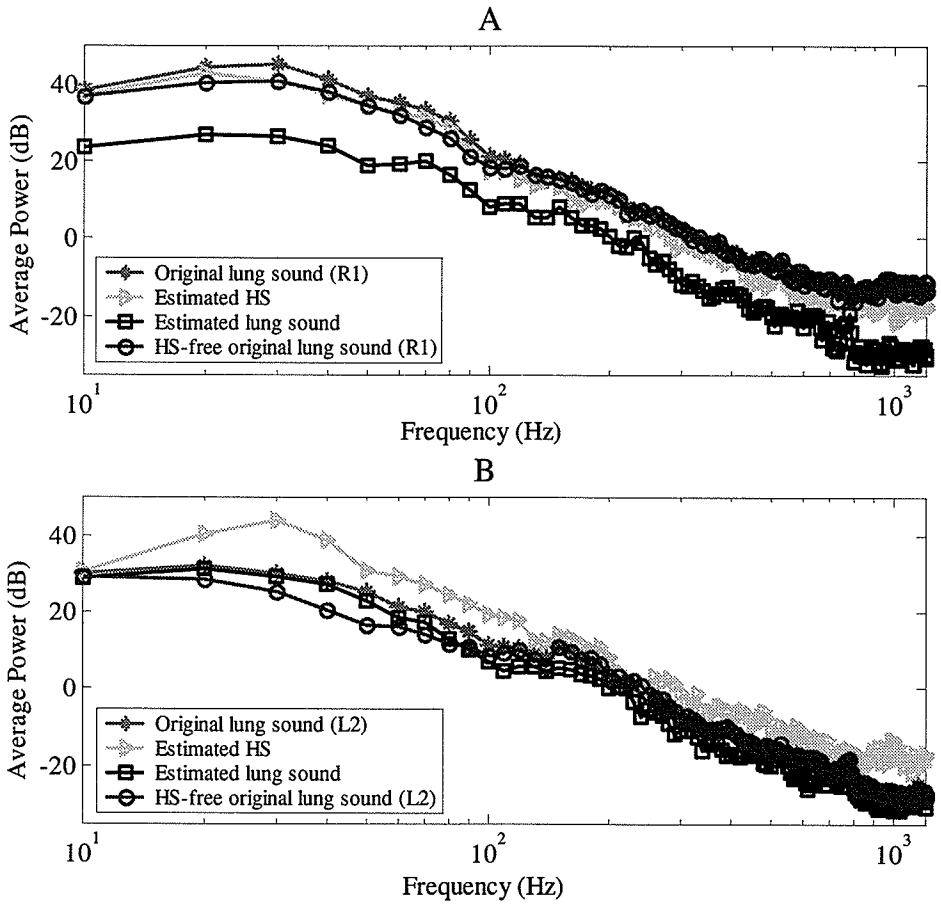


Figure 4.3 Comparison between the average PSD of the estimated HS and lung sound signals using spectrogram ICA-based method and that of original lung sound signal with and without HS, recorded at R1 (A) and L2 (B) locations of the chest at low flow.

As illustrated in Tables 4.3 and 4.4, the difference between the average PSD of the estimated source signals and that of original HS-free recorded signals varies among different locations on the chest for all frequency bands at both low and medium flow rates. The spectrogram ICA-based method claims to estimate the source signals from the

observed mixed signals on the skin, which are convolutive mixtures due to the delays and reflections of the lung tissues. Thus comparing the average PSD of the estimated source signals with the average PSD of the recorded signals on the skin does not seem to be a suitable approach for assessing the estimated signals. Consequently from the difference between the average PSD of the estimated signals and the HS-free signals, one cannot conclude for the spectrogram ICA-based method, which locations on the chest are the most appropriate recording site, in terms of HS reduction.

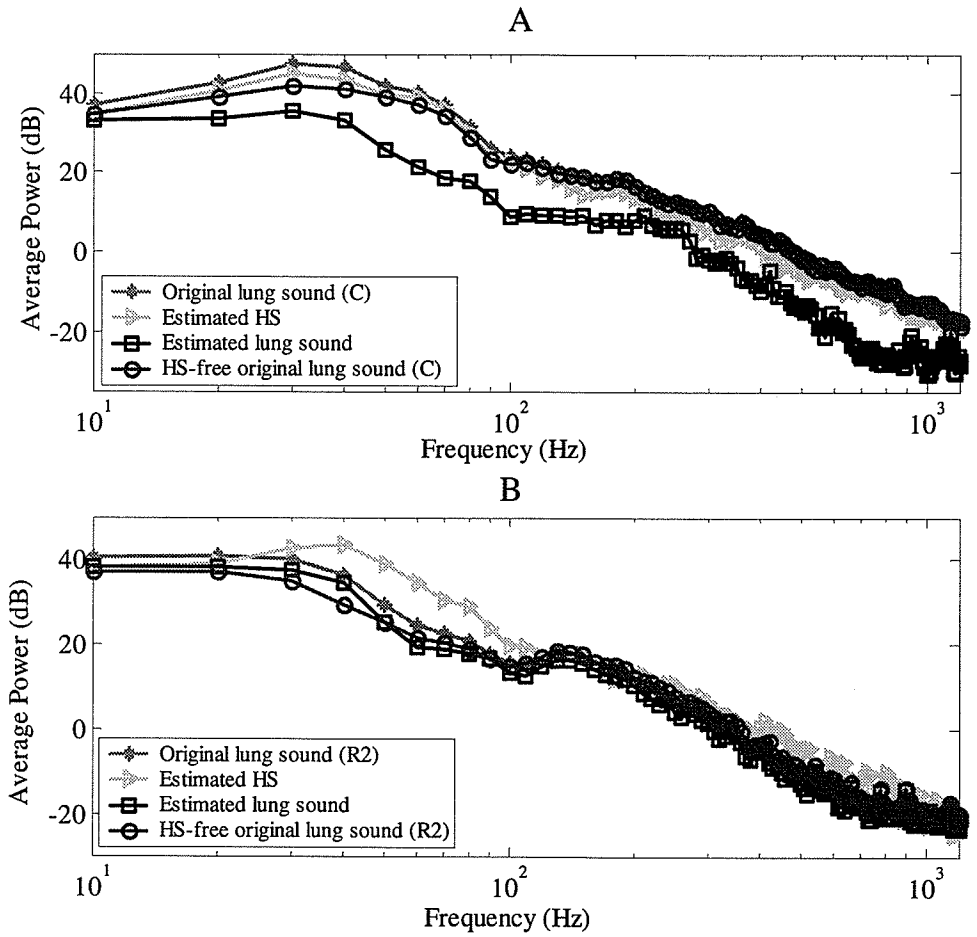


Figure 4.4 Comparison between the average PSD of the estimated HS and lung sound signals using spectrogram ICA-based method and that of original lung sound signal with and without HS, recorded at C (A) and R2 (B) locations of the chest at medium flow.

Low Flow	Estimated Signal & Corresponded Sensor	Mean Difference between Average Power Spectra of the Estimated Signals and HS-free Original Lung Sound Signal (in dB)			
		20-40 (Hz)	40-70 (Hz)	70-150 (Hz)	150-350 (Hz)
L1L2	HS at L1	2.51	1.45	-1.90	-2.41
	Lung Sound at L1	-5.10	-7.05	-8.33	-10.99
	HS at L2	5.76	10.89	5.95	5.07
	Lung Sound at L2	1.89	4.93	1.07	-3.33
L1R1	HS at L1	6.60	3.35	-0.30	-4.20
	Lung Sound at L1	2.23	1.48	-2.19	-2.23
	HS at R1	2.44	1.90	-0.59	-2.24
	Lung Sound at R1	-5.62	-2.22	-2.73	-2.29
L1R2	HS at L1	0.66	-0.07	-1.92	-2.99
	Lung Sound at L1	4.23	-8.81	-8.27	-9.50
	HS at R2	-1.19	4.34	0.84	0.69
	Lung Sound at R2	1.44	0.34	-3.06	-2.08
CL1	HS at C	6.91	6.19	4.20	1.04
	Lung Sound at C	-1.02	-0.88	-0.61	-0.44
	HS at L1	3.39	5.00	0.83	-4.51
	Lung Sound at L1	2.37	1.39	-2.24	-1.76
R1L2	HS at R1	1.94	1.44	-0.80	-2.64
	Lung Sound at R1	-12.22	-11.38	-10.05	-10.85
	HS at L2	11.08	11.90	7.61	4.83
	Lung Sound at L2	2.54	4.17	1.67	-2.84
R2L2	HS at R2	1.35	1.71	-1.79	-2.75
	Lung Sound at R2	-5.43	-2.38	-2.68	-6.02
	HS at L2	5.18	4.69	-0.85	0.41
	Lung Sound at L2	1.10	2.55	1.19	-2.65
CL2	HS at C	4.86	4.25	3.77	0.09
	Lung Sound at C	-6.83	-10.27	-7.51	-9.56
	HS at L2	11.77	17.00	9.24	3.37
	Lung Sound at L2	0.97	3.15	0.65	-2.10
R2R1	HS at R2	7.61	10.98	6.48	1.62
	Lung Sound at R2	2.10	1.76	-2.73	-3.30
	HS at R1	1.20	-0.46	-1.09	-3.46
	Lung Sound at R1	-6.25	-9.50	-10.43	-7.71
CR1	HS at C	6.36	4.61	3.27	-0.11
	Lung Sound at C	5.65	0.06	1.00	-0.60
	HS at R1	-0.63	3.34	-0.32	-4.30
	Lung Sound at R1	0.59	1.83	-1.52	-2.94
CR2	HS at C	5.84	1.85	2.61	-1.07
	Lung Sound at C	-5.74	-9.96	-10.30	-5.69
	HS at R2	8.54	15.16	7.35	-1.11
	Lung Sound at R2	2.69	4.38	-0.95	-2.57

Table 4.3 Mean difference (in dB) between average power spectra of the estimated HS and lung sound signals and original HS-free lung sound signal recorded at different locations on the chest, at low flow rate in the four frequency bands.

Medium Flow	Estimated Signal & Corresponded Sensor	Mean Difference between Average Power Spectra of the Estimated Signals and HS-free Original Lung Sound Signal (in dB)			
		20-40 (Hz)	40-70 (Hz)	70-150 (Hz)	150-350 (Hz)
L1L2	HS at L1	-1.08	-1.93	-3.35	-3.90
	Lung Sound at L1	-7.68	-11.26	-11.19	-11.84
	HS at L2	4.79	8.55	4.73	3.88
	Lung Sound at L2	-0.04	0.53	-2.22	-2.81
L1R1	HS at L1	3.27	0.07	-1.29	-3.78
	Lung Sound at L1	0.46	-1.14	-2.26	-3.02
	HS at R1	2.54	0.63	-1.73	-2.46
	Lung Sound at R1	-6.73	-3.51	-4.15	-3.57
L1R2	HS at L1	-2.50	-4.51	-2.39	-3.70
	Lung Sound at L1	-4.30	-10.97	-10.13	-7.98
	HS at R2	1.34	6.38	2.41	1.36
	Lung Sound at R2	0.62	-2.55	-2.45	-2.68
CL1	HS at C	3.11	2.75	0.89	-2.42
	Lung Sound at C	-1.71	-3.69	-4.25	-1.36
	HS at L1	1.26	1.52	-2.64	-6.05
	Lung Sound at L1	0.28	-0.87	-2.29	-2.71
R1L2	HS at R1	1.38	-0.07	-1.73	-2.22
	Lung Sound at R1	-11.48	-13.00	-13.07	-11.01
	HS at L2	11.34	10.27	6.20	4.40
	Lung Sound at L2	0.69	0.64	-1.80	-2.91
R2L2	HS at R2	-0.32	1.40	-2.53	-3.23
	Lung Sound at R2	-5.23	-1.76	-4.76	-6.34
	HS at L2	6.59	2.17	-1.39	0.43
	Lung Sound at L2	0.06	-1.15	-1.79	-3.98
CL2	HS at C	3.54	2.77	1.22	-2.16
	Lung Sound at C	-10.32	-13.27	-11.14	-8.69
	HS at L2	9.10	11.12	5.15	2.21
	Lung Sound at L2	1.17	0.97	-1.34	-3.04
R2R1	HS at R2	4.09	10.20	5.83	1.08
	Lung Sound at R2	2.01	2.54	-2.99	-4.05
	HS at R1	2.05	0.11	-1.69	-3.71
	Lung Sound at R1	-7.67	-10.18	-12.46	-7.83
CR1	HS at C	3.24	2.29	1.09	-2.61
	Lung Sound at C	1.54	-0.25	-1.44	-1.23
	HS at R1	-1.97	1.53	-3.92	-6.36
	Lung Sound at R1	3.29	1.66	-1.03	-2.54
CR2	HS at C	2.94	1.70	0.85	-3.24
	Lung Sound at C	-5.12	-11.04	-10.31	-5.91
	HS at R2	4.87	10.78	4.02	-1.20
	Lung Sound at R2	1.17	-0.20	-2.33	-2.52

Table 4.4 Mean difference (in dB) between average power spectra of the estimated HS and lung sound signals and original HS-free lung sound signal recorded at different locations on the chest, at medium flow rate in the four frequency bands.

Fig. 4.5 illustrates the estimated and the original signals recorded at C and L1 locations in the time domain at low flow rate. Fig. 4.6 demonstrates the same but for the original signals recorded from R2 and R1 locations at medium flow rate. Comparing the signals before and after filtering shows the amplitude of the estimated lung sounds has been reduced which implies a reduction of HS. Note that the HS-free portions of data have also been slightly altered, that may be due to the reduction of the third and forth HS, which are not usually perceptible but exist in the original lung sounds.

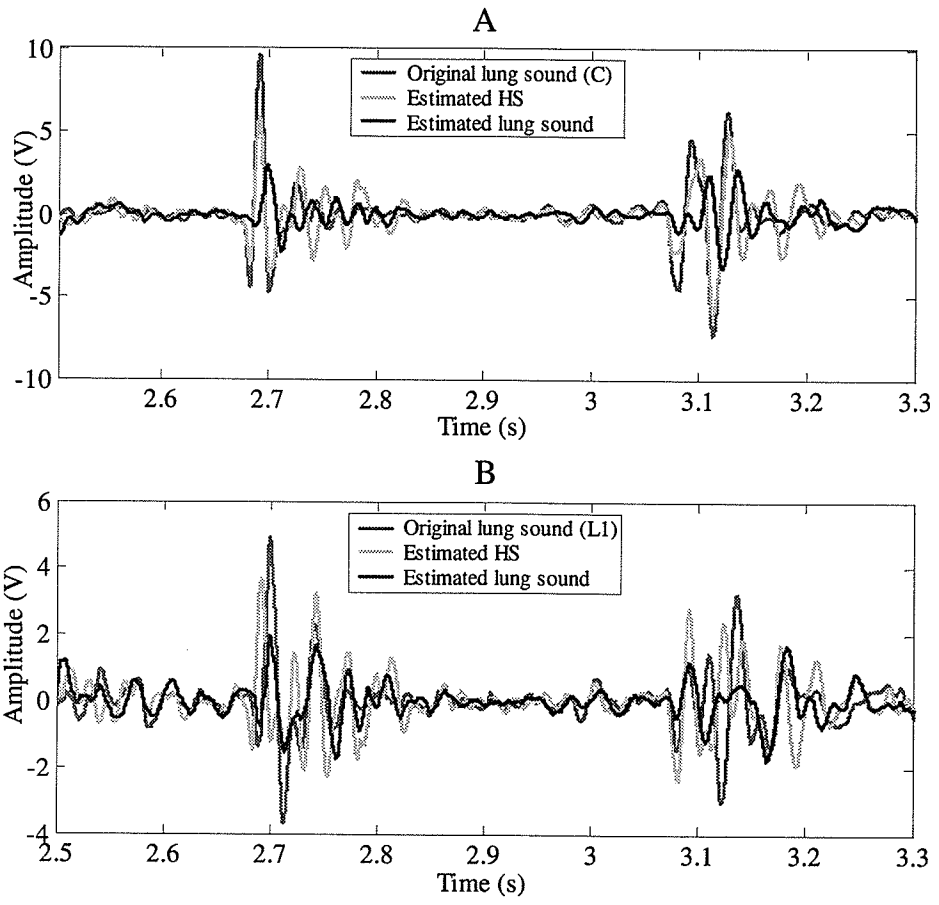


Figure 4.5 Time domain representation of original lung sounds recorded at C (A) and L1 (B), the estimated HS and the estimated lung sound, at low flow rate for a typical subject.

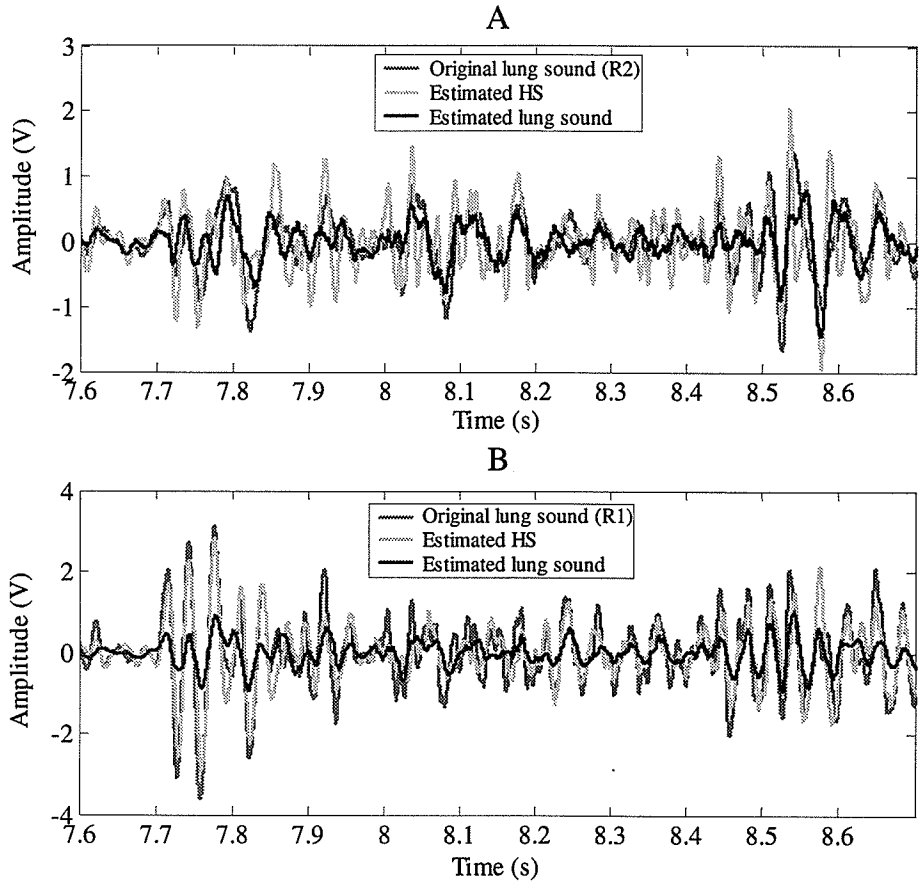


Figure 4.6 Time domain representation of original lung sounds recorded at R2 (A) and R1 (B), the estimated HS and the estimated lung sound, at medium flow rate for a typical subject.

In this study, it is assumed the recorded lung sound signals are stationary within a short time (i.e., 100 ms). Thus the length of each segment was chosen 1024 samples (100 ms) with 85% overlap between adjacent segments. The length of the segment has a strong relationship with the transfer function of the mixing process [Murata et al., 2001]. The major factors causing convolutions between sources are reflections and delays due to the lung tissues. If all the reflections are not included within a segment, the spectrogram ICA-based technique will not work well [Murata et al., 2001]. Therefore in our study, a longer segment length may be required for better performance. However, the segments'

length should not go beyond the stationarity duration of the signal. For finding the optimum segment length, advance studies are required on the lung tissues' reflection strength at different frequencies.

In the present study, two simultaneous recordings have been used for separating the lung sounds from HS without considering the noises at sensors. However, as suggested in [Cichocki et al., 2002], additive noises can be interpreted as unknown source signals. Thus, more simultaneous recordings may improve the results by separating additive noises beside HS from lung sounds. Although this may be considered as a drawback, but in cases where multiple simultaneous recordings are required for diagnosis purposes, applying spectrogram ICA-based method for HS removal with more simultaneous recordings may be beneficial. Also the sampling rate plays a role in the performance of the algorithm. In our experiments, the sampling rate of 10240 Hz was used for data recordings. From the sampling theorem, it fallows that the acquired data includes signals with frequencies less than 5120 Hz. The spectrogram ICA-based method is applied to each frequency component but if even one component does not have any power at that frequency, the algorithm will fail [Murata et al., 2001]. Considering the fact that the lung sound signals' average power is negligible at frequencies over 2000 Hz, sampling rate of 10240 Hz may not be appropriate for implementing spectrogram ICA-based method. Thus lower sampling rate is recommended for data recording to enhance the performance of this method.

4.3.2 Spectrogram Filtering-Based Method

Fig. 4.7 shows the spectrogram of the TF-filtered signal, shown in Fig. 3.4, using the bilinear spline interpolation approach after removing the entire HS-included segments (approach b), which were detected by wavelet coefficient-based technique, at low flow rate. As it can be observed, the segments with higher density that correspond to the HS in Fig. 3.4 were removed entirely and reconstructed in Fig. 4.7 without any alteration in other segments. As it was mentioned in the Section 3.3.4, we used two different filtering approaches based on a) stop-band filtering in the range of 20-300 Hz; and b) entire removal of HS-included segment. When the stop-band filtering approach was implemented, the spectrogram of the TF filtered signal showed some impulses regardless of the type of interpolation approach that was applied for reconstruction. These impulses were heard as click sounds. The reason for the creation of these extra impulses is due to the phase of the HS. Although the main components of the HS are in the range of 20-100 Hz, its phase spans considerably in the entire frequency domain. Therefore, partial approximation of data after stop-band filtering was not effective as the HS phase still remained at higher frequencies. Hence, the second approach of filtering is recommended for HS cancellation from lung sound.

Fig. 4.8 shows the average PSD of an original signal and TF-filtered signal after applying different approaches for interpolation in comparison to the average PSD of the original signal free of HS for the ±20% of the target flow, at low and medium flow rates. The average PSD of the TF-filtered signal closely matched with the average PSD of the original lung sounds free of HS, for both low and medium flow rates.

Comparing the graphs in Fig. 4.8, confirmed that the most appropriate method for estimating the data of the removed HS-included segments is bilinear spline interpolation. Bilinear spline interpolation was also the most efficient approach among the three used approaches in terms of CPU time.

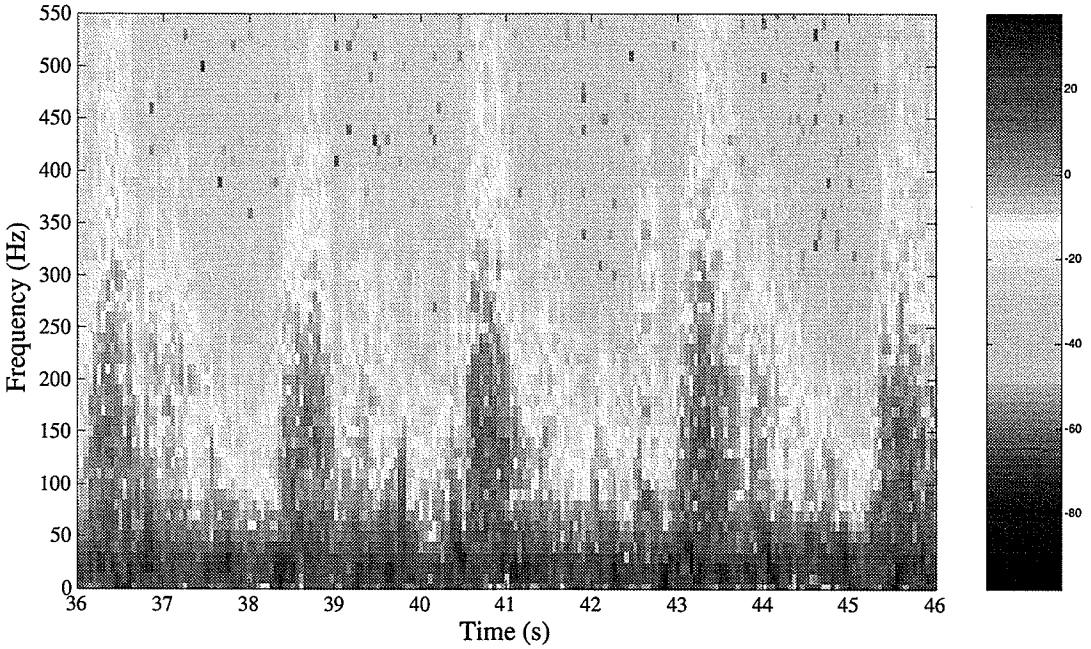
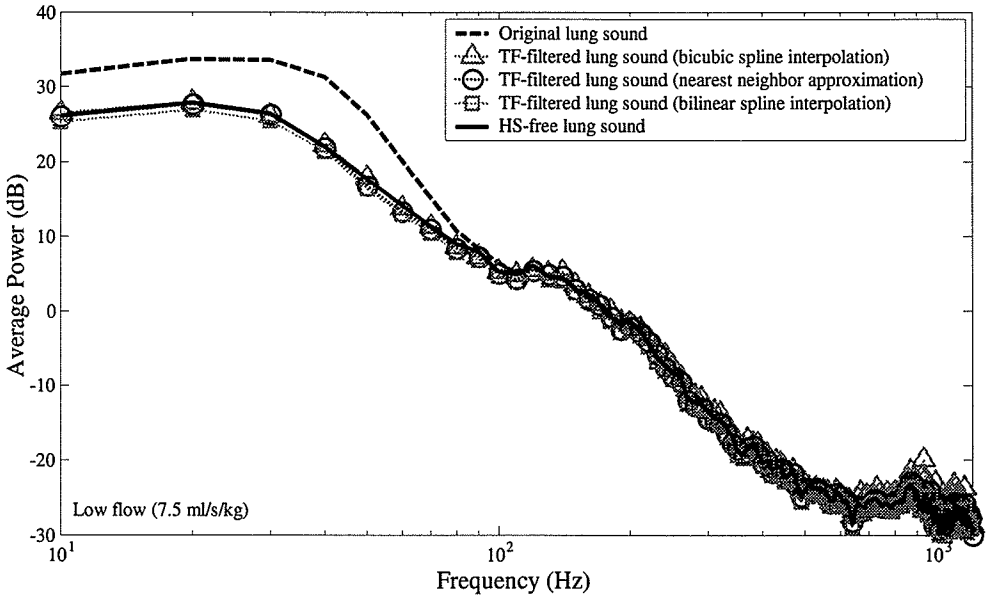


Figure 4.7 Spectrogram of the TF-filtered of the signal of Fig. 3.4.

Statistical analysis, as illustrated in Table 4.5, verified that there was no significant difference between average PSD of HS-free original lung sounds and TF-filtered signal using bilinear spline interpolation for all frequency bands at both low and medium flow rates. These results confirmed the hypothesis that the average PSD of TF-filtered signal is close to the average PSD of lung sound signal free of HS. Table 4.5 elaborates on the result of the statistical test.

A



B

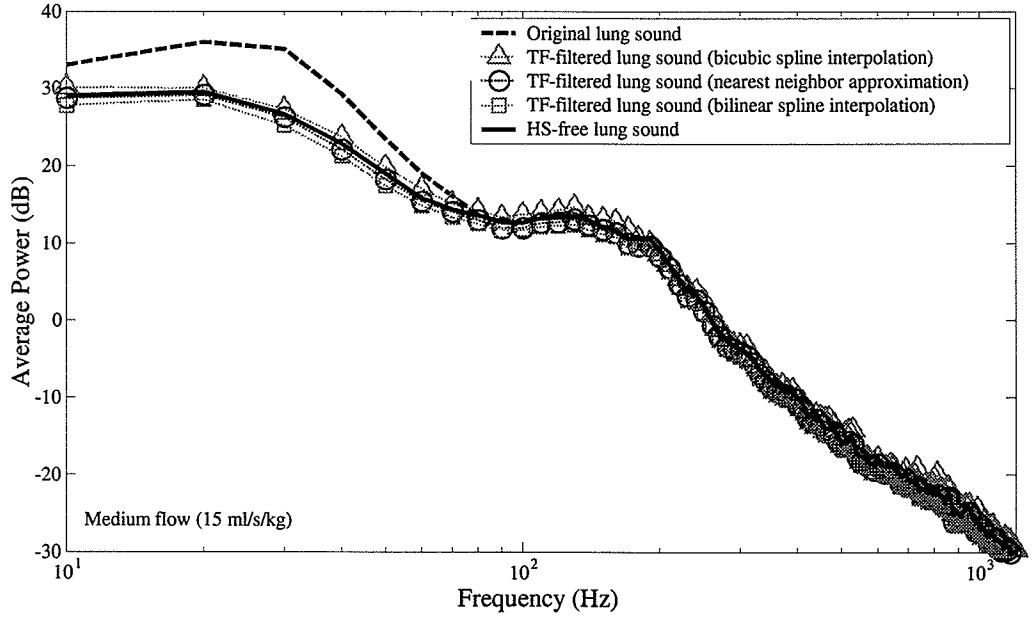


Figure 4.8 Comparison between the average PSD of the original lung sound signal including HS, TF-filtered signal using different interpolation approaches, and original lung sound signal free of HS at low (A) and medium (B) flow rates for a typical subject.

Flow Rate	20-40 Hz	40-70 Hz	70-150 Hz	150-300 Hz
Low (7.5 ml/s/kg)	0.94 ± 0.58	0.94 ± 0.4	0.95 ± 0.33	0.79 ± 0.29
Medium (15 ml/s/kg)	0.82 ± 0.31	0.78 ± 0.38	0.8 ± 0.22	0.82 ± 0.17

Table 4.5 Difference ($\mu\pm\sigma$, in dB) between average power spectra of the TF-filtered and original lung sound signal free of HS, at low and medium flow rates and four frequency bands (p>0.05).

Although in the proposed method, the HS-included segments are entirely removed and reconstructed by applying bilinear spline interpolation in time-frequency plane, the power spectrum estimation still remains asymptotically unbiased. Recall that if the segments of a signal are considered stationary, a linear combination of those segments is also stationary. Therefore, it can be concluded that the reconstructed segments by applying bilinear spline interpolation are also stationary. In this study, the power spectrum was estimated using the Welch method and calculated as:

$$p_{xx}^{W}(f)=\frac{1}{L}\sum_{i=0}^{L-1}\tilde{P}_{xx}^{(i)}(f), \tag{4.1}$$

where L is the number of segments and $\tilde{P}_{xx}^{(i)}(f)$ is the estimated periodogram of the ith segment [Proakis et al., 1996]. The mean value of the Welch estimate is:

$$E[p_{xx}^{W}(f)]=\frac{1}{L}\sum_{i=1}^{L}E[\tilde{P}_{xx}^{(i)}(f)]=\frac{1}{L}\times L\times E[\tilde{P}_{xx}^{(i)}(f)]=E[\tilde{P}_{xx}^{(i)}(f)]. \tag{4.2}$$

As it can be observed from Eq. 4.2, the mean value of the Welch power spectrum estimate is independent of the number of segments (i.e., L). Therefore, although some of the segments are eliminated in the proposed method, the estimation still remains asymptotically unbiased. In the case of 50% overlap between successive segments, the

variance of the Welch power spectrum estimation with the Bartlett (triangular) window is given as:

$$\text{var}[p_{xx}^{W}(f)] \approx \frac{k}{L}\Gamma_{xx}^{2}(f), \qquad k = \frac{9}{8}, \tag{4.3}$$

where $\Gamma_{xx}^{2}(f)$ is the Fourier transform of the autocorrelation function of the signal [Proakis et al., 1996]. If another data window (i.e. Hanning window) were used, only the constant k in the equation would change. As Eq. 4.3 indicates, the variance of the Welch power spectrum increases as the number of segments, L, decreases. Therefore we can expect by removing a number of HS-included segments, the variance of the power spectrum estimation of the reconstructed signal would increase slightly. Thus the type of HS localization technique should be carefully selected.

Fig. 4.9 illustrates the original and-TF filtered signals in the time domain (low and medium flow rates, respectively). Comparing the signals before and after filtering suggests that HS-free portions of data have not been altered by the proposed HS cancellation method.

This method relies on the stationarity of lung sounds over a few segments, as it uses interpolation between the adjacent segments of the HS-included segments. In cases where HS occurs in the vicinity of the breath onset, stationarity is not preserved. To investigate its consequences these cases were detected manually and the impact of non-stationarity on signal estimation by interpolation was tested. Fig. 4.10 and Fig. 4.11 demonstrate two typical examples of these cases. As can be observed from the graphs in Fig. 4.10 and Fig. 4.11, two adjacent HS-included segments have occurred in the vicinity of the breath onset when respiratory phase changes, one from expiratory to inspiratory and the other from inspiratory to expiratory.

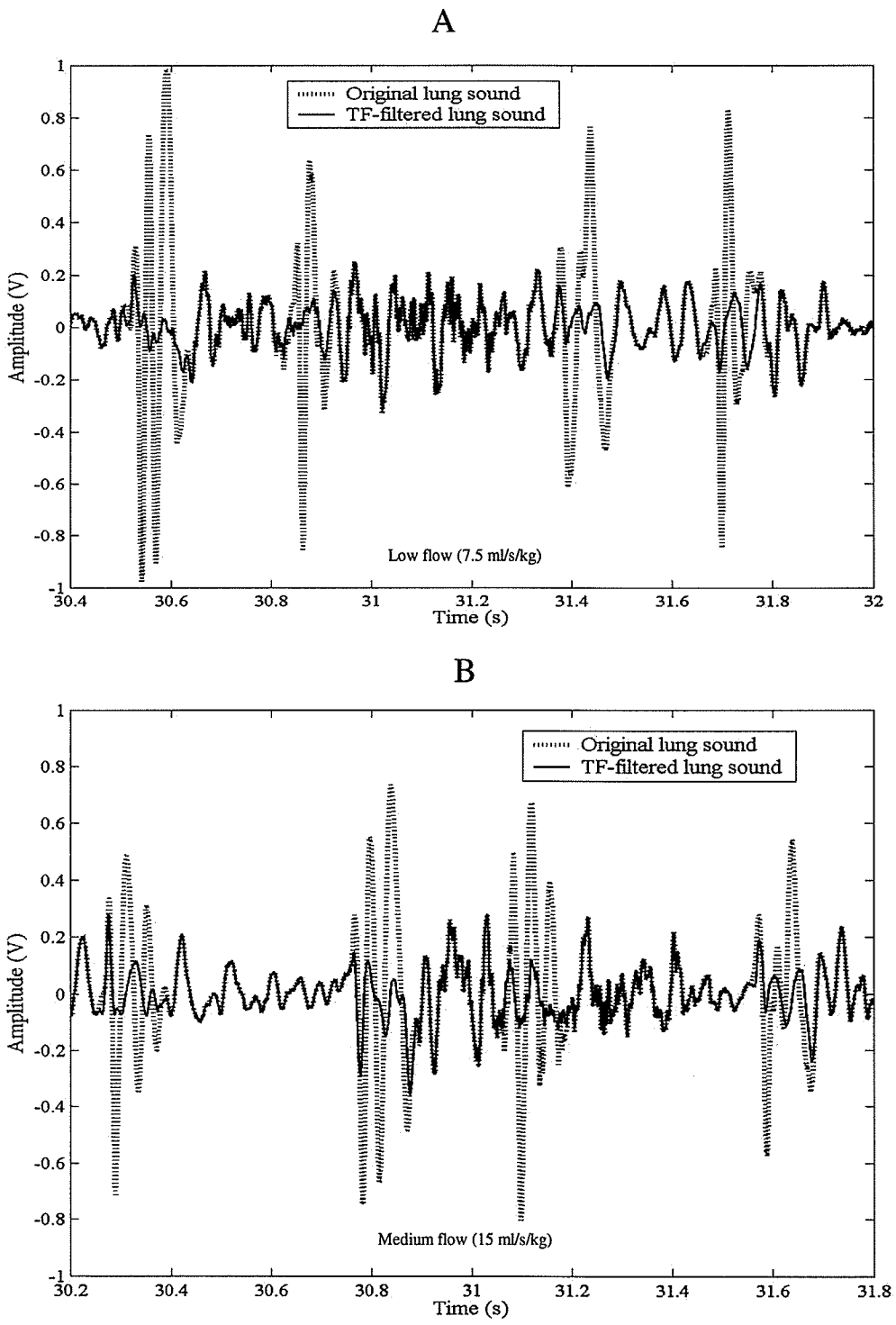


Figure 4.9 Time domain original lung sound record and TF-filtered signal, at low (A) and medium (B) flow rates, for a typical subject.

In order to analyze the impact of non-stationarity on signal estimation by interpolation the spectrogram filtering-based method was applied to the aforementioned cases and the average power of the two adjacent HS-included segments, which had occurred in the vicinity of breath onset, was compared before and after TF-filtering. As it can be observed from Fig. 4.12 and Fig. 4.13, the average power of these segments after TF-filtering is reduced at low frequencies which was expected. There is also a slight change at high frequencies since the filtered segments are estimated by interpolating the adjacent HS-free segments from different respiratory phases. In cases when the respiratory phase changes from expiratory to inspiratory, the average power of the estimated segments will increase slightly at high frequencies. On the other hand when the respiratory phase changes from inspiratory to expiratory, there will be a slight decrease in the average power of the estimated segments at high frequencies.

The spectrogram filtering-based method has shown promising results mainly in terms of the preservation of lung sound characteristics and almost complete removal of HS. The use of TF-filtering based on the short-time Fourier transform did not add any noticeable clicks or artifacts in the reconstruction of the signal. Manual inspection by visual and auditory means and quantitative analysis of the reconstructed signals confirmed that lung sounds were the dominant sounds with no perceptible heart sounds in the background. Furthermore, the proposed method is far more efficient than other methods such as ANC-RLS [Gnitecki et al., 2003] and ANC-FOS [Hadjileontiadis et al., 1997] for HS cancellation in terms of computational load and speed.

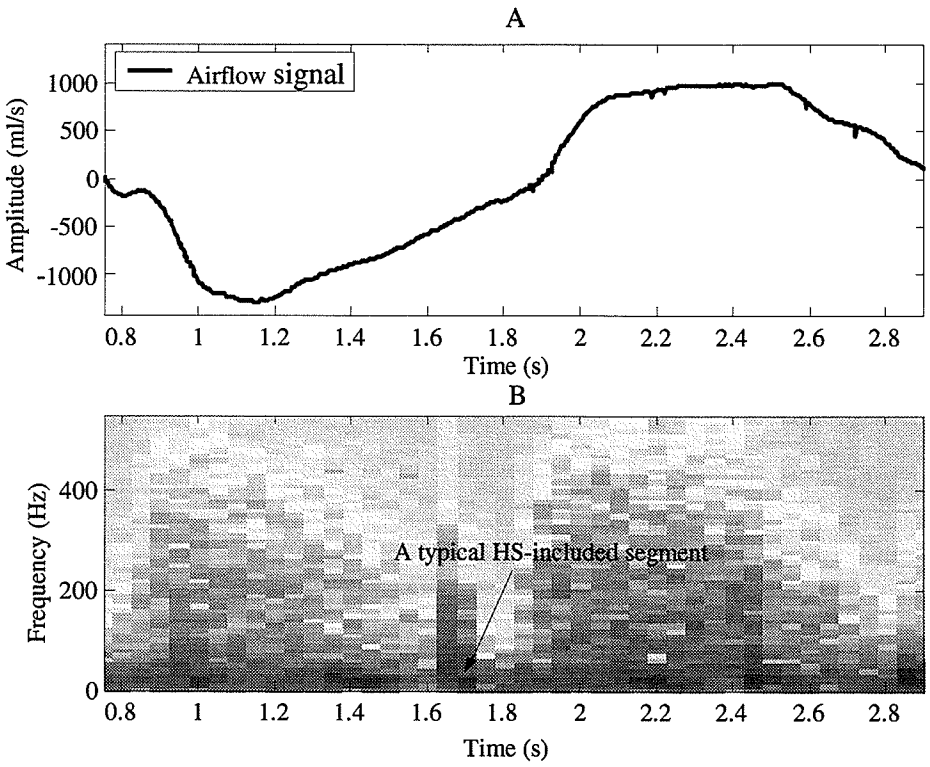


Figure 4.10 Airflow signal (A), spectrogram of original lung sound signal (B), where respiratory phase changes from expiratory to inspiratory.

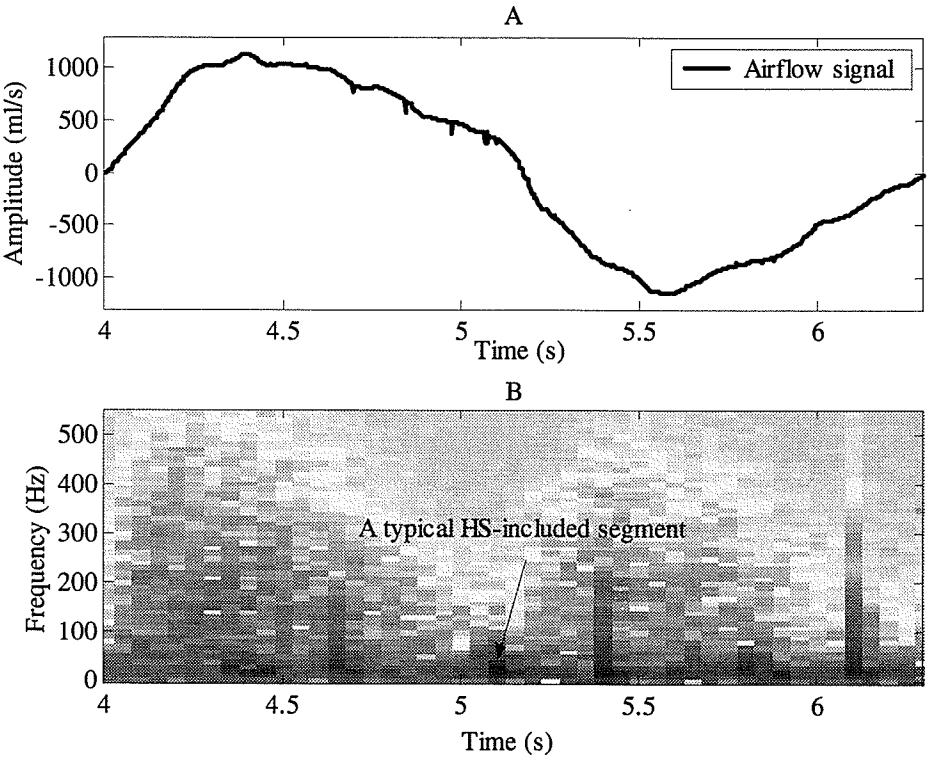


Figure 4.11 Airflow signal (A), spectrogram of original lung sound signal (B), where respiratory phase changes from inspiratory to expiratory.

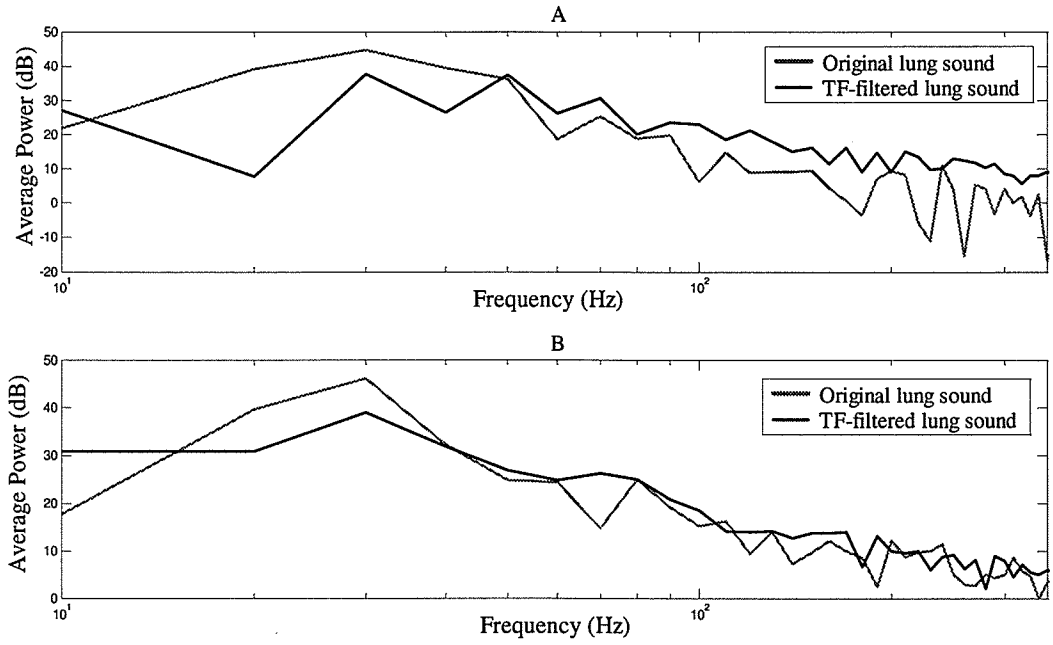


Figure 4.12 Comparison between the average power of two adjacent HS-included segments of lung sound signal in Fig. 4.10, which have occurred at the vicinity of breath onset, before and after TF-filtering. A) the average power of the first segment; B) the average power of the second segment.

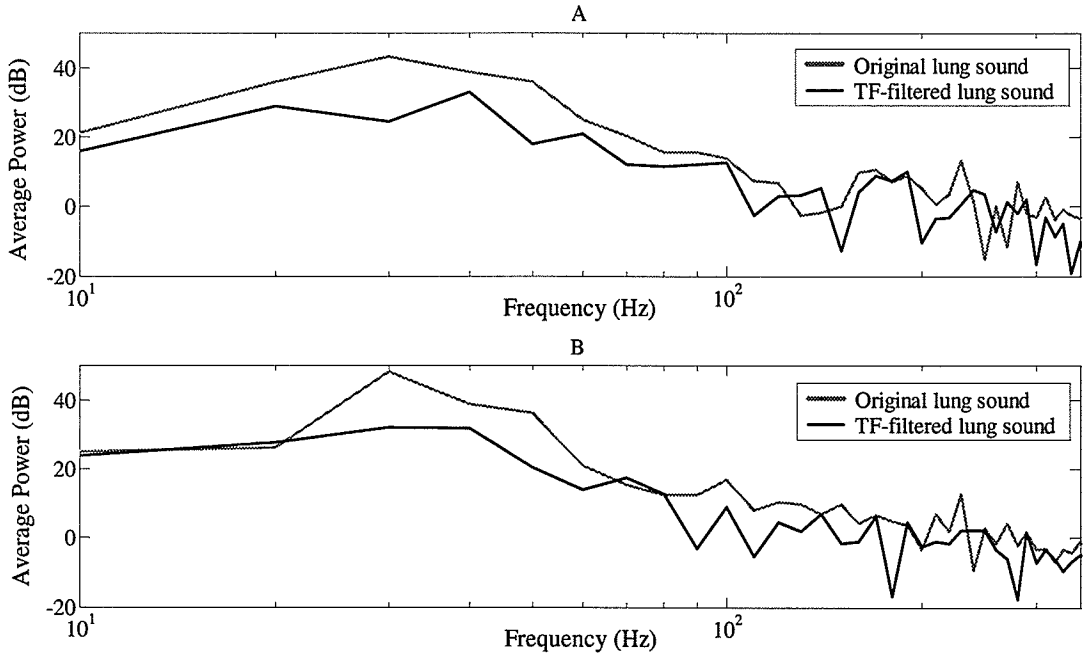


Figure 4.13 Comparison between the average power of two adjacent HS-included segments of lung sound signal in Fig. 4.11, which have occurred at the vicinity of breath onset, before and after TF-filtering. A) the average power of the first segment; B) the average power of the second segment.

4.4 Chapter Summary

This chapter presented the results of this study and provided further discussions on the efficiency of the proposed methods. The performance of HS localization techniques were evaluated and compared with each other. The wavelet coefficient-based technique was found superior than the others in terms of showing the most accuracy for HS-included segments detection.

The results of spectrogram-ICA based and spectrogram filtering-based methods were reported in time and frequency domain. The average power of result signals were compared with the average power of HS-free original lung sounds over four frequency bands. In addition, the results of spectrogram filtering-based method were statistically examined. The results show that the spectrogram ICA-based method is promising in term of HS reduction from lung sound recordings but it also changes the original lung sounds. On the other hand the spectrogram filtering-based method successfully removed HS from lung sound signals, while preserving the original fundamental components of the lung sounds. The computational cost and the speed of both proposed methods were found to be much more efficient than other HS reduction methods. The following chapter concludes on this study and provides some recommendations for future research in this area.

CHAPTER 5

CONCLUSION & RECOMMENDATIONS

5.1 Conclusion Remarks

The heart and lung signals overlap in the time and frequency domain, therefore removing heart sounds interference from respiratory sound recordings is a challenging task. This thesis presented two novel methods to cancel HS from lung sound recordings in TF domain, using spectrogram ICA-based technique and spectrogram filtering-based method in TF domain along with three new techniques for HS localization in respiratory sound recordings.

The spectrogram ICA-based method implements the ICA algorithm independently at each frequency on the spectrogram of two lung sound signals recorded simultaneously from two different locations on the anterior chest, and yields the independent components at that frequency. Then the proper independent components from each frequency is chosen and combined with each other to produce the spectrogram of separated signals. Afterwards, using ISTFT, the separated signals are reconstructed in the time domain. The difference between the average PSD of the estimated result signals and that of original HS-free recorded signals was examined for different locations on the chest over four frequency bands from 20 to 300 Hz for the ±20% of the target flow, at low and medium flow rate. The results showed some HS reduction but the method was not able to remove HS from lung sounds completely. Furthermore the separated lung sounds showed some differences with the original lung sounds free of HS.

HS localization procedure is required for performing the spectrogram filtering-based method. In this study three different HS localization algorithms were suggested: average power-based technique, wavelet coefficient-based technique and second order statistical-based technique. The performance of proposed HS localization techniques was quantitatively evaluated and compared with each other. It was found that the wavelet coefficient-based technique is superior than the others in terms of showing the most accuracy for HS-included segments detection.

Afterwards, the spectrogram filtering-based method was applied to remove the detected HS-included segments from the spectrogram of the recorded lung sound signal and estimates the missing data via a 2D interpolation in the TF domain. Finally, the signal is reconstructed into the time domain. To assess the efficiency of the spectrogram filtering-based method, in addition to the qualitative manual validation by visual and auditory means, quantitative analysis was also performed by comparing the average PSD of the filtered signals over four frequency bands from 20 to 300 Hz with that of the original lung sound segments with and without HS. The results showed no significant difference between average PSD of the HS-free original lung sounds and the TF-filtered signal using bilinear spline interpolation for all frequency bands at both low and medium flow rates. This proves the hypothesis that the method was successful in terms of HS removal without hampering the main components of the lung sounds.

Overall, by visual and auditory inspections it was found that although the spectrogram ICA-based method could not completely cancel the HS from lung sound recordings but showed promising results in term of HS reduction at the cost of altering lung sound slightly. On the other hand, the spectrogram filtering-based method

successfully removed HS from lung sound signals, while preserving the original fundamental components of the lung sounds. The computational cost and the speed of both proposed methods were found to be much more efficient than other HS reduction methods such as ANC-RLS [Gnitecki et al., 2003] and ANC-FOS [Hadjileontiadis et al., 1997].

5.2 Suggestions for Future Work

Although the spectrogram filtering-based method is reliable in HS cancellation from lung sound recordings and fulfills the goals of this study, more research, testing and refining are still required. The performance of the spectrogram filtering-based method relies on the quality of HS localization procedure. Therefore to minimize any inconsistency, the performance of HS localization technique should be improved by combining the proposed HS localization techniques. Although this method is computationally efficient, further development is still needed to optimize it for real time applications.

This study paved the way towards application of spectrogram ICA-based technique in respiratory sound analysis. Due to the inherent potential of spectrogram ICA-based method, further study and research is recommended to improve its performance in respiratory acoustics. Since the spectrogram ICA-based method will not work if all the reflections are not included within a segment [Murata et al., 2001], it is important to know more about the statistics of the respiratory sounds and heart sounds and also the reflection of the lung tissues to select a suitable sampling rate and proper segment length for calculating the spectrogram of the respiratory signals. For better performance, trying on more than two simultaneous recordings from different locations on the chest is also

suggested. Due to the fact that the spectrogram ICA-based method estimates the source signals from the observed mixed signals on the skin, this method may provide novel clinical information for physicians. Hence, further investigations on patient subjects whose respiratory sounds encompass the adventitious sounds such as crackle and wheezes will be beneficial.

REFERENCES

[Abella et al., 1992]

Abella, M., Formolo, J., Penney, D.G.,"Comparison of the acoustic properties of six popular stethoscopes," *J. Acoust. Soc. Am.*, vol. 91, pp. 2224-28, 1992.

[Cardoso et al., 1993]

Cardoso, J.F., Souloumiac, A., "Blind beamforming for non Gaussian signals," *IEE Proc-F*, vol. 140, no. 6, pp. 362-70, 1993.

[Cardoso, 1999]

Cardoso, J.F., "High-order contrasts for independent component analysis," *Neural Computation*, vol. 11, no. 1 , pp. 157-92, 1999.

[Charleston et al., 1996]

Charleston, S., Azimi-Sadjadi, M.R., "Reduced order Kalman filtering for the enhancement of respiratory sounds," *IEEE Trans. Biomed. Eng.*, vol. 43, no. 4, pp. 421-24, Apr. 1996.

[Cheney et al., 1999]

Cheney, W., Kincaid, D., *Numerical Mathematics and Computing.* 4th edition, Books/Cole Publishing Company, ch. 7, 1999.

[Cichocki et al., 2002]

Cichocki, A., Amari, S., *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. New York, NY: John Wiley & Sons, Inc., ch. 1, 2002.

[Gavriely et al., 1995]

Gavriely, N., Nissan, M., Rubin, A.H.E., Cugell, D.W., "Spectral characteristics of chest wall breath sounds in normal subjects," *Thorax*, vol. 50, pp. 1292-300, 1995.

[Gnitecki et al., 2003]
Gnitecki, J., Moussavi, Z., Pasterkamp, H., "Recursive least squares adaptive noise cancellation filtering for heart sound reduction in lung sounds recordings," *Proc. 25th Ann. Int. Conf. IEEE EMBS*, pp. 2416-19, Sep. 2003.

[Guangbin et al., 1992]
Guangbin, L., Shaoqin, C., Jingming, Z., Jinzhi, C., and Shengju, W., "The development of a portable breath sounds analysis system," *Proc. 14th Ann. Int. Conf. IEEE EMBS*, pp. 2582-83, 1992.

[Hadjileontiadis et al., 1997]
Hadjileontiadis, L.J., Panas, S.M., "Adaptive reduction of heart sounds from lung sounds using fourth-order statistics," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 7, pp. 642-48, Jul. 1997.

[Hadjileontiadis et al., 1998]
Hadjileontiadis, L.J., Panas, S.M., "A wavelet-based reduction of heart sound noise from lung sounds," *Int. J. Med. Infor.*, vol. 52, pp. 183-190, 1998.

[Hidalgo et al., 1991]
Hidalgo, H.A., Wegmann, M.J. Waring, W., "Frequency spectra of normal breath sounds in childhood," *Chest*, vol. 100, 999-1002, 1991.

[Hossain et al., 2003]
Hossain, I., Moussavi, Z., "An overview of heart-noise reduction of lung sound using wavelet transform based filter," *Proc. 25th Ann. Int. Conf. IEEE EMBS*, pp. 458-61, Sep. 2003.

[Hyvarinen et al., 2001]
Hyvarinen, A., Karhunen, J., Oja, E., *Independent Component Analysis*. John Wiley & Sons, Inc., ch. 7 and 13, 2001.

[Iyer et al., 1986]
Iyer, V.K., Ramamoorthy, P.A., Fan, H., Ploysongsang, Y., "Reduction of heart sounds from lung sounds by adaptive filtering," *IEEE Trans. Biomed. Eng.*, vol. 33, no. 12, pp. 1141-48, Dec. 1986.

[Jain, 1989]
Jain, A.K., *Fundamentals of Digital Image Processing*. Prentice Hall, Englewood Cliffs, NJ, 1989.

[Kanga et al., 1986]
Kanga, J.F., Kraman, S.S., "Comparison of the lung sound frequency spectra of infants and adults," *Pediatr. Pulmonol*, vol. 2, pp. 292-95, 1986.

[Kompis et al., 1992]
Kompis, M., Russi, E., "Adaptive heart-noise reduction of lung sounds recorded by a single microphone," *Proc. 14th Ann. Int. Conf. IEEE EMBS*, pp. 691-92, 1992.

[Kraman et al., 1995]
Kraman, S.S., Wodicka, G.R., Oh, Y., Pasterkamp, H., "Measurement of respiratory acoustic signals: Effect of microphone air cavity width, shape and venting," *Chest*, vol. 108, pp. 1004-08, 1995.

[Lee, 2000]
Lee, T.W., *Independent Component Analysis: Theory and Applications.* 2th edition, Kluwer Academic Publishers, Netherlands, ch. 2 and 4, 2000.

[Lewicki et al., 1998]
Lewicki, M., Sejnowski, T.J., "Learning nonlinear overcomplete representations for efficient coding," In M. I. Jordan, M. J. Kearns, and S. A. Solla, editors, MA, 1998. MIT Press, 1998.

[Mersereau et al., 1974]
Mersereau, R.M., Oppenheim, A.V., "Digital Reconstruction of Multidimensional Signals From their Projections," *Proc. IEEE*, vol. 62, no. 10, pp. 1319–38, 1974.

[Mewett et al., 2001]
Mewett, D.T., Nazeran, H., Reynolds, K.J., "Removing power line noise from recorded EMG," *Proc. 23th Ann. Int. Conf. IEEE EMBS*, pp. 2190-93, 2001.

[Murata et al., 2001]
Murata, N., Ikeda, S., Ziehe, A., "An approach to blind source separation based on temporal structure of speech signals," *Neurocomputing*, vol. 41, pp. 1-24, 2001.

[Nawab et al., 1988]
Nawab, H., Quatieri, T.F., "Short-Time Fourier Transform," In Chapter 6 of *Advanced Topics in Signal Processing*, Lim, J.S., Oppenheim, A.V., (Eds.), Prentice Hall, NJ, pp. 289-337, 1988.

[Pasterkamp et al., 1985]
Pasterkamp, H., Fenton, R., Tal, A., and Chernick, V., "Interference of cardiovascular sounds with phonopneumography in children," *Am. Rev. Respir. Dis.*, vol. 131, no. 1, pp. 61-64, Jan. 1985.

[Pasterkamp et al., 1993]
Pasterkamp, H., Kraman, S.S., DeFrain, P.D., Wodicka, G.R., "Measurement of respiratory acoustical signals: Comparison of sensors, " *Chest*, vol. 104, pp. 1518-25, 1993.

[Pasterkamp et al., 1997]
Pasterkamp, H., Kraman, S.S., Wodicka, G.R., "Respiratory sounds: Advances beyond the stethoscope," *Am. J. Respir. Crit. Care Med.*, vol. 156, no. 3.1, pp. 974-87, 1997.

[Pourazad et al., 2003]
Pourazad, M.T., Moussavi, Z., Thomas, G., "Heart sound cancellation from lung sound recordings using adaptive threshold and 2D interpolation in time-frequency domain," *Proc. 25th Ann. Int. Conf. IEEE EMBS*, pp. 2586-89, 2003.

[Proakis et al., 1996]
Proakis, J.G., Manolakis, D.G., *Digital Signal Processing: Principles, Algorithms, and Applications.* 3th edition, Printice-Hall, Upper Saddle River, NJ, ch.12, 1996.

[Rushmer, 1970]
Rushmer, R.F., *Cardiovascular Dynamics*, 3rd ed. Philadelphia: Saunders, 1970.

[Schuttler et al., 1996]
Schuttler, F., Penzel, T., Wichert, P., "Digital recording and computer-based analysis of lung sounds," *Proc. 18th Ann. Int. Conf. IEEE EMBS*, pp. 2301-02, 1996.

[Son et al., 2001]
Son, J.S., Thomas, G., Flores, B.C., *Range-Doppler Radar Imaging and Motion Compensation.* Artech House Publishers, Boston. London, ch. 9 and 13 and software, 2001.

[Spath, 1995]
Spath, H., *Two Dimensional Spline Interpolation Algorithms.* A K Peters, Ltd. Wellesley, Massachusetts, ch. 2 and 4, 1995.

[Vannuccinni et al., 2000]
Vannuccinni, L., Earis, J., Helisto, P., Cheetham, B., Rossi, M., Sovijarvi, A.R.A., Vanderschoot, J., "Capturing and preprocessing of respiratory sounds," *Eur. Respir. Rev.,* vol. 10 (77), pp. 616-20, 2000.

[Webster, 1998]
Webster, J.G., *Medical Instrumentation: Application and Design*. New York, NY: John Wiley & Sons, Inc., ch. 7, 1998.

[Wodicka et al., 1994]
Wodicka, G.R., Kraman, S.S., Zenk, G.M., Pasterkamp, H., "Measurement of respiratory acoustic signals: Effect of microphone air cavity depth," *Chest*, vol. 106, pp. 1140-44, 1994.

[Yang-Sheng et al., 1988]
Yang-Sheng, L., Wen-Hui, L., Guang-Xia, Q., "Removal of the heart sound noise from the breath sound," *Proc. 10th Ann. Int. Conf. IEEE EMBS*, pp. 175-76, 1988.

[Yip et al., 2001]
Yip, L., Zhang, Y.T., "Reduction of heart sounds from lung sound recordings by automated gain control and adaptive filtering techniques," *Proc. 23th Ann. Int. Conf. IEEE EMBS*, pp. 2154-56, 2001.

[Ziehe, 1998]
Ziehe, A., *Statistische Verfahren zur Signalquellentrennung*, Master's Thesis, Humboldt Universitat, Berlin, 1998.

[Ziehe et al., 2000]
Ziehe, A., Muller, K.-R., Nolte, G., Mackert, B.-M., Curio, G., "Artifact reduction in magnetoneurography based on time-delayed second-order correlations," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 1, pp. 75-87, 2000.

APPENDIX A

A.1 Centering

Centering is one of the most basic and necessary preprocessing stages before applying an ICA algorithm on the observed data. Observation matrix $X(t)$ can always be centered by subtracting the mean vector $m = E\{X\}$ from that [Hyvarinen et al., 2001]. This implies that source matrix S is zero-mean as well, since by taking expectations on both sides of Eq. 3.5:

$$X = A * S \Rightarrow E\{X\} = A * E\{S\}, \tag{1}$$

$$X - E\{X\} = A * S - A * E\{S\} = A * (S - E\{S\}). \tag{2}$$

This preprocessing is made solely to simplify the ICA algorithms: It does not mean that the mean could not be estimated. After estimating the filter matrix A with centered data, we can complete the estimation by adding the mean vector of S back to the centered estimates of S [Hyvarinen et al., 2001]. The mean vector of S is given by $A^{-1}m$, where m is the mean that was subtracted for centering.

A.2 Whitening

Another useful preprocessing strategy in ICA is to first whiten the observation matrix, $\hat{X}$. This means that before the application of the ICA algorithm $\hat{X}$ to be transformed *linearly* so that, a new vector $\hat{\hat{X}}$, which is white (i.e., its components are uncorrelated and their variances equal unity), is obtained [Hyvarinen et al., 2001]. In other words, the covariance matrix of $\hat{\hat{X}}$ equals the identity matrix:

$$E\left\{\hat{\hat{X}}\hat{\hat{X}}^T\right\} = I, \tag{3}$$

The whitening transformation is always possible. One popular method for whitening is to implement the Eigen-Value Decomposition (EVD) of the covariance matrix $E\left\{\hat{X}\hat{X}^T\right\} = EDE^T$, where E is the orthogonal matrix of eigenvectors of $E\left\{\hat{X}\hat{X}^T\right\}$ and D is the diagonal matrix of its eigenvalues, $D = diag(d_1,...,d_n)$. Whitening can now be done by:

$$\hat{\hat{X}} = ED^{-\frac{1}{2}}E^T\hat{X}, \tag{4}$$

where the matrix $D^{-\frac{1}{2}}$ is computed by a simple component-wise operation as: $D^{-\frac{1}{2}} = diag\left(d_1^{-\frac{1}{2}},...,d_n^{-\frac{1}{2}}\right)$. Whitening transforms the mixing matrix into a new one, $\hat{\hat{A}}$:

$$\left.\begin{aligned}\hat{\hat{X}} &= ED^{-\frac{1}{2}}E^T\hat{X}\\ \hat{X} &= \hat{A}\hat{S}\end{aligned}\right\} \Rightarrow \hat{\hat{X}} = ED^{-\frac{1}{2}}E^T\hat{A}\hat{S} = \hat{\hat{A}} * \hat{S} \Rightarrow \hat{\hat{X}} = \hat{\hat{A}}\hat{S}, \tag{5}$$

The utility of whitening resides in the fact that the new mixing matrix $\hat{\hat{A}}$ is orthogonal. This can be seen from:

$$E\left\{\hat{\hat{X}}\hat{\hat{X}}^T\right\} = \hat{\hat{A}}E\left\{\hat{S}\hat{S}^T\right\}\hat{\hat{A}}^T = \hat{\hat{A}}\hat{\hat{A}}^T = I. \tag{6}$$

Since $\hat{\hat{A}}$ is orthogonal, the number of parameters to be estimated is reduced, i.e., instead of estimating the n^2 parameters that are the elements of the original matrix $\hat{A}$, we only need to estimate the new, orthogonal mixing matrix $\hat{\hat{A}}$ with $n(n-1)/2$ elements [Hyvarinen et al., 2001]. In large dimensions, an orthogonal matrix contains only about half of the number of parameters of an arbitrary matrix. Thus one can say that whitening solves half of the problem of ICA and reduce the complexity of the problem.

A.3 Overlap Selection Criteria

There are two parameters that have to be defined for calculating the spectrogram of the signal: segment length (l) and overlap (ov). In spectrogram ICA-based method, there is also another parameter (r), which is the number of matrices to be diagonalized simultaneously, that should be determined as well. The length of the segment is decided based on the stationarity duration (St) of the signal and is independent of other two parameters. In return, as can be observed from the following criteria ov and r are strongly related to each other [Murata et al., 2001]:

- $r \times (l - ov) < St,$
- $r \geq 2,$
- $(l - ov) \times (2M + 1) < St, \qquad M \geq 1.$

where M in Eq.3.19 is the number of samples contributes in defining moving average operator $\Im$, for estimating the time series envelope.

Assuming St=100 ms, l=100 ms, it was found the overlap should be more than 66 ms. Considering the fact small r gives an unstable solution and large r leads to a wrong solution [Murata et al., 2001], the following values were assigned to the parameters in this study:

ov= 85 ms,

r= 6,

M= 2.