MEASUREMENT AND ANALYSIS

OF BREATH SOUNDS

YIHONG QIU

A THESIS SUBMITTED TO
THE DEPARTMENT OF MECHANICAL ENGINEERING OF
THE FACULTY OF ENGINEERING OF
THE UNIVERSITY OF GLASGOW FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

To my family

Acknowledgements

I would like to thank the following people who made this thesis possible. First and foremost I would like to express my deep gratitude to my supervisor Dr. Arthur Whittaker for his continuous guidance and help. I'm deeply indebted to Dr. Margaret Lucas for her valuable advice, help, and review of this thesis which improved technical quality. I would like to express my sincere thanks to Drs. Kenneth Anderson, Kevin Rooney, Magnus Garrioch and Mr. John Sandbach for their collaborations.

I would also like to thank electronics technicians Ian Russell, Gordon Hicks and Bernard Hoey for their help. My thanks also go to Dr. Zhihong Huang, Andrea Cardoni and Mark Davie for their sharing of lab devices. I am grateful to Annabel and David Blenkinson for their friendship and proof-reading part of this thesis. Thanks also go to all the other people who have helped me during my study at the University of Glasgow.

I would like to extend my thanks to my parents, my brother and his family for their encouragement and support. Finally, special thanks go to my family: my husband Chunming and son Jiasheng, for their love and strong support.

Contents

List of figures

List of Tables

Abstract

Previous research has shown that computer-aided breath sounds analysis can be a non-invasive objective method to help in evaluating respiratory system conditions. However, some problems need to be solved. One is that investigators use various measurement systems without calibration, thus it is impossible to compare results directly. Another is that there is not yet a reliable means of quantification of adventitious sounds. For continuous adventitious sounds—wheeze, which is a sign of airway obstruction, the previous automatic detection algorithms are not reliable enough. In addition, those algorithms have only been validated by subjective methods.

This thesis aims at investigating these problems. Existing breath sound measurement systems and possible new methods have been critically investigated. The frequency response of each part of the measurement system has been studied. Emphasis has been placed on frequency response of acoustic sensors; especially, a method to study a diaphragm type air-coupler in contact use has been proposed. Two new methods of breath sounds measurement have been studied: laser Doppler vibrometer and mobile phones. It has been shown that these two methods can find applications in breath sounds measurement, however there are some restrictions.

A reliable automatic wheeze detection algorithm based on auditory modelling has been developed. That is the human's auditory system is modelled as a bank of band pass filters, in which the bandwidths are frequency dependent. Wheezes are

treated as signals additive to normal breath sounds (masker). Thus wheeze is detectable when it is above the masking threshold. This new algorithm has been validated using simulated and real data. It is superior to previous algorithms, being more reliable to detect wheezes and less prone to mistakes.

Simulation of cardiorespiratory sounds and wheeze audibility tests have been developed. Simulated breath sounds can be used as a training tool, as well as an evaluation method. These simulations have shown that, under certain circumstance, there are wheezes but they are inaudible. It is postulated that this could also happen in real measurements. It has been shown that simulated sounds with predefined characteristics can be used as an objective method to evaluate automatic algorithms.

Finally, the efficiency and necessity of heart sounds reduction procedures has been investigated. Based on wavelet decomposition and selective synthesis, heart sounds can be reduced with a cost of unnatural breath sounds. Heart sound reduction is shown not to be necessary if a time-frequency representation is used, as heart sounds have a fixed pattern in the time-frequency plane.

Chapter 1 Introduction

1.1 Background

Since Laënnec invented the stethoscope and published his book describing the meaning of sounds, auscultation has been one of the routine clinical methods to assess pulmonary situations. Doctors can distinguish between normal and abnormal breath sounds characteristics depending on their experience. The advantage of auscultation is that it is quick, non-invasive, and needs minimum co-operation, thus it is especially useful for all those who could not perform conventional respiratory function tests. However, there are a few drawbacks. It lacks a method of recording, has insufficient sensitivity, offers no quantitative description, and is prone to observer variability.

Early objective measurement and analysis of breath sounds was attempted in the 1920s. More recently, the success of radiography, which can provide more accurate information of structural abnormalities in the lung, led to the decline in the status of auscultation (Forgacs 1978). However, obstructive diseases which affect the airways are not easily diagnosed by radiography (Druzgalski *et al.* 1980). The research of Forgacs (1967; 1969; 1971; 1978) intrigued many investigators. With the development of electronic devices, computer technology, and signal processing, investigators are active in improving the understanding of the mechanisms for the production of breath sounds and developing automatic systems to classify breath sounds. However, there are some problems that need to be solved. One is that investigators use various measurement systems without calibration, thus it is impossible to compare results directly. Another is that

quantification of adventitious sound is important. For a specific type—wheeze, which is a sign of airway obstruction, the previous automatic detection algorithms are not reliable enough. In addition, those algorithms are only validated by subjective methods—visual inspection of expanded waveform and/or listening.

1.2 Aims and Objectives

The aims of this project are:

- to critically investigate existing breath sounds measurement systems and possible new methods,
- to develop a reliable automatic wheeze detection algorithm,
- to investigate the necessity of heart sounds reduction,
- to develop breath sound simulation procedures,
- to validate the wheeze detection algorithm using real and simulated data.

Investigating the performance of a breath sounds measurement system is important. Each part of a system may distort the signal. Studying the frequency response of each part can help in choosing suitable devices. The frequency response can also be used to calibrate a measured signal for comparison purposes.

Two new methods have been investigated. One is the laser Doppler vibrometer, which is a non-contact sensor. The other is the mobile phone.

Automatic wheeze detection algorithms can provide a parametric description of wheeze characteristics, such as frequency, duration, number etc. These values

have the potential to be incorporated in a patient's record, which will facilitate patient monitoring and management.

1.3 Outline of the Thesis

Previous research is briefly reviewed in chapter 2. Stress is put on work which is closely related to this thesis. The methodology to study frequency characteristics of each part of a measurement system is presented in chapter 3. The focus is on the acoustic sensors' frequency responses. Heart sounds reduction using wavelet decomposition and selective synthesis is described in chapter 4. Simulation of cardiorespiratory sounds and wheeze audibility tests are discussed in chapter 5. An automatic wheeze detection algorithm based on auditory modelling is proposed in chapter 6. Both simulated data and real data have been used to validate the algorithm.

1.4 Originality

In chapter 3, measurements using LDV and mobile phones, study of contact frequency response of a diaphragm type air-coupler, and added noise to improve mobile phone recording sound quality are all original. In chapter 5, simulation of wheezes and wheeze audibility tests are novel. The automatic wheeze detection algorithm in chapter 6 is original too.

Chapter 2 Review of Previous Research

Previous research carried out in the field of acoustic study of breath sounds is presented briefly in this chapter. More detail is given where the research relates closely to this thesis.

2.1 History of Acoustic Research in Breath Sounds

Loudon (1985) described research in respiratory sounds over three successive time periods: immediate auscultation, stethoscope auscultation and electronic observation.

Immediate auscultation involves directly listening to the patient's chest, which could date back to its use by Hippocrates in the 4th century BC (Gavriely and Cugell 1995). This shows that auscultation was considered important long ago. But obviously this method was not always practicable.

Then immediate auscultation was replaced by the stethoscope auscultation after Laënnec invented the stethoscope in 1816 (Gavriely and Cugell 1995). The stethoscope has greatly improved our understanding of the sounds heard from the body surface with pulmonary conditions. The stethoscope can assist cheap, simple, non-invasive diagnosis, and is therefore still today the most frequently used medical device (Weitz and Mangione 2000).

In the 1920s, researchers (Cabot and Dodge 1925; Hannon and Lyman 1929) started to record and analyse breath sounds using electronic systems, which led to

the modern objective observation. Big advances have been made since early use of purely analogue techniques to today's computer assisted techniques. Today qualitative and quantitative analysis are able to provide more objective and accurate information, which forms the basis of our understanding of the mechanism of respiratory sounds generation and propagation, and also helps assess pulmonary situations. With continuous advances of analogue and digital devices and signal processing techniques, further improvements in resolution and accuracy can be expected.

2.2 Categories of Breath Sounds

2.2.1 Normal Breath Sounds

Normal lung sounds (vesicular sounds). A lung sound is detected through the chest wall of a healthy subject. This is characterised by a faint low-frequency noisy sound, heard throughout inspiration and at the beginning of expiration. The expiration has a lower pitch, lesser intensity and shorter duration than inspiration (Forgacs 1978; Gavriely and Cugell 1995; Pasterkamp *et al.* 1997).

Normal tracheal sounds (bronchial sounds). A tracheal sound is detected on the neck. This is characterised by a broader spectrum of noise than the normal lung sounds from the chest wall, audible throughout inspiration and expiration (Forgacs 1978; Gavriely and Cugell 1995; Pasterkamp *et al.* 1997).

2.2.2 Adventitious Sounds

Wheezes. These are continuous sounds with a musical character. Acoustically, they are characterised by periodic waveforms, with a dominant frequency usually

over 100Hz and with duration greater than about 100ms. They can be classified according to pitch (high or low), complexity (monophonic or polyphonic), duration (long or short) and timing (inspiration and expiration; early and late; random and sequential) (Forgacs 1978; Pasterkamp *et al.* 1997; Sovijarvi *et al.* 2000b).

Crackles. Crackles are discontinuous, short, explosive non-musical sounds. There are two types: fine crackles and coarse crackles, which can be classified according to waveform, duration, and timing. A fine crackle has a high pitch, low amplitude and short duration, whereas a coarse crackle has a low pitch, high amplitude and long duration (Forgacs 1978; Pasterkamp *et al.* 1997; Sovijarvi *et al.* 2000b).

Other adventitious sounds have been identified, such as rhonchus and squawks, but are less studied, so they are not defined here.

2.3 Production and Propagation of Breath Sounds

Much research has been based on lung sounds, and relatively recently research based on tracheal sounds has increased. Typical sites of recording breath sounds on chest and neck have been recommended (Rossi *et al.* 2000; PixSoft Inc. and Medi-wave Inc. 2001). As investigators intended to relate the recording sounds to the physiological or pathological respiratory system, the origin of normal and abnormal breath sounds and their propagation is one of the research focuses. The precise anatomical origin of breath sounds and mechanisms of generation are not known, and this area is still under investigation.

2.3.1 Production of Breath Sounds

Normal Breath Sounds

Some doctors believe that the apparent normal lung sound originates in the larger airways due to turbulent flow and is subsequently transmitted to the chest wall. The study by Austrheim and Kraman (1985) compared the amplitudes of breathing air and He-O_2 gas (20% oxygen and 80% helium) at trachea and chest sites under the target flow rate. They suggested that tracheal and expiratory lung sounds were produced by turbulent flow, which is density dependent; while inspiratory lung sounds were produced by some other unknown mechanism. Pasterkamp and Sanchez (1996) repeated the above experiment but extended to the higher frequencies. Their findings draw the conclusion that flow turbulence is the dominant mechanism for the production of lung sounds at higher frequencies. These experiments seem to infer that production of the normal tracheal sounds and lung sounds are not identical.

Normal tracheal sounds. According to Fahr (1927), Martin and Muller (1923) recorded graphically the vibrations set up in the airway system up to certain branch generations. They demonstrated that bronchial sounds were produced in airways with a diameter of 4 mm or above. Fahr confirmed these observations, and showed that the glottis played a minor part in the production of tracheal sounds. The turbulent flow has been observed within cast models of central airways from larynx to segmental bronchi above critical flow rate (West and Hugh-Jones 1959; Dekker 1961; Olson *et al.* 1973). This turbulent flow has been suggested as the dominant source of tracheal sounds (Forgacs *et al.* 1971; Olson *et al.* 1984; Austrheim and Kraman 1985; Pasterkamp and Sanchez 1996).

Normal lung sounds. As early as 1884, Bullar (1884) designed a series of experiments to study the production of respiratory sounds. He concluded that sounds were produced at those parts of the respiratory tract where the air passes from a narrower to a wider space, and pointed out that vesicular sound was produced in the lungs.

Later studies suggest that at least the inspiratory component of the lung sound is produced locally, within each lung and probably within each lobe (Ploy-Song-Sang *et al.* 1977; Kraman 1980; 1981; 1985b; Pasterkamp *et al.* 1997; Kompis *et al.* 2001). The expiratory component of the lung sound originates from the larger central airways in comparison with the inspiratory sound (Kraman 1980; 1981; Gavriely 1983; Kompis *et al.* 2001). Hardin and Patterson (1979) speculated that the mechanism to produce normal lung sounds was the unsteady movement of vortices in the lung.

Adventitious Sounds

Crackles. Crackles are usually of a recurrent rhythmical pattern (Forgacs 1967; 1978; Nath and Capel 1974a; Mori *et al.* 1980), to be heard much more often during inspiration than during expiration, and are sensitive to posture changes (Forgacs 1967; 1978). Forgacs hypothesised that the quick equalisation of gas pressures that follows the reopening of previously closed small airways produces implosive sound waves (Forgacs 1967; 1978). However Fredberg and Holford (1983) stated that their stress relaxation quadruploe model may be preferable. Their theory is that the elastic stress in and near the airway wall in transition between static equilibrium after a closed airway opens suddenly can be detected

as a crackle via parenchyma. Experiments do seem to provide evidence that at least the inspiratory crackles are due to opening of small airways (Nath and Capel 1974a; Mori *et al.* 1980; Munakata *et al.* 1986), but crackles are probably produced by more than one mechanism (Loudon and Murphy 1984; Olson and Hammersley 1985; Raymond and Murphy 1985). In addition to the above mechanism, crackles may result from air bubbling through secretions. This is based on the observation of crackles from patients with secretion in their airways or pulmonary oedema (Urquhart *et al.* 1981; Banham *et al.* 1984). In these cases crackles appear in both phases of respiration and at random (Forgacs 1969).

Wheezes. It seems likely that a coupling of airflow and airway wall vibration causes wheezing. Forgacs (1967; 1978) speculated that wheezes were produced by a mechanism similar to the reed in a toy trumpet, with the mass and elasticity of airway walls contributing to the pitch of the wheezes. However the two prerequisites of producing wheezes are: (i) the airway calibre should be reduced to closure and (ii) the flow velocity should reach a critical value. Based on a mathematical fluid dynamic flutter model (Grotberg and Davis 1980; Grotberg and Gavriely 1989), wheezes are produced by the interaction of fluid forces and friction and airway wall elastic-restoring forces and damping. This theory has been examined by Gavriely and associates (Gavriely *et al.* 1984b), and can explain the results well. The model showed that wheezes will be always accompanied by flow limitation[1], but flow limitation will not always produce wheezes. These are supported by experimental observations (Gavriely *et al.* 1984a; 1987; 1989; Spence *et al.* 1996; Doherty *et al.* 1998). The theory inferred that two types of

[1] Flow limitation appears when the flow rate is independent on the driving pressure (pressure difference between upstream pressure and downstream pressure).

airway vibrations may result from the interaction of the flow and the elastic airway wall. One is stable airway oscillation, which may account for end expiratory wheezes. The other is unstable oscillation, which may account for all other wheezes. The model also predicted that gas density has a minor effect on wheeze pitch, which is consistent with observations (Forgacs 1978; Shabtai-Musih *et al.* 1992) .

2.3.2 Transmission of Breath Sounds

The tracheal sounds are not conducted within the airways where internal diameters are less than 3mm (Fahr 1927). At high frequencies the larger airways are rigid because of their inertance and sound travels at relatively high speed (Rice 1980). At very low frequencies, sound propagation is much slower (Rice 1985).

The lung is a bad sound conductor (Bullar 1884). Sound travels very slowly in parenchyma as the lung is a homogeneous mixture of gas and tissue (Kraman 1983b; Rice 1983; Bergstresser *et al.* 2002). Sound attenuation increases with higher frequency and is suggested to occur in the parenchyma, which is supported by experimental observation (Hannon and Lyman 1929; Gavriely *et al.* 1981; 1983) and theoretical models (Rice 1983; Wodicka *et al.* 1989; Vovk *et al.* 1995).

The chest wall exhibits strong fundamental resonance over different lung regions and at different lung volumes (Ploy-Song-Sang *et al.* 1977) according to body size (Wodicka and Shannon 1990). This resonance may have particularly important

effects on sound transmission at low frequencies, where the parenchyma attenuation seems to be small (Wodicka and Shannon 1990) .

As the parenchyma acts as a low pass filter, lung sounds have little high frequency component. The thorax further attenuates the lung sound transmission (Rice 1983; Vovk *et al.* 1995). In contrast, tracheal sounds have richer high frequencies and louder intensity (Kraman and Austrheim 1983; Wodicka and Shannon 1990).

Conditions, such as consolidation or large pleural effusion, decrease the absorption in parenchyma, and thus increase the measured sound intensity over the chest at higher frequencies.

In short, abnormal pulmonary situations differ from normal ones in the production mechanism and transmission properties of breath sounds. It is postulated that these differences can be reflected in the measured breath sounds.

2.3.3 Acoustic Models of the Respiratory System

Acoustic models of the respiratory system have been proposed based on the understanding of the production mechanism and the transmission property of breath sounds and appropriate simplification.

The respiratory tract consists of the vocal tract and subglottal airways. The branching airway structures within the thorax have been modelled to study transmission and reflection properties (Ishizaka *et al.* 1976; Jackson *et al.* 1989; Wodicka *et al.* 1989) and to predict tracheal sounds characteristics (Harper *et al.*

2001; 2003). The tract system can be simply modelled as a soft-walled tube of effective length with an open-end (Ishizaka *et al.* 1976; Jackson *et al.* 1989). Large airways vibrate in response to airway pressure oscillations; thus a significant acoustic energy is coupled directly from within large airways to surrounding parenchyma (Wodicka *et al.* 1989). The important airway geometry and properties and glottal open size determine the characteristics of tracheal sounds (Harper *et al.* 2001; 2003).

The lung parenchyma surrounds the intrathoracic tracts. It can be modelled as a homogeneous medium of air bubbles in a liquid in the audible frequency range (Rice 1983) with sound attenuation effect. The chest has been acoustically modelled as a large cylinder (Wodicka *et al.* 1989; Vovk *et al.* 1994). A further refined model included muscular-rib and muscular-fatty layers (Vovk *et al.* 1995). The mismatch of acoustic impedances of the chest wall and parenchyma leads to further attenuation of breath sounds detected at the chest surface (Rice 1983; Vovk *et al.* 1995).

These models reflect the respiratory sound transmission from central airway to the chest surface. The predicted results from the models compared reasonably well with experiment data.

2.4 Measurement System

Although the mechanism of respiratory sounds generation is not exactly understood, it does not prevent investigators from measuring and analysing these sounds.

Measurement is the basis for analysis. The measurement systems used in breath sounds analysis have been improved consistently with the development of electronic devices.

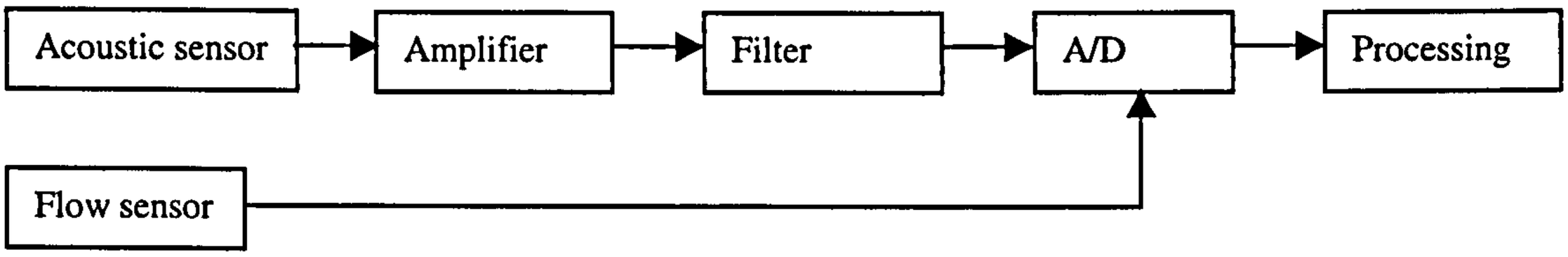

Figure 2.1 Structure of measurement system.

Figure 2.1 schematically shows the common structure of a digital data acquisition system, though various instruments used by different research groups. From the viewpoint of an engineer it seems that investigators have not paid much attention to the construction of a particular measurement system. In many medical and engineering papers, analysis methods and clinical applications are the most important parts of a paper. Little detail is given about why that particular system is used. Relatively few studies were focused on measurement techniques, such as sensor properties (Druzgalski *et al.* 1980; Pasterkamp *et al.* 1993; Gavriely and Cugell 1995) , air-coupler properties (Wodicka *et al.* 1994; Kraman *et al.* 1995), and filter characteristics (Sun *et al.* 1998).

2.4.1 Acoustic Sensor

The whole measurement system determines the quality of measured data. The first critical stage is the sound sensor. Two kinds of acoustic sensors are commonly used: one is an air-coupled sensor and the other is surface vibration sensor.

Air-coupled sensors. Air-coupled sound sensors measure the sound pressure. Various types and models of microphones have been used by researchers. According to Druzgalski (1980), an ideal microphone for recording should have a dynamic range greater than 40-50dB, with high sensitivity and signal-to-noise ratio, and flat frequency response. Electret condenser microphones have a flat frequency response, light-weight, good sensitivity and wide dynamic range, as well as being relatively cheap, so they were widely used in breath sounds recording (Druzgalski *et al.* 1980; Charbonneau *et al.* 1983; Austrheim and Kraman 1985; Spence *et al.* 1992; Malmberg *et al.* 1995; Doherty *et al.* 1998).

However electret microphones cannot be put on the body surface to collect the breath sound signal directly because of the impedance mismatch between body surface and microphone membrane. Some kind of rigid chamber is needed to couple the microphone to the body surface. An air cavity between the microphone and body surface will affect the characteristics of breath sounds. Work done by Wodicka et al (1994) and Kraman (1995) suggested that an optimal air coupler should be light-weight, conical, shallow depth, and either not vented or vented with a very thin tube.

Electronic stethoscopes have an air cavity of the shape and size of an ordinary stethoscope (diaphragm or bell) but integrate the microphone and following amplifier and filter stages. Telephones and mobiles can also be considered in this category.

Surface vibration sensors. Compared to air-coupled sensors, surface vibration sensors are put on the body surface directly without air couplers. Vibration on the body surface can be measured by a vibrometer or an accelerometer. Quite a few groups preferred these surface type sensors, such as Gavriely's group (Gavriely *et al.* 1981; 1984b; 1992; 1995), Pasterkamp's team (Pasterkamp *et al.* 1984; 1989; 1992; 1996; Harper *et al.* 2003) and some others (Fenton *et al.* 1985; Vovk *et al.* 1995) .

Not all the industrial vibrometers or accelerometers are suitable for detecting respiratory sounds (some preliminary unreported studies at Glasgow University). Heavy sensors or those needing firm contact could not be used in respiratory sounds recording, as they induce distortions by changing the transmission characteristics of the chest.

The choice between microphones and contact sensors may be decided by the availability, cost and maintenance. Though Vovk's (1994) model preferred surface type transducers to detect vibration velocity or acceleration on the body surface, they are more expensive and fragile.

Studies of sound sensors (Druzgalski *et al.* 1980; Pasterkamp *et al.* 1993) showed that the frequency responses of the sensors were not flat. It's understandable that frequency characteristics differ among different sensors, but even sensors of the same model showed significant difference. Others also measured frequency responses of acoustic sensors (with coupler) (Kraman 1980; Gavriely *et al.* 1981).

2.4.2 Amplifiers

Signals from sound sensors are normally too small to reach the level that the successive stage (A/D or tape recorder) requires. Thus amplifiers are used to amplify the output signal from the sound sensor. Various types of amplifiers were used among researchers. Quite a few researchers didn't mention what kind of amplifier was used (Austrheim and Kraman 1985; Anderson *et al.* 1990; Gavriely and Herzberg 1992; Shabtai-Musih *et al.* 1992). Others chose from built-in amlifiers in electronic stethoscopes (Baughman and Loudon 1984; Lessard and Wong 1986; Bohadana *et al.* 1994) , built-in amplifiers incorporated with filters (Homs-Corbera *et al.* 2000; Fiz *et al.* 2002), and instrumentation amplifiers (Benedetto *et al.* 1983; Shykoff *et al.* 1988; Malmberg *et al.* 1995). The important parameters to choose an amplifier should be the flatness of frequency response and bandwidth.

2.4.3 Filters

High pass filters were used to reduce low frequencies (such as heart sounds), where their high energy may saturate the sampling channel or reduce high frequency resolution. Low pass filters were used to provide an anti-aliasing function, in which the cut-off frequency will depend on the sampling rate. Sometimes band-pass filters were used to achieve both the above functions. Various filter types and orders as well as different cut-off frequencies were used. For example: Fenton and colleagues (1985) used a 5^{th} order elliptic low pass filter with cut off frequency of 1000Hz and a sampling rate of 2560Hz. Anderson and associates (1990) employed a 4^{th} order Chebychev low pass filter with cut off frequency of 3kHz and a sampling rate of 9600Hz. Pasterkamp and co-workers

(1996a) utilised a 6th order Butterworth low pass filter with cut off frequency of 2400Hz and a sampling rate of 10240Hz. Broersen and de Waele (2000) used a 4th order Bessel band pass filter with cut off frequencies of 100Hz and 1500Hz, and a sampling rate of 5000Hz.

Sun and associates (1998) proposed several methods to study the frequency response of high pass filters. This is because a high pass filter would significantly affect the signal wave shapes in the time domain (Katila et al. 1991; Sun *et al.* 1998), which would affect values of parameters based on the waveform.

2.4.4 Digitisation

Purely analogue techniques were used in early studies. For example, Cabot and Dodge (1925) compared sound intensity and frequency character with and without electric filters. In some cases, an oscilloscope was used to show results directly (Hannon and Lyman 1929; Forgacs *et al.* 1971; Kraman 1980; 1981; Ploy-Song-Sang *et al.* 1977; 1983). In most current work the signals are either recorded on a tape then digitised, or digitised and stored on a computer disk directly. The most commonly used A/D resolution was 12-bit (Fenton *et al.* 1985; Anderson *et al.* 1990; Gavriely *et al.* 1995; Broersen and de Waele 2000), some were 10-bit (Gavriely *et al.* 1984b; Bohadana *et al.* 1994) and 13- bit (Malmberg *et al.* 1994; Sovijarvi *et al.* 1996; Vanderschoot *et al.* 1998). The higher the resolution, the wider the dynamic range, and the more accurate the discrete data; however higher resolution is more expensive.

2.4.5 Flow Sensor

Flow rate affects the sound intensity (Leblanc *et al.* 1970; Banaszak *et al.* 1973; Dosani and Kraman 1983; Kraman 1984; Olson *et al.* 1984; Shykoff *et al.* 1988). From most of the studies, the relation between sound amplitude a and flow rate f can be expressed as $a = cf^b$, where b and c are constants. But the value b is not consistent among investigators (Kraman 1984; Olson *et al.* 1984; Shykoff *et al.* 1988; Gavriely and Cugell 1996). The spectral components of breath sounds are dependent on flow rate below a certain flow rate (Charbonneau *et al.* 1983; Lessard and Wong 1986; Schreur *et al.* 1994), but are independent above that critical flow rate (Kraman 1986; 1998; Mussell *et al.* 1990; Gavriely and Cugell 1996; Harper *et al.* 2003). For these reasons, many investigators made their study under monitored flow rate conditions by using a flow sensor.

The most frequently used flowmeter is a pneumotachograph with a differential pressure transducer (Dosani and Kraman 1983; Shykoff *et al.* 1988; Soufflet *et al.* 1990; Gavriely *et al.* 1995). However studies with and without a flow sensor (Urquhart 1983; Mussell *et al.* 1990) showed that the flowmeter had significantly distorted the 'true' breath sounds. It can be presumed that different flow sensors may have distinct distortions.

2.4.6 Calibration

As mentioned above in sections 2.4.1 to 2.4.5, various sensors, amplifiers, filters etc were used by different research groups. It would be sensible to expect that characteristic variations exist among measurement systems.

For comparison purposes, calibration of the system is necessary. A few calibrations have been done when two or more identical measurement channels were used (Ploy-Song-Sang *et al.* 1977; 1979; Dosani and Kraman 1983; Kraman 1984; Austrheim and Kraman 1985; Gross *et al.* 2000; Kompis *et al.* 2001). In breath sounds measurement, relative calibration is sufficient. However, maybe the lack of standard calibration signals hampered the investigators calibrating their system. When direct comparisons are impossible, indirect comparisons are made via clinical parameters.

2.5 Analysis Methods

Except when displaying the results directly on an oscilloscope (Forgacs *et al.* 1971; Ploy-Song-Sang *et al.* 1977; 1983; Kraman 1980; 1981), almost every signal processing method has been used in analysing breath sounds.

2.5.1 Time-domain Methods

Phonopneumography

Phonopneumography is the visual display of breath sounds, either alone or with other simultaneously measured signals, usually airflow (Kraman 1985a). Phonopneumography has found applications in teaching and training medical students (Cugell 1971; Weiss and Carlson 1972). It was also used to study the relationship between sound amplitude and airflow (Banaszak *et al.* 1973) and breath sound generation (Kraman 1980; 1981).

Phonopneumography could not provide quantitative parameters, but it is still useful as an assistant visual display to show the profile of breath sounds in the time domain.

Time-expanded Waveform Analysis

Murphy (1977) first proposed the time-expanded waveform method, which rescaled the original signal time axis to show the locally detailed waveform. Although this is quite an easy method, it shows the fast transient events such as crackles (Murphy, Jr. *et al.* 1977; 1984; Mori *et al.* 1980; Workum *et al.* 1982; Munakata *et al.* 1991; al Jarad *et al.* 1993) and monophonic wheezes quite well (Murphy, Jr. *et al.* 1977; Forgacs 1978). Detailed characteristics of crackles such as initial deflection width (IDW)[1] and two-cycle duration (2CD)[2] can be measured by this method.

The drawback is that the segments of interest should be manually located, maybe with the help of listening, before they could be visually examined. Thus it is time consuming with possible large inter-observer variability.

Time series models

Theoretically, an unknown linear stationary stochastic process can be modelled by at least one of the three time series models: Autoregressive (AR), Moving Average (MA), and mixed ARMA. AR-based classification of breath sounds was achieved by Cohen and Landsberg (Cohen and Landsberg 1984; Cohen 1990) and Sankur et al (Sankur *et al.* 1994; Kahya *et al.* 1999). Iyer and associates (1989) and

[1] IDW is the duration of the first deflection in a crackle waveform.
[2] 2CD is the duration of a crackle from the beginning of the initial deflection to the end of two cycles.

Hadjileontiadis and Panas (1997b) used AR modelling to estimate a lung sounds' source and transmission characteristics. Gavriely and Herzberg (1992) described normal lung sounds with a 6-8 orders AR model and normal tracheal sounds with a 12-16 orders AR model. Vanderschoot and Schreur (1991; 1992; 1994) found that AR parameters of normal lung sounds depend strongly on the flow and volume. Broersen and de Waele (2000) used AR and ARMA models to detect methacholine.

Respiratory sounds are non-stationary stochastic signals. While using time series models which are applied to stationary stochastic processes, the choice of which segments and how long the segments of the respiratory sounds should be to represent a stationary period is important. If the segments are too long, then the assumption that the signal is locally stationary is invalid; but if they are too short, then the estimated parameters will be erroneous and have big variance. Otherwise, some methods of removing the non-stationary component are necessary.

2.5.2 Frequency Domain Methods

Most research has been done in the frequency domain. Based on the discrete Fourier transform (DFT) or fast Fourier transform (FFT), amplitude spectra or power spectra were used to represent the frequency characteristics. Parameters extracted from the spectrum such as median frequency, F_{50}[1], (Anderson *et al.* 1990; Spence *et al.* 1992; Malmberg *et al.* 1994; 1995; Sovijarvi *et al.* 1996; Fiz *et al.* 1999), maximum frequency, F_{max}[2], (Malmberg *et al.* 1994; Lenclud *et al.* 1996;

[1] F_{50} is defined as the frequency below which 50% of the energy of the signal lies.
[2] F_{max} is defined as the frequency of maximum power

Gross *et al.* 2000), slopes of regression lines of spectrum (Gavriely *et al.* 1984b; 1995), index composite of spectrum based parameters (Charbonneau *et al.* 1983) were used to characterise normal and abnormal breath sounds, or for comparison before and after some challenge or medicine.

The advantage of spectrum-based parameters is that one or a few indices can describe or compare frequency characteristics. But as pointed by Whittaker *et al.*, (2000), these indices are dependent on the frequency range to be calculated as well as the number of FFT points and the type of windowing function.

However, DFT or FFT don't display time resolution, i.e. they cannot reflect frequency changes with time. Breath sounds are non-stationary signals, especially when adventitious sounds appear. On the one hand, when adventitious events are short compared to the respiratory cycle, they may not have enough bearing on the spectral shape. On the other hand, the timing of those adventitious sounds has clinical importance. Thus time-frequency representation, which can show frequency content evolution with respect to time, such as short time Fourier transform (STFT) or Wigner distribution (WD) will be more suitable.

STFT based spectrographs have shown different patterns of normal and abnormal breath sounds obtained at different sites in the time-frequency plane (Pasterkamp *et al.* 1989; Sovijarvi *et al.* 2000a; PixSoft Inc. and Medi-wave Inc. 2001). WD based mapping of cardiovascular and respiration signals reflect their instantaneous frequency changes (Novak and Novak 1993). STFT based automatic wheeze detection algorithms (Waris *et al.* 1998; Homs-Corbera *et al.*

2000) and WD based crackle analysis methods (Pasika and Pengelly 1994) have been developed.

2.5.3 Wavelets

Since the wavelet emerged, this time-scale multiresolution method has been widely used in signal processing applications. The wavelet is capable of detecting transient events, so it has been used to detect crackles (Sankur *et al.* 1996; Hadjileontiadis and Panas 1997c) and to reduce heart sounds interference (Charleston *et al.* 1997; Hadjileontiadis and Panas 1998). Wavelet packet decomposition is an extension of the wavlet transform. Pesu and associates (1996; 1998) used a wavelet packet to classify respiratory sounds. Ademovic and colleagues utilised a wavelet packet to optimally segment respiratory sounds (1998a) and segment wheezes (1998b). And Bahoura and co-workers (1998) employed a wavelet packet to de-noise respiratory sounds.

The discrete wavelet analysis is an orthogonal decomposition. The choice of mother wavelet is somewhat empirical, and is data orientated.

2.5.4 Others

Other signal processing methods include pattern recognition (Urquhart *et al.* 1981; 1983; Banham *et al.* 1984; Cohen and Landsberg 1984; Anderson *et al.* 1986; Sankur *et al.* 1994; Oud *et al.* 2000), neural networks (Pesu *et al.* 1996; 1998; Malmberg *et al.* 1996; Rietveld *et al.* 1999; Waitman *et al.* 2000), high-order statistics (Hadjileontiadis and Panas 1997a; 1997b), fuzzy inference systems (Tolias *et al.* 1997; 1998; Mastorocostas *et al.* 2000), fractal dimension (De

Oliveira *et al.* 1998), digital filtering (Ono *et al.* 1989; Plante *et al.* 1998), and imaging processing (Waris *et al.* 1998).

2.5.5 Wheeze Detection

As mentioned above in 2.5.1 to 2.5.4, various signal-processing methods have been used in analysing respiratory sounds. One of the aims of the current work is automatic wheeze detection; so more details are described here about previous wheeze detection algorithms.

In the frequency domain

Baughman and Loudon (1985) analysed respiratory sounds recorded on the chest. 50 segments of 250ms during 5 minutes were chosen for analysis using FFT. Each spectrum was analysed for the presence or absence of a sharp peak at a frequency between 150 and 1000Hz with amplitude more than 3 times larger than the baseline. The occupation of wheezes T_w/T_{tot} was estimated as the ratio of number of spectrums with peaks and number of total spectrums over a five-minute period. Later, they used similar criterion on 100ms segments (1989). A peak was associated with wheezing when its frequency was higher than 200Hz and its amplitude was 3 times greater than the baseline and lasted for more than 200ms. The former criterion was also used by Schreur et al. (Schreur *et al.* 1994).

Fenton and associates (1985) also proposed a similar automatic wheeze detection algorithm. Segments of 100ms over the trachea and chest were analysed. Wheezes were detected as a large peak, which was above 200Hz, at least 15 times large than average power between 110-1200Hz. Then sequential 100ms

data were classified as normal or wheezing, and were used to derive the duration of wheezes.

These segment-by-segment peak-pick methods are computationally simple, but the threshold for a sharp peak is data-dependent. Thus louder wheezes will be much more easily identified than quieter wheezes. As for quieter wheezes, their corresponding peak amplitudes won't be high. Wheeze duration estimation is a useful parameter to relate to asthma severity. But it could not answer the timing of the wheezes, and their evolutions with time.

In time-frequency plane

As early as 1955, McKusic and associates (McKusic *et al.* 1955) presented a novel analogue sonographic imaging technique for respiratory sounds by playing respiratory sounds repeatedly through an adjustable filter. The filter's central frequency and bandwidth was tuned progressively with the repeated playback. The intensities of various frequency components were recorded as different grey shades. In 1989, Pasterkamp and colleagues (Pasterkamp *et al.* 1989) proposed a computer-calculated spectrogram which was based on STFT. Respiratory sounds were presented in the time-frequency plane, and the sound intensities displayed on a colour scale. This kind of spectrogram can provide an image of time-frequency localisation of wheezes.

Ademovic et al. (1998b) used Malvar's wavelet to optimally represent a signal in the time-frequency plane. The Malvar's wavelet decomposition achieves an adaptive segmentation in the time domain and a spectral analysis on each

segment. Under optimal parameters, the results provide optimal visual clues of the stationary (wheezes) and non-stationary (normal breath sounds) components of a signal.

Visual clues provide useful information about wheeze duration, frequency location, and its evolution, but they are not enough to provide quantitative parameters automatically.

Waris and associates (1998) proposed an image processing scheme to detect wheezes automatically. First, based on a spectrogram, an edge detection method was used to test every pixel. An edge was found at a pixel when the absolute value of the gradient exceeded a threshold. Wheezes were detected by searching for horizontal or nearly horizontal edges. Some of the edges represented different parts of a single wheeze. In the next step, a labelling method was used to group these components together. Then an exponential criterion was added, that was, the shorter the labelled component, the stronger it had to be. Finally two components, which were less than 30Hz and 150ms, were connected together. The results, which were validated by a pulmonary physician, had a best average sensitivity[1] of 68% and predictivity[2] of 70%. It had problems in detecting wheezes shorter than 125ms, which are quite easy for human ears to detect.

Shabtai-Musih and co-workers (1992) proposed a scoring scheme. After normalising the spectrum, peaks were picked up. Then according to 8 rules the

---

[1] sensitivity=true positives/(true positives + false negatives)
[2] predictivity=true positives/(true positives + false positives)

peaks were assigned a score. The peaks which had a score bigger than 3 were regarded as wheezing peaks.

Homs-Corbera and associates (Homs-Corbera *et al.* 2000; Fiz *et al.* 2002) improved Shabtai-Musih's method by adding a grouping algorithm. They first picked up peaks; then scored the peaks according to six empirical rules with empirical values. Peaks with a score greater than 3 were considered wheezing peaks. Then the grouping method grouped individual peaks, which are less than 25.6 to 51.2ms and 50 to 65Hz, to the nearest wheezing peaks. The results, which were validated by a pneumology specialist, had reported accuracy from 100%, 87.8% to 71%, at high, middle and low flow rate respectively. Parameters concerning wheeze numbers, mean frequencies, and duration (percentage) were derived.

These methods can detect the time and frequency characteristics of wheezes in the time-frequency plane. Such results may provide useful information in diagnosis. As their results showed, at lower flow rate, when the respiratory sounds were quieter, the accuracy decreased. And when wheezes were short, the accuracy degraded. The peaks or edges detection is dependent on FFT numbers and windowing. Thus the sensitivity could be changed under different conditions, making it inferior to human ears. For the scoring scheme, the score rules to determine a wheezing peak were empirical; it could be better to use different parameter values under different flow rates, as those parameters were data-dependent.

Classification

A lot of methods were proposed to classify normal and abnormal respiratory sounds (Cohen and Landsberg 1984; Sankur *et al.* 1994; Malmberg *et al.* 1996; Pesu *et al.* 1996; 1998; Rietveld *et al.* 1999; Waitman *et al.* 2000). Most of them used an artificial neural network method. Those that could classify asthmatics (wheezes) are discussed here.

Malmberg and associates (1996) classified respiratory sounds based on FFT spectra and a self-organising map (an artificial neural network). The correct classification of asthmatics was moderate.

Pesu et al. (1996; 1998) classified respiratory sounds based on wavelet packet decomposition and an artificial neural network. First, wavelet packet decomposition was performed and the best wavelet packet coefficients were searched. Secondly, the feature vectors were constructed based on the results of the first stage. Then a neural network was used to classify respiratory sounds. For wheeze detection, their results had average sensitivity of 58.7% and predictivity of 37.3%.

By using neural networks, training data play an important role. Results may be improved by using typical training data. However, inter- and intra-subject variations of respiratory sounds make it not an easy task.

Oud and associates (2000) explored an algorithm to classify degrees of airway obstruction. First, based on DFT normalised power spectra or Welch power

spectra in the frequency range of 100-1300Hz, a Box-Cox transformation was performed. Then k-nearest neighbour classification with k=1 was applied to classify spectral vectors into classes of airways obstruction. Under optimal parameters, about 60%-90% of data can be classified correctly according to their corresponding FEV_1[1]-value.

Although these classification methods avoided the clinic interests of wheeze characteristics, such as pitch, duration and timing, they proved that airway obstruction could be inferred from respiratory sounds in asthmatics.

Pasterkamp and colleagues (1987) compared assessments of wheezes between professionals and computer analysis. They found objective analysis was more reproducible than subjective assessment. Rietveld and associates (1999) studied the classifying capacity of untrained human examiners according to spectrograms vs. artificial neural networks. Three groups were classified: asthma in exacerbation, asthma in remission, and normal. They found the examiners incapable of differentiating the spectrograms in accordance with three groups. While for a supervised neural network, almost all training vectors and half of the test vectors could be classified correctly.

The objective analysis and classification of wheeze can provide more reproducible qualitative and quantitative performance than subjective evaluation.

[1] FEV_1 is defined as the forced expiratory volume in one second.

2.6 Miscellaneous

The mainstream literature has focused on the mechanisms of respiratory sounds production and propagation, relations between measured respiratory sounds and physiological and pathological pulmonary conditions, and various analysis algorithms. Some work has been done on other aspects.

2.6.1 Noise Reduction

Ambient noise, heart sounds, muscle contraction sounds, and noise due to friction of skin and hair can easily contaminate breath sounds measurement.

Heart sounds reduction

During recording of breath sounds, heart sounds, which are louder than breath sounds in most of the chest and neck sites, are a perpetual noise source. High-pass filtering can remove heart sounds effectively, as most energy of heart sounds is below 100Hz. However it distorts the breath sounds due to the inherent overlapping of the low frequency region. So adaptive filtering such as the least-mean-square adaptation algorithm (Iyer *et al.* 1986; Kompis and Russi 1992), Kalman filtering (Charleston and Azimi-Sadjadi 1996), high-order statistics (Hadjileontiadis and Panas 1997a), and wavelets (Hadjileontiadis and Panas 1998; Charleston *et al.* 1997) were proposed. Results of moderate to good heart sounds reduction performance were reported.

Ambient noise reduction

Except for a few measurements which were done in a sound-proof room, most tests were done under quiet conditions. Ambient noises can interfere with

respiratory sound measurement. Traditional filtering is ineffective due to the frequency overlap between respiratory sounds and ambient noise. Adaptive filtering (Suzuki *et al.* 1995; Patel *et al.* 1998) and wavelet packet based denoising were much more effective (Bahoura *et al.* 1998). If clean breath sounds can be obtained under general conditions, it will benefit routine clinical usage and remote monitoring where quiet circumstances are not always available.

2.6.2 Simulation

Computer-based respiratory sounds simulation (Kompis and Russi 1997; Cardionics Inc 2002) has been developed as an educational tool. Such a tool can produce various simulated normal and abnormal respiratory sounds, with easily changed parameters. These well-defined and reproducible data could also be useful to evaluate respiratory sounds analysis algorithms.

Kiyokawa et al. questioned if auscultation is a reliable reference standard for crackles (2001). They added simulated crackles into recorded normal sounds. Then these synthesised sounds were played for auscultation. Under some circumstances, simulated crackles were inaudible due to masking[1] .

2.7 Clinic Applications

2.7.1 Applications Based on Normal Breath Sounds

As described in 2.3, normal lung sounds heard on the chest wall contain information about their origin and transmission. Even in normal individuals, the

---

[1] Masking is defined as the process by which the threshold of audibility for one sound is raised by the presence of another (masking) sound.

intensity and spectrum vary according to various factors: flow rate (Kraman 1984; Shykoff *et al.* 1988; Gavriely and Cugell 1996), pick-up sites (O'Donnell and Kraman 1982; Dosani and Kraman 1983; Kompis *et al.* 2001), volume (Kraman 1986; Vanderschoot and Schreur 1991; 1992; 1994; Kiyokawa and Pasterkamp 2002), respiration phases (Dosani and Kraman 1983; Gavriely *et al.* 1995; 1996), age (Kanga and Kraman 1986; Hidalgo *et al.* 1991; Pasterkamp *et al.* 1996a; Gross *et al.* 2000) , sex (Gavriely *et al.* 1995; Gross *et al.* 2000), as well as intersubject variations and temporal variations (Mahagnah and Gavriely 1994).

The compensated lung sound intensity has been used to evaluate regional ventilation, distribution of regional ventilation and sequences of regional ventilation (Ploy-Song-Sang *et al.* 1977; 1978; 1979; 1983).

For the same flow rate, no significant difference in lung sound intensity was found between emphysema patients and normal subjects (Schreur *et al.* 1992); while the apparent normal lung sounds from stable symptoms-free asthmatics were distinguishable from those of normal subjects (Schreur *et al.* 1994). Mild degrees of obstruction during bronchial challenges were detectable (Anderson *et al.* 1990; Bohadana *et al.* 1994a; 1995; Malmberg *et al.* 1994).

Recently, microphone arrays have been used to map the chest wall acoustically, which can reflect the regional structural and functional properties (Kompis *et al.* 2001). Similar arrays have been used to study breath sound transit time, which can reflect tissue property (Bergstresser *et al.* 2001).

There is a strong relation between flow rate and tracheal sound intensity and spectrum. Usually the higher the flow rate, the higher the intensity, but the relationship is not linear. The spectral shape does not change with different flow rates if they are above a critical value (Mussell *et al.* 1990; Kraman *et al.* 1998; Harper *et al.* 2003), and there is no significant difference between respiration phases (Kraman *et al.* 1998). However the spectral characteristics are dependent on body height (airway length) (Sanchez and Pasterkamp 1993; Kraman *et al.* 1998), airway geometry and properties (Pasterkamp *et al.* 1996b; Harper *et al.* 2001; 2003), and gas density (Pasterkamp and Sanchez 1996).

Tracheal sounds have been used to examine apnoea (Cummiskey *et al.* 1982; East and East 1985; Pasterkamp *et al.* 1996b) and to diagnose tracheal obstruction and upper airway dysfunction (Pasterkamp and Sanchez 1992; Yonemaru *et al.* 1993). Flow rate can be estimated according to tracheal sound intensity if the relationship has been derived (Charbonneau *et al.* 1987b; Soufflet *et al.* 1990).

2.7.2 Diagnosis Based on Adventitious Sounds

Wheezes can be heard in several diseases (Waring *et al.* 1985; Meslier *et al.* 1995). Healthy people can also produce wheezes during forced expiration (Gavriely *et al.* 1987; Charbonneau *et al.* 1987a; Ploysongsang *et al.* 1988; Beck and Gavriely 1990; Shabtai-Musih *et al.* 1992). However, more wheezes were found in obstructive airway patients than that of normal subjects during forced expiration (Fiz *et al.* 2002). Wheezes are clinical signs of obstructive airway diseases if they are heard during spontaneous respiration or during induced

airway narrowing. Various patterns of wheezes may relate to different obstructions (Forgacs 1969; 1978; Waring *et al.* 1985).

The relationship between wheezes and the severity of airway obstruction has been widely studied (McFadden *et al.* 1973; Marini *et al.* 1979; Baughman and Loudon 1984; Pasterkamp *et al.* 1985; Bohadana *et al.* 1994b; Kiyokawa *et al.* 1999). No prediction could be made between absence or presence of wheeze and the degree of airflow obstruction (McFadden *et al.* 1973; Marini *et al.* 1979). No correlation was found between the intensity of wheezing and degree of obstruction (Baughman and Loudon 1984; Marini *et al.* 1979). Also no relation existed between pitch and pulmonary function (Pasterkamp *et al.* 1985). But the proportion of breath sounds occupied by wheezing corresponded to the severity of airway obstruction (Baughman and Loudon 1985; Fenton *et al.* 1985; Pasterkamp *et al.* 1985).

The relationship between wheezes and the severity of airway obstruction during bronchial provocation has also been extensively studied (Avital *et al.* 1988; ; Noviski *et al.* 1991; Beck *et al.* 1992; Spence *et al.* 1992, 1996; Bohadana *et al.* 1994b; Sprikkelman *et al.* 1996; Yong *et al.* 1999). In the majority of the patients, a correlation of presence of wheeze and a 20% fall in FEV_1 was found.

In summary, breath sounds contain enough information for airway obstruction estimation. Automatic wheeze detection algorithms facilitate the wheeze analysis (Fenton *et al.* 1985; Baughman and Loudon 1985; 1989; Shabtai-Musih *et al.*

1992; Ademovic *et al.* 1998b; Waris *et al.* 1998; Homs-Corbera *et al.* 2000; Fiz *et al.* 2002), so it is possible to assist the assessment of airway obstruction.

Wheeze and asthma. Most but not all asthmatics wheeze during deterioration of their disease. In early stages, wheezing may be heard over central airways during expiration. As asthma becomes more severe, it can be detected over the chest both in inspiration and expiration. When asthma becomes even more severe, wheezing may disappear (Loudon and Murphy 1984; Bohadana *et al.* 1995). The transmission of wheezes through airways is better than through the lung to the chest wall. So in most asthmatics, the trachea is superior to the chest for detection of wheezes (Pasterkamp *et al.* 1984; Fenton *et al.* 1985). Extensive investigation has been performed to detect wheezes in asthmatics (Baughman and Loudon 1984; Spence *et al.* 1992; 1996; Rietveld *et al.* 1995; 1996; Sprikkelman *et al.* 1996; Springer *et al.* 2000; Anderson *et al.* 2001). As the symptoms of asthma may be worse during the evening (BTS 2003), nocturnal wheeze detection has been attempted to trace the bronchoconstriction changes (Baughman and Loudon 1985; Lenclud *et al.* 1996; Kiyokawa *et al.* 1999).

Crackles usually correlate to different diseases (Raymond and Murphy 1985), but they can also be heard in healthy people during slow inspiration from residual volume (Workum *et al.* 1982; Thacker and Kraman 1982). It is generally accepted that fine and coarse crackles are associated with different conditions and so have diagnostic importance.

Many investigators have explored the potential correlation between crackle features and certain diseases (Nath and Capel 1974b; Epler *et al.* 1978; Shirai *et al.* 1981; Piirila *et al.* 1991; al Jarad *et al.* 1993; Vanderschoot *et al.* 1998). Nath and Capel pointed out that early inspiratory crackles are related to obstructive lung disorder while late inspiratory ones to restrictive dysfunction (Nath and Capel 1974b). Properties of crackles have been reported to change as the course of pneumonia progressed (Piirila 1992; Vanderschoot *et al.* 1998). Spectral shapes were distinct between asbestosis and pulmonary oedema patients (Urquhart *et al.* 1981; Banham *et al.* 1984). Various algorithms to detect and characterise crackles automatically have been developed (Urquhart *et al.* 1981; Pasika and Pengelly 1994; Sankur *et al.* 1996; Hadjileontiadis and Panas 1997c; Tolias *et al.* 1998; Vannuccini *et al.* 1998).

2.7.3 Self and Remote Monitoring

Self-monitoring and remote monitoring (Bhat *et al.* 1985; Rietveld *et al.* 1999; Scanlon 1999; Anderson *et al.* 2001; Woodward *et al.* 2001; Young *et al.* 2001) may benefit patients at work or at home in several aspects. In breath sounds monitoring there are two important factors. One is the measurement device, which should be compact and portable. The other is the quality of sounds, as these sounds are captured in spontaneous breath without registering flow rate under general conditions.

2.8 Standardisation Need

As mentioned above in 2.4 and 2.5, it is not possible to compare results from different research groups directly, because the data are often measured using

different measurement systems without calibration and analysed using different parameters. Murphy and Kramann were aware of the standardisation need (Murphy, Jr. *et al.* 1977; Kraman 1983a). This need was emphasised by Mussell later on (1992). The European project of CORSA (Computerised Respiratory Sound Analysis) aimed to provide a guideline for standardisation. The guideline was published in the European Respiratory Review (2000,vol 10).

Chapter 3 Study of Measurement System

This chapter focuses on the frequency characteristics of the measurement system. As described in section 2.4, investigators may choose different measuring systems to measure breath sound. Due to lack of standardisation of equipment, results may be difficult to compare directly between research groups.

Investigators don't care what the absolute S.P.L. (RMS sound pressure level) or the absolute vibration speed is at the body surface. Under repeatable conditions, they are interested in what frequency components occur and how much power is contained in measured breath sounds, so relative calibration is sufficient.

The aims of this chapter include:

1. Using a simple experimental set-up to test each component of the measurement system.
2. Using such equipment to build confidence in measurement data.
3. Verifying the stability of the measurement system.

3.1 General Setup of Measurement System

Figure 3.1 shows schematically the system structure in the study. Three measurement systems were investigated.

1. Two commercial electronic stethoscopes with an A/D device. The two stethoscopes have built-in amplifiers and filters. The outputs from the stethoscope were input to a sound card or a data acquisition device for digitisation.

2. A customised electronic stethoscope with or without filter and an A/D device. The customised electronic stethoscope was made from a high fidelity microphone with a matched amplifier; using an ordinary stethoscope's head as its air-coupler (refer to Figure 3.2). The signals from the microphone amplifier were either fed to a filter or not, then input to a sound card or a data acquisition device for digitisation.

3. Landline telephones and mobile telephones. They comprised the whole measurement system. By using a voice mail service (http://www.yac.com/), digitised signals were received directly.

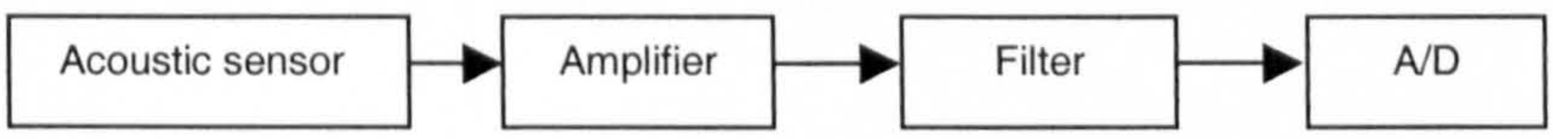

Figure 3.1 Schematic illustration of measurement system

Figure 3.2 Electronic stethoscopes: left-Escope; middle-Hp Stetho; right-customised.

3.2 Linear Time-invariant System

Each part of the measurement system was studied as a linear time-invariant system. The class of linear systems is defined as obeying the principle of superposition, stated as (Oppenheim 1989)

$$y[n] = f\{ax_1[n] + bx_2[n]\} = af\{x_1[n]\} + bf\{x_2[n]\} \tag{3.1}$$

where $y[n]$ is a discrete-time output sequence from the system; $x_1[n]$ and $x_2[n]$ are discrete-time input sequences to the system; and a, b are constants.

Any discrete-time sequence can be expressed as a sum of scaled, delayed impulses:

$$x[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k] \tag{3.2}$$

where

$$\delta[n] = \begin{cases} 0, & n \neq 0 \\ 1, & n = 0 \end{cases} \tag{3.3}$$

Specifically, let $h_k[n]$ be the response of the system to $\delta[n-k]$, i.e.,

$$h_k[n] = f\{\delta[n-k]\} \tag{3.4}$$

then for system $y[n] = f\{x[n]\}$, from equation (3.2)

$$y[n] = f\left\{\sum_{k=-\infty}^{\infty} x[k]\delta[n-k]\right\} \tag{3.5}$$

and from equations (3.1) and (3.4)

$$y[n] = \sum_{k=-\infty}^{\infty} x[k] f\{\delta[n-k]\} = \sum_{k=-\infty}^{\infty} x[k]\, h_k[n] \tag{3.6}$$

According to equation (3.6), the system response to any input can be described in terms of the response of the system to $\delta[n-k]$.

Time invariance means the properties do not vary with time. A time-shift in the input leads to a time-shift in the output. If $h[n]$ is the response to $\delta[n]$, then the response to $\delta[n-k]$ is $h[n-k]$. Thus equation (3.6) becomes

$$y[n]=\sum_{k=-\infty}^{\infty} x[k]\,h[n-k]=x[n]*h[n] \tag{3.7}$$

where * denotes convolution.

As a result of equation (3.7), a linear time-invariant (LTI) system is completely characterised by its impulse response, $h[n]$. That is, given $h[n]$, it is possible to use equation (3.7) to compute the output $y[n]$ due to any input $x[n]$.

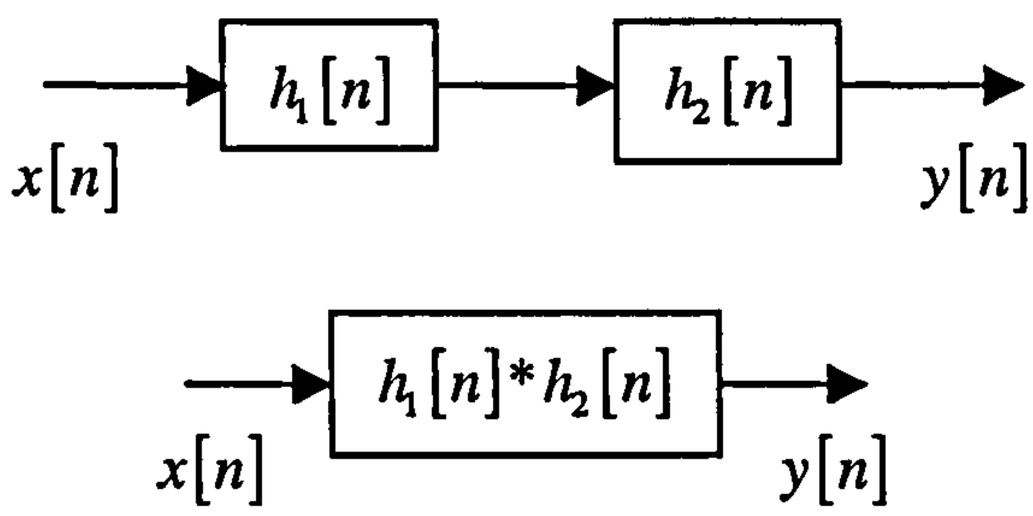

Figure 3.3 Two LTI systems with same responses

In a cascade connection of systems, the output of the previous system is the input to the next. The output of the last system is the overall output. Two linear time-invariant systems in cascade correspond to a linear time-invariant system with an impulse response that is the convolution of the impulse responses of the two systems. This is illustrated in Figure 3.3.

Since the frequency response and impulse response are directly related through the Discrete Time Fourier Transform (DTFT), the frequency response, assuming it exists, provides an equally complete characterisation of LTI systems.

The FT of the system input and output are related by

$$Y(\omega) = H(\omega) X(\omega) \tag{3.8}$$

where $X(\omega)$ and $Y(\omega)$ are the FT of the system input $x[n]$ and output $y[n]$ respectively, $H(\omega)$ is the transfer function (TF) or frequency response. With the frequency response expressed in polar form, the magnitude and phase of the FT of the system input and output are related by

$$|Y(\omega)| = |H(\omega)||X(\omega)| \tag{3.9}$$

$$\angle Y(\omega) = \angle H(\omega) + \angle X(\omega) \tag{3.10}$$

where $|\,|$ and $\angle$ denotes magnitude and phase respectively. $|H(\omega)|$ is the magnitude response or gain of the system, and $\angle H(\omega)$ is the phase response or phase shift of the system.

Frequency-preservation is a major property of linear systems. That is the output can only contain the same frequency components as the input.

The true transfer function of a LTI system is not always known, so the transfer function can be estimated by analysing the relationship between input and output. If the input signal can properly excite the system, then the TF can be estimated reasonably. Normally used input signals include swept sinusoidal signals, white noise etc which contain all frequency components of the LTI system.

3.3 Frequency Characteristics of Data Acquisition Device

An A/D converter is the last stage to digitise analogue data; its bandwidth and accuracy should not be a limit to the measurement of respiratory sounds. The device under investigation was a DAQPad-6200E, a product of National Instruments.

Manually swept sinusoidal signals of known amplitude and frequency from an analogue signal generator were used to test the device. The magnitude response at each frequency was calculated by the ratio of the root-mean-squared (RMS) value of the measured data to the RMS of the input signal. Details of the test process are in the appendix A.

Results

The test results (refer to Figure A.1, A.2 in the appendix) are comparable with the device's specifications (National Instruments 1998). Thus this device is suitable to sample respiratory sound signals.

3.4 Frequency Characteristics of Sound Card

3.4.1 TBS-2000 sound card as an A/D

TBS-2000 can act as a 16-bit digital sound recording device with selectable sampling rate from 4kHz to 48kHz.

Materials and methods

A 1V 1Hz square wave from the signal generator (SERVOMEX) was used. The square signal was more convenient than manually adjusted sinusoidal signals.

And the Fourier transform of the square signal contains all frequency components below the Nyquist frequency. The signal was fed to the line-in socket of the sound card and to one of the channels of the DAQPad simultaneously. The lineout signal from the sound card was connected to another channel. The sound card volume control was tuned to make the recording amplitude near full scale but avoid signal clipping. An existing LabView program was utilised to control and sample DAQPad signals. The sampling rate was 11025Hz. The sample length was about 10s.

Then 100mv, 10mv square waves were measured sequentially under the same volume level.

Next, a 100mV 1Hz square wave signal was input to the sound card. This time the sound card volume control was adjusted to make the recording amplitude near full scale again. Then signals of amplitude 10mV were tested under this volume level.

Results

The DC offset was removed before calculation. The magnitude response and phase response based on Welch's method (equation 3.11) (The Mathworks 1999) are shown in Figure 3.4 and Figure 3.5 respectively. Calculation using 1048 samples with a Hanning window of the same length with 50% overlap was undertaken in Matlab.

$$H(\omega) = \frac{P_{xy}(\omega)}{P_{xx}(\omega)} \tag{3.11}$$

Where P is the Welch's method of estimation of the power spectral density (Welch 1967). x is the input signal and y is the output signal. P_{xy} is the cross power spectral density of x and y; and P_{xx} is the power spectral density of x.

The coherence between x and y was calculated to act as a signal-to-noise ratio (SNR) index.

$$C_{xy} = \frac{|P_{xy}(\omega)|^2}{P_{xx}(\omega) P_{yy}(\omega)} \tag{3.12}$$

That is, $C_{xy}(\omega) \geq 0.75$ means good SNR at the frequency ω. So the estimated $H(\omega)$ at ω from equation 3.11 is acceptable.

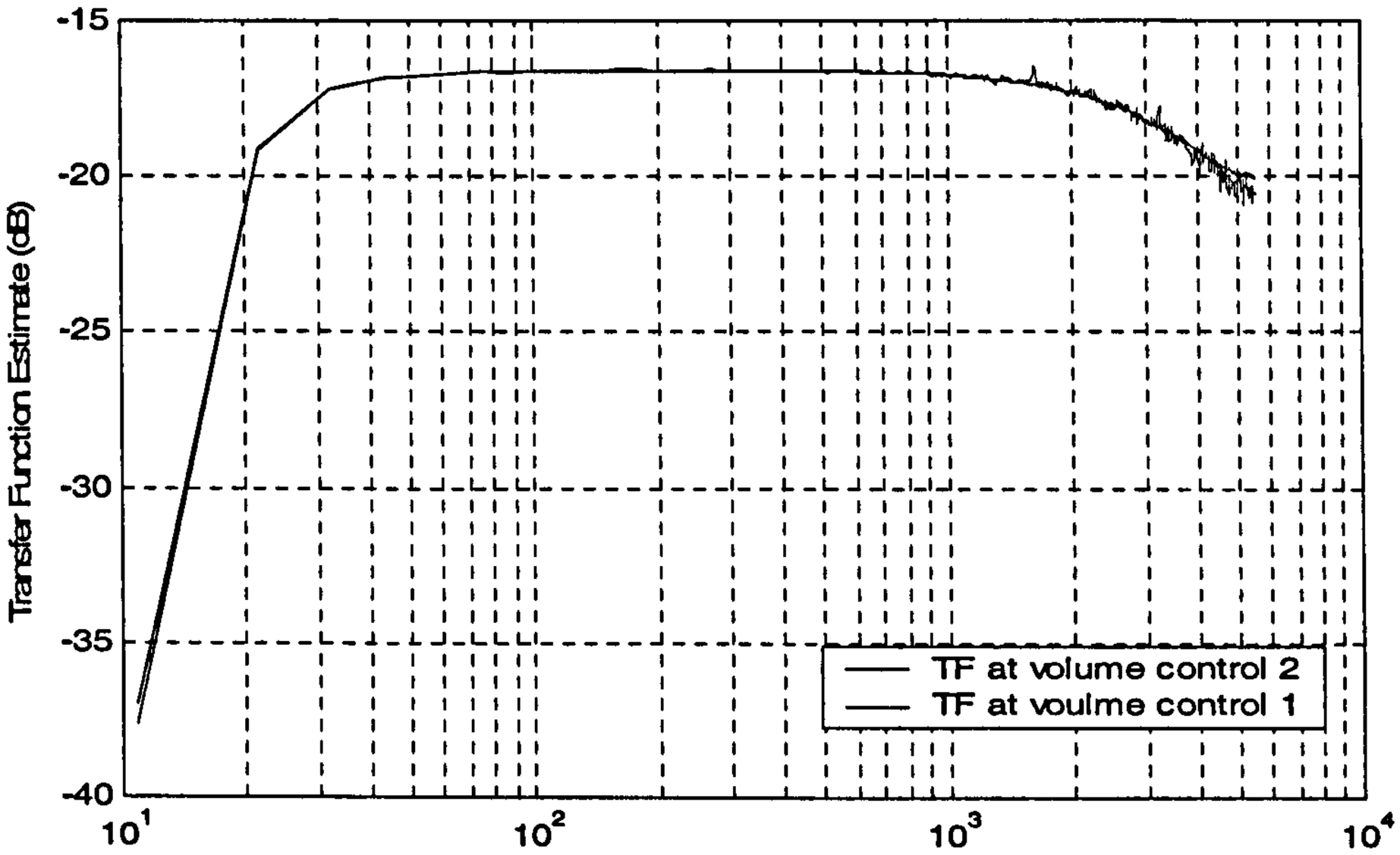

Figure 3.4 Magnitude response of TBS-2000 under different volume

It can be seen from Figure 3.4 that the TBS-2000 had a flat gain from 20Hz to the Nyquist frequency with ± 3dB. The two lines were almost superimposed. Figure

3.5 shows the phase response. Figure 3.6 shows that the measurement is linear at the same volume level. Thus the volume control acts like a linear amplifier.

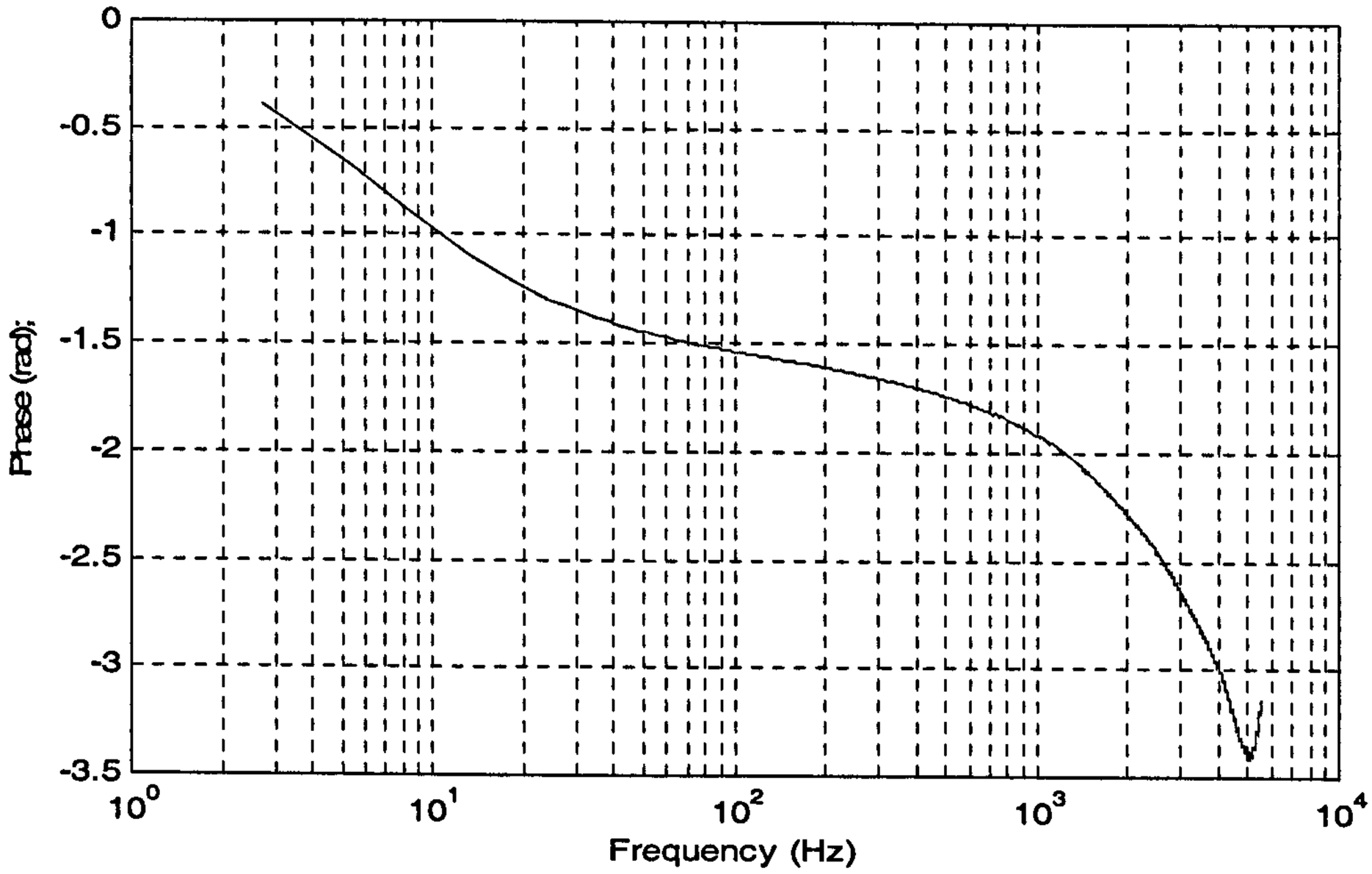

Figure 3.5 Phase response of TBS-2000.

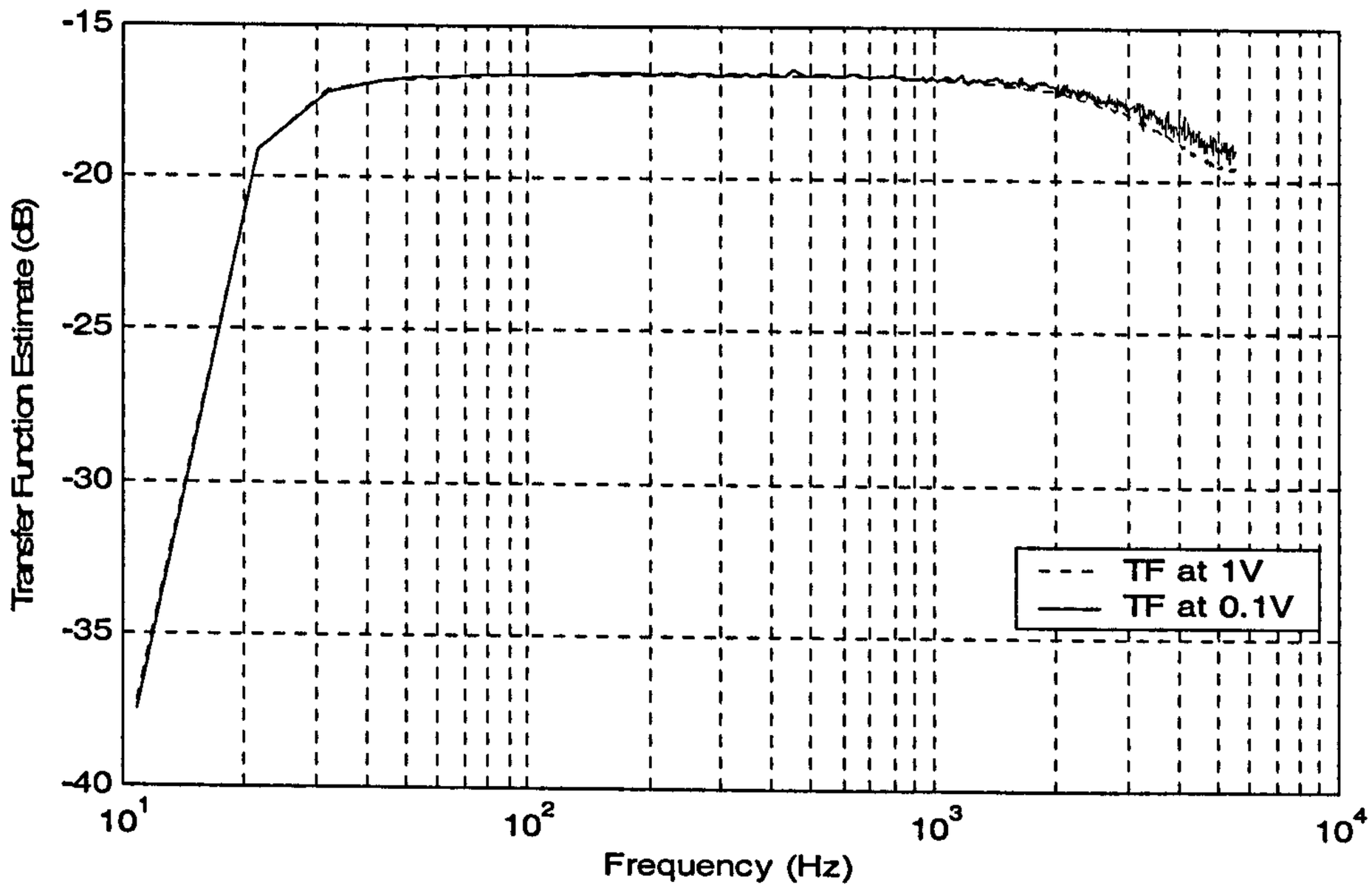

Figure 3.6 Magnitude response of TBS-2000 under same volume

The test results are comparable with the device's specifications. Therefore the sound card is suitable to sample signals above 20Hz at a certain volume level.

3.4.2 Sound card as an audio output

The TBS-2000 also can act as an 8-bit playback device. In the later sections, the sound card is connected to a loudspeaker, which is used to produce a sound field for sensor tests.

Materials and methods

A programmed 0-4kHz band limited white noise wave[1] was played via the sound card. The signal from the line-out socket was sampled by the DAQPad-6200E. Then the programmed wave was played under different playback volumes.

Results

The TBS-2000 had a flat gain from 20Hz to the Nyquist frequency with ±3dB (refer to Figure A.3). The volume control acts like a linear amplifier.

3.5 Frequency Characteristics of Filters

A low pass filter and a high pass filter (both BARR and STROUD) were studied. Both filters featured tuneable cut-off frequency.

Materials and methods

A square wave from the above-mentioned signal generator was input to the filter under test. Two channels of the DAQPad-6200E were used to sample input and output signals simultaneously at 11025Hz sampling rate.

---

[1] All programmed wave files were in PCM format.

Results

An example measurement of the low pass filter with cut-off frequency 3000Hz is shown in the appendix (Figure A.4). An example of the high pass filter with cut-off frequency 100Hz is shown in Figure A.5.

3.6 Sensors

Figure 3.7 shows schematically the sensor test system. Wave files were played via the TBS-2000 sound card. The line out signal from the sound card was input to a DAQPad-6200E channel, and the speaker out signal from the sound card was connected to an 8-ohm loudspeaker. The loudspeaker was the sound source. A membrane was stuck on the loudspeaker frame to be a contact surface. When a sensor was placed onto or near the loudspeaker, it measured vibration speed or sound pressure. Signals from the sensor were fed to another DAQ channel. Sampled data were saved in text files for subsequent processing.

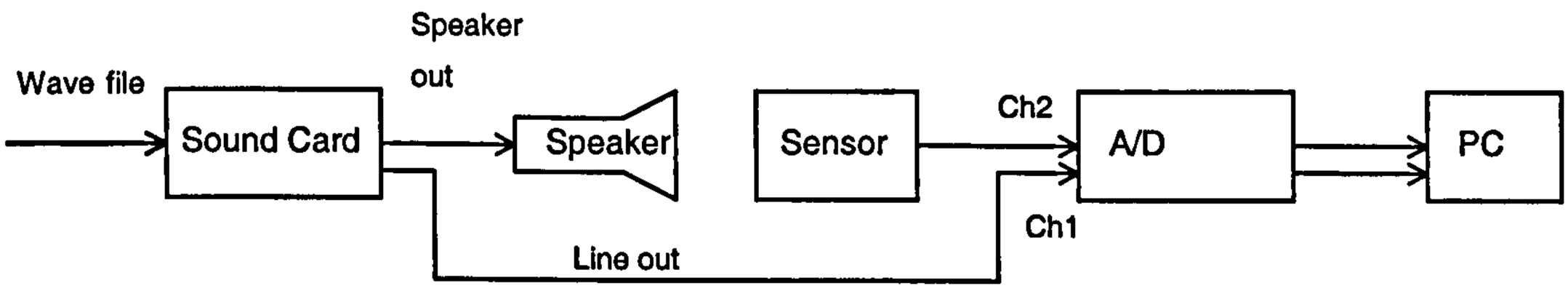

Figure 3.7 Schematic of sensor test system.

3.6.1 Loudspeaker Frequency Response

The sensor test system was considered as a cascade LTI system. So the loudspeaker's frequency response was studied first.

Materials and methods

Programmed wave files containing 0-4kHz band limited white noise were played via the sound card under fixed volume control. The loudspeaker (JVC SP-MX1BKE) was covered with a membrane to form a contact surface.

A high quality microphone (CANFORD MB-C550) was placed close to the loudspeaker to measure the sound pressure. A laser Doppler vibrometer (LDV, Polytec OFV 303) was used to measure the vibration velocity at the membrane centre. As the membrane is non-reflecting, a tiny piece of reflective tape was stuck on the membrane centre.

Results

For a one-dimensional sound wave travelling along in a tube, the relationship between pressure and velocity is (Turner and Pretlove 1991):

$$p = \rho c \frac{\partial s}{\partial t} \tag{3.13}$$

where p is pressure, ρc is characteristic impedance, s is displacement, $\partial s / \partial t$ is velocity.

Assuming the microphone was measuring the sound pressure at the membrane surface, then when the input signal (lineout signal from soundcard) is the same, the estimated transfer functions (TFs) of the loudspeaker should have the same shape.

But as the placement of microphone influenced the free field sound pressure, the velocity was measured at one point, and due to the influences of immeasurable noises, the two curves in Figure 3.8 did not have the same shape. They are similar in the frequency range of 100Hz-2kHz.

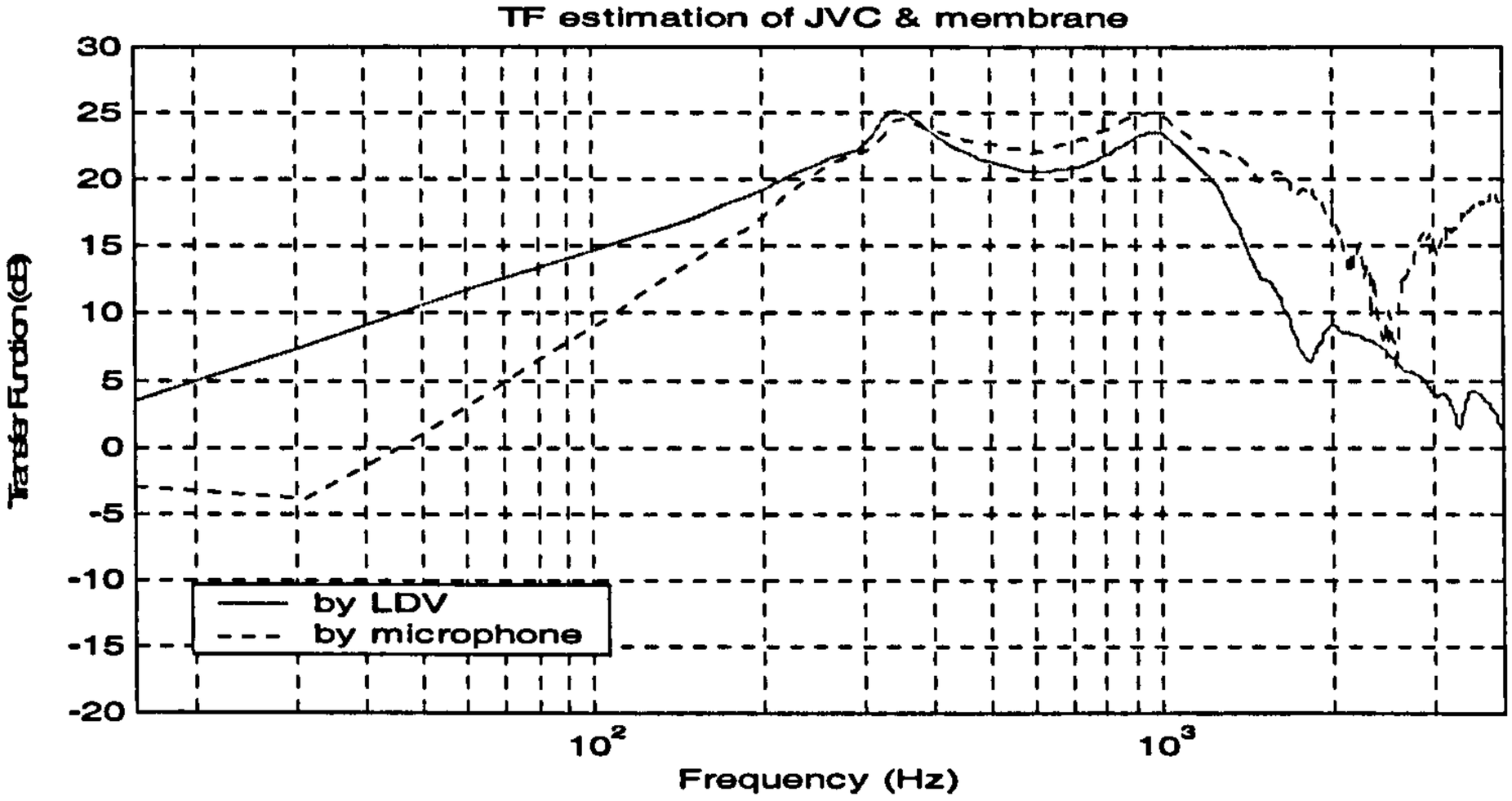

Figure 3.8 Estimated transfer function of the loudspeaker.

In later experiments, the TF of the loudspeaker estimated using the microphone signal was used, because this TF relates the pressure produced by the loudspeaker to its input signal, and the tested sensors were sound pressure transducers.

The linearity of the relationship between vibration speed and sound card volume was also checked. There is a linear relationship (refer to Figure A.6).

3.6.2 Non-contact Frequency Response

Materials and Methods

Programmed wave files containing band-limited white noise were played via the sound card under a fixed volume control. The sensor under test was placed close to the loudspeaker to measure the sound pressure.

Two methods were tried. 1. The input signal was 'coloured' to make the membrane vibration signal a white noise. Thus the input signal (sound pressure) to the sensor is a white noise. The magnitude response of the sensor could then be estimated by:

$$\left|H(\omega)\right| = \frac{\sqrt{P_{yy}(\omega)}}{c} \tag{3.14}$$

where P_{yy} is the power spectral density of the sensor output signal, and c is a constant which represents the white noise spectral level.

When the TF of the loudspeaker is obtained, say $H_1(\omega)$, then the colouring noise at the input of the loudspeaker is calculated as:

$$x[n] = s\,\mathrm{Re}\left\{ifft\left[\left(\frac{1}{H_1(\omega)}\right)e^{j\theta}\right]\right\} \tag{3.15}$$

where θ is a random phase distribution between 0 and 2π, $ifft$ is the inverse fast Fourier transform, Re is real operator, and s is a constant scale to make $\left|x[n]\right| \le 1$. (The amplitude of a wave of PCM format should not be greater than 1).

2. Considering the loudspeaker as cascaded with the sensor, the TF of the whole is estimated, which is $H(\omega)$. Because the TF of the loudspeaker is known, i.e. $H_1(\omega)$, the TF of the sensor $H_2(\omega)$ could be deduced from:

$$H_2(\omega) = {H(\omega)}/{H_1(\omega)} \tag{3.16}$$

Although the two methods were different, the physical meaning was the same. Experiments showed that when the coloured noise was played via the loudspeaker, the sound pressure the loudspeaker produced was roughly 'white noise'. But using equation 3.15 to estimate the TF would induce some errors, as the actual noise spectrum level was not a constant. So the second method was employed. The following sensors were tested: three electronic stethoscopes- Escope (Cardionics), Hp Stetho (Hewlett Packard), and a microphone (Canford MB-C550) with a Littmann chestpiece; and a Nokia 3310e mobile phone.

The lineout signal from the sound card was the input signal, and the output from the sensor was the output signal. The sampling rate was 8000Hz to be consistent, because the GSM 06.10 sampling rate is 8000Hz.

The pathway of signals from a mobile phone was different from the electronic stethoscopes. By dialling a number provided by a voicemail service (http://www.yac.com/), the acoustic signal captured by the mobile phone was sent to an email address as an attached file. The attachment was a wave file in GSM 06.10 format (13 bit sampling at 8000Hz) with a time stamp.

Results

The transfer function was estimated by using equation 3.11. Figure 3.9 shows the non-contact TFs of the sensors. None of the sensors has a flat response. The Hp Stethos transfer function in diaphragm mode is comparable with its manual specification (Hewlett Packard 1999). A mobile phone's frequency bandwidth is from 300Hz to 3400Hz.

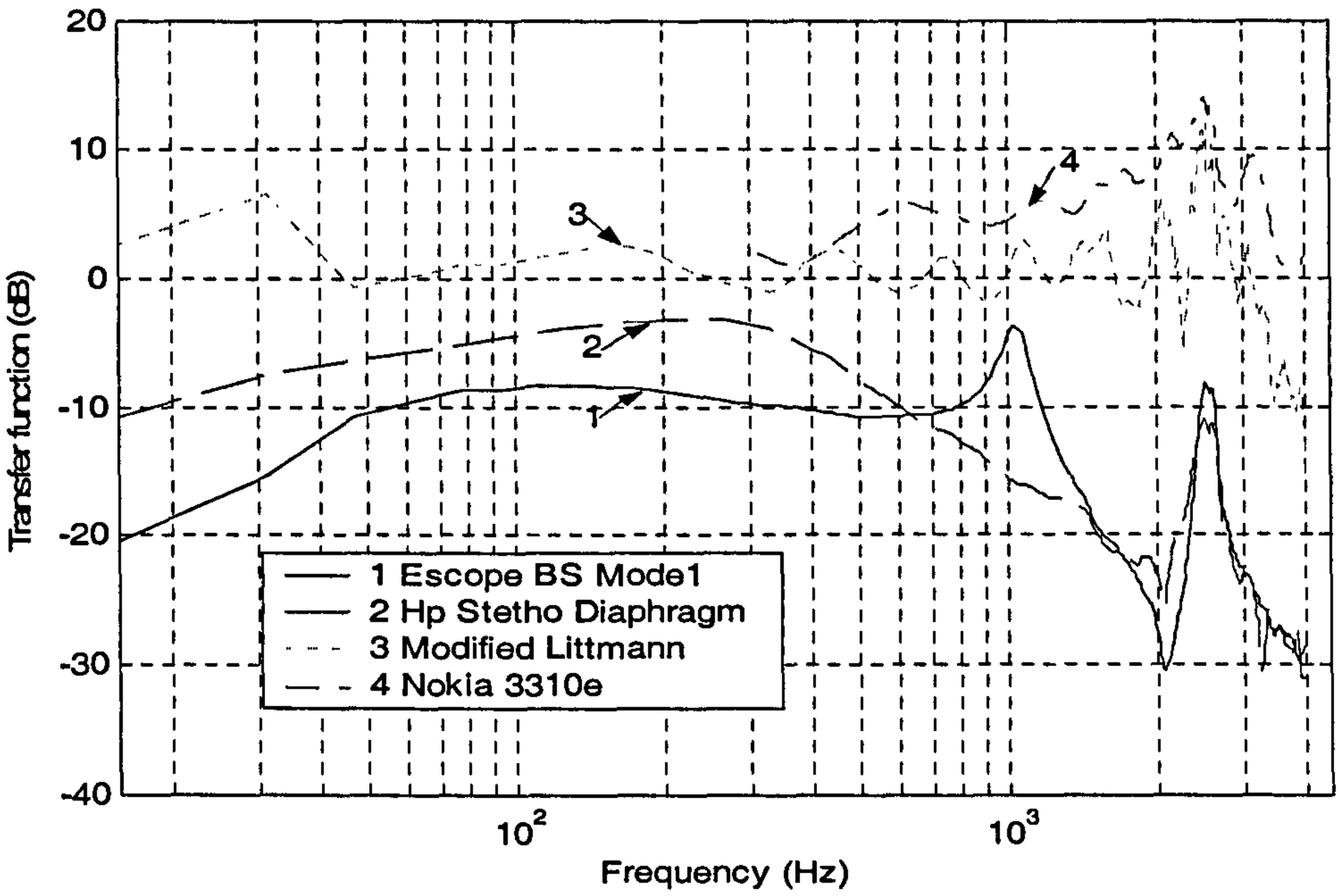

Figure 3.9 Estimated TFs.

3.6.3 Validation of non-contact frequency response

Materials and methods

A procedure was performed to see by calibration whether the calibrated spectrum would be same as the true spectrum. A wave file containing normal tracheal breath sounds was played via the sound card. The sensors were again placed in front of the loudspeaker to collect signals.

Results

Figure 3.10 shows that the power spectral density (PSD) after calibration is more similar to the true PSD than that before calibration. The peak around 300Hz in the PSDs of the measured signal was mainly due to the loudspeaker (refer to Figure 3.8).

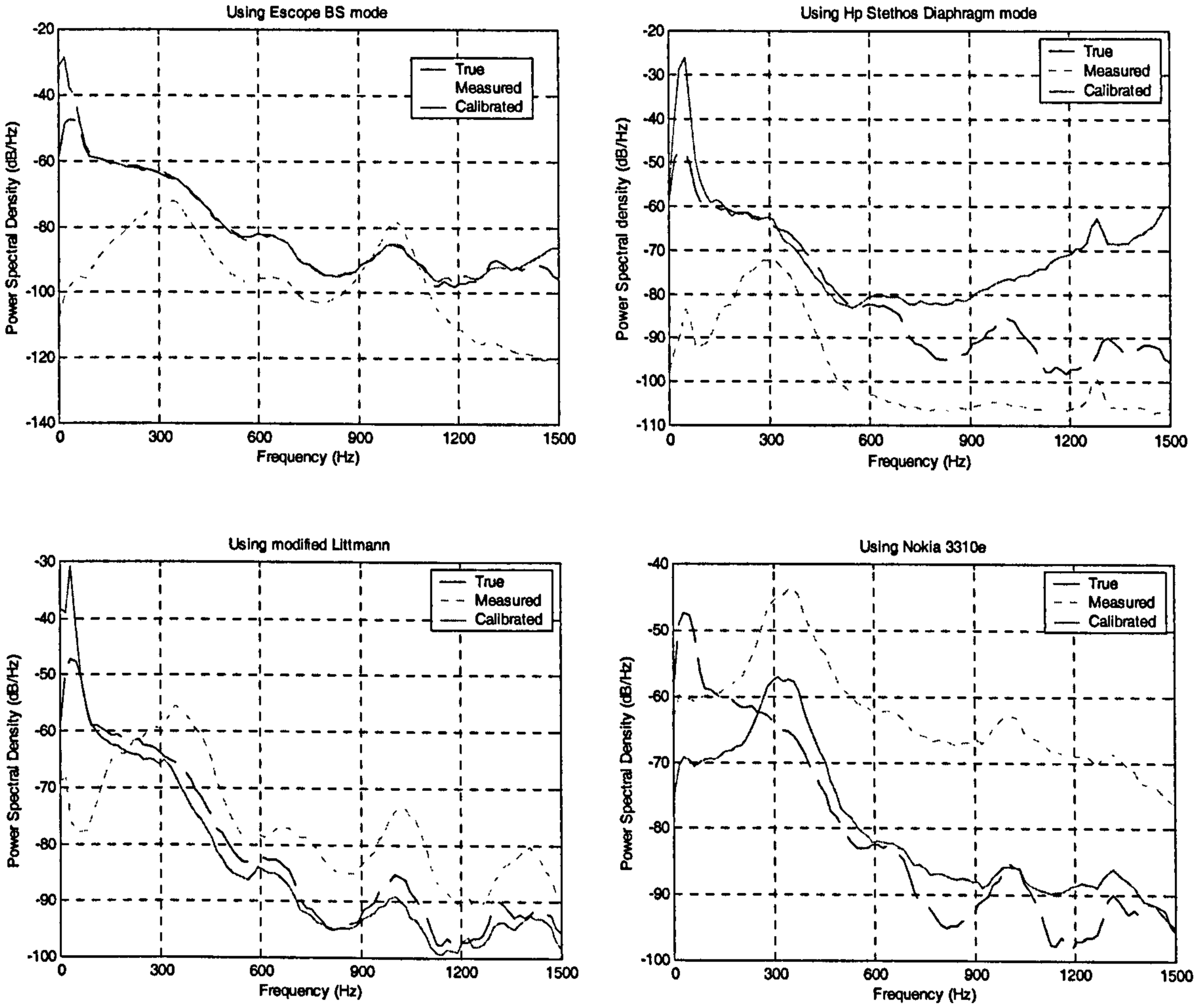

Figure 3.10 True, measured and calibrated PSDs.

3.6.4 Simplified Model of Chest Wall as a SDOF System

Gavriely (Gavriely and Cugell 1995) used a simplified single degree of freedom (SDOF) model of the chest wall to study the mechanical impedance and natural frequency of the chest wall. This model was used here to study the sensor's load effect on the chest (or trachea).

1. Without a sensor

The part of chest wall to be under sensor is modelled as below.

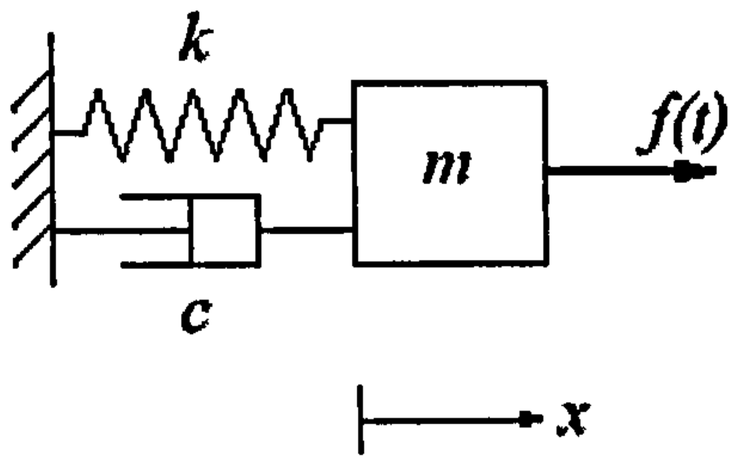

Figure 3.11 SDOF model of chest wall. k is stiffness, c is viscous damping, m is mass, f is acoustic force acting at the chest wall, and x is displacement.

The equation of motion is expressed as below when the excitation force is harmonic $f(t) = F_0 \sin(\omega t)$,

$$F_0 \sin(\omega t) = m\ddot{x} + c\dot{x} + kx \tag{3.17}$$

The solution to this equation is

$$x(t) = c_1 x_c(t) + c_2 x_p(t) \tag{3.18}$$

where c_1 and c_2 are constants, $x_c(t)$ is the complimentary solution, and $x_p(t)$ is the particular solution.

$$x_c(t) = e^{-\zeta\omega_n t}\left[x_0 \cos(\omega_d t) + \frac{v_0 + \zeta\omega_n x_0}{\omega_d}\sin(\omega_d t)\right] \tag{3.19}$$

x_0 and v_0 are the initial (t=0) position and velocity respectively.

$$\zeta = \frac{c}{2m\omega_n} \tag{3.20}$$

$$\omega_n = \sqrt{\frac{k}{m}} \tag{3.21}$$

$$\omega_d = \omega_n \sqrt{1-\zeta^2} \tag{3.22}$$

$$x_p(t) = \frac{F_0}{\sqrt{\left(k-m\omega^2\right)^2 + (c\omega)^2}} \sin\left(\omega t - \tan^{-1}\left(\frac{c\omega}{k-m\omega^2}\right)\right) \tag{3.23}$$

2. With a sensor

The chest wall and the sensor are modelled as below.

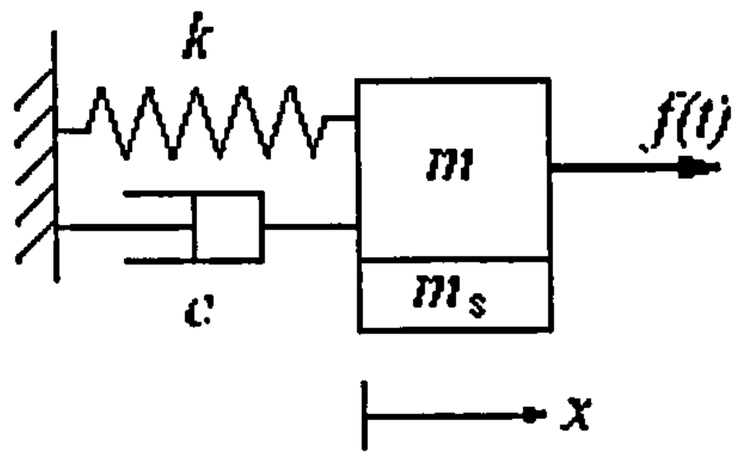

Figure 3.12 SDOF model of chest wall with a sensor, m_s is mass of sensor.

From the above equations, it can be deduced that the additional sensor mass will change $x(t)$ from that without the sensor. That is, the measured signal on the surface of the chest wall will be distorted by the presence of the sensor.

3.6.5 Hydraulic Transmission of an Air-coupler

The air-coupler (stethoscope chestpiece) between body surface and microphone is regarded as a hydraulic system. Movement of the diaphragm causes a pressure change inside the air chamber. According to Boyle's law:

$$V_1 P_1 = V_2 P_2 = C \tag{3.24}$$

where V and P are volume and pressure respectively, C is a constant.

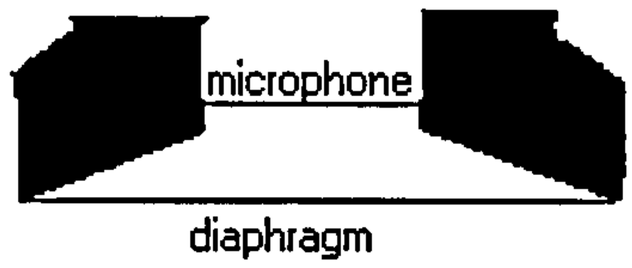

Figure 3.13 Graph of stethoscope's head.

Assuming that the volume of the air-coupler is V_0 and pressure is P_0 when it is not placed on the body, that when it is placed on the body, the body surface vibration induces the diaphragm motion. The amplitude of the diaphragm motion, x_d, involving its area, A_d, causes a volume change of Δ_v and a pressure change of Δ_p. From the above equation 3.24,

$$V_0 P_0 = (V_0 + \Delta_v)(P_0 + \Delta_p) \tag{3.25}$$

Expanding the right side of the equation and ignoring the product $\Delta_v \Delta_p$,

$$\Delta_p = -P_0 \frac{\Delta_v}{V_0} = -\frac{CA_d}{V_0^2} x_d \tag{3.26}$$

The pressure change at the microphone is equal to Δ_p, where the wave length is much larger than the coupler length (Gavriely and Cugell 1995). The microphone output is proportional to Δ_p, thus is proportional to diaphragm displacement x_d. When taking Δ_p as the output and x_d as the input, the transfer function of the electronic stethoscope is determined.

When the stethoscope contacts the body, the initial volume is V_i but not V_0 (contact air) due to the diaphragm deformation under force. Thus, the transfer function of the stethoscope in air (non-contact) is different from that in contact. However, if diaphragm is very rigid, then transfer functions are the same.

3.6.6 LDV Measurements

As the LDV is a non-contact vibrometer, its application to the breath sounds measurement was attempted.

Materials and methods

Measurements were undertaken on the chest wall and over the trachea of sitting normal subjects. On any measurement site, a tiny piece of reflective tape was stuck on the skin to reflect the laser beam. The subjects tried hard to hold their posture still during deep breathing. The LDV output was sampled by the DAQPad with 11025Hz sampling rate.

Results

During breathing the subject's unconscious body movement usually induced strong 'noise'. Meantime two factors, the body expanding and contracting with the breath cycle and heart-beating impulses, led to great energy in the low frequency region. When played back, the breath sounds could be heard clearly. An example of a measured tracheal signal in the time domain is shown in Figure 3.14. The digital high pass filtered counterpart is shown in Figure 3.15, which shows the normal breath pattern (spikes correspond to noise).

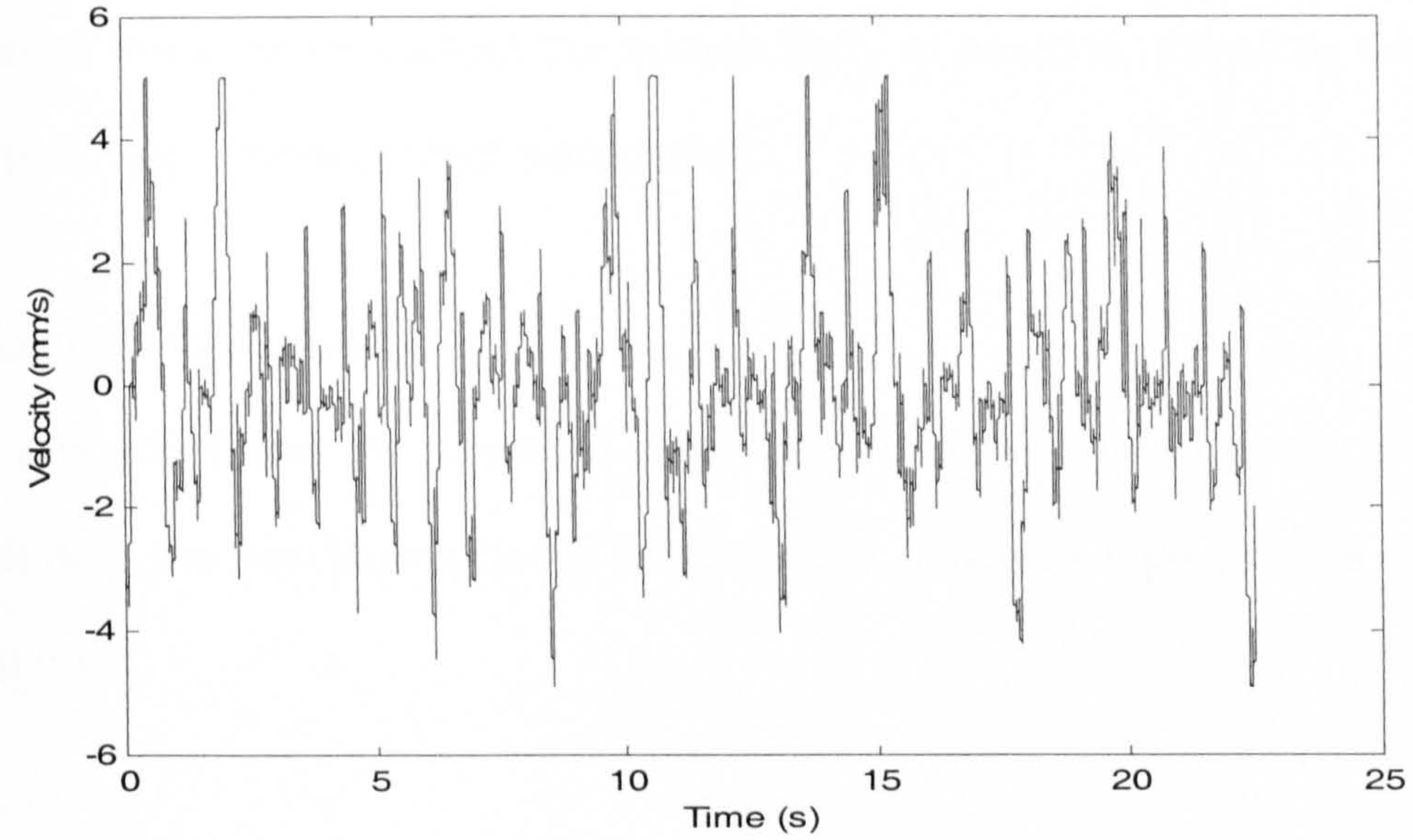

Figure 3.14 LDV signal of tracheal sounds.

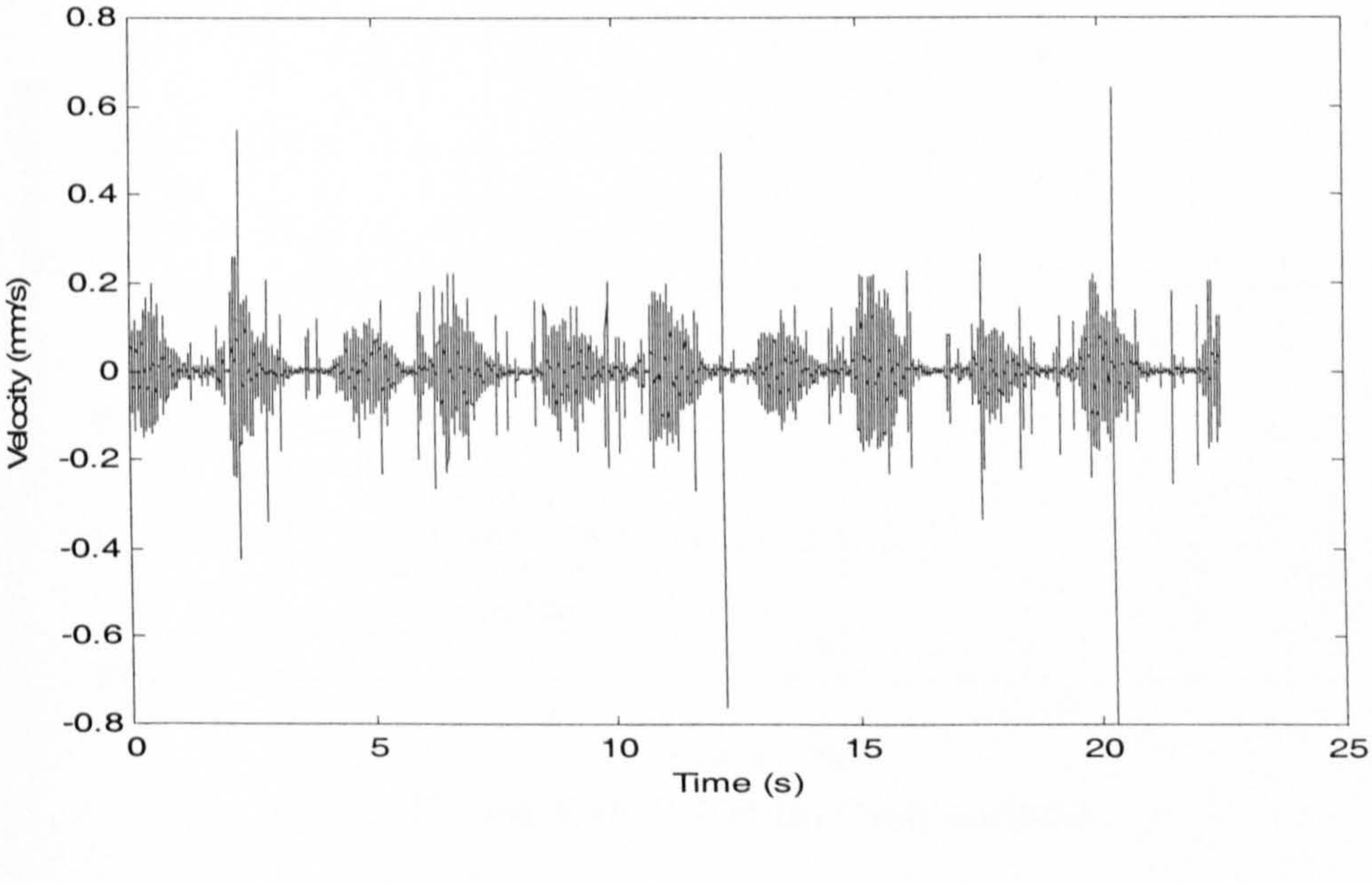

Figure 3.15 Digital high pass filtered LDV signal.

3.6.7 Contact Frequency Response

Because the sensors contact the human body in practice, therefore the sensors' characteristics under contact were tested.

Materials and methods

The procedure was the same as described in 3.6.2 except that the sensor was stuck onto the membrane using double-sided adhesive tape (mobile phone was hand held).

Results

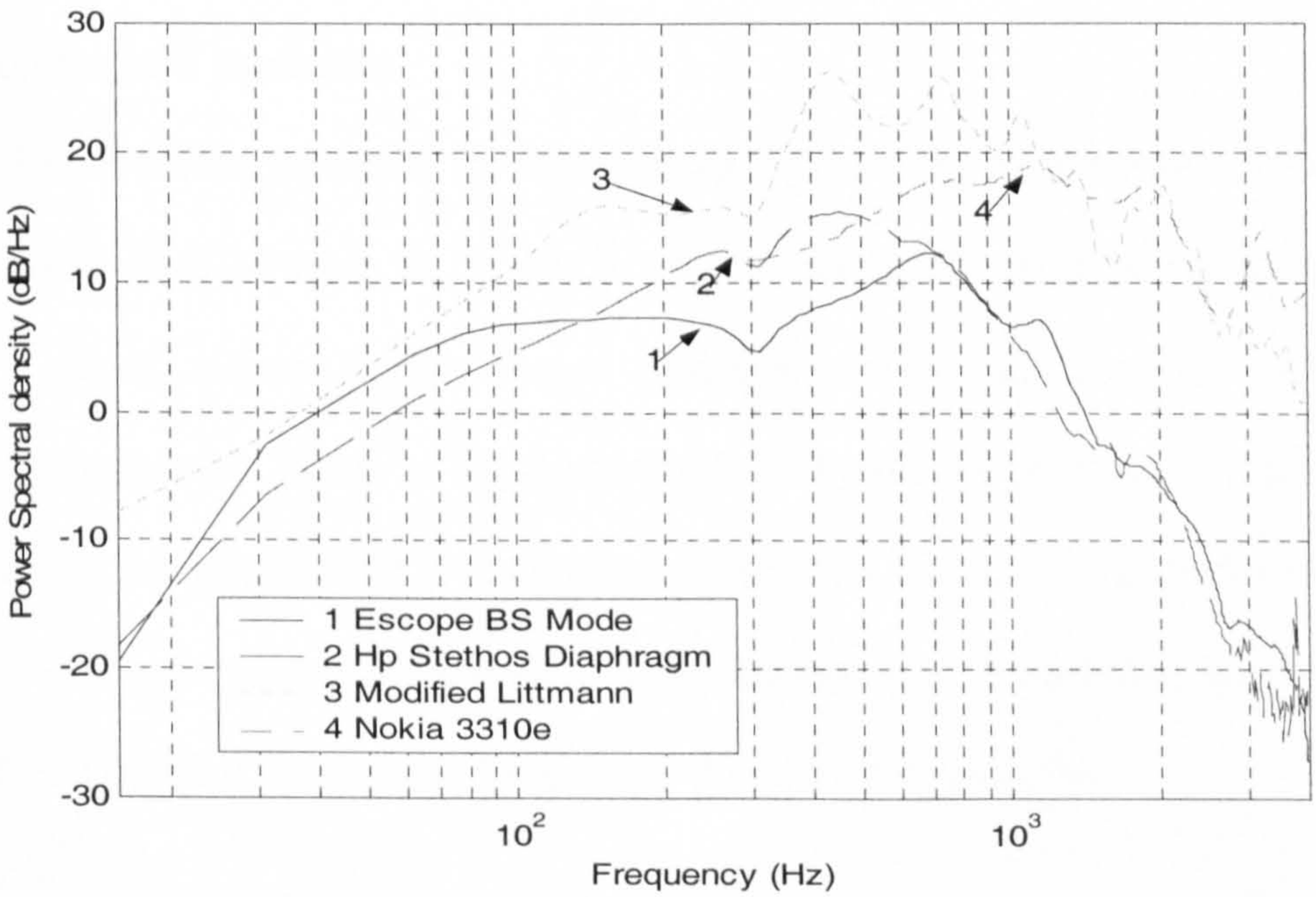

Figure 3.16 PSDs on thick surface.

The power spectrum of each sensor on a thick surface was calculated (see section 3.6.8 for a description of the "thick" surface). Figure 3.16 shows the outcomes. It is not unexpected that the results show differences. One reason is that different

sensors were used. The other reason is that when the membrane is loaded with a sensor its vibration response will change. This effect is reflected in the results.

3.6.8 Effects of Contact

There are two effects between an air-coupled sensor in contact with the membrane. One is on the membrane and the other is on the sensor. For the former, when the membrane is loaded with a sensor, its vibration response is changed, thus its sound transmission properties changed. For the latter, when an air-coupler sensor is in contact with the membrane, the pressure between them changes the acoustic properties of the air-coupler as well.

Materials and methods

The Escope was selected for the test as it could be disassembled easily. Three materials were chosen to be stuck on the loudspeaker: a piece of loudspeaker cover paper, a 3mm thick piece of synthetic rubber with one side of stiff plastic, and a 5mm thick piece of synthetic rubber. Again programmed white noise was played as an input signal to the loudspeaker. The Escope was stuck onto the three materials surfaces consecutively. Its output signal was sampled by the DAQPad-6200E. The LDV was used to measure the vibration velocity.

First, the velocity at the centre of the thick surface was measured by the LDV. Secondly, the chestpiece of the Escope without the diaphragm was stuck onto the middle of the surface. The velocity at the centre of the thick surface was measured again. Then the chestpiece with the diaphragm was stuck onto the middle of the

surface. The velocities of the surface centre and diaphragm centre were measured respectively.

Then the thick surface was replaced by the thin surface, and the above procedures repeated.

Results

Experiments showed that the Escope's contact TF with paper was almost the same as that of non-contact, as shown in Figure A.7. This indicates that the paper was 'acoustically transparent'. Thus, the paper was excluded from further study.

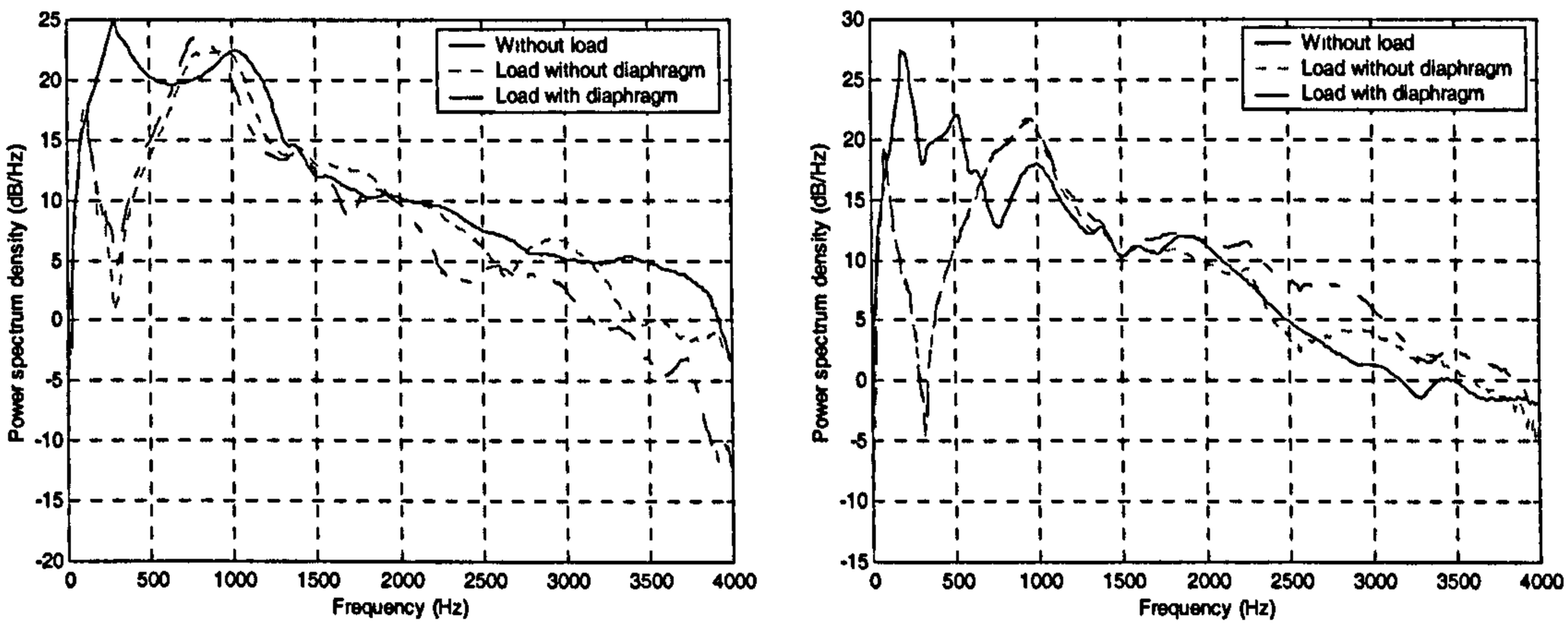

Figure 3.17 PSDs without and with load. Left- on thick; right- on thin.

Figure 3.17 shows the spectrum difference with and without load, especially in the frequency range of about 200 - 1200 Hz. The mass loading of the surface changed the surface vibration. Figure 3.18 shows the effects of contact between the Escope and different surfaces. The results also show that the diaphragm could almost follow the vibration of the surface (refer to Figure A.8).

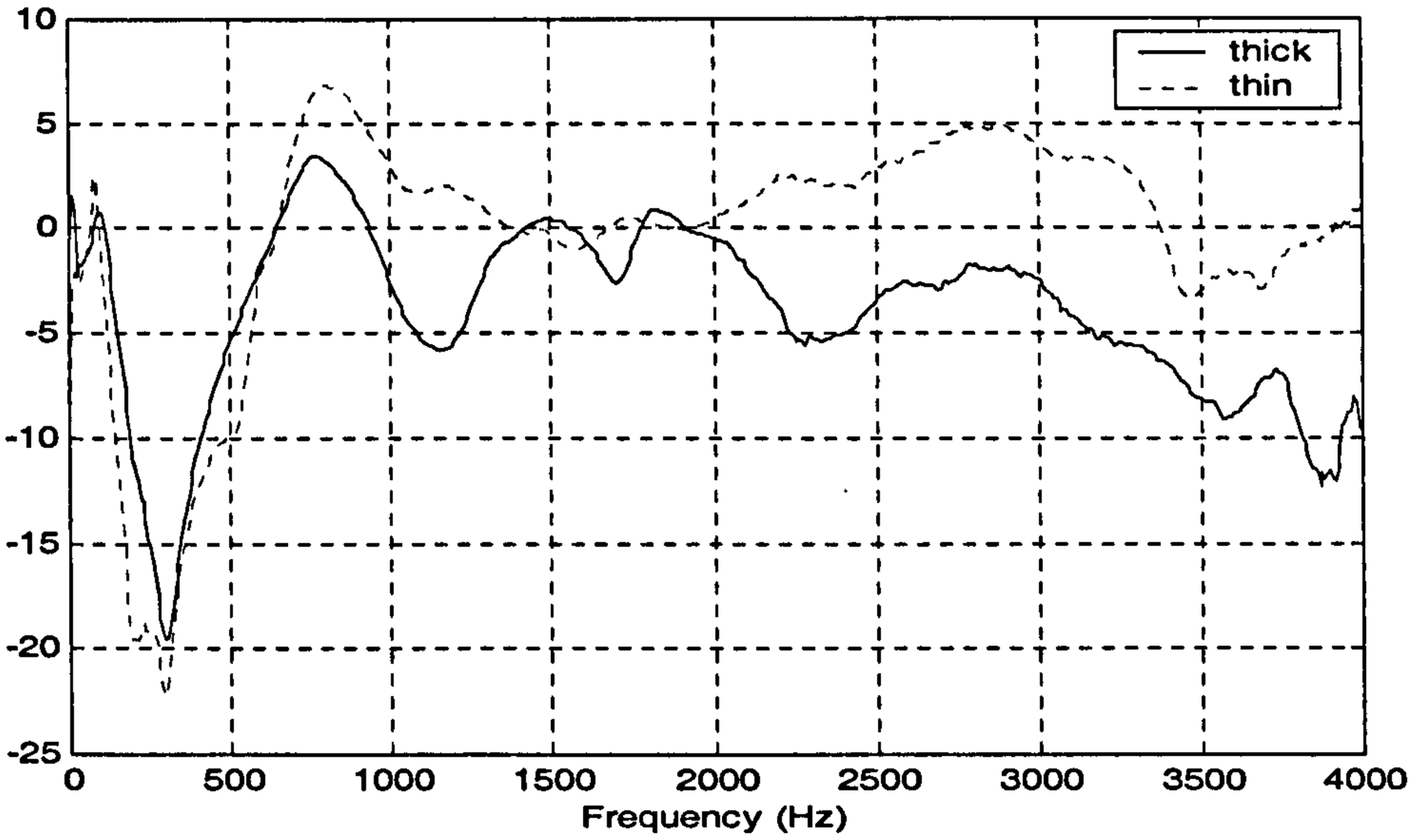

Figure 3.18 Effects of contact between Escope and surface.

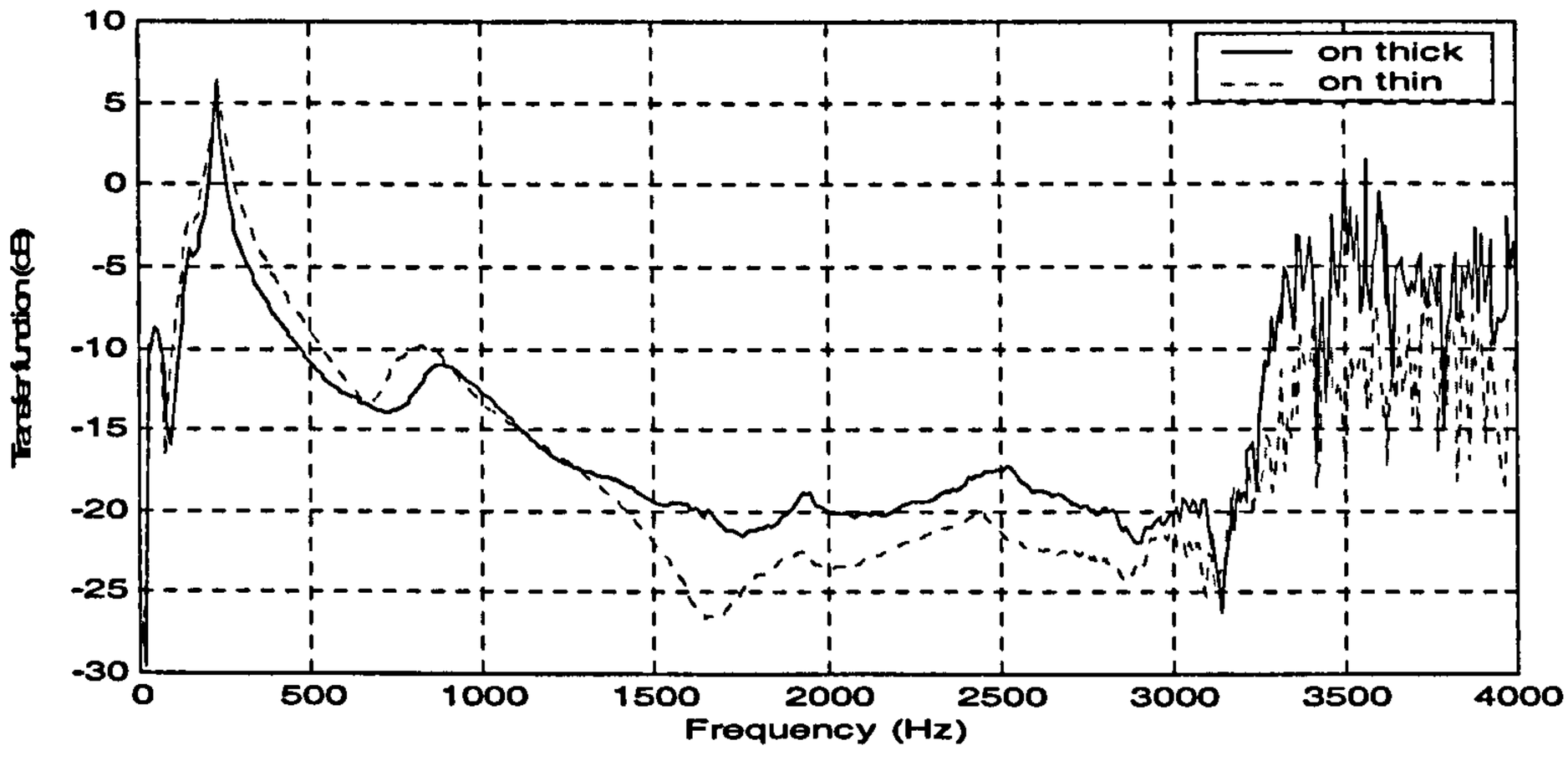

Figure 3.19 Escope contact TFs estimated on thick and thin materials.

Figure 3.19 shows the contact TFs of the Escope, excluding the contact effects. The two lines have similar shape below 3kHz. Due to bad signal-to-noise ratio above 3kHz, the TFs estimated above 3kHz are unreliable.

3.6.9 Validation of contact measurements

Although LDV measurement data is not a perfect reference, it contains information on vibration velocity at the body surface without mass loading. As the Escope's contact frequency response has been studied, it was possible to compare LDV data with calibrated Escope data.

Materials and methods

The LDV and Escope were used simultaneously to measure tracheal breath sounds of a sitting normal subject. The Escope was stuck on the right of the neck (refer to Figure 3.23 point 3), a tiny piece of reflective tape was stuck on the left symmetric position (point 4). The subject undertook deep breathing. These two channel signals were sampled by the sound card at a sampling rate of 8000Hz with 16-bit accuracy.

Results

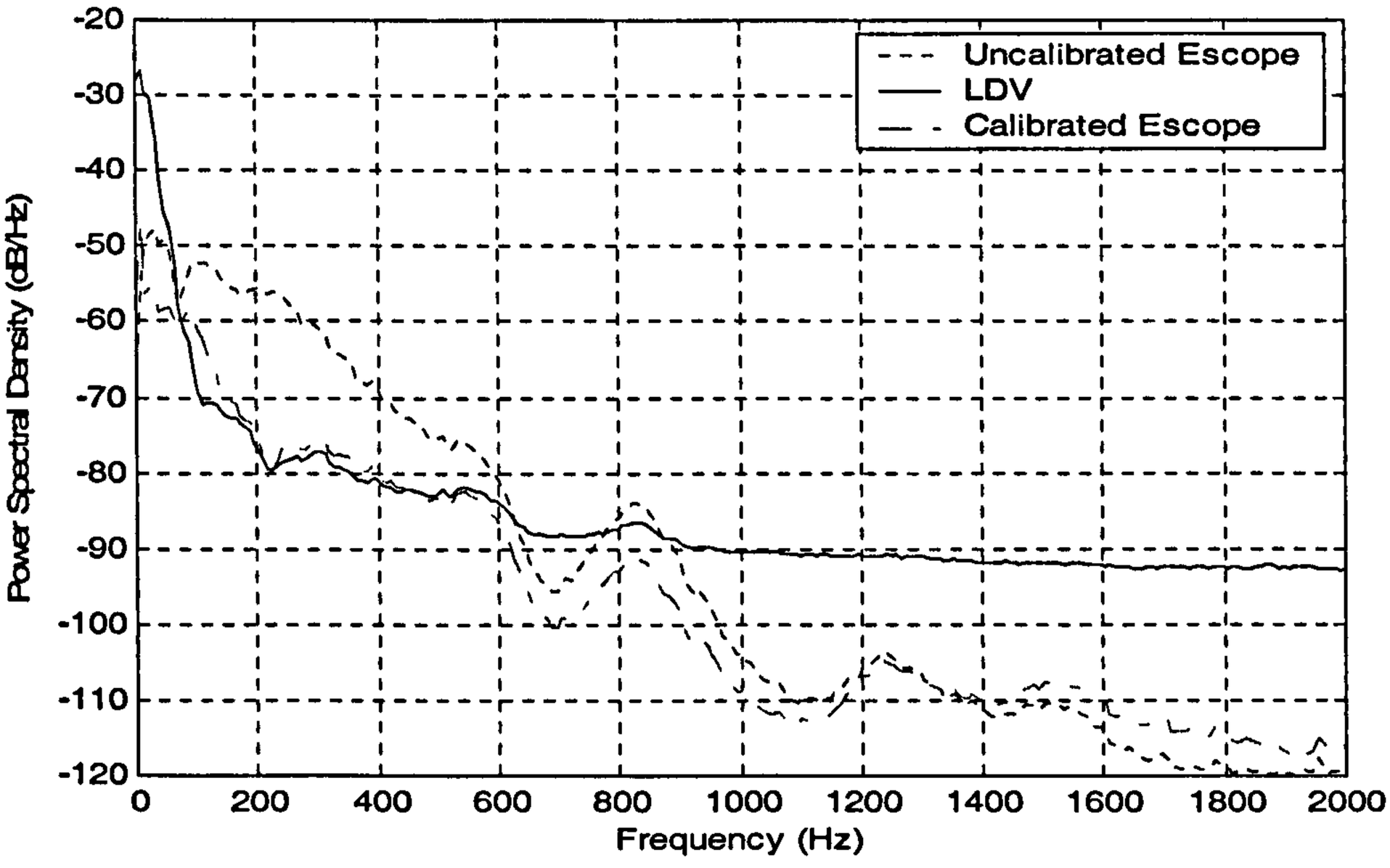

Figure 3.20 PSDs of LDV, uncalibrated Escope and calibrated Escope data.

It can be seen in Figure 3.20 that the calibrated Escope spectrum differs from the uncalibrated one mainly in a frequency range of 100 to 500Hz. The biggest difference is around 200Hz, where there is a trough in the calibrated spectrum, while it is flat in the uncalibrated spectrum. The calibrated spectrum is more similar to the LDV spectrum than the uncalibrated one.

3.7 Repeatability

The following experiments were done to study the factors that could influence repeatability. The recording site was on a sitting healthy subject's extrathoracic trachea as this site was relatively easy to measure.

3.7.1 Flow Rate

Materials and Methods

The Escope was stuck on the subject's anterior cervical triangle (refer to Figure 3.23 point 1) with double-sided adhesive tape. The subject tried to keep each breath cycle repeatable. Two measurements were taken for quiet breathing and deep breathing. For each breathing method, 5 cycles of breath sounds were sampled by the sound card with a sampling rate of 8000Hz.

Results

The PSDs were calculated based on 512 samples, using a Hanning window with 256 samples overlap. It can be seen from Figure 3.21 that flow rate (quiet breath, deep breath) influences spectrum level and shape. Deep breath produces higher power and more peaks than quiet breath.

Under the same conditions, the measurements are acceptably repeatable. An example is shown in Figure A.9.

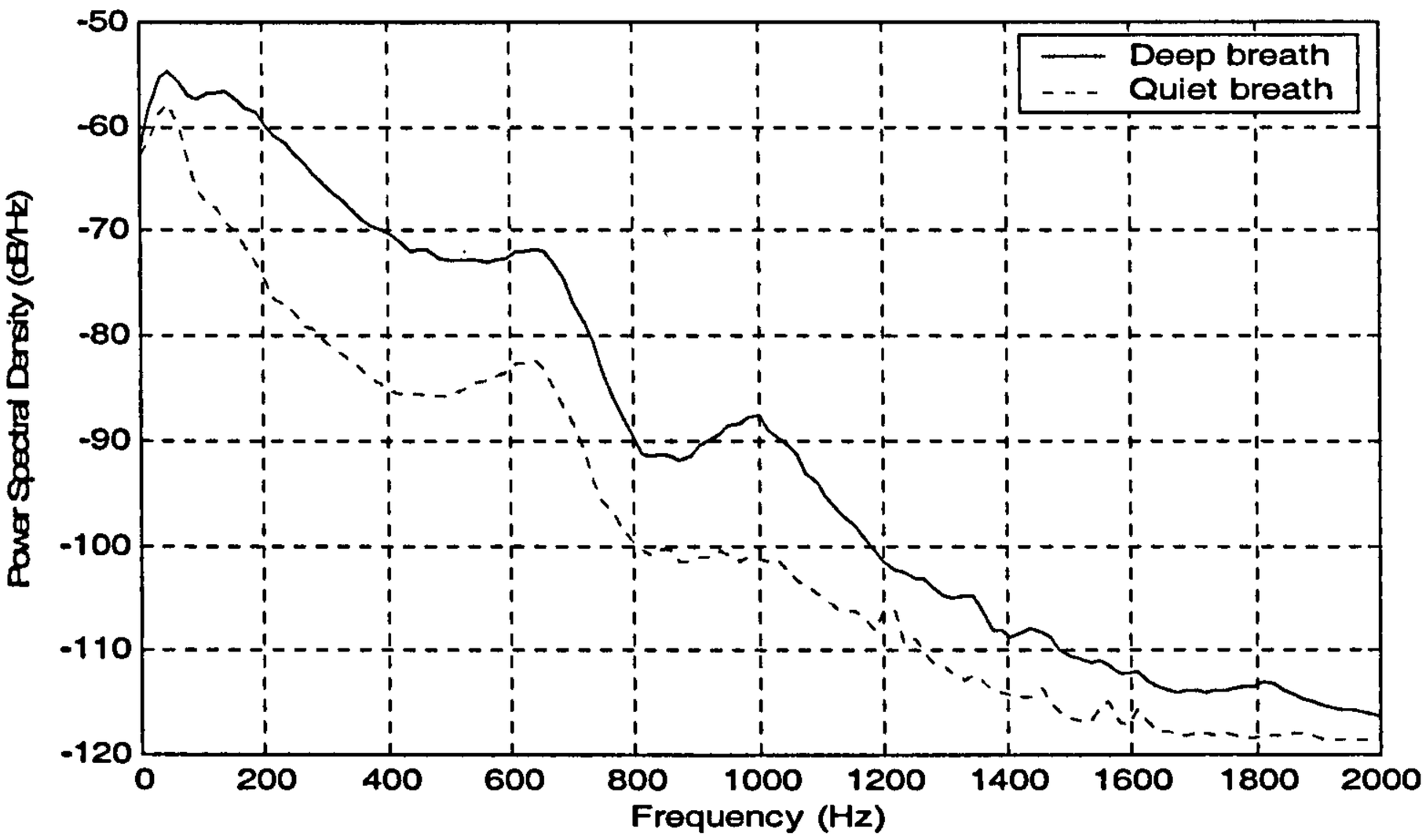

Figure 3.21 PSDs of deep breath and quiet breath.

3.7.2 Pressure

Materials and Methods

Three measurements of deep breath were taken under light pressure (no hand pressure), middle pressure (moderate hand pressure) and heavy pressure (higher hand pressure). Other conditions were the same as described in 3.7.1.

Results

Figure 3.22 shows that the spectrum shapes are similar but there is some change in frequency and level of the peaks. Assuming flow rates were the same during continuous measurements, then the pressures between the Escope and trachea

affected the spectrum shape. The results of the repeatability tests are shown in Figure A.10.

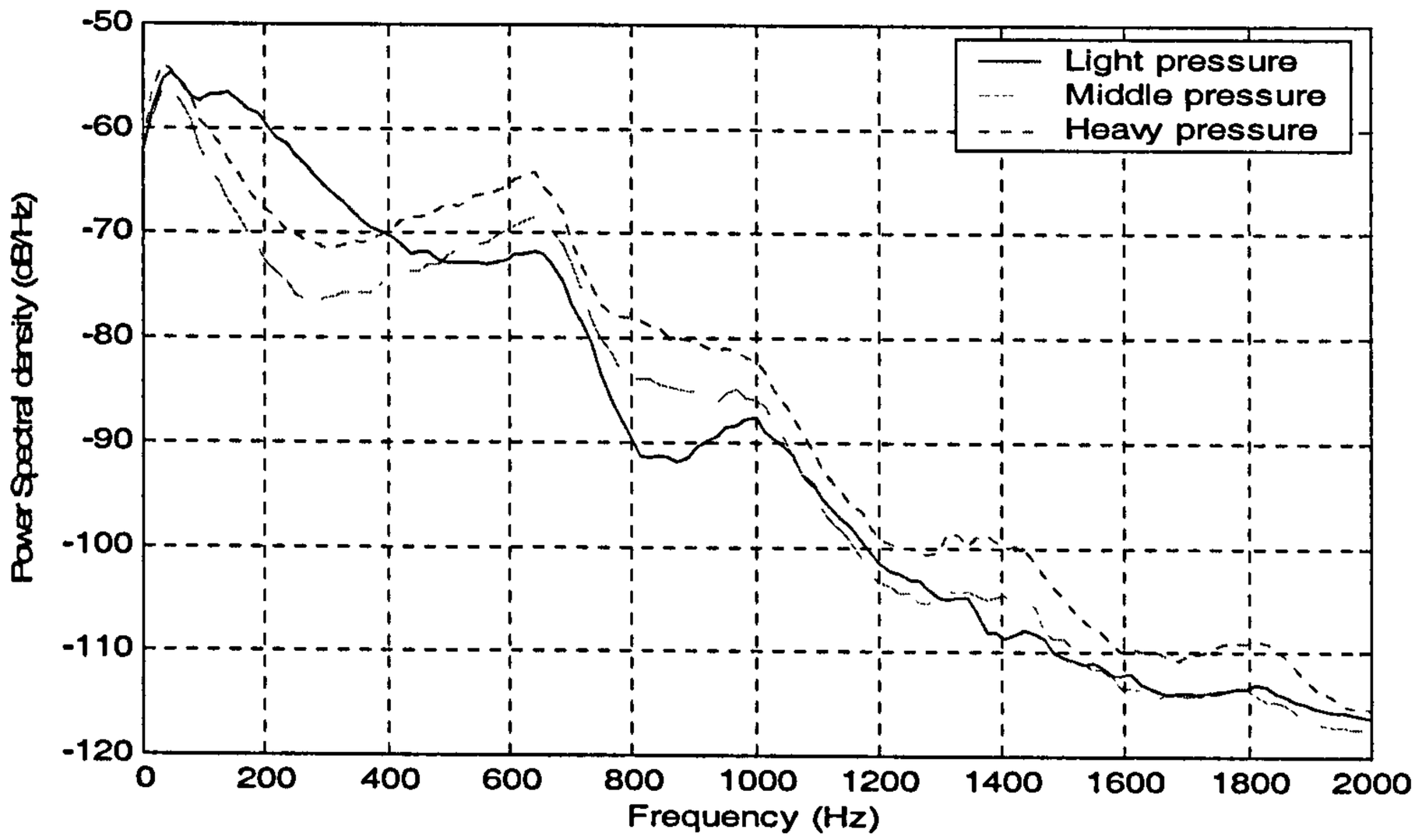

Figure 3.22 PSDs of different pressure at same position during deep breath.

3.7.3 Position

Materials and Methods

The conditions were the same as described in 3.7.1. except 3 more positions were tested. Point 1 was at the anterior cervical triangle; point 2 was above point 1. Point 3 was beside the right artery and parallel with point 1; and point 4 was beside the left artery and parallel with point 1. That is point 3 and 4 were symmetrical (refer to Figure 3.23).

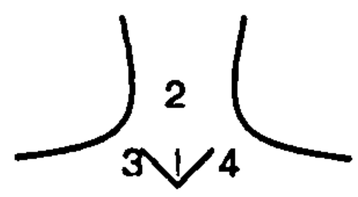

Figure 3.23 Test positions on trachea.

Results

It can be seen from Figure 3.24 that the spectrum shape and level differ a little at the different positions; assuming the flow rates were the same during measurements. The spectral shapes are similar when measurements were undertaken at symmetric positions.

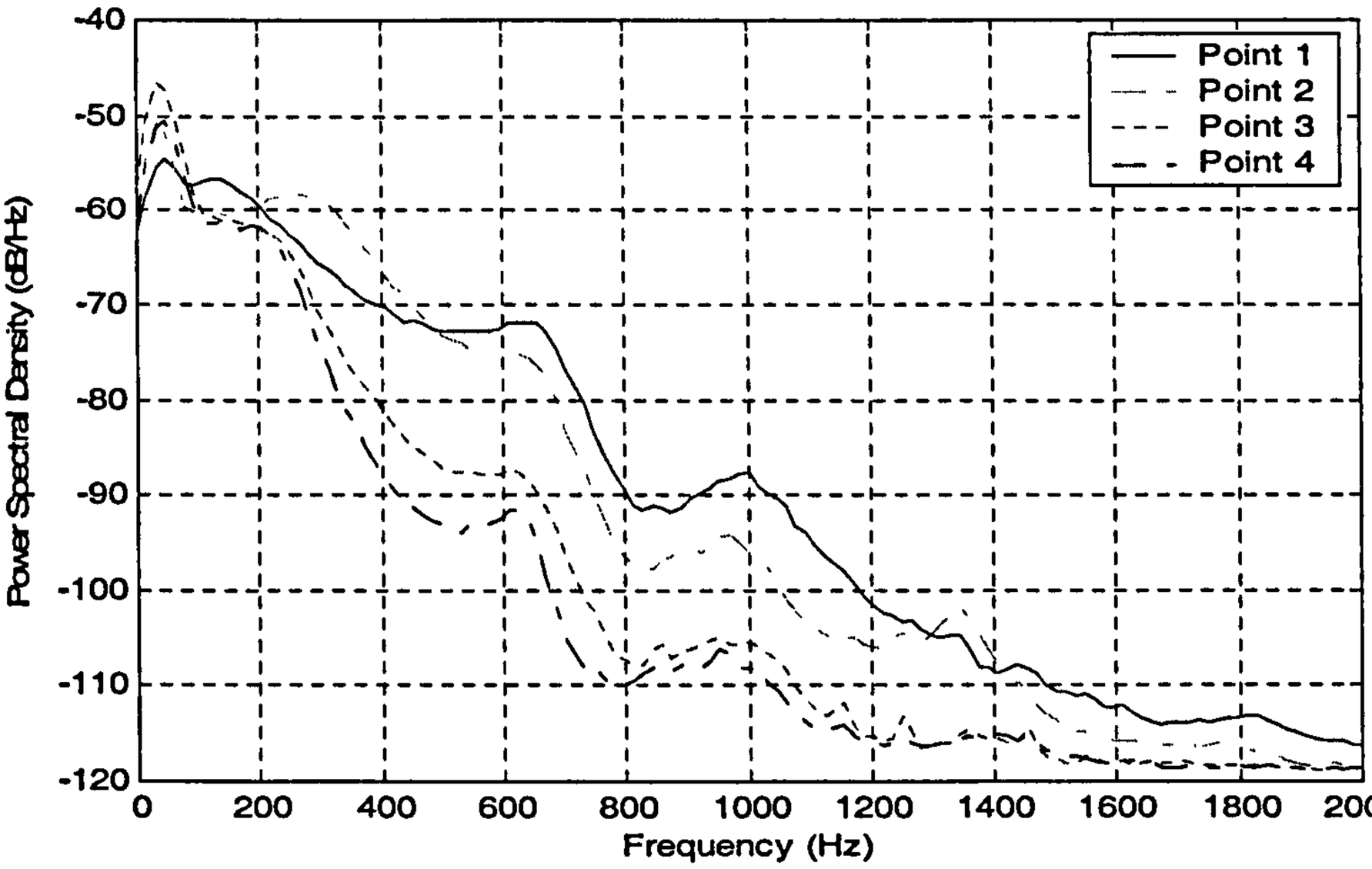

Figure 3.24 PSDs at different positions.

3.8 Landline phones and mobile phones in practical use

Using voicemail as mentioned in 3.6.2, 16 patients were asked to take a two week monitoring of breath sounds and peak expiratory flow rate using their mobile phones (unpublished data, refer to table A.1 in the appendix). When listening to these data, it was found that recordings from some mobile phones were quiet noises, which didn't correspond in any way to normal or abnormal breath sounds.

Due to the algorithms implemented between the sending end and receiving end, breath sounds collected by a mobile phone could be either treated as background acoustic noises or as speech signals. For the former, the received signals could lose their time-variant properties, with the resulting signals not being consistent with breath sounds. For the latter, the received signals are consistent with normally recorded breath sounds. The following experiments were therefore designed to avoid tracheal sounds being treated as noise.

Materials and methods

A purposely-produced noise (high-frequency narrow band noise) was played from the JVC loudspeaker to accompany the tracheal sounds recording using a landline telephone or mobile phone.

Results

Under certain situations, i.e., when the mobile phone or telephone was put on the trachea, the breath sounds reaching the mobile phone were stronger than the sounds from the loudspeaker, and measurements were improved. Figure 3.25 shows the spectrogram of an example. For comparison, an example without purposely-produced noise is shown in Figure 3.26.

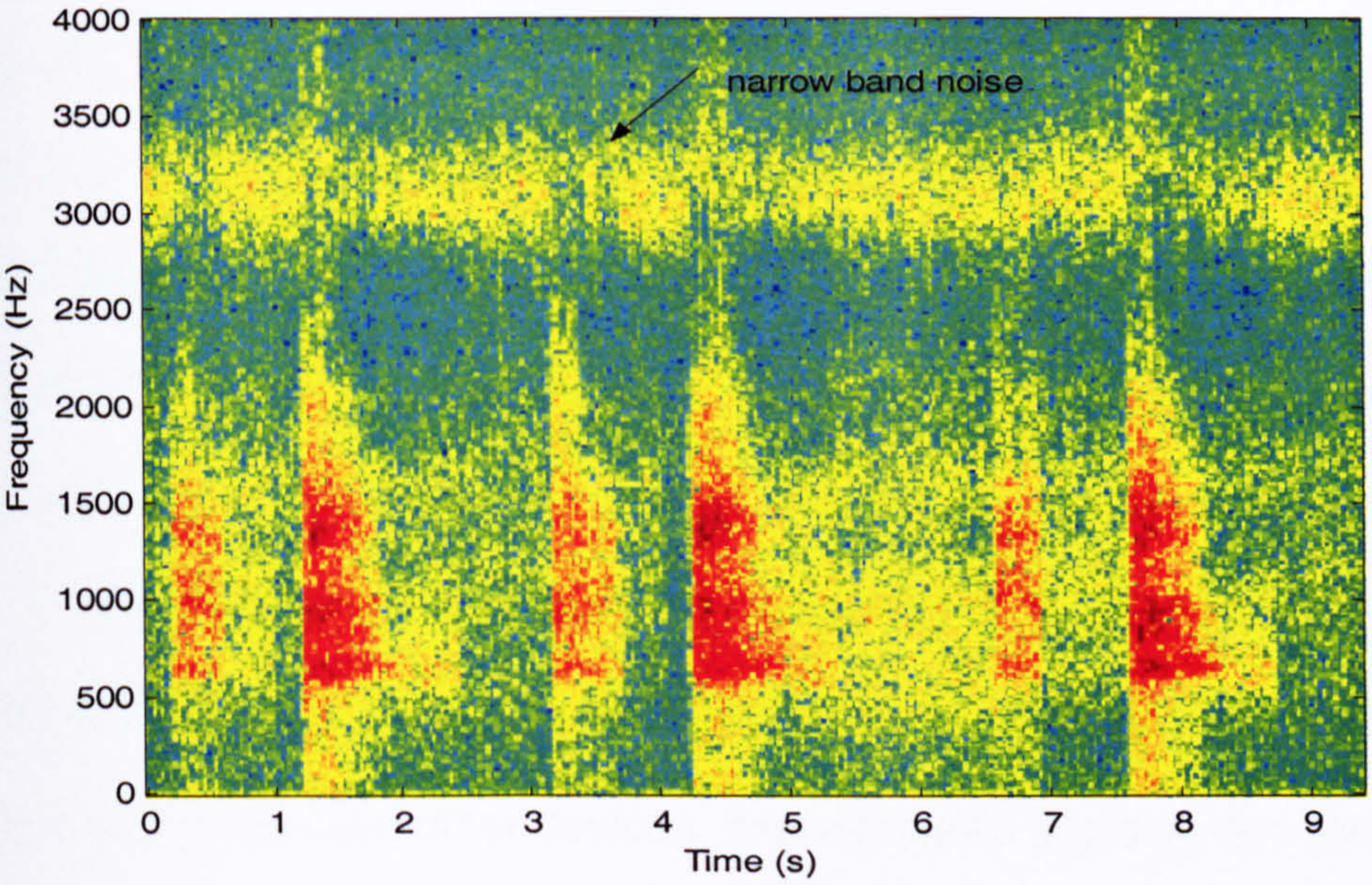

Figure 3.25 Recorded tracheal sounds by Nokia 3310e with a noisy background.

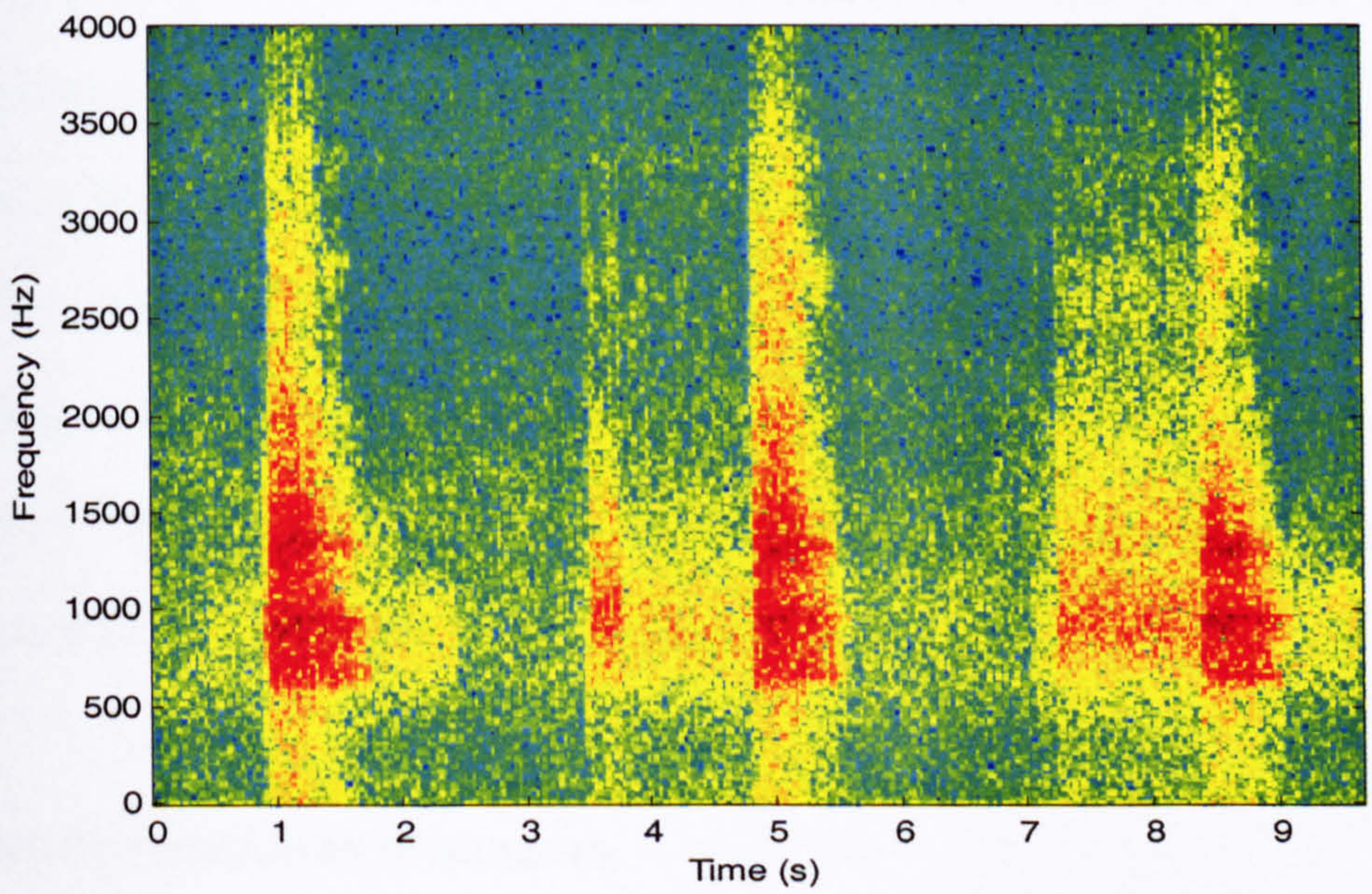

Figure 3.26 Recorded tracheal sounds by Nokia 3310e without a noisy background.

3.9 Discussion

DAQ and filters test

Section 3.3 to 3.5 dealt with the TF estimation of the data acquisition devices and filters. Different input signals were used. For the DAQPad-6200E, manually stepped sinusoidal signals were used, so that the amplitude and frequency of the signal could be controlled. But use of this signal source was time-consuming. First, confidence in the accuracy and bandwidth of the DAQPad is obtained, so that it will not distort the measured signals. It was then used to sample the input and output signals from other devices; the measured signals were treated as 'true' signals.

A square signal was used to test the sound card and filters, as it was more convenient. Both the magnitude response and phase response are obtained by applying Welch's method. This is different from Sun and associates' step-response method (1998). They reported that the method was sensitive to the window position and length. By using equation 3.12, there is no effect from window position, although window length affects the frequency resolution. The results are comparable with available manufacturer's specifications.

Frequency response of sensors

Most of the work was focused on the sound sensors' properties. In the first stage, non-contact transfer functions of sensors were studied. The results were compared with available manufacturer's specifications.

This part of the experimental work was similar to previous studies (Ertel *et al.* 1966a; 1966b; 1969; Druzgalski *et al.* 1980; Abella *et al.* 1992). Ertel and

associates used a sound stage as a sound source to a stethoscope's chest piece. The sound pressure in the sound stage was maintained at a constant level. One microphone was placed beside the chestpiece to measure the input signal; another microphone was positioned at the earpiece with an artificial ear to measure the output signal. Thus, comparing the output and input signals, the stethoscope's acoustic property was obtained. Druzgalski and colleagues applied a similar method to study the air-coupler's characteristics. They used a loudspeaker to produce sound. A laboratory microphone was used as a reference (input), and the signals from a tested microphone with and without an air-coupler were recorded (output). Abella and coworkers used two identical microphones. One was placed beside the stethoscope chestpiece in front of a white noise sound source from a sound coupler (input), and the other was placed at the earpiece with an artificial ear (output). Thus they obtained the acoustic transfer function of the tested stethoscope.

Druzgalski's method was also used to study the Escope and Hp Stethos frequency response. The input sound pressure at the chestpiece was measured by a microphone (CANFORD MB-C550). The result was very similar to that shown in Figure 3.9. As the microphone was then used with a Littmann chestpiece, so the method described in 3.6.2 was used to maintain consistency.

A white noise sound source is preferred. However, if a white noise generator is used, a sound stage or a sound coupler should be properly designed to make the sound source a white noise (Ertel *et al.* 1966a; 1966b; 1969; Abella *et al.* 1992). As the contact usage of the electronic stethoscope was under investigation, a

contact surface was used. When a surface was attached to the white noise sound source, the sound source was no longer a white noise because the surface affected the sound source.

A loudspeaker was employed to produce a sound field in the experiments. Although programmed white noise signals were played via the loudspeaker, the output sound pressures were not at a constant level because of the loudspeaker's frequency response. The loudspeaker and the sensor were regarded as a cascaded linear time-invariant system. First, the loudspeaker's frequency response was obtained. Then the compound TF of the loudspeaker and the sensor was estimated. Next the TF of the loudspeaker was deduced to obtain the sensor's TF.

A validation procedure was proposed to verify the usefulness of the TF estimation results. The outcomes of the validation show that the designed system is able to estimate the sensors' non-contact frequency response.

In the second stage, the contact response of the sensors was studied using the same testing system, but contacting a surface. Results show that the power spectra of the sensors are different from each other when measuring the same signal (Figure 3.16). It is not surprising that different results are obtained from different sensors. Druzgalski *et al.* (1980) reported the spectrum differences of measured data when they put different sensors on a membrane, which was on a speaker. Significant spectrum differences were reported by Gavriely (1984) when he recorded signals over the trachea; spectral diversities were also reported by

Pasterkamp *et al.* (1993) when they sampled signals from a human thorax using different sensors.

Although some investigators were aware of the contact effects, these effects have not been measured. When Ertel and associates studied the diaphragm type chestpieces, they deliberately avoided contact between the diaphragm and sound stage (Ertel *et al.* 1966b; 1969). Wodica and colleagues assumed that the properties of the diaphragm and its interaction with both the chest wall and air cavity dominated the overall frequency response (Wodicka *et al.* 1994). Vovk's acoustic model of the chest wall predicted that the presence of a sensor would alter the vibration velocity at the point of contact. They suggested that a transducer should be small in area and light in weight to minimise the distortion (Vovk *et al.* 1994; Vovk *et al.* 1995).

The effects of contact were therefore further studied. Experiments show that under mass loading, the vibration of the contact surface changes. This is consistent with the modelling prediction in section 3.6.4. Although the model is very simplified, it provides a qualitative prediction of the mass effect. A 2-DOF model, which also involves the stethoscope chestpiece elastic and damping properties, is probably more suitable. A quantitative description is possible only if all the model's parameters are known.

The electronic stethoscope's contact frequency response changes compared to that of its non-contact characteristics. This is also predicted in section 3.6.5. When in the free field, no extra pressure acts on the diaphragm. When in contact, an extra pressure is applied on the diaphragm. Thus, the initial air volumes of the

chestpiece are different. This is likely to explain the differences between the Escope's non-contact TF and contact TF. While under the same pressure, the estimated TFs were almost the same (Figure 3.19). Similarly, when different pressures are applied on the surface, the frequency response will also be different. This is supported by the results obtained using different pressures (3.7.2).

The diaphragm's property may also change under pressure. Howell and Aldridge (1965) speculated that increasing strain within the diaphragm would suppressed low frequency sounds. This effect was not studied in the current experiments.

Thus when the mutual effects of contact are removed, the 'true' contact TF of a stethoscope could be obtained. The comparison between a LDV spectrum and calibrated Escope spectrum (Figure 3.20) shows the similarity, especially in the frequency range of 100-500Hz. However, the difference may be explained by the mass of the Escope changing the sound pressure beneath it.

The above conclusion is obtained from limited experiments. On the one hand, for the Escope materials having similar properties to human skin and flesh would provide more persuasive results. On the other hand, experiments described in 3.6.8 could not be applied to those air-coupler sensors that were not easily disassembled.

Using the LDV, which is capable of non-contact measurement, made direct experimental study of the mass loading effects possible. The LDV has been reported as being used to study the sound-induced motion of the body walls of

some kinds of amphibian (Hetherington 2001). Experiments show that the LDV can also be used to measure breath sounds. When the LDV signals were played back, the breath sounds could be heard clearly. Although the time domain waveforms didn't look like normally recorded breath sounds, the digital filtered counterparts did have the similar patterns. However, there are some limitations in using the LDV. It needs patients to restrict their body movement during breathing, which needs quite a few practices, and the LDV is expensive. For these reasons, the LDV is not likely to be a common choice of transducer for breath sounds measurements. Nevertheless, as the LDV is capable of non-contact measurement, it could be used to monitor a patient's cardiorespiratory system in special situations, such as during anaesthesia.

Even in-vivo study of the contact effects between the human body and the sensor could be carried out as described in 3.6.8 with some modifications. Band-limited white noise could be introduced into the mouth of human subjects (Kraman and Austrheim 1983; Wodicka and Shannon 1990). However, the difficulty is that the subjects should keep the posture still for quite a long time for the described procedure in 3.6.8 to be completed. The results shown in Figure 3.18 seem to imply that even if using the same sensor on different subjects the effects of contact may be different. This would increase the difficulties of making measurements comparable.

Repeatability

Factors that could influence repeatability are also studied. Flow rates affect the intensity and spectral shape (Figure 3.21), which is consistent with previously

reported results (Charbonneau *et al.* 1983; Lessard and Wong 1986; Soufflet *et al.* 1990; Gavriely and Cugell 1996). Pressure (Figure 3.22) and position (Figure 3.24) also play a role in affecting measurements; the latter factor has been briefly reported before (Gavriely 1984). Under the same conditions, the measurements were repeatable (Figure A.9, A.10). Thus keeping measurement conditions as constant as possible is a prerequisite to obtaining repeatable results.

Mobile phone usage

In the previous study (Anderson *et al.* 2001), the mobile phones showed the potential to be an easily available tool to capture breath sounds on the trachea. Tracheal sounds contain richer frequency components and are louder than lung sounds and they may hold useful information about asthmatics.

The signal pathway from a mobile phone to the received end point could be viewed as follows. The acoustic signals at the sent end were coded as background frames and speech frames, which contained frequency and level information. Whether it would be a background frame or a speech frame was decided by a Voice Activity Detection algorithm (ETSI 2000c). A speech signal is coded and updated continuously. The encoding was an adaptive speech compressive algorithm (ETSI 2000a). At the received end the signals were synthesised according to received speech frames with added comfort noise (ETSI 2000b) according to background noise frames.

So when a mobile phone is used to collect tracheal sounds, the signals could be either treated as background acoustic noises or speech signals. For the former, the received signals could lose the time-variant property, which makes the signals

sound different from breath sounds. For the latter, the received signals sound like those using normal recording techniques. (Wheezes are more likely than normal sounds to be treated as speech signals).

Under certain conditions, i.e. when the mobile phone or telephone was put on the trachea, the breath sounds reaching the mobile phone were stronger than the other purposely-added stationary background noise, and measurements were improved. If the purposely-added background noise can be easily separated from the breath sound signal, then the true breath sounds can be acquired. In this study, the narrow band high frequency noise is above the tracheal sound frequency range. Thus low pass filtering the received signals will recover the tracheal sounds. However, insufficient experiments have been carried out to check whether this method would work on all those mobile phones that treat normal breath sounds as noises. If it is a general solution, then it is the simplest way to improve measurement quality without many extra devices. Accordingly, it may be possible to make and provide to patients a small portable device that can play the purposely-designed background noise in the future. Patients will then be able to play the noise to accompany the recording of the tracheal breath sounds. At the processing side, incorporating a pre-processing method by digital filtering can extract the measured breath sounds.

Alternatively, the scheme of Woodward *et al.* (2001) can be considered, that is, using a mobile phone's data transmission capability. However, this method requires acoustic sensors to capture the breath sounds. Then the signals should

be pre-processed by functions, such as store, compress, frame etc, before they are sent to a data transmission capable mobile phone via an infrared port.

It is worth pointing out that when doing self or remote monitoring of breath sounds necessary training should be provided to patients, so that measured signals are carried out under similar conditions for which the influential factors are minimised. Should there be any variations between the recordings they will be explained as manifesting the variations of the pulmonary system.

In summary, each part of the available measurement system has been studied. The non-contact frequency responses of electronic stethoscopes and a mobile phone have been obtained, using easily set-up devices. The estimated TFs have been verified, and demonstrate that the procedure of estimating TFs is correct. The contact frequency response of the Escope has been estimated when mutual contact effects have been isolated. This contact frequency response has also been validated. Under the same measurement conditions the measured data were repeatable. Mobile phones could be a convenient tool to monitor breath sounds if sound quality could be guaranteed.

Chapter 4 Heart Sounds Reduction

Heart sounds are an inherent interference in lung sound analysis. They may mask lung sound in auscultation and alter the power spectral density (PSD) distribution in the spectrum. Normally, there are two methods used to minimise heart sounds interference. One is to choose measurement positions where the heart sounds are faint. The other is filtering, including conventional high pass filtering (Shykoff *et al.* 1988; Gavriely *et al.* 1995; Homs-Corbera *et al.* 2000) and adaptive filtering (Iyer *et al.* 1986; Kompis and Russi 1992; Charleston and Azimi-Sadjadi 1996; 1997; Hadjileontiadis and Panas 1997; 1998). Adaptive filtering is better because it retains the overlapped low frequency part of the lung sounds. In this chapter, an attempt to reduce heart sounds by using Hadjileontiadis' (1998) method is described.

4.1 Wavelet Analysis

4.1.1 Continuous Wavelet Transform (CWT)

For any square-integrable function $f(t)$, the continuous wavelet transform is defined as (Vidakovic 1999):

$$C(a,b) = \int f(t)\, \psi^{*}_{a,b}(t)\, dt \tag{4.1}$$

where $*$ denotes the complex conjugate, and $\psi_{a,b}(t)$ is a family of wavelets, which can be derived from the mother wavelet $\psi(t)$:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \tag{4.2}$$

a, b are scaling and translation parameters respectively, and are continuously variable. $1/\sqrt{a}$ is for energy normalisation across the different scales.

4.1.2 Discrete Wavelet Transform (DWT)

The CWT maps a one-dimensional time signal to a two-dimensional time-scale domain that is highly redundant. To improve the efficiency, scaling and translation parameters in the wavelet can be discretised:

$$\psi_{mn}(t) = \frac{1}{\sqrt{a_0^{\,m}}}\psi\left(\frac{t - nb_0a_0^m}{a_0^m}\right) \tag{4.3}$$

where m, n are integers, and $a_0 > 1$, $b_0 \neq 0$.

Usually $a_0 = 2$, $b_0 = 1$ are chosen, which is called dyadic sampling. It is possible to construct $\psi(t)$ such that the set of scaled and translated versions of $\psi_{mn}(t)$ forms an orthonormal basis so that no redundancy exists (Vetterli and Herley 1992). That is:

$$\int_{-\infty}^{\infty}\psi_{mn}(t)\psi_{jk}^{*}(t) = \delta_{mj}\delta_{nk} \tag{4.4}$$

By discretisation of time, the discrete wavelet transform is:

$$D_{mn} = 2^{\frac{m}{2}}\sum_{k} f(k)\psi\left(2^{-m}k - n\right) \tag{4.5}$$

For signal processing, a wavelet is a band-pass filter, and a set of scaled and translated wavelets is a band-pass filter bank. By introducing a low-pass filter, scaling function, whose spectrum covers the left space of wavelets, the number of wavelets used to analysis a signal is limited (Vetterli and Herley 1992).

4.1.3 Wavelet Decomposition and Reconstruction

The DWT can be achieved effectively by using a filtering algorithm. By using half-band high pass and low pass decomposition filters and downsampling, the DWT coefficients of a signal can be obtained. Reversibly, by using associated reconstruction high pass and low pass filters and upsampling, a signal can be reconstructed perfectly from the DWT coefficients.

4.1.4 Wavelet Thresholding

For selective wavelet reconstruction, a threshold λ should be defined to distinguish large and small coefficients. There are two kinds of thresholding; hard thresholding and soft thresholding, which are defined in equations (4.6), (4.7) respectively (Ogden 1997).

$$\delta^h(x) = \begin{cases} x, & |x| > \lambda \\ 0, & |x| \le \lambda \end{cases} \tag{4.6}$$

$$\delta^s(x) = \begin{cases} x-\lambda, & x > \lambda \\ 0, & |x| \le \lambda \\ x+\lambda, & x < -\lambda \end{cases} \tag{4.7}$$

Coefficients larger than λ are kept (hard thresholding) or shrunk (soft thresholding), but smaller ones are set to 0. Via this non-linear operator on the wavelet coefficients, selective (denoised) wavelet reconstruction can be achieved. λ is application dependent. Choosing a very large one will result in an oversmoothing; while choosing a very small one will result in undersmoothing (Ogden 1997).

4.2 Materials and Methods

Lung sounds were recorded on a sitting healthy subject's anterior left and right chest simultaneously by using two Escopes, sampling at 11025Hz using the TBS-2000 sound card. The idea of the Hadjileontiadis' method was that the mixture of lung sounds and heart sounds was considered as a non-stationary signal (heart sounds) with a stationary noise (lung sounds). Daubechies wavelet of order 4 was used for the calculation, since this wavelet family has been shown to be appropriate in early work (Hadjileontiadis and Panas 1998; Hall et al. 2000). Applying wavelet decomposition and hard thresholding, large coefficients were used to reconstruct hearts sounds, while small coefficients were used to reconstruct breath sounds. This procedure was repeated several times until the difference between updated lung sounds and previous lung sounds was very small.

4.3 Results

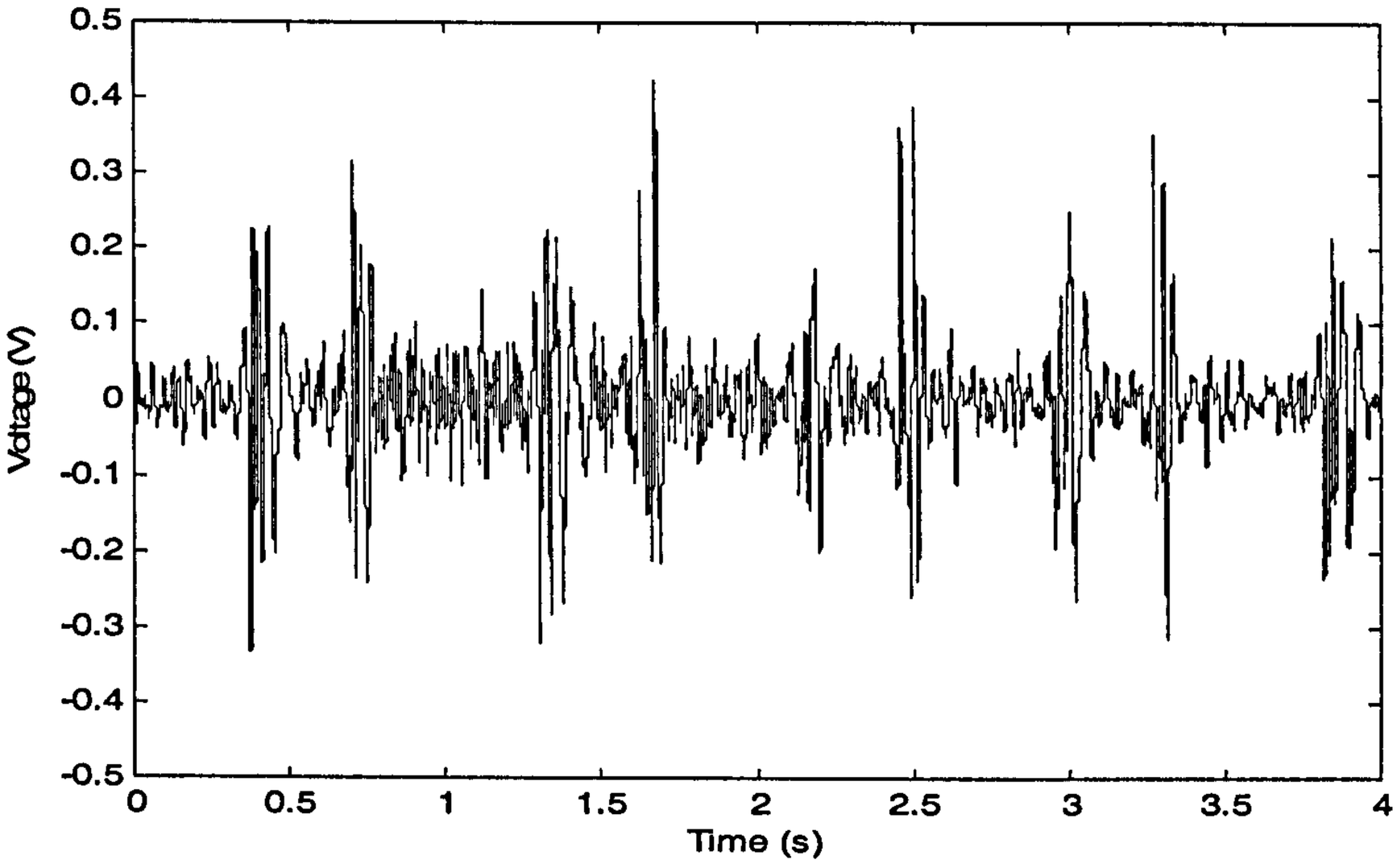

Figure 4.1 Measured signal at the left chest.

Two results are shown in Figures 4.1 to 4.6. Figure 4.1 is a measured signal at the left of the chest. Figure 4.2 and 4.3 are reconstructed heart sounds and lung sounds. Figure 4.4 is a measured signal at the right of the chest. Figure 4.5 and 4.6 are reconstructed heart sounds and lung sounds respectively.

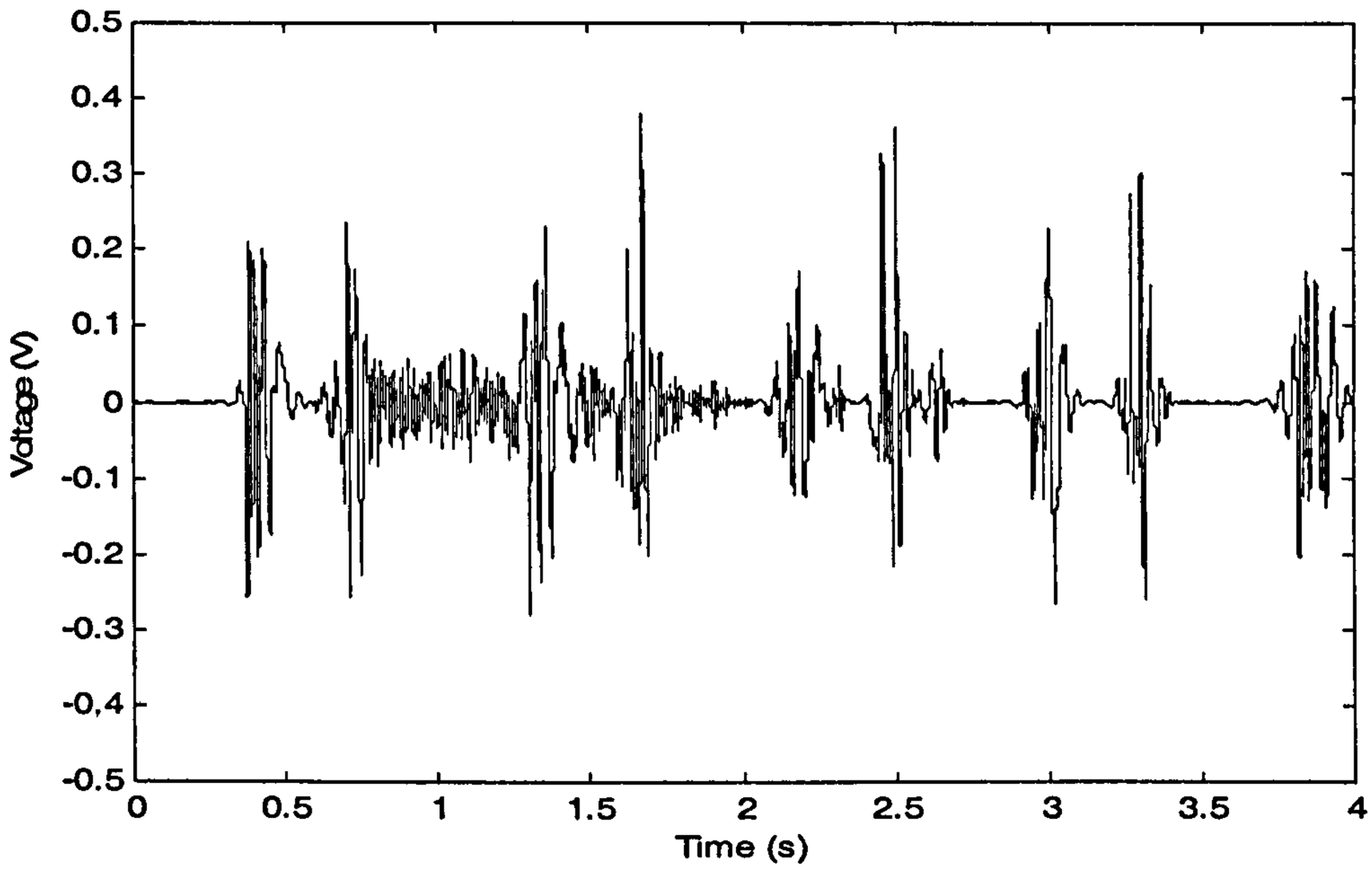

Figure 4.2 Reconstructed heart sounds.

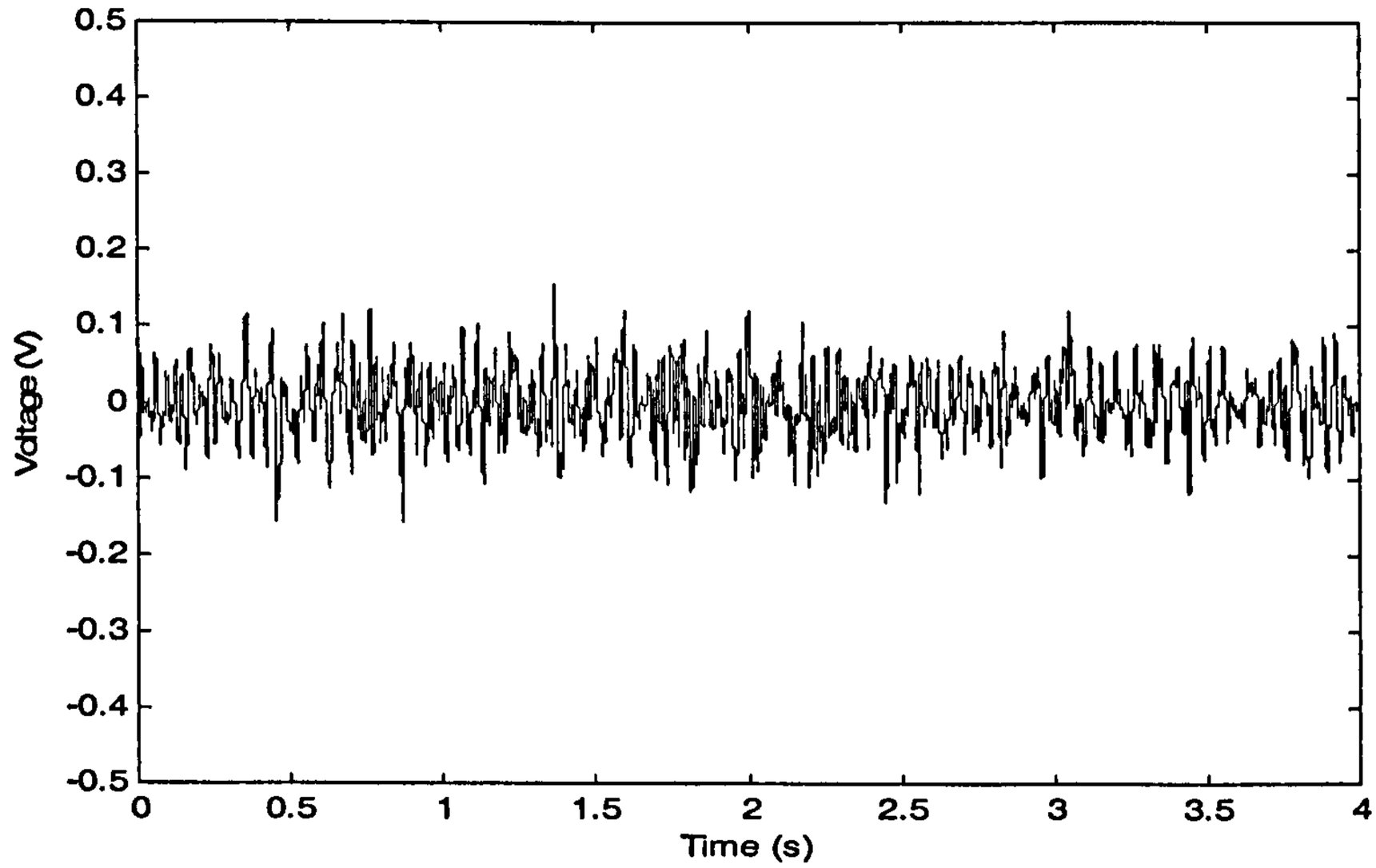

Figure 4.3 Reconstructed lung sounds.

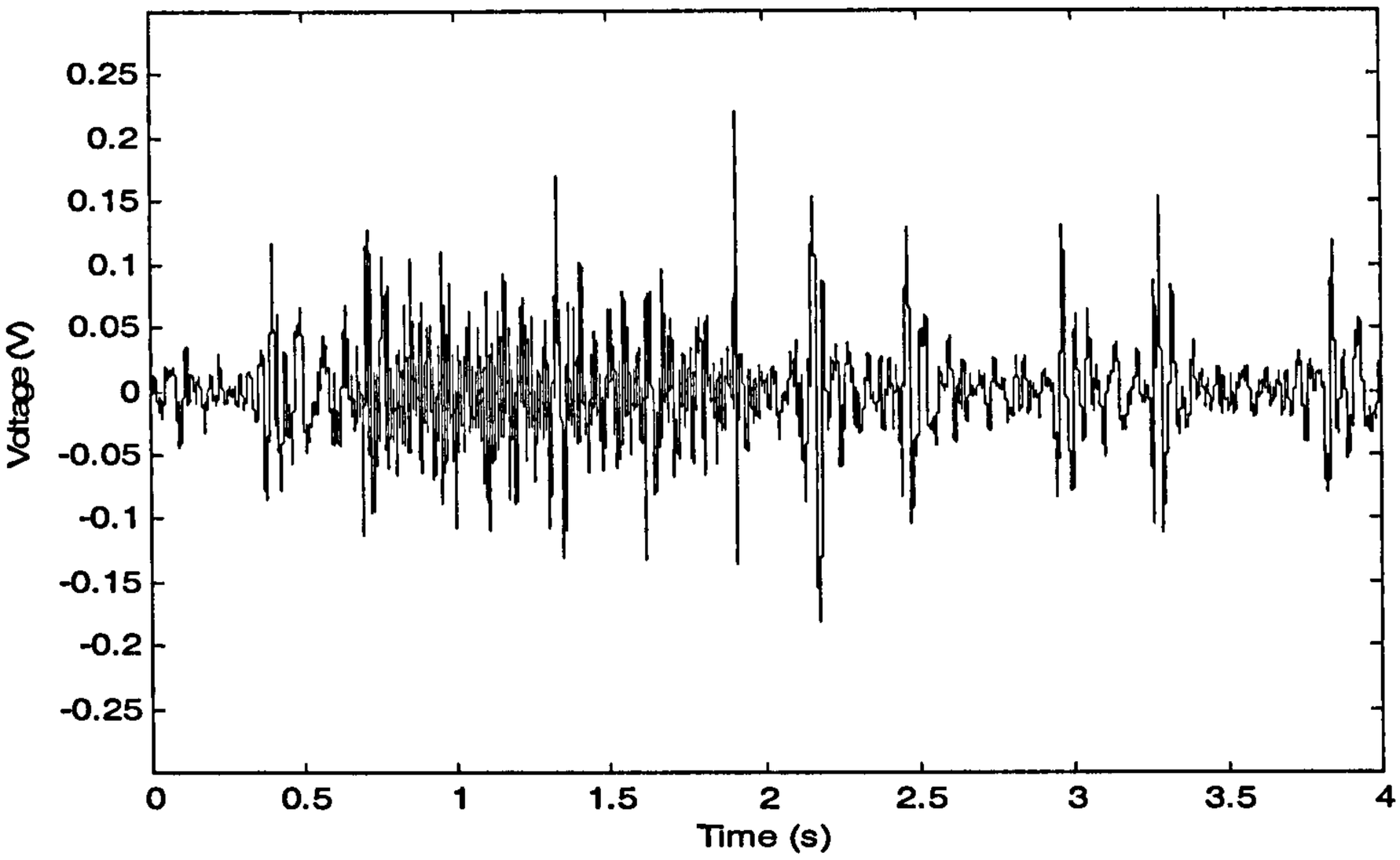

Figure 4.4 Measured signal at the right chest.

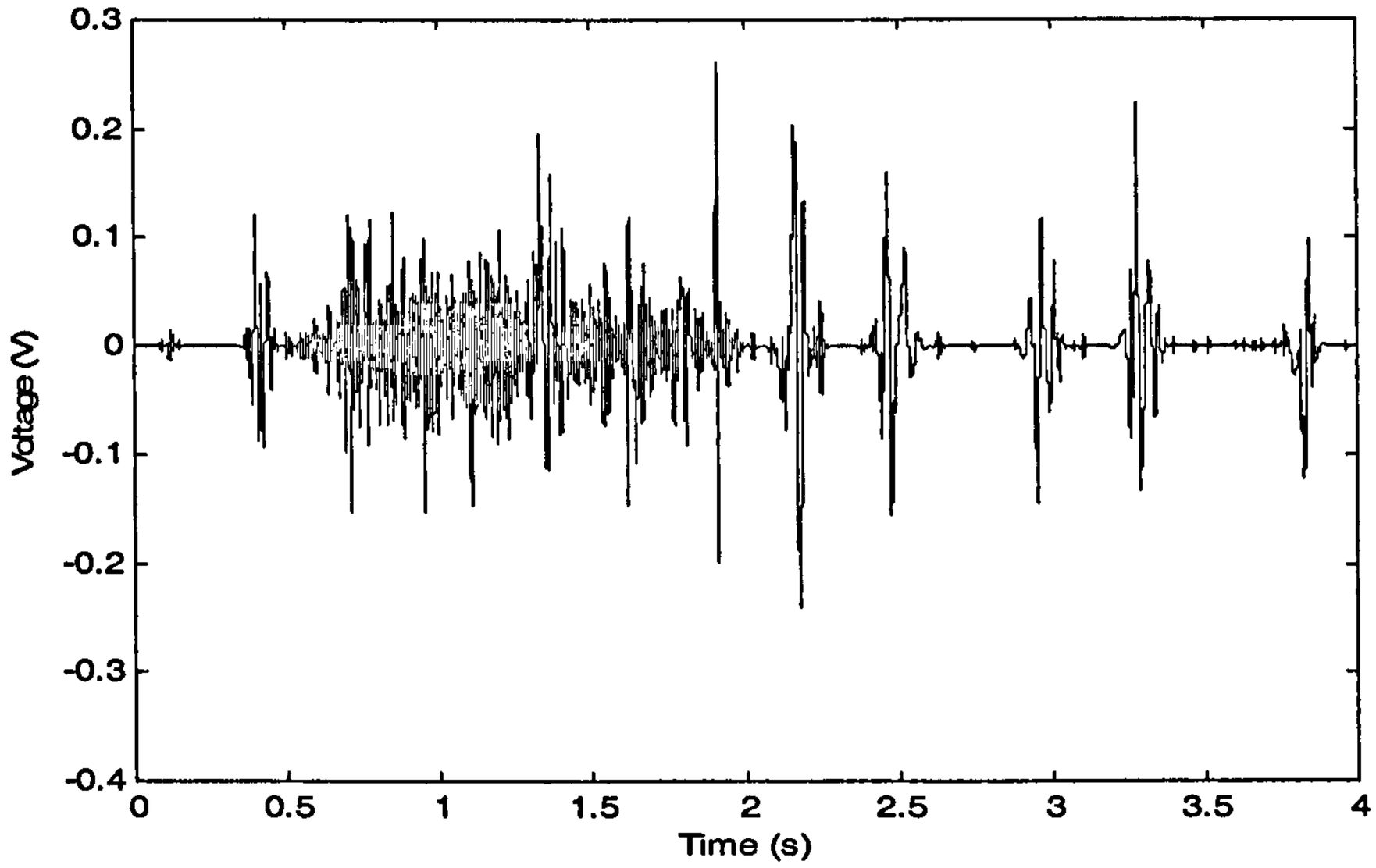

Figure 4.5 Rreconstructed heart sounds.

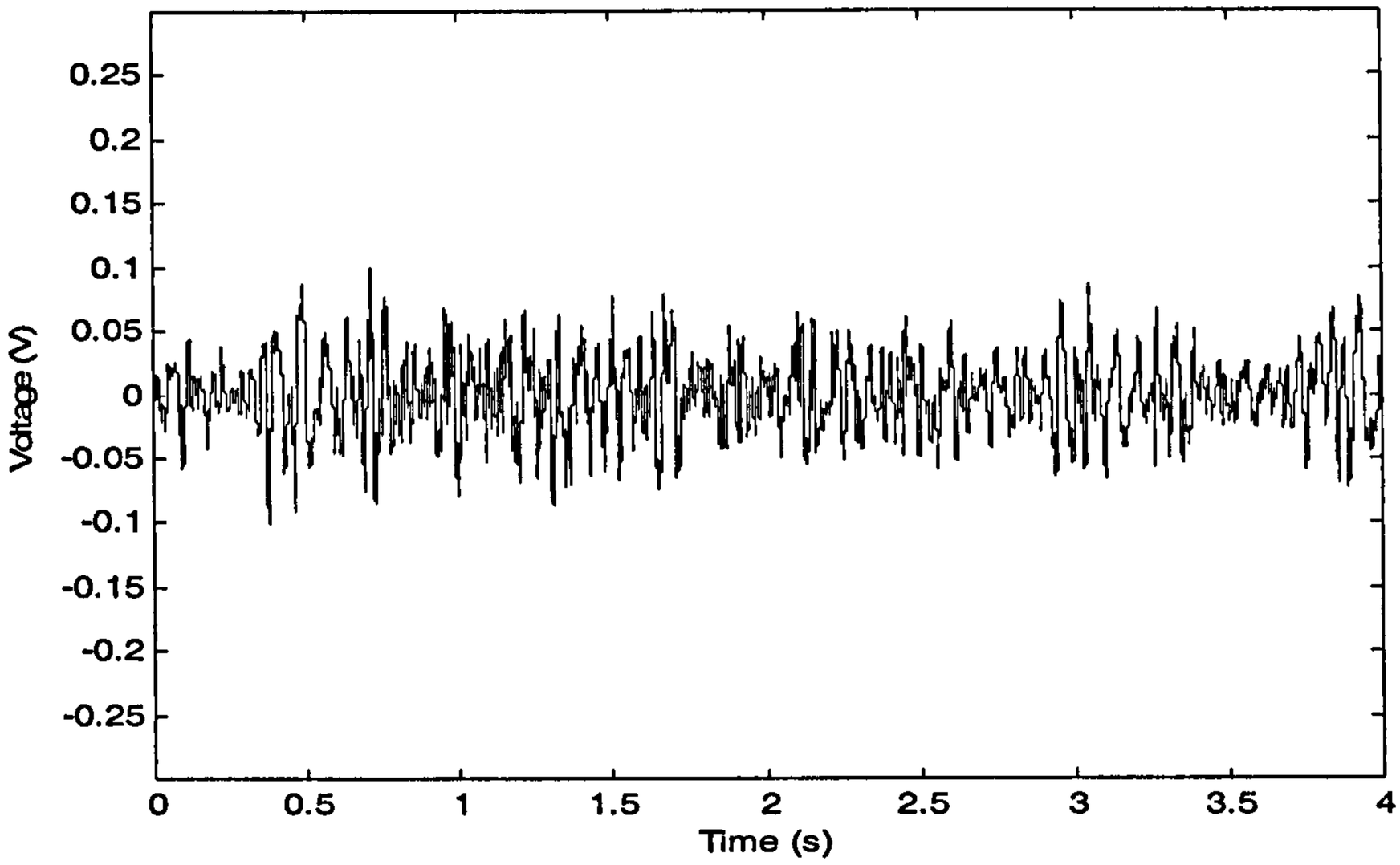

Figure 4.6 Reconstructed lung sounds.

It can be seen from Figures 4.3 and 4.6 that for measured locations corresponding to strong heart sounds, the heart sounds peaks are reduced.

4.4 Discussion

This wavelet based decomposition and selective synthesis method was reported to have good heart sounds reduction when the signals to be processed were recorded at positions where heart sounds were most intensive (Hadjileontiadis and Panas 1998). This method was chosen to be investigated because it does not require a reference signal. In practice the recording positions are not where the heart sounds are most intensive. For this situation, Figures 4.2 and 4.5 show that some content of lung sounds are contained in heart sounds.

Playing back the extracted sounds, the heart sounds sound almost natural. But the lung sounds do not, because part of their content is lost. Actually, the more intensive the lung sounds in the measured signals, the worse the processed results (refer to Figures 4.5 and 4.6), which is expected because of the thresholding.

Heart sounds have low frequency components with high intensity. Their reduction is necessary when measured indices are based purely on the breath sound spectrum. An example of such an index is F_{50}[1], where the low frequency heart sound components would have a significant influence on its value.

If time-frequency (or time-scale) representation is chosen, then heart sound reduction is not necessary, as heart sounds have a fixed pattern in the time-frequency plane. For breath sounds where the abnormalities do not appear in regions of the spectrogram occupied by heart sounds, the abnormalities can be detected by visual examination of the spectrogram or by extracting quantitative data. Also when breath sounds abnormalities do occupy the regions of the heart sounds, they still can be detected visually or quantitatively, but will be more difficult.

So if spectrum based indices are chosen, the measurement sites should be selected carefully to avoid strong heart sounds. Time-frequency analysis is the best choice wherever possible.

---

[1] F_{50} is defined as the frequency below which 50% of the energy of the signal lies.

Chapter 5 Breath Sounds Simulation

Although real breath sounds data can be obtained experimentally or commercially, simulation can produce characteristic predefined breath sounds, which can be used to evaluate algorithms for automatic computer-aided breath sounds analysis. Simulation work is described in this chapter.

5.1 Autoregressive (AR) Process

A time series is a set of data obtained sequentially in time. Assuming the present value of a stationary time series depends on the immediate past values together with a random error, this time series is an autoregressive (AR) process, expressed as (Chatfield 1989):

$$x_n = \sum_{i=1}^{p} a_i x_{n-i} + e_n \qquad (5.1)$$

where x_n is the current value, p is the order, a_i are coefficients, x_{n-i} are previous values, and e_n is the error.

From a signal processing point of view, the signal x modelled by AR of order p can be realised by filtering white noise e through a pth order all-poll infinite impulse response (IIR) filter.

5.2 Methods

The Kompis' method (1997) with modification was used for simulation. The idea was that a normal breath sound could be modelled as an AR process. Heart sounds, continuous (wheezing) and discontinuous (crackle) sounds could be

added to the normal sounds to simulate lung or tracheal normal and abnormal breath sounds.

5.2.1 Normal Breath Sounds

AR modelling should be applied on stationary processes, but a real normal breath sound is not stationary as its variance changes with flow rate. Using Kompis' method of segmentation of the breath sounds, that is, assuming within each segment the signal is stationary, then the AR modelling could apply to that segment. Eight segments in each inspiration and expiration were made with equal length. Each segment was modelled as a 4^{th} order AR process, i.e., 8 sets of IIR filter coefficients and white noise variance for each inspiration and expiration.

Transitions between successive filter coefficient sets caused abrupt changes in data, which contaminated the signal with a 'clicking' sound. This effect was smoothed in a way different from Kompis' method. Using the final conditions from preceding filter data as the initial conditions to the successive filter assured data continuity.

Simulation for lung sounds and tracheal sounds uses different coefficient sets, as these two sounds have different characteristics.

5.2.2 Wheezes

In the time domain, wheezes have periodical waveforms. In the time-frequency domain, wheezes look like a nearly straight line or a sinusoidal curve. Thus wheezes can be simulated by frequency modulation. A very low frequency (say 1Hz) sinusoidal signal x with small random amplitude variations was first

produced. Then the signal x was frequency modulated with a carrier frequency (mean wheezing frequency) f_w, thus simulation of wheezes could be achieved.

$$x(t) = A_t \cos(2\pi t) \tag{5.2}$$

$$y(t) = A\cos\left\{2\pi\left[f_w + x(t)\right]t\right\} \tag{5.3}$$

where A_t is a small random amplitude variation, A is amplitude, f_w is carrier frequency, and t is the time point.

5.2.3 Crackles

Kiyokawa's method (2001) for simulation of crackles was employed. The crackle signal $y(t)$ could be simulated as:

$$y_0(t) = \sin\left(4\pi t^{\frac{\log(0.25)}{\log(t_0)}}\right) \tag{5.4}$$

$$A(t) = 0.5\left\{1 + \cos\left[2\pi\left(t^{0.5} - 0.5\right)\right]\right\} \tag{5.5}$$

$$y(t) = A(t)\, y_0(t) \tag{5.6}$$

where $0 \le t \le 1$, t_0 is first positive t-intercept.

Tuning t_0 can achieve fine, middle and coarse crackles. Randomised crackle events made each cycle sound similar but not exactly the same.

5.2.4 Heart Sounds

Tran's (1995) method for simulation of heart sounds was used. Only normal heart sounds were simulated in this work. The heart sound signal $y(t)$ could be produced by:

$$y_0(t) = \sin\left(2\pi\left(f_0 + bt\right)t\right) \tag{5.7}$$

$$e_a = \left[\left(1 - e^{\frac{-5t}{T}}\right)\sin\left(\frac{0.5\pi t}{T}\right)\right]^{\left(1-\frac{a}{11}\right)} \tag{5.8}$$

$$e_d = \left[e^{\frac{-2.5t}{T}}\cos\left(\frac{0.5\pi t}{T}\right)\right]^{\left(1-\frac{d}{11}\right)} \tag{5.9}$$

$$y(t) = e_a e_d y_0(t) \tag{5.10}$$

where $y_0(t)$ is a chirp signal, whose instantaneous frequency changes with time from f_0 with a slope of b. T is heart sounds duration, a and d are attack and decay constants respectively.

The first and second heart sounds are produced using the same equations but with different parameters.

5.3 Results

The simulation has been implemented in Matlab. A graphical user interface (Figure 5.1) has been developed to facilitate the parameter changes. Changeable parameters are:

1. breath cycle parameters-breath rate (how many seconds per breath cycle), what percentage is occupied by inspiration (the rest occupied by expiration), how many cycles to be simulated;
2. heart sounds parameters (selective)-heart rate (how many beats per minutes) and loudness;
3. site-lung or trachea;
4. type-normal, wheeze or crackle;

5. abnormal sounds parameters-in which phase (inspiration/expiration), at which part (early/middle/late), frequency and loudness. (These parameters are changeable only when 'wheeze' or 'crackle' is chosen.)
6. operation-construct, play and save simulation data as .wav files.

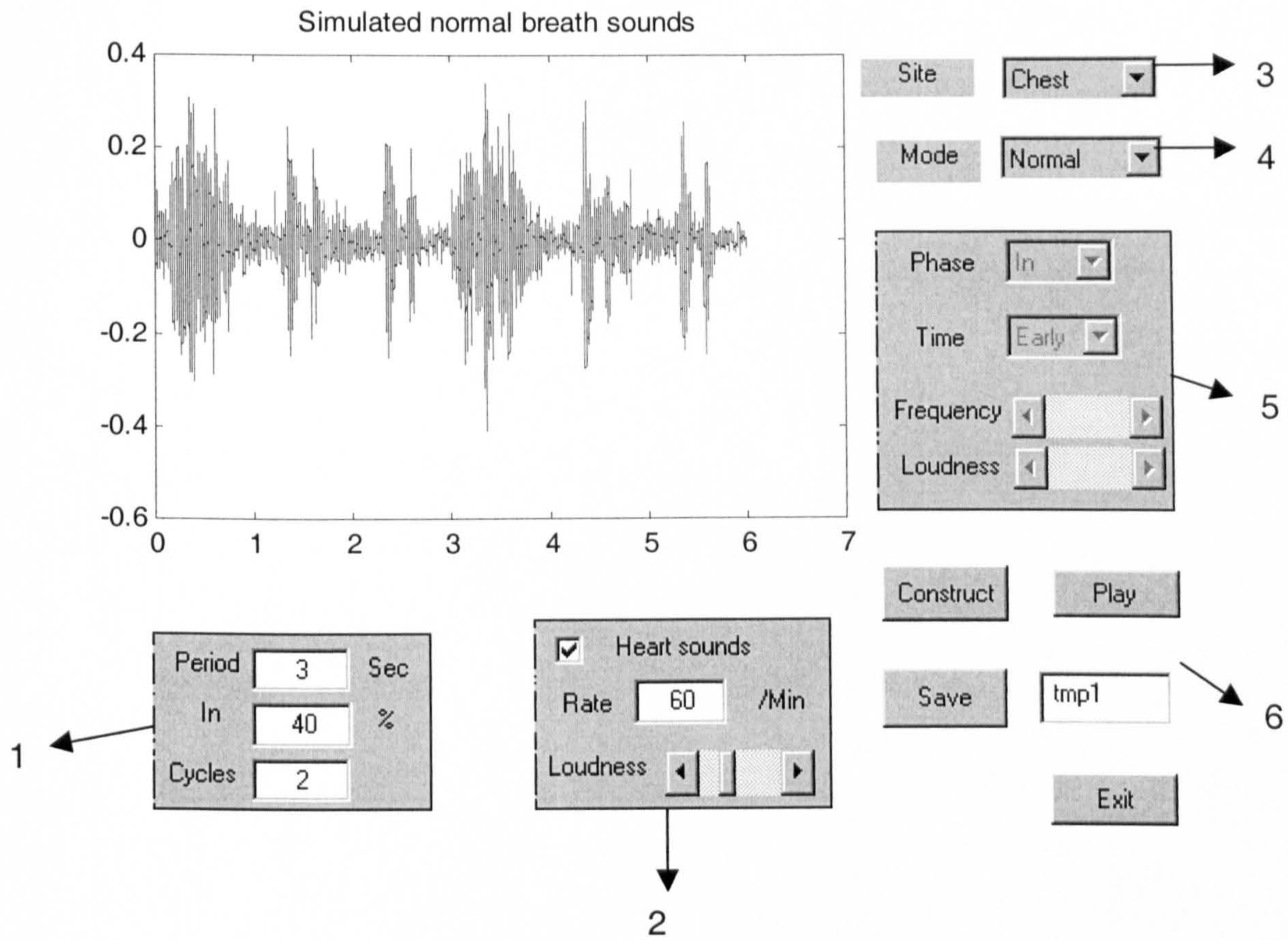

Figure 5.1 Graphical User Interface of simulation

5.4 Audible Tests of Wheezes

Simulated tracheal late expiratory wheezes were produced to test their audibility. All other parameters were kept constant, except that wheezing frequencies and amplitudes were altered. The loudness of wheezing was represented as a percentage of RMS[1] of normal tracheal sound amplitude (NTSA).

---

[1] RMS—root mean squared

Figure 5.2 shows an example of part of a simulated wheeze signal around 280Hz with amplitude of 15% RMS of NTSA. It was audible. (It was inaudible when its amplitude is less than 4% RMS of NTSA.) When this part was embedded in simulated normal tracheal sounds, wheezes were inaudible. By visually examining the wave expanded part of the corresponding segments (refer to Figure 5.3), the periodical waveforms do not appear.

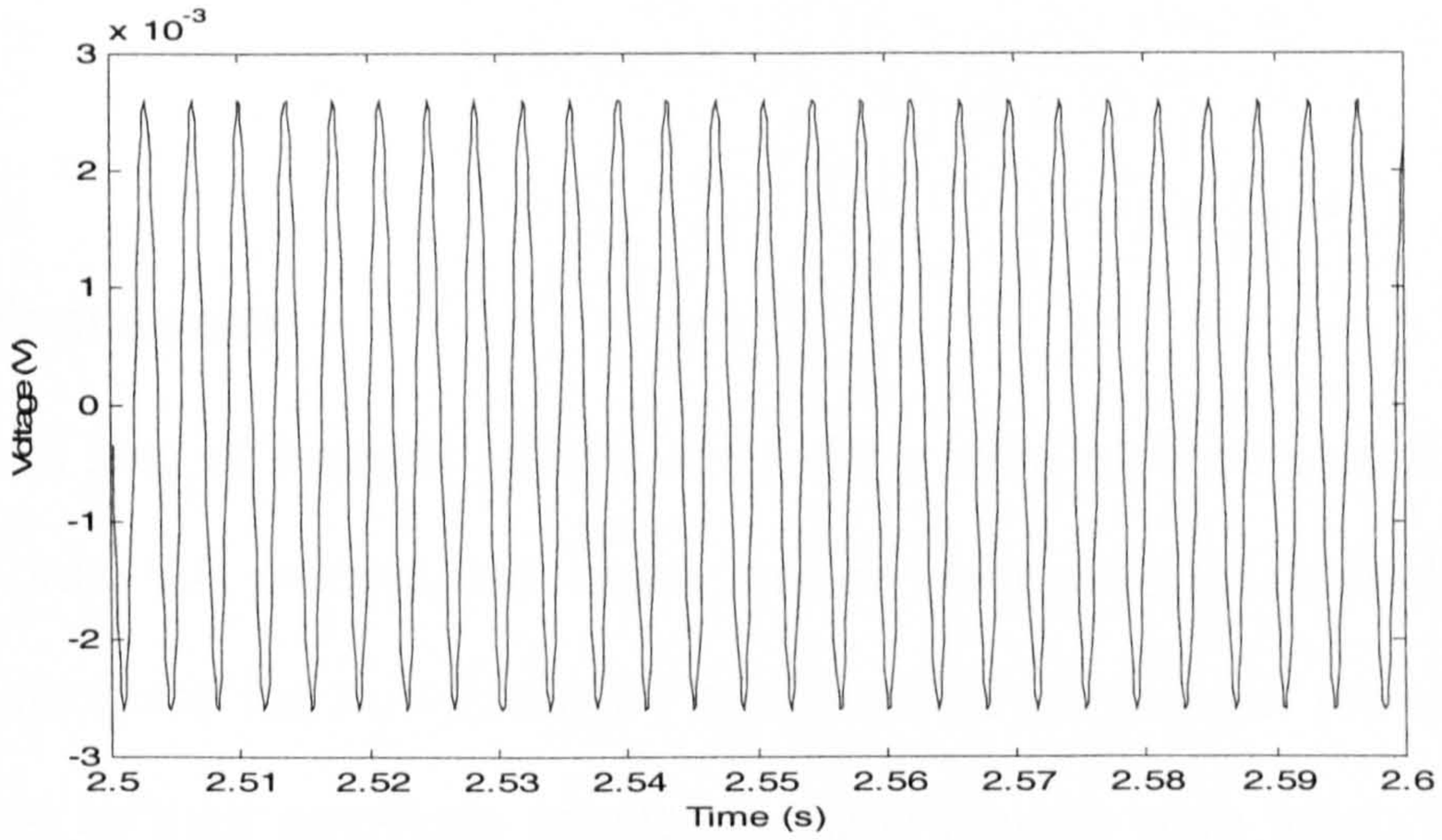

Figure 5.2 Waveform of simulated wheezes.

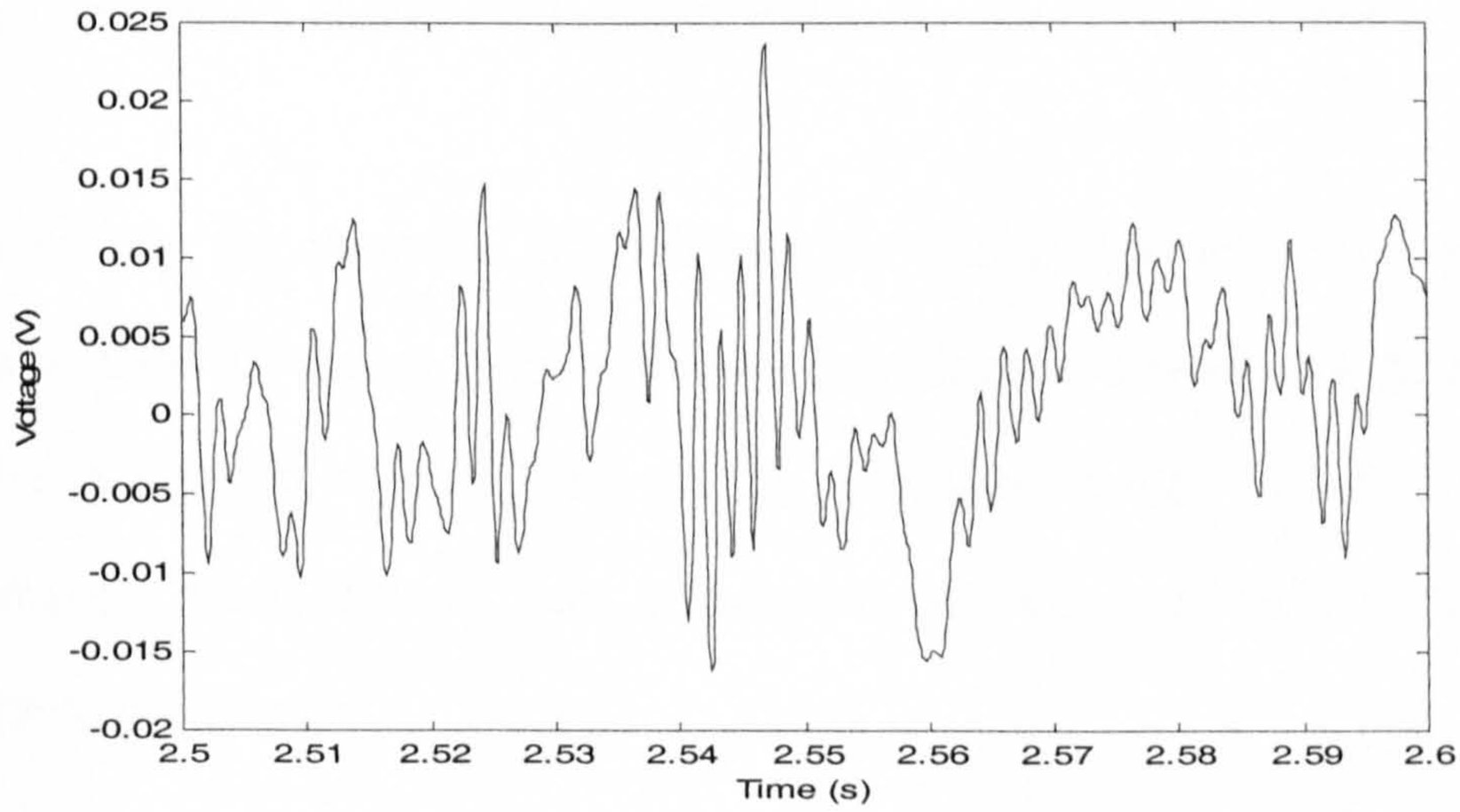

Figure 5.3 Expanding waveform of a wheezing segments.

When the average wheezing frequency was within the normal tracheal sound frequency range and remote from the frequency peaks, it was audible when its amplitude was greater than 15% RMS of NTSA. When the average wheezing frequency was near the peaks, it was audible when its amplitude was greater than 20% RMS of NTSA. When the average wheezing frequency was above the normal frequency range, it was audible when its amplitude was greater than 2% RMS of NTSA.

5.5 Discussion

The normal lung or tracheal sounds simulations were realised by modelling them as a 4^{th} order AR process. But higher orders were used by Gavriely (1992). In order to minimise the non-stationary nature of breath sounds, the sounds were normalised by their amplitude envelopes. Lung sounds were found to require at least order 6, and tracheal sounds require at least order 12. The higher the order, the more detailed spectral information can be modelled, but the greater the calculation burden. Higher orders have not been used in the current work.

Wheezes can be classified as monophonic and polyphonic. Only monophonic wheezes were simulated in the current work. Harmonic type polyphonic wheezes can be easily implemented by adding harmonic parts, that is, multiples of fundamental frequencies. But inharmonic wheeze simulation needs more adjustable parameters.

The phase and timing of abnormal sounds are of importance. In this work, limited combinations were provided. More combinations are possible by providing more

selectable items. In some diseases, both crackles and wheezes can be heard. Currently, either crackle or wheeze events could appear, but their combination is also possible.

Only normal heart sounds were simulated. Abnormal heart sounds can be simulated by selecting suitable parameters in equations (5.7)-(5.10).

The breath sounds simulation has found applications in training (Kompis and Russi 1997; Cardionics Inc 2002). Both advantages and disadvantages are that the simulated parameters can be altered freely. For training purposes, a great variability in breath sounds should be available. This could be achieved easily by altering simulation parameters properly. The simulated sounds should sound as natural as possible, thus some restrictions should be put on the simulation parameters. Otherwise, some exaggerated parameters will make the simulated sounds unnatural.

The breath sounds simulation can also be used for evaluation purposes. Breath sounds with pre-defined characteristics could be produced and reproduced without much effort. Simulated crackles embedded in recorded real breath sounds have been used to test their audibility. In some situations the crackle alone is audible but inaudible when embedded in normal sounds (Kiyokawa *et al.* 2001).

Similar to the above phenomena, wheezes which are of small amplitude and frequency components within normal sounds frequency range, are audible alone but inaudible when added to the normal sounds. This happens when wheeze amplitude is less than about 15% RMS of normal sound amplitude.

When Gavriely and associates (1989) studied the mechanism for wheeze generation, they found oscillations were generated that produced a loud honking noise (wheeze) under selected pressure-flow conditions. It is unclear whether wheezes always appear with strong sound energy. Even if this is the case, it is possible that wheezes are inaudible at the recording site due to transmission loss or by being masked by normal sounds.

In addition, these simulated sounds could be used to test the usefulness and effectiveness of signal processing algorithms for a computer-aided breath sounds analysis system. Simulated normal and wheezy sounds have been used to test analysis algorithms in this work. Details are described in the next chapter.

In summary, this flexible simulation environment integrates simulated algorithms of normal, crackle, wheeze, and heart sounds. It can be used to evaluate abnormal sounds' audibility and automatic algorithms. Though at present it is not a perfect cardiorespiratory sound simulator, further performance improvement may be achieved with the aid of medical experts.

Chapter 6 Automatic Wheeze Detection

A wheeze may be a sign of some diseases, but the presence or absence may not reflect the severity of the associated disease. The clinical status of the patient may be better defined by parametric characteristics of wheezes, such as duration, timing, dominant frequencies.

As described in 2.5.5, some studies of wheeze detection were validated by a physician listening. In this chapter, the possibility of detecting wheezes based on auditory system modelling is investigated.

6.1 Some Hearing Principles

6.1.1 Hearing Mechanism

The human hearing system consists of three sections: the outer ear, the middle ear and the inner ear. The outer ear helps to locate sound sources and enhances some frequencies with respect to others (Howard and Angus 2001). The major function of the middle ear is to achieve the efficient transfer of sound from the air to the inner ear, i.e. cochlea (Moore 1997). The cochlea converts mechanical vibrations into nerve firings that will be processed by the brain (Howard and Angus 2001). The cochlea is divided along its length by two membranes; one of them is the basilar membrane. This membrane is responsible for carrying out a frequency analysis of input sounds. In a normal, healthy ear each point on the basilar membrane is sharply tuned, responding with high sensitivity to a limited range of frequencies (Moore 1997).

6.1.2 Auditory Sensitivity

Absolute thresholds of loudness

The minimum detectable level of a sound in the absence of any other sounds is the absolute threshold of that sound. In the audible frequency range of 20Hz to 20kHz, the absolute threshold is a non-linear function of frequency. In the interesting frequency range of up to 4kHz, the ear is most sensitive around 3-4kHz in free field, below which the thresholds increase; while using an earphone, the ear is most sensitive around 1 or 2kHz depending on the earphone type (Moore 1997; Yost 2000).

Temporal integration

A signal must have some critical amount of energy to be detectable. The process of integration of the power is completed in 300ms. If the duration is less than 300ms and greater than 10ms, then the power of the signal must be increased to make it audible (Yost 2000).

6.1.3 Masking and Critical Band

It is intuitive that a sound (signal) should be more intensive to be heard when another sound appears. Sometimes when the other sound is much more intensive than the signal, the phenomena of masking occurs. Masking has been defined as the process by which the threshold of audibility for one sound is raised by the presence of another sound. The peripheral auditory system behaves as if it contains a bank of band pass filters (also named as critical band, auditory filters), with overlapping pass-bands. When a listener attempts to detect a signal in a noisy background, he is assumed to use a filter with a centre frequency close to

that signal. Thus ignoring noise outside the filter band, only the components in the noise that pass through the filter have a masking effect on that signal. It is assumed that the threshold for the signal to be detected corresponds to a certain signal-to-noise ratio at the output of the filter (Moore 1997).

The bandwidth of such an auditory filter is called 'critical bandwidth'. Practically, the equivalent rectangular bandwidth (ERB) of the critical band can be calculated as (Moore 1997)

$$ERB_f = 24.7(4.37f + 1) \tag{6.1}$$

where f is centre frequency in kHz.

6.1.4 Co-modulation Masking Release (CMR)

Co-modulation means the components of a masker have the same amplitude modulation pattern in different frequency regions. In this situation, the masking threshold will decrease (Moore 1997).

6.2 Short-Time Fourier Transform (STFT)

6.2.1 Fourier Transform and Discrete Fourier Transform (DFT)

The Fourier transform of a function $f(t)$ is defined by

$$F(\omega) = \int f(t)\, e^{-j\omega t} dt \tag{6.2}$$

The discrete Fourier transform of a sequence $f[n]$ $(n = 0,1,\ldots N-1)$ is defined by

$$F[k] = \sum_{n=0}^{N-1} f[n]\, e^{-jkn\frac{2\pi}{N}} \tag{6.3}$$

6.2.2 Short-time Fourier Transform (STFT)

If a signal changes over time, then the classic Fourier transform could not reflect the signal's time-varying nature. An easy way to overcome the shortcomings of the normal Fourier transform is to compare the signal with elementary functions located in time and frequency simultaneously, that is (Qian and Chen 1996),

$$STFT(t_0, \omega) = \int f(t) h^*(t - t_0) e^{-jwt} dt \qquad (6.4)$$

where $f(t)$ is the signal, and $h(t)$ is a window function.

The window function $h(t)$ has short time duration. The Fourier transform of the signal $f(t)$ windowed with $h(t)$ shifted by t_0 is calculated. Thus the signal's local frequency characteristics could be reflected.

If the time duration and frequency bandwidth of $h(t)$ are Δ_t and Δ_ω respectively, then according to Heisenberg's uncertainty principle (Qian and Chen 1996),

$$\Delta_t \Delta_\omega \geq 0.5 \qquad (6.5)$$

That is, good time resolution and frequency resolution can not be obtained simultaneously. There must be a trade-off between the time resolution and frequency resolution.

Discrete calculation of the STFT is realised by breaking $f[n]$ into windowed frames and applying the DFT to each frame.

6.3 Spectrogram

Normal breath sounds are time-dependent signals as they are changing with flow rate. In a breath cycle the flow rate increases from 0 to a maximum then decreases to 0 again, both in inspiration and expiration. When abnormal sounds appear, their timing and/or evolution with time are important.

The analogue spectrograph (McKusic *et al.* 1955) and later on the digital spectrograph based on STFT (Kraman 1983; Pasterkamp *et al.* 1989) provided a useful visual tool to distinguish normal and abnormal sounds after some training. Nowadays the digital spectrogram is relatively easy to realise due to the progress of computing capacity, and is normally used as a representation of the results.

A spectrograph represents a breath sound in the time-frequency plane, with sound intensity displayed with a grey or colour scale. Figure 6.1 shows a normal tracheal sound of a breath cycle. Figure 6.2 shows an abnormal tracheal sound containing wheezes in expiration, where the wheezes are indicated by arrows. Shown in Figures 6.1 and 6.2, inspiration and expiration phases are clearly distinguished. Compared to the normal breath sounds, the abnormal breath sounds show a prolonged expiration phase, and especially some irregular wavy lines (wheezes).

Although the frequency and duration of wheezes can be read from the spectrograph, it is time consuming when samples are huge. The realisation that wheezes can be detected and quantified automatically will relieve investigators from tedious work and overcome subjective variations.

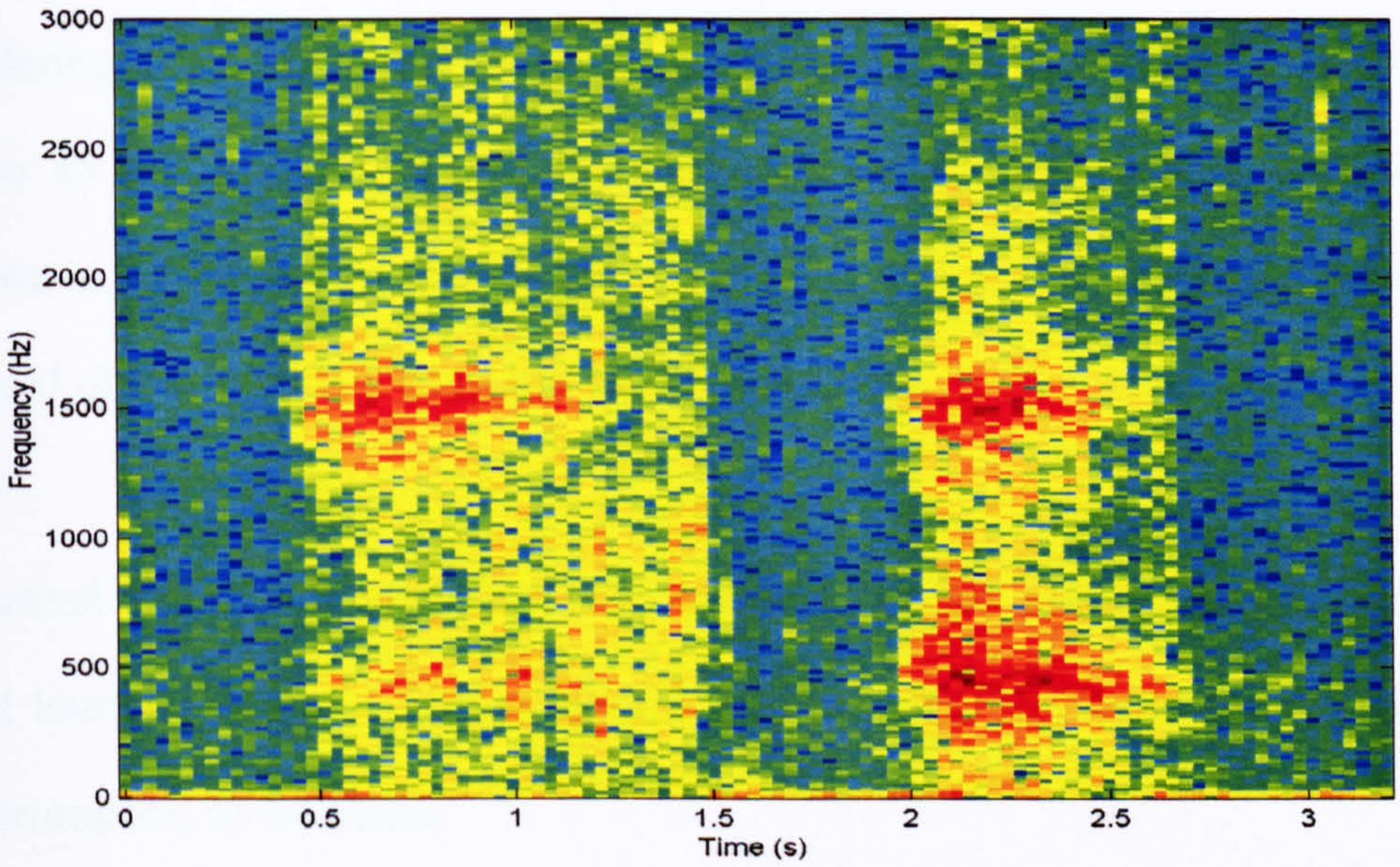

Figure 6.1 Spectrogram of a normal tracheal sound.

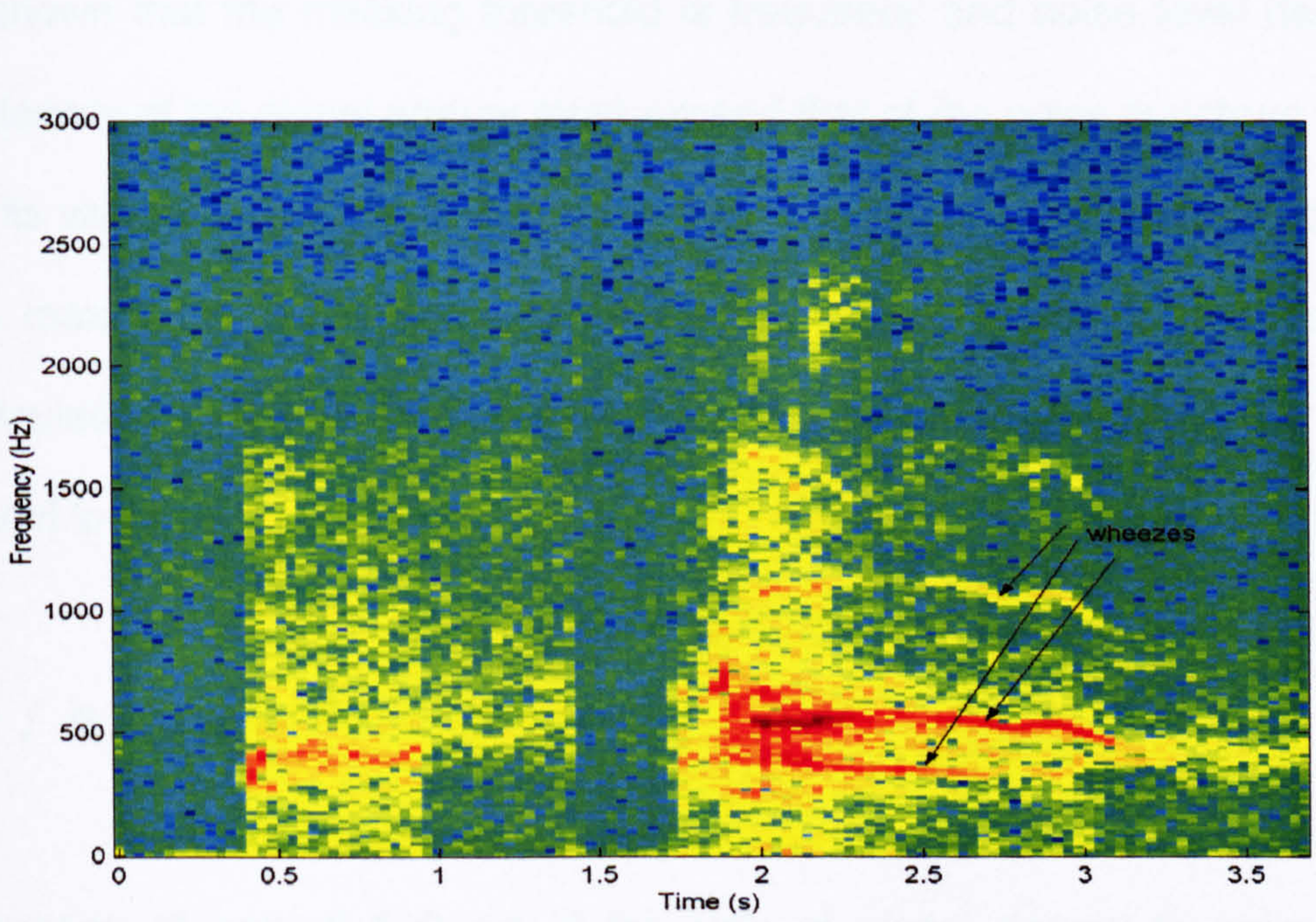

Figure 6.2 Spectrogram of a wheezing tracheal breath sound.

6.4 Algorithm

By the definition in 2.2, wheezes are audible signals that have a 'musical' nature. Considering the continuous adventitious sounds are the tonal sounds (signal) additive to the apparent normal breath sounds (noise), then the detection of wheezes is like detection of a signal in a masker (noises). If the signal is above the threshold of the masking, then the signal is audible.

The overall aim of this algorithm is to detect wheezes automatically. This is done by first identifying audible tonal signals and then sorting them to find the signals that correspond to wheezes.

The threshold is based on the results of Reed and Bilger (1973). Their results have shown that the masking threshold is frequency and noise level dependant. The intensity of the signal energy must exceed that of the noise spectrum level[1] by amounts varying from 8 dB at 250 Hz to 14 dB at 4 kHz. The effect of noise level is less, making the above threshold values change by only around 1dB. To simplify the calculation, only three threshold levels are used in the algorithm. The baseline threshold level th is given by:

$$th = 9.5+3.48\log10(f)-10\log(ERB_f) \qquad (6.6)$$

where f is the signal frequency in kHz.

For duration of around 100 ms, if the ratio of signal energy to noise energy, (E_S / E_N expressed in dB) is greater than or equal to th, then the signal is audible.

---

[1] Noise spectrum level is the average noise power per Hz.

In this work, the STFT time resolution (Δ_t) is 32ms, so 96ms ($3\Delta_t$) is used as the baseline duration. Two more threshold levels are defined: $th+3$, corresponding to 64ms ($2\Delta_t$) and $th+4.8$, corresponding to 32ms (Δ_t). See table 6.1 for a summary of this information.

Table 6.1 Threshold

Duration (ms)	$10\log E_S - 10\log E_N$ (dB)
96	$th(=9.5+3.48\log(f)-10\log ERB_f)$
64	$th+3$
32	$th+4.8$

The procedure is illustrated in Figure 6. 3 and details are described in the following steps:

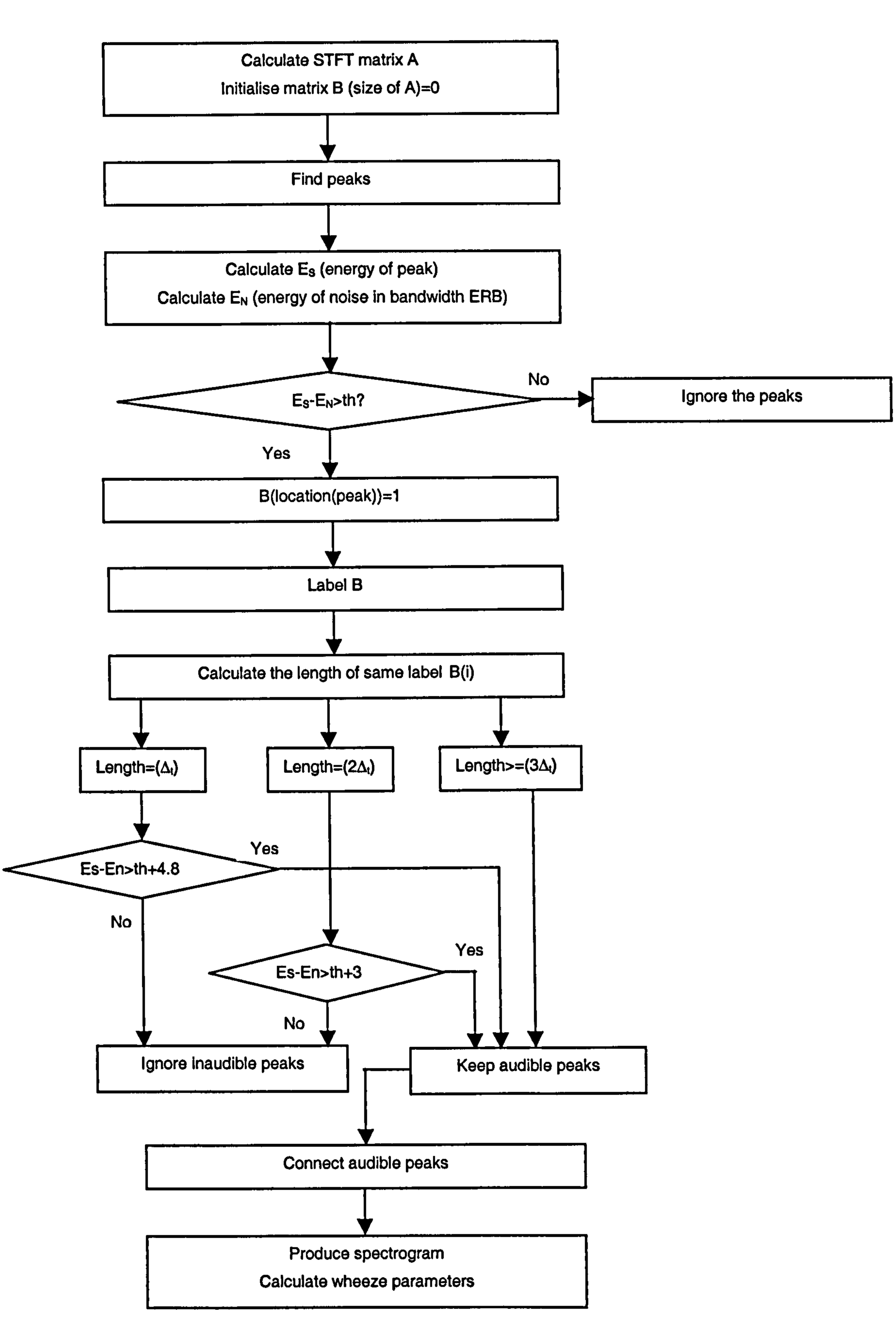

Figure 6.3 Schematic representation of wheeze detection algorithm.

1. Construct a time-frequency matrix A by calculating the STFT with the following parameters: 512 samples, using a Hanning[1] window with the same length, 50% overlap. The sampling rate was 8000Hz. Each cell of the matrix contains the power spectrum value ($\left|STFT(t_0,\omega)\right|^2$) at $[t_0-\Delta_t, t_0+\Delta_t]$ and $[\omega-\Delta_\omega, \omega+\Delta_\omega]$. Initialise to zero all elements of a binary matrix B (the same size as the matrix A). Matrix B is used as a way of 'tagging' the identified peaks.

2. Find peaks along the columns (along time axis). The peaks may either belong to wheezes (signal) or noises.

3. Regarding a peak as corresponding to a signal component, calculate the energy of the signal, and the energy of the noises that pass through the auditory filter that centres the signal. The filter bandwidth is calculated according to equation (6.1). If the signal-to-noise ratio is above the threshold, mark that peak with 1 in matrix B.

4. Label B using a 'connected component labelling' algorithm (Seul *et al.* 2000) with minor modifications. (Peaks in a column separate at least two frequency bands as a result of step 3.)

   Search the matrix column by column and assign a value greater than 1 (a label) to each non-zero cell. The value is decided by the labels of three neighbours, as shown in Figure 6.4.

---

[1] $h(n)=0.5[1-\cos(\frac{2\pi(n-1)}{N-1})]$, where $n=1,2,...N, N$ is window length.

- If all the neighbours of c (i, j) are 0, c (i, j) is assigned a new unused label.
- If there is one neighbour cell with a non-zero label, assign the label to the cell c (i, j).

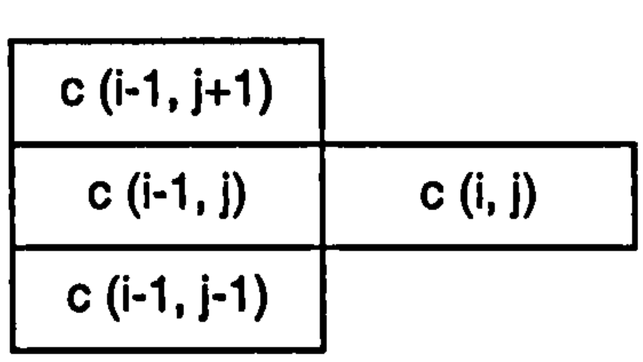

Figure 6.4 Neighbours definition in labelling algorithm.

5. According to the length (continuity in time) of the same label (continuity in frequency), identify signals of less than 96ms and use the threshold values in table 1 to remove the tags in the B matrix corresponding to shorter signals which do not exceed the higher threshold values.

6. Normal breath sounds have a strong relationship with flow rate (Kraman 1984; Shykoff *et al.* 1988; Soufflet *et al.* 1990; Gavriely and Cugell 1996; Harper *et al.* 2003). That is, all frequency components of the spectrum change with flow rate: the higher the flow rate, the higher the spectral power. So the signal (wheezes) and the noise (apparent normal sounds) are considered to be co-modulated by the flow. The co-modulation masking release phenomena has the effect of decreasing the masking threshold (Moore 1997). This effect may mean that separately identified wheezes in step 5 are actually two components of the same wheeze. To compensate for this, detected peaks from step 2 which are neighbours of a wheezing component are relabelled so that they

each have the same label, thereby connecting two wheeze signals (neighbour definition is the same as in step 4).

7. Produce a spectrograph with wheezes marked. Calculate the average frequency (a wheeze usually has a frequency change from the beginning to the end), standard deviation of wheezing frequency, duration of each wheeze, and percentage occupation of wheeze in each respiratory phase. (The duration of each phase was estimated because of lack of reference signals such as flow rate.) Only average wheezing frequency and occupation in each phase are displayed to keep the results display compact.

6.5 Implementation

The algorithm described above was implemented in Matlab. A GUI shown in Figure 6.5 was developed to incorporate displays of phonopneumography, spectrogram, spectrum, and the following adjustable parameters and operations:

1. Upper frequency—this parameter is used to limit the shown frequency range of the spectrogram and the spectrum (up to the Nyquist frequency).

2. Load signal—a standard Windows open file dialog box appears for the user to load PCM format wave files.

3. Select/Restore—visible alternately. If *Select* is visible, clicking this button brings out a cross cursor. Moving and clicking the left mouse button to select start and end points of the signal segment of interest, then corresponding

expanded phonopneumograph, spectrogram and spectrum will be displayed. If *Restore* is visible, click this button to restore displays to their originals.

4. Analysis—perform automatic wheeze detection, results are shown in a pop up window. The mean wheezing frequency and occupation (%) in each phase are shown. Results can be exported to an Access file if required. In case of no wheeze detected, 'no wheeze' will be displayed.

5. Play sound—playing sound with cursor tracking the signal part being played. Only the displayed part of the signal is played.

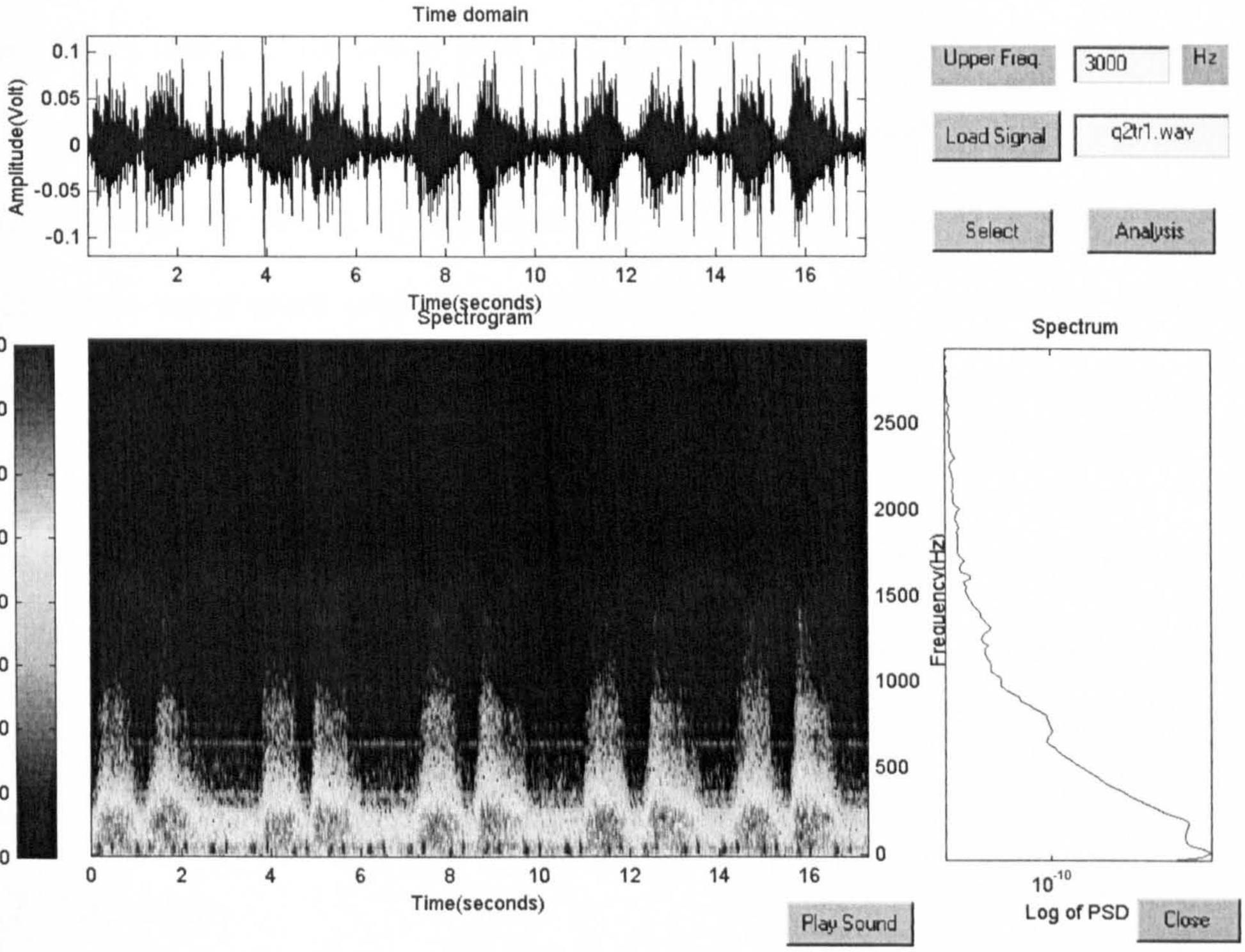

Figure 6.5 A GUI for display and analysis of breath sounds.

6.6 Samples

First, simulated data were used to validate the algorithm. Then three sets of real data were employed.

One set of real sound samples was recorded on the trachea from 20 patients using the same mobile phone. The ages of the subjects ranged from 12 to 61 years, ten of them were women, and seven of them had asthma.

Another set of real data was recorded on the trachea from 16 adult patients using their own mobile phones. Three of them were men. They were asked to take part in a two weeks monitoring programme, in which they reported their peak expiratory flow (PEF) rate and recorded their tracheal sounds for 5 cycles twice a day. Two of the recording sets totally failed. The first contained nothing and the other was contaminated by strong ripples, possibly due to electrical interference. Nine of the recordings were clear and of normal quality. Five of the recordings were audible but did not sound like normal recordings (refer to table A.1 in the appendix).

The third set of measured sounds was recorded on the chest from 6 patients using the Escope in a hospital ward. The ages of the subjects ranged from 54 to 81 years, four of them were women.

6.7 Results

Figure 6.6 shows an example of simulated normal lung sounds with moderate heart sounds. No wheezes are detected in this sample. Figure 6.7 shows an example of simulated monophonic wheezy tracheal sounds, with wheeze of mean

frequency around 280Hz and duration about 600ms. Detected wheezes are marked with black pixels. Analysed results are shown in figure 6.8.

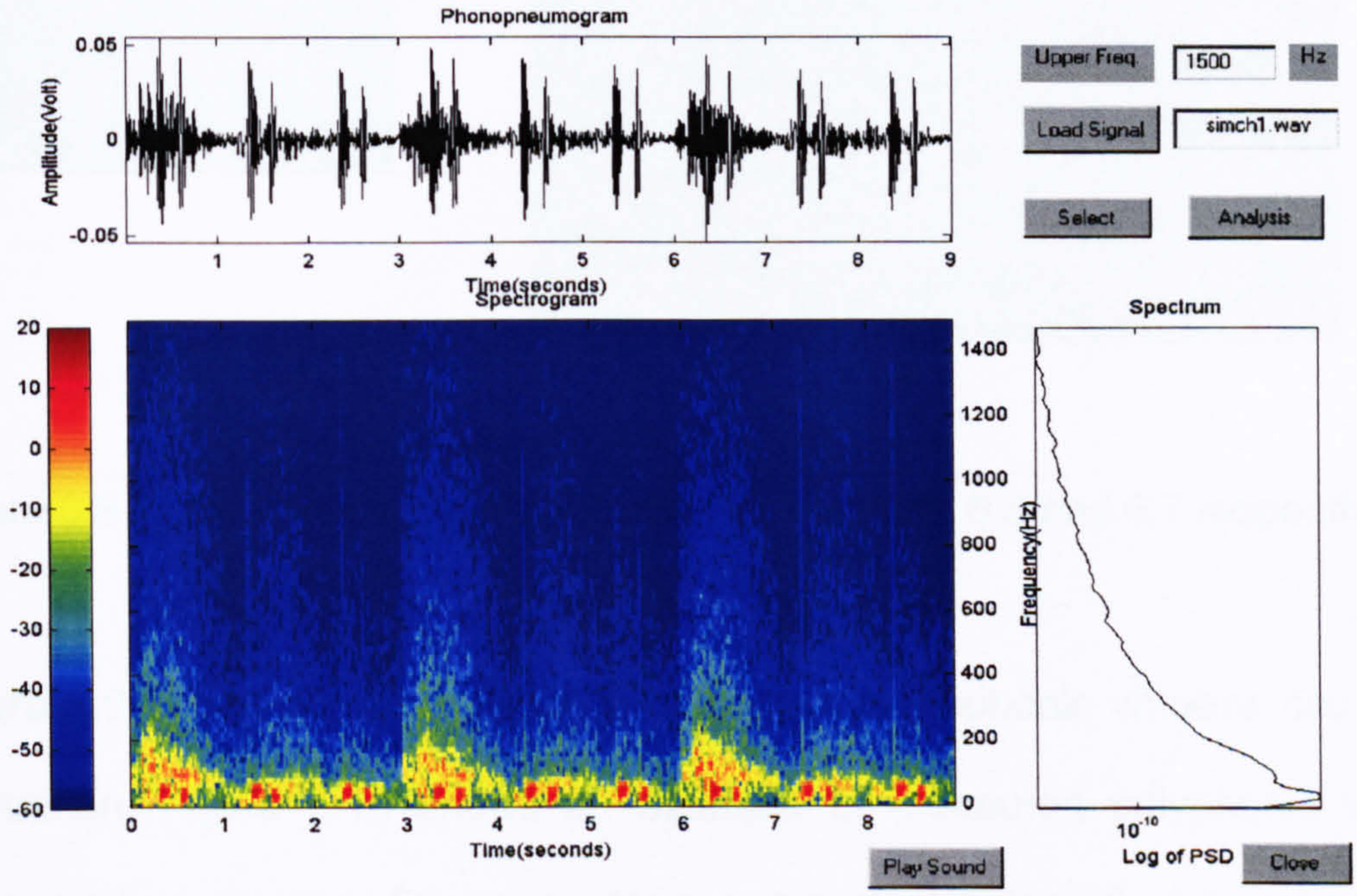

Figure 6.6 Simulated normal lung sounds with moderate heart sounds.

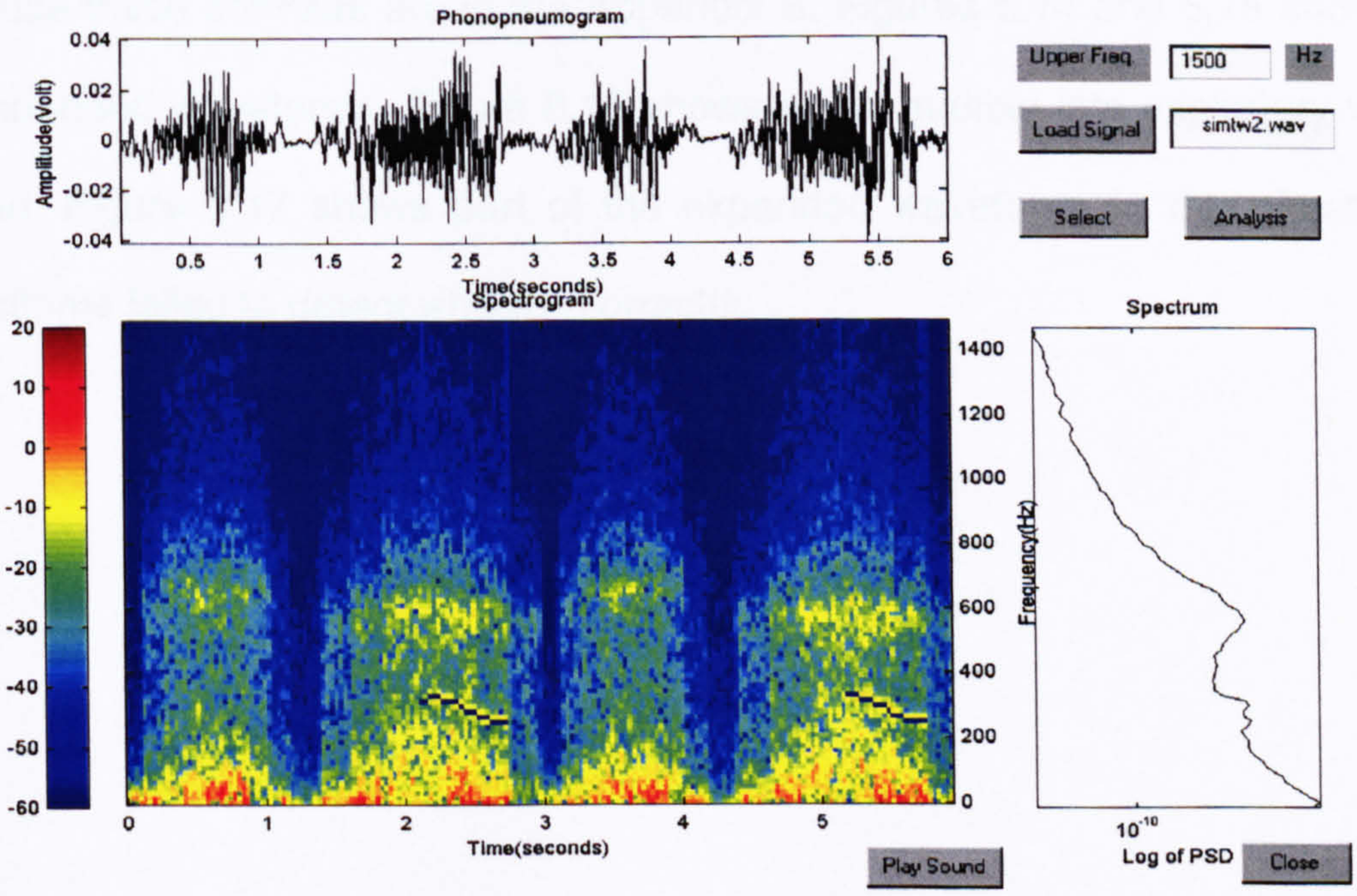

Figure 6.7 Simulated wheezy tracheal sounds. Black lines trace wheezes.

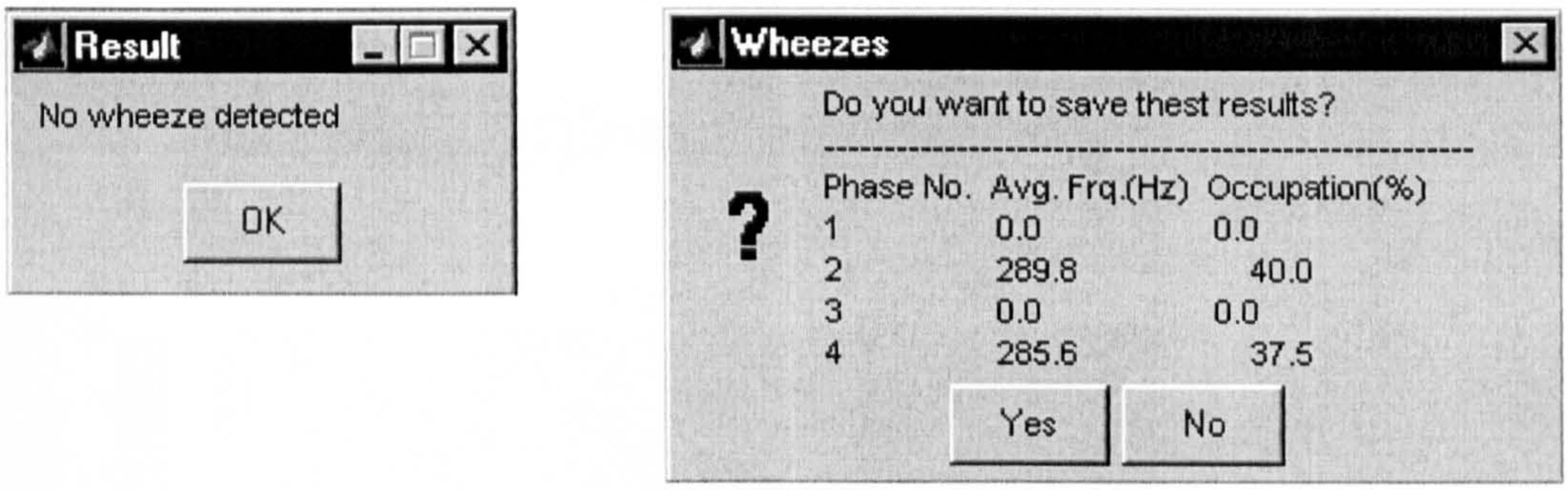

Figure 6.8 Analysis results of data displayed in Figure 6.6 and 6.7 respectively.

Figure 6.9 shows an example of a measured monophonic wheeze occurring in inspiration. Figure 6.10 shows an example of measured polyphonic wheezes occurring in expiration. For comparison, contours of detected wheezes using this new method and previous methods are shown in Figures 6.11- 6.13. Programs to produce these contours are in the Appendix B. Figures 6.14 and 6.15 show parts of expanded waveforms. Figure 6.16 shows a just audible late expiratory wheezy sound. Figure 6.17 shows part of the expanded waveform. In this situation, all algorithms failed to detect wheeze correctly.

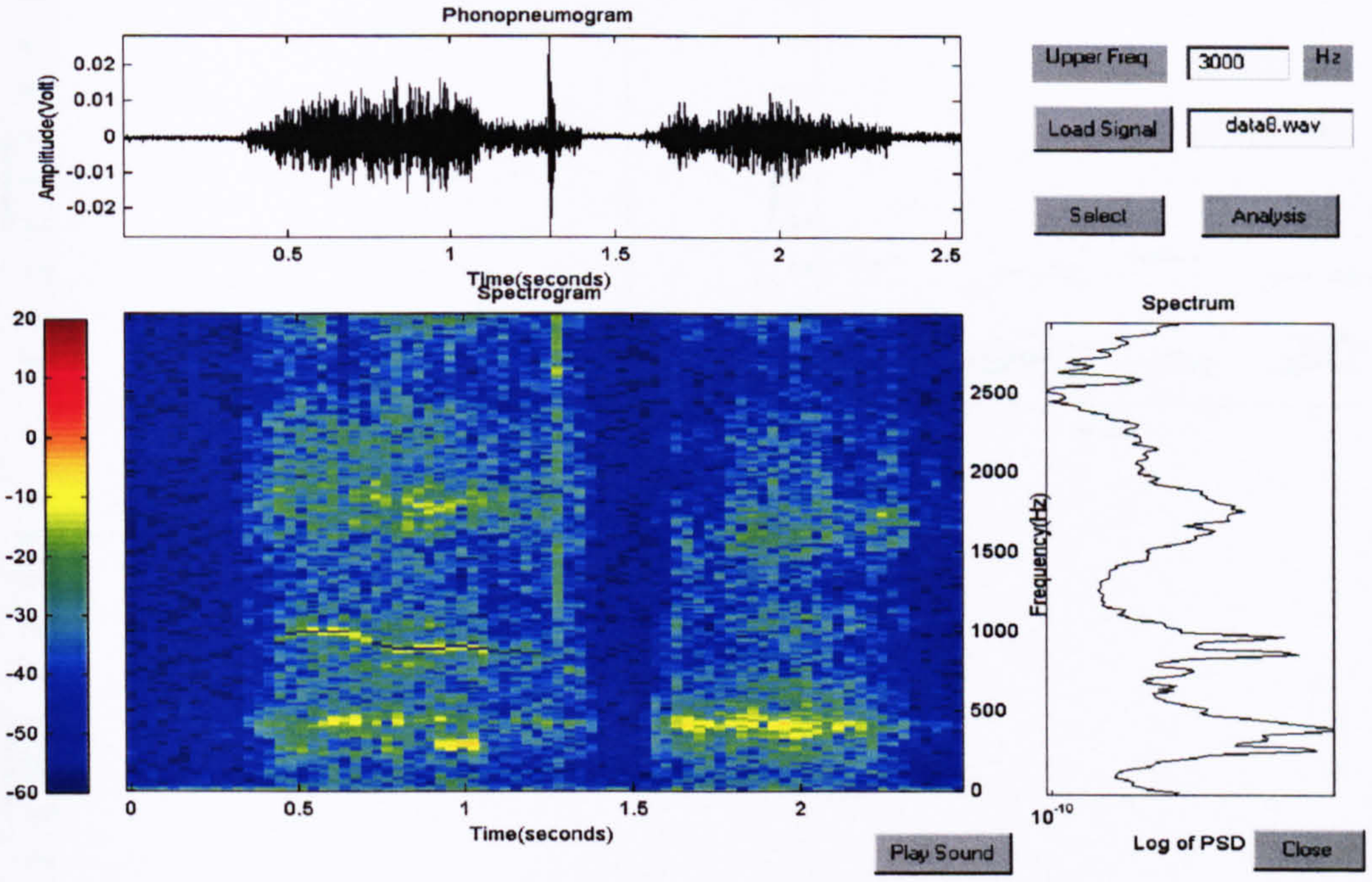

Figure 6.9 Monophonic wheeze in inspiration.

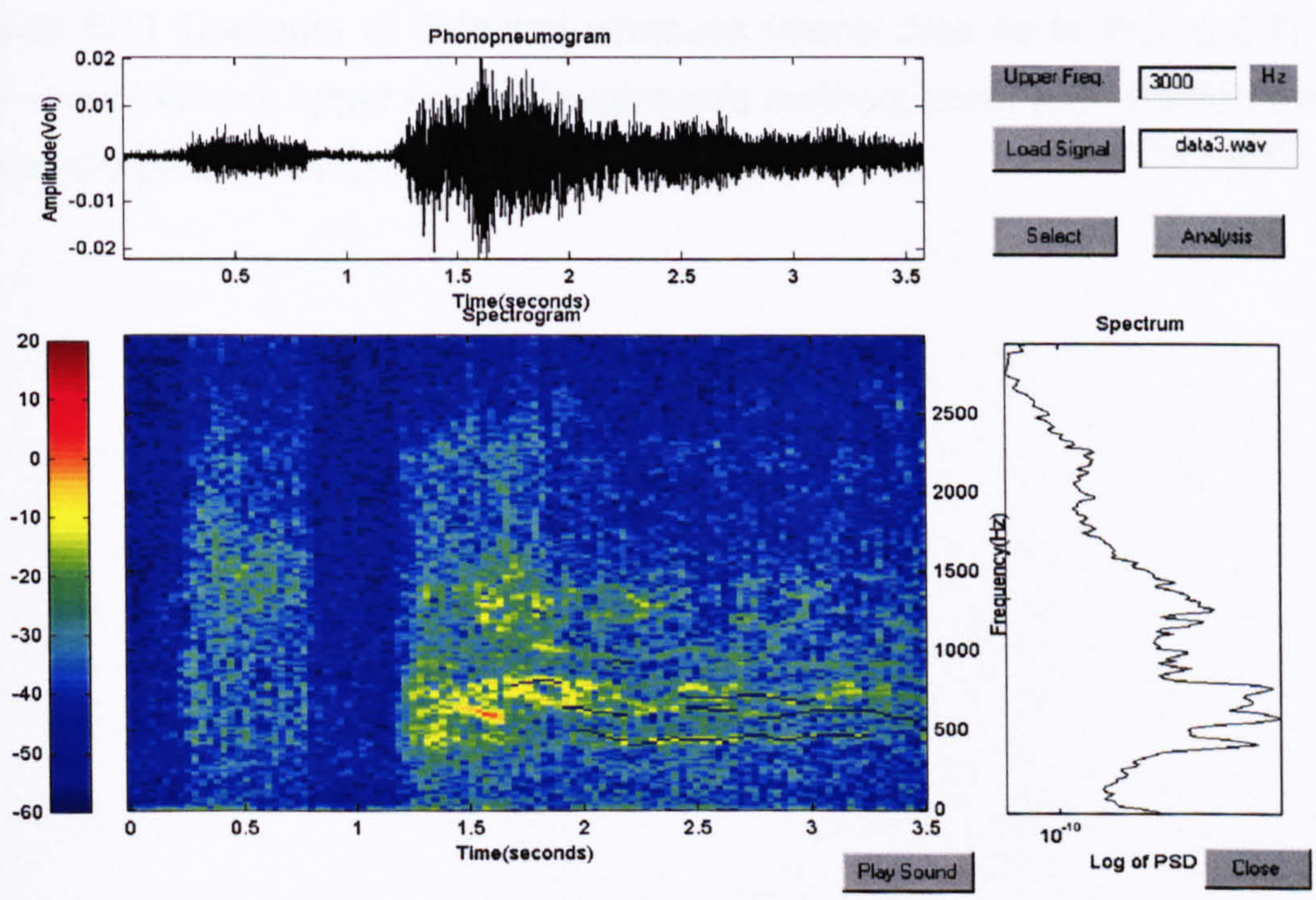

Figure 6.10 Polyphonic wheezes in expiration.

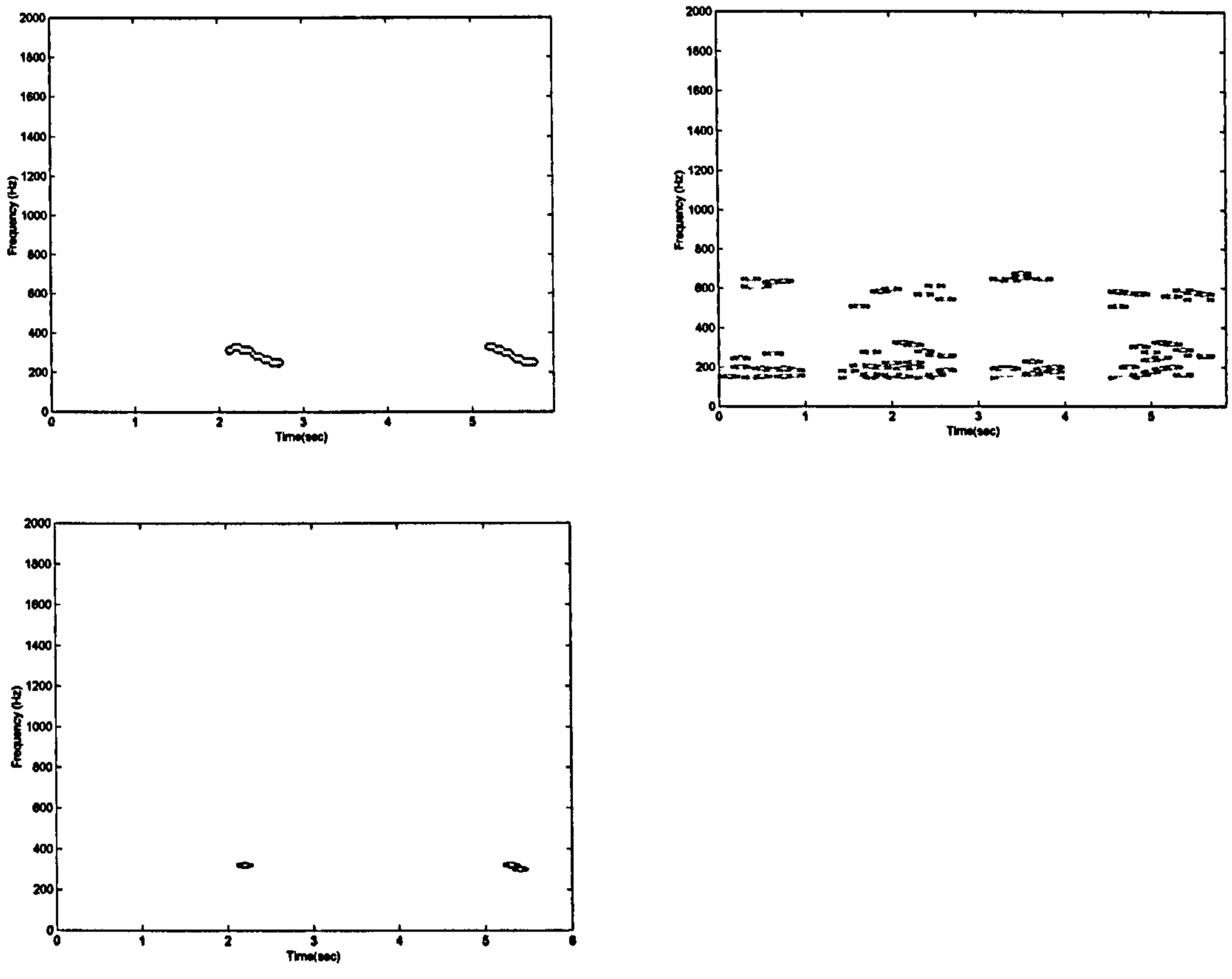

Figure 6.11 Contours of detected wheezes (same data as in Figure 6.7). Upper left—new method; upper right—Baughman's method; lower left—Fenton's method. Homs-Corbera's method failed to detect any wheezes.

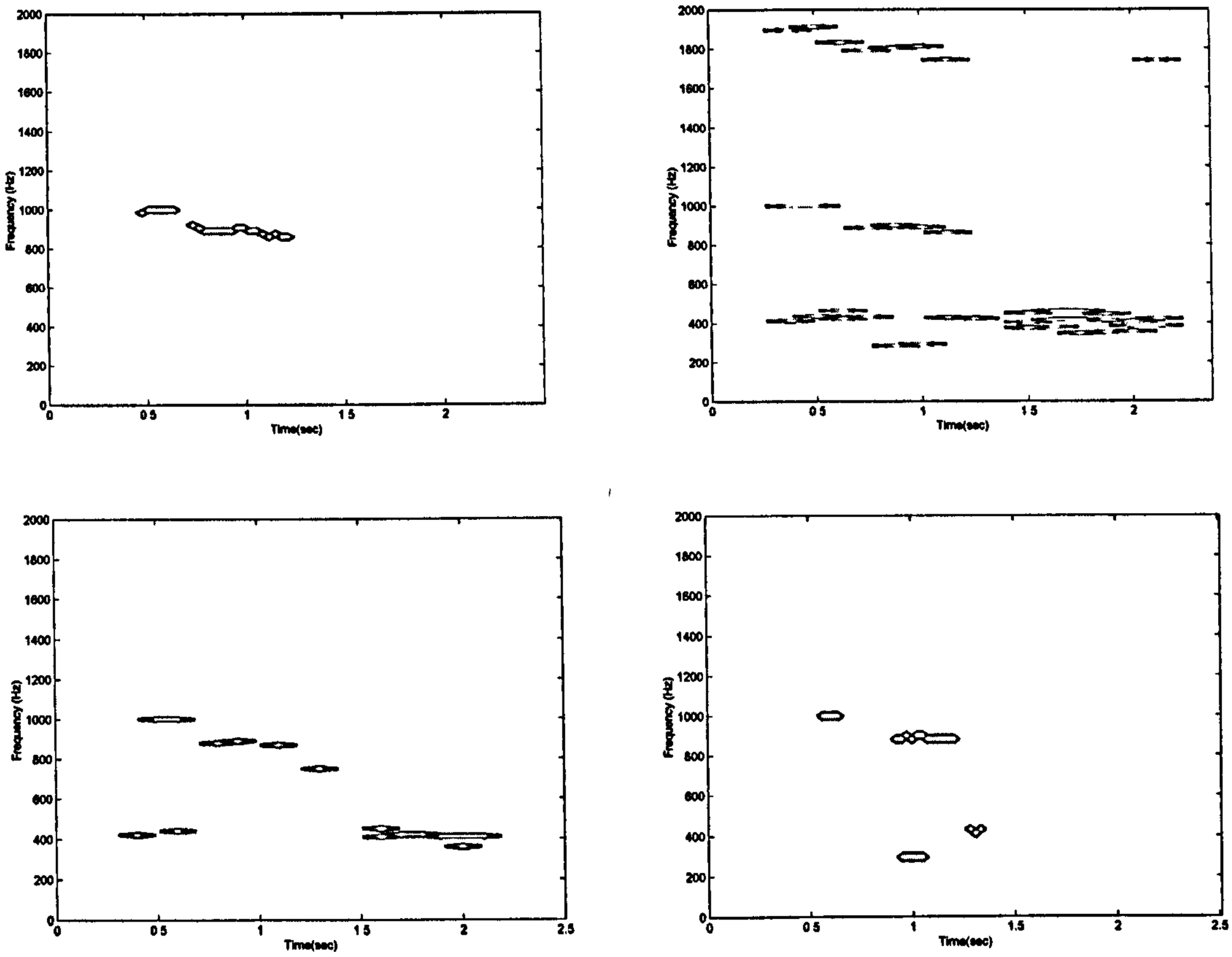

Figure 6.12 Contours of detected wheezes (same data as in Figure 6.9). Upper left—new method; upper right—Baughman's method; lower left—Fenton's method; lower right—Homs-Corbera's method.

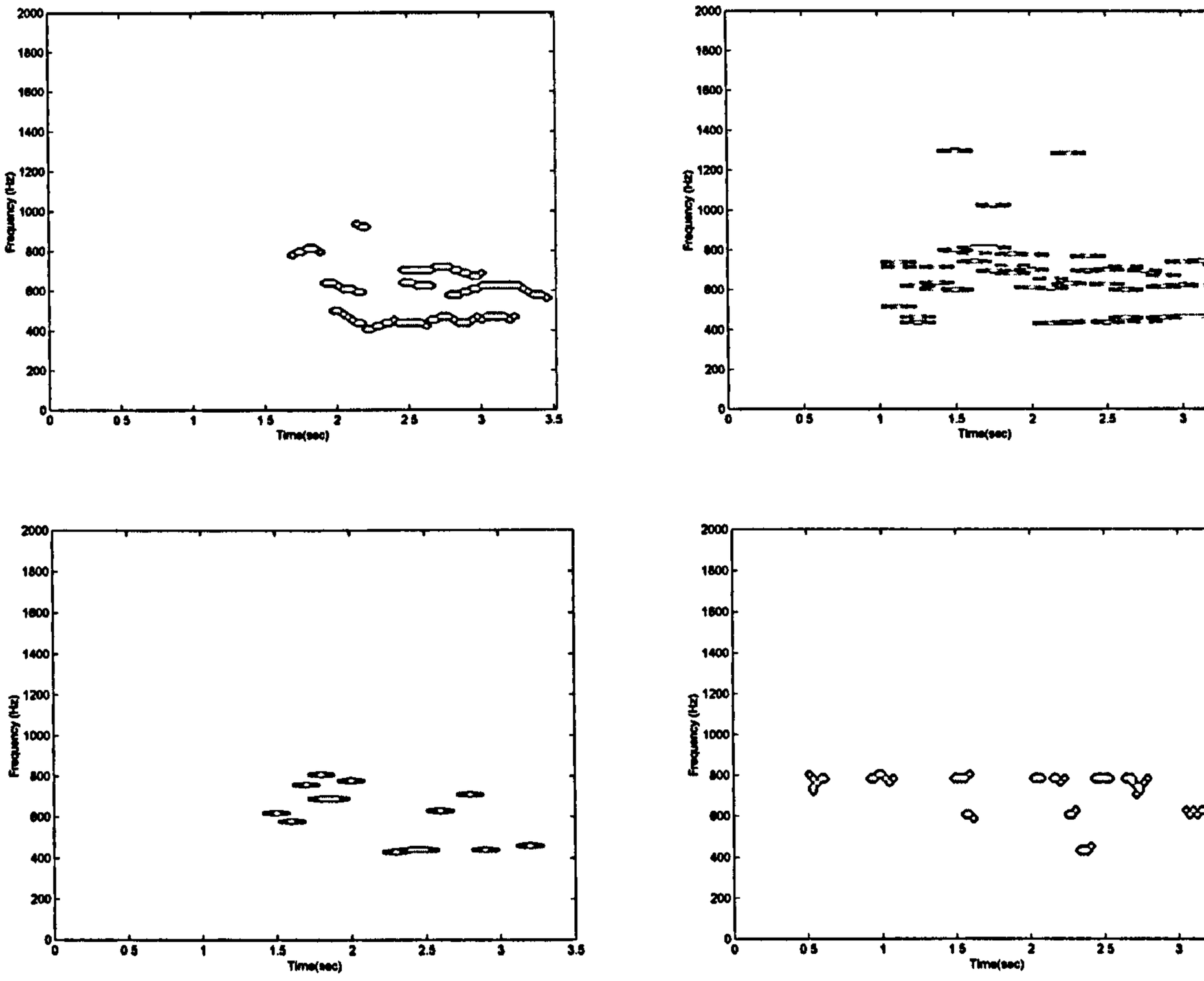

Figure 6.13 Contours of detected wheezes (same data as in Figure 6.10). Upper left—new method; upper right—Baughman's method; lower left—Fenton's method. Lower right—Homs-Corbera's method.

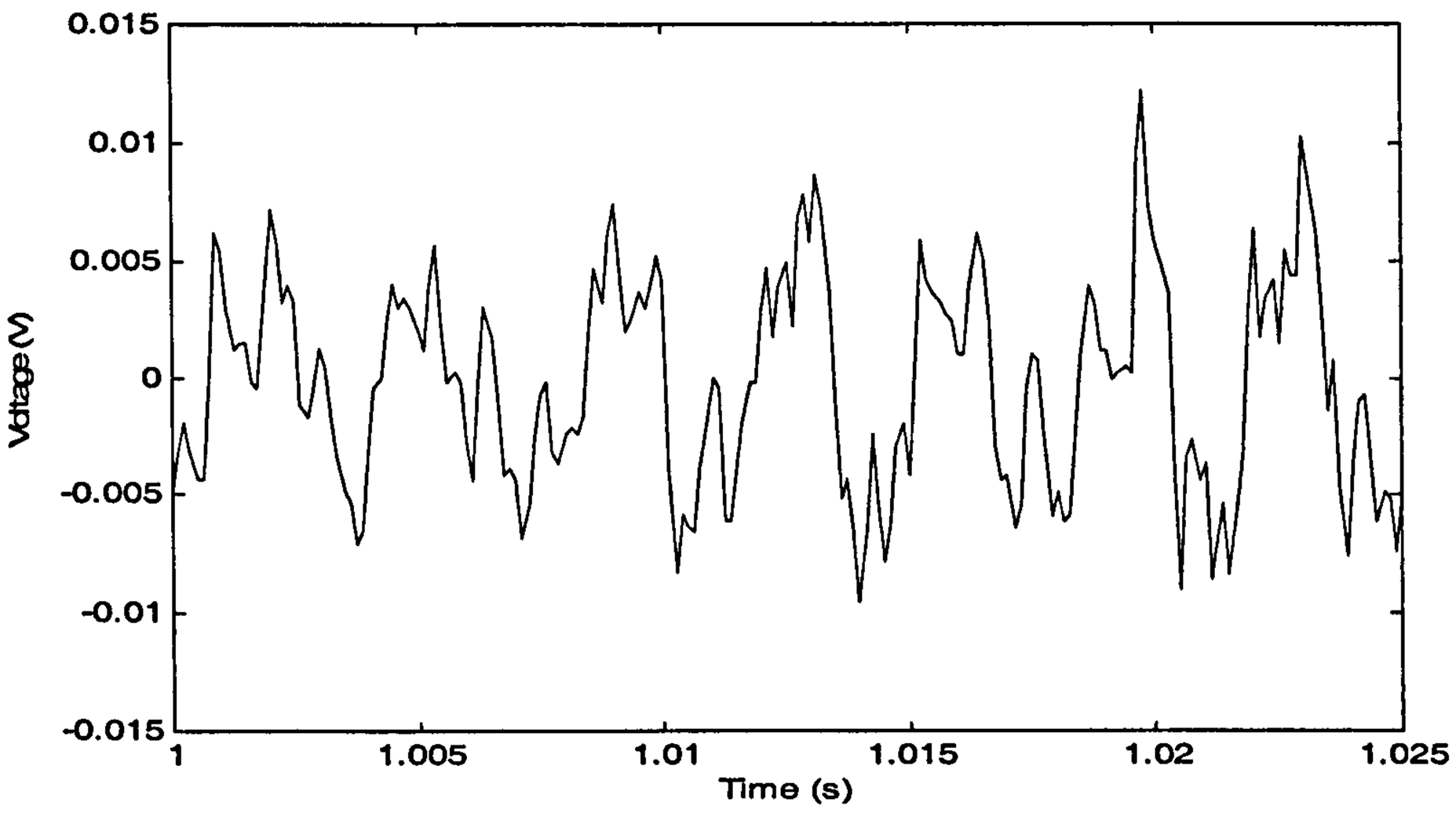

Figure 6.14 Wave-expanded part of monophonic wheezes (same data as in Figure 6.9).

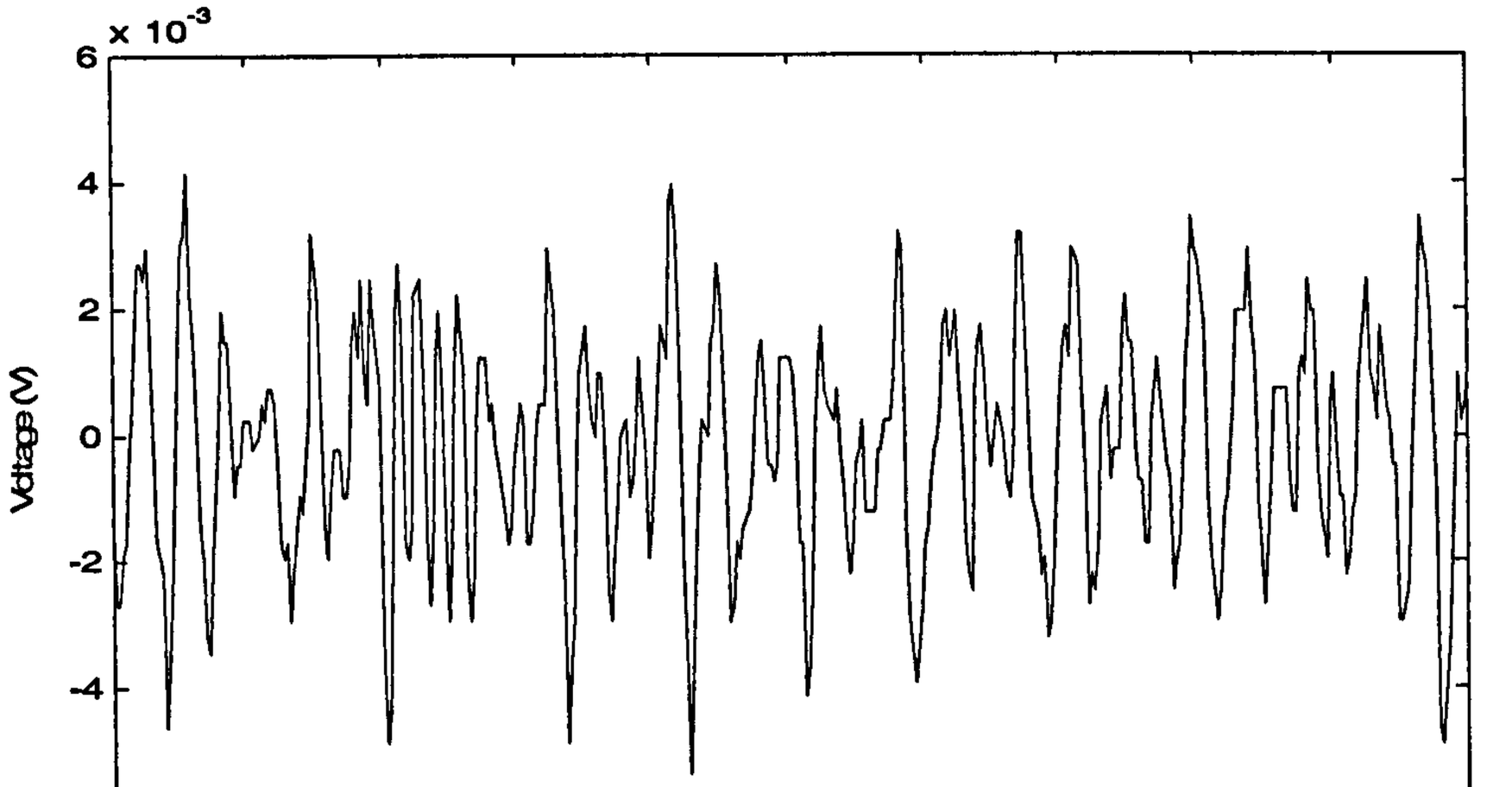

Figure 6.15 Wave-expanded part of polyphonic wheezes (same data as in Figure 6.10).

Figure 6.18 shows an example of data from a patient during the two weeks monitoring programme. It shows cycle to cycle differences of tracheal wheezy sounds. Day to day variations also existed. During the monitoring period, this patient had persistent wheezes. Figure 6.19 shows the trends of peak expiratory flow rate (PEFR) and average expiratory wheeze percentage occupation, which is averaged on 2 to 6 expiratory phases of each recording.

Another example is shown in Figure 6.20. This patient had intermittent wheezes during the monitoring period. It shows the trends of PEFR and average expiratory wheeze percentage occupation, which is averaged on 2 to 4 expiratory phases of each recording.

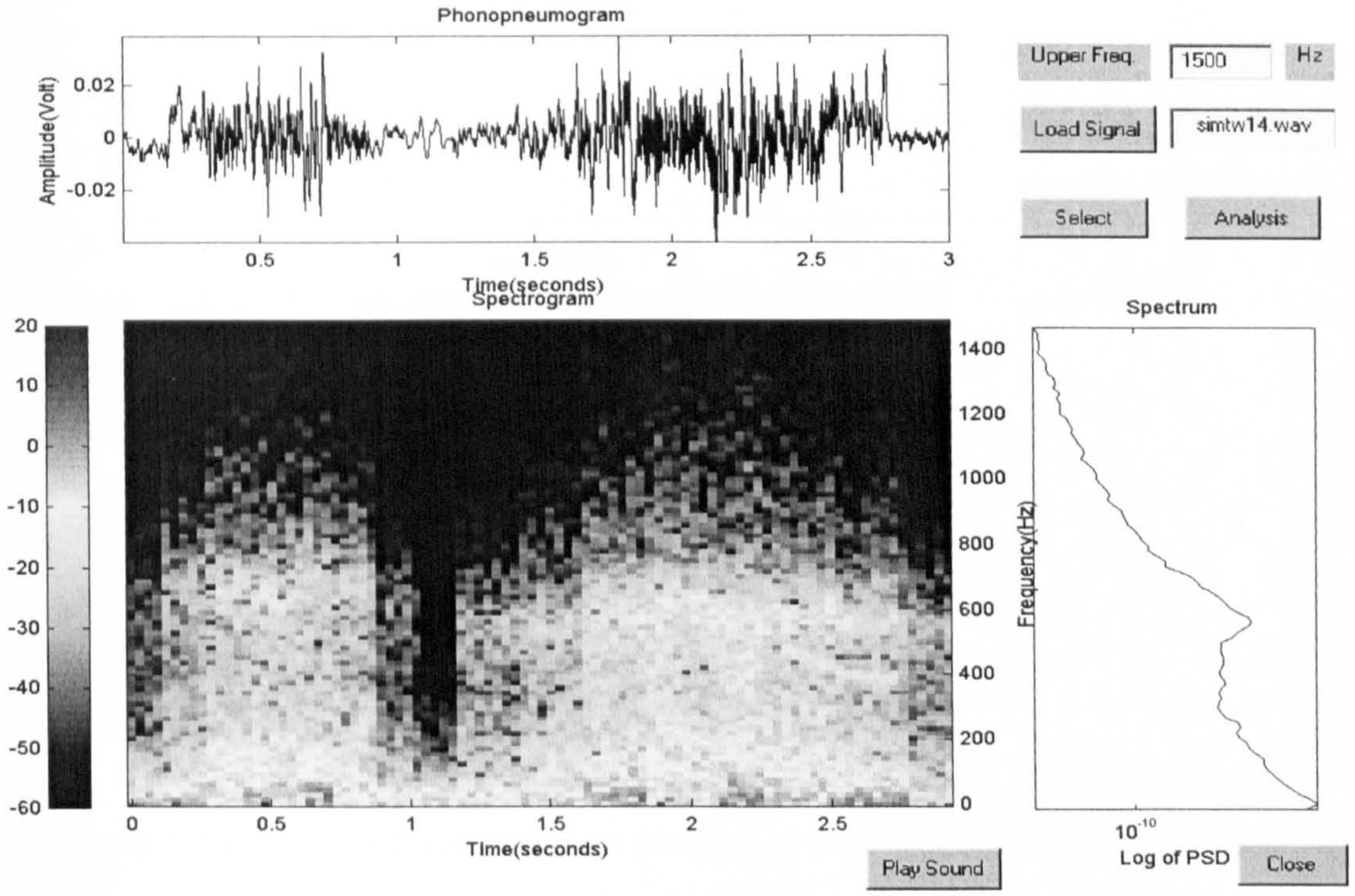

Figure 6.16 Just audible wheezy tracheal sound.

Figure 6.17 Expanded waveform of part of data in Figure 6.16.

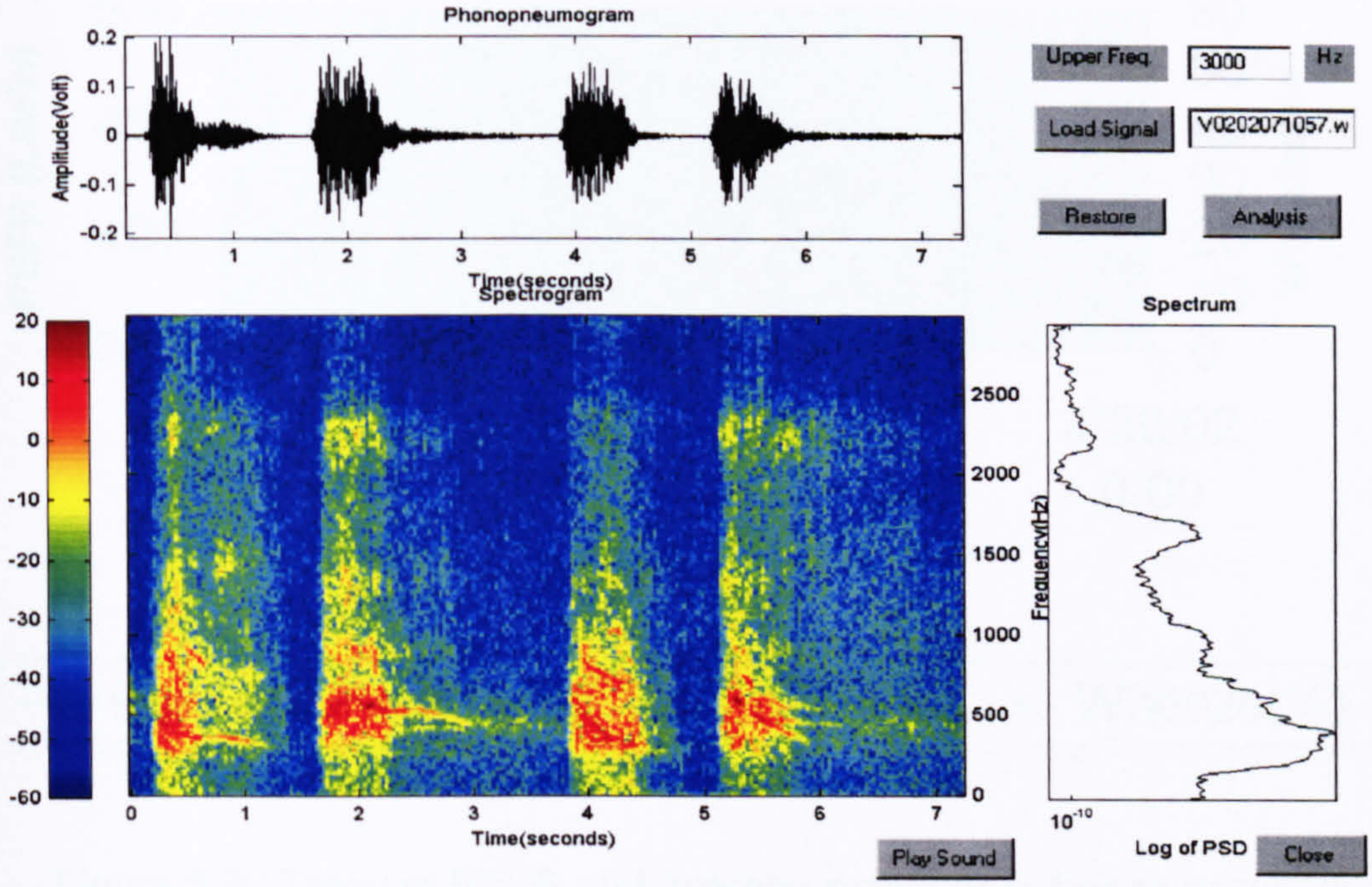

Figure 6.18 Cycle to cycle variation of wheezy tracheal breath sounds captured by a mobile phone.

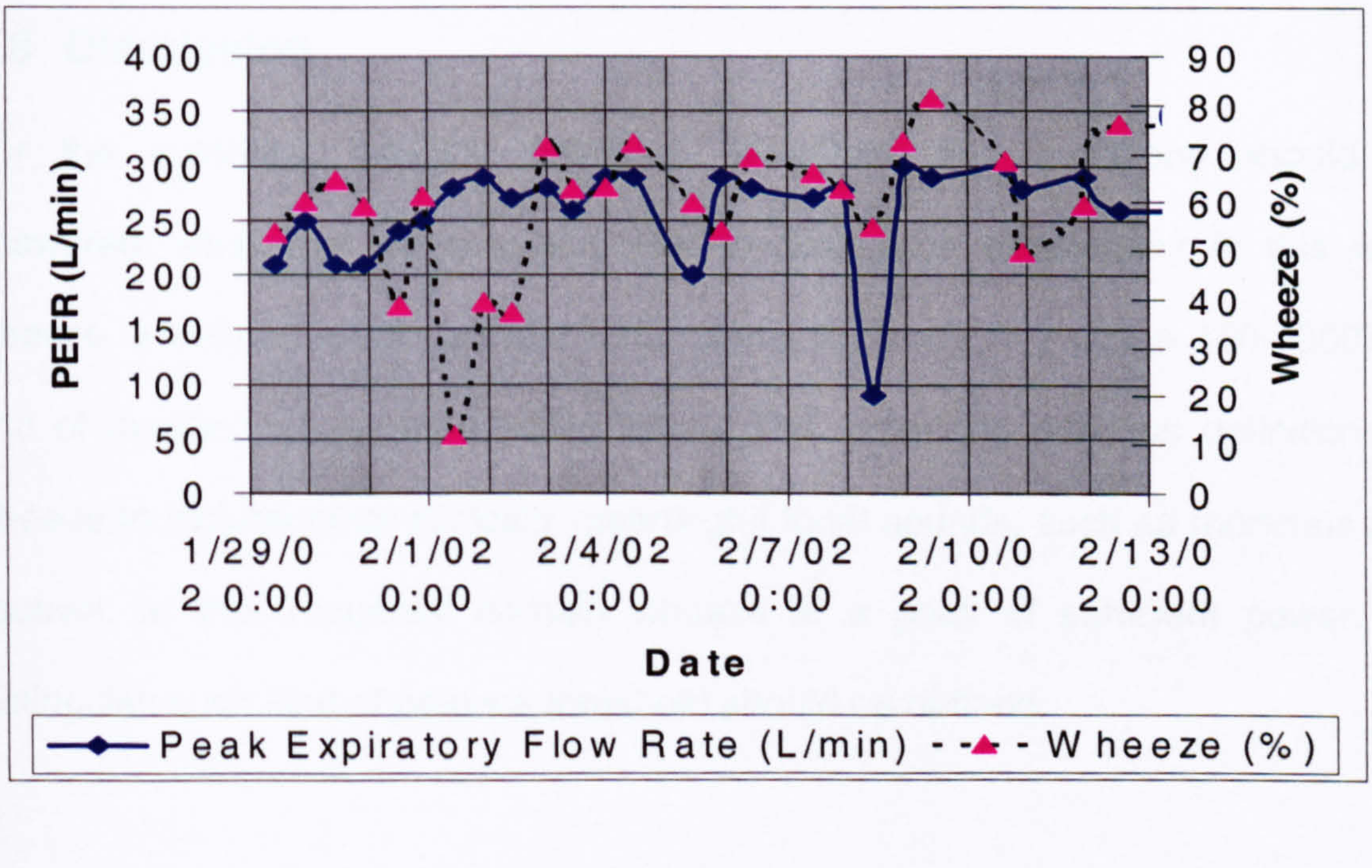

Figure 6.19 Trend of peak flow rate and average expiratory wheeze occupation.

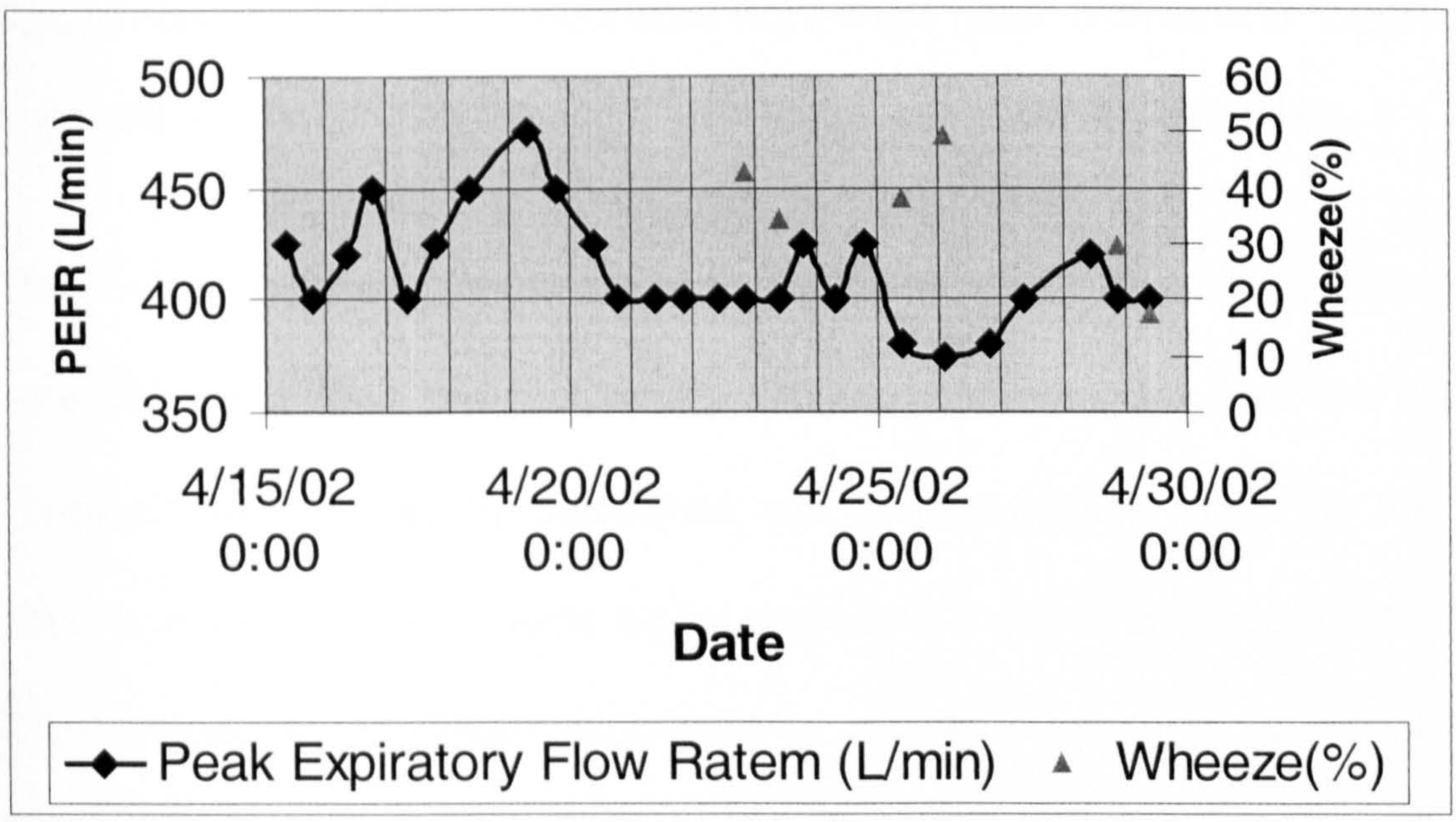

Figure 6.20 Trend of PEFR and average expiratory wheeze occupation.

6.8 Discussion

For the automatic wheeze detection algorithms two questions should be answered: what is a wheeze and how to distinguish a wheeze? In this work wheeze is defined as an audible tonal signal (in frequency range 100-4000Hz), and of duration longer than about 30ms. This broadens previous definitions of wheeze to include other clinically meaningful tonal sounds, such as rhonchus and squawk. In the frequency domain wheeze is a peak of sufficient power. To distinguish such kind of peaks a threshold should be defined.

A peak is found when its value is a local maximum. Using the Fourier transform, a complex signal is decomposed as a sum of sinusoidal signals. Wheezes are considered as signals, which are added to the noise (apparent normal sounds).

Due to noise fluctuations, a peak does not always mean a sinusoidal signal. So a threshold is defined to distinguish a 'true' signal peak from a noise peak.

Threshold was defined by Fenton *et al.* (1985) as $15P_a$ in the 200-1000Hz range for each 100ms signal segment, where P_a is average power between 110-1200Hz. Threshold was defined by Baughman and Loudon (1985) as $3P_b$ in the 150-1000Hz range for each 250ms signal segments , where P_b is baseline power. Charbonneau *et al.* (2000) suggested to define a threshold based on total spectrum energy for a given duration. Threshold definition by Homs-Corbera *et al.* (2000) was much more complicated, and was a modification of the work of Shabtai-Musih *et al.* (1992). First the power spectrum was normalised in every 100Hz band. Then six rules were used to score a peak, which applied 4 empirical constants. When a peak was scored more than 3, it was identified as a wheezing peak.

The threshold definition in this study is based on the results of masking experiments (Reed and Bilger 1973). Thus, the threshold is not a constant but frequency dependent, which is in contrast with the above mentioned threshold definitions. The signal energy is compared with noise energy, which passes through the auditory filter centering the signal. This is also different from other methods, by comparing the signal power with a reference power, which is based on global average power or baseline power.

This algorithm has been validated using simulated data and real data. Results show that this algorithm works successfully and is superior to previous algorithms,

being more reliable in detecting 'true' wheezes with fewer mistakes. Overall, Baughman's method had more false detection; Fenton's algorithm missed some wheezes; while Homs-Corbera's method was the worst of all methods, detecting very few wheezes correctly. One possible explanation is that their thresholds are data dependent.

However, when wheezes are faint, for example just audible, this algorithm failed, as did the other algorithms. So human auditory ability is still better than this algorithm at detecting faint but audible wheezes. However, the human ear has difficulty in quantifying parameters such as wheeze numbers, sequence, frequencies, and occupation. Improvement of this algorithm could be achieved by refining the technique.

Frequency resolution is dependent on the sample length (time duration). That is why the contours shown in Figures 6.11-6.13 look different. The time resolution in this algorithm is 32ms, i.e., the minimum detectable wheeze duration. The window function can also affect the frequency resolution (Harris 1978). According to Baraniuk and Jones (1993), when the window resembles the signal components, excellent time-frequency representation could be achieved. A cosine (Hanning) window was used to calculate the time-frequency representation in this algorithm, meaning the frequency resolution is reasonably good.

There is no standard procedure to validate a wheeze. The common procedures are by listening, by expanded the waveform, or both. For the example shown in Figure 6.9, the wheeze appeared almost throughout the inspiration, which made

the visual validation impossible (refer to Figure 6.14). For the example shown in Figure 6.10 the wheezes were inharmonic polyphones, which also made the visual validation difficult (refer to Figure 6.15).

Wheeze is one of the symptoms of asthma. The presence of wheeze is a cardinal sign of asthma (BTS 2003). Wheezes detected in the asthmatic sound samples include different patterns. Wheezes appeared either in inspiration, expiration, or both. Wheezes were either monophonic or polyphonic; for the latter, either harmonic or not. Cycle-to-cycle as well as day-to-day variations were existent, which phenomena were also found by previous investigators (Kiyokawa *et al.* 1999).

The proportion of signal occupied by wheezing has been reported to relate inversely to FEV_1[1] (Baughman and Loudon 1985; Fenton *et al.* 1985; Pasterkamp *et al.* 1985). The number of patients who were wheezing during the mobile phone tests was not great enough to come to any real conclusion about the equivalence of peak expiratory flow rate (PEFR) measurement and wheeze percentage occupation at this stage. It is assumed the patients breathed spontaneously as asked during all the recordings, so the absence and presence of wheezes were not regarded as being due to breath manoeuvre changes. Because wheezes can be produced during forced expiration even in normal subjects. The presence of wheezes was considered to reflect airway obstruction. Results in Figure 6.20 show that an asthmatic patient had intermittent wheezes during the tests. Wheezes occurred when PEFR was below her average value. So the presence and

---

[1] FEV_1 is defined as the forced expiratory volume in one second.

occupation of wheezes in this instance perhaps reflects variations of airway obstruction (BTS 2003). Results in Figure 6.19 show that an asthmatic patient had persistent wheezes during the tests. Wheeze occupation was not found to correlate with peak flow rate in this patient. A possible explanation for this is due to one problem in mobile phone recording. That is, when the signal is weak the signal will be treated as noise. In the recording of an asthmatic patient who has a prolonged expiratory phase with very weak sound intensity at the end of the expiration, the end part is treated as weak noise. Thus the wheezing components in that part are undetectable as well as inaudible. This effects the results in that the occupation of the wheezes is shorter than that of the actual case. The estimation of respiration duration could also induce some errors. The relationship between wheeze occupation and PEFR is worth of further investigation. A large sample of wheezy patients is required. For self or remote monitoring, measurement quality can be improved by providing proper training to patients on measuring PEFR and recording breath sounds.

In summary, the algorithm developed can detect wheezes automatically when wheezes are not faint. Parametric results have the potential to be integrated in patients' records for management and comparison. Mobile phones may have the potential to be a breath sound monitoring tool.

Chapter 7 Conclusions

7.1 Conclusions

Computer-aided breath sounds analysis is a non-invasive objective method to aid assessment of the respiratory system. This thesis has fulfilled the proposed aims and contributed knowledge in the following aspects.

A general and improved methodology to study a breath sound measurement system is described in chapter 3. The frequency response of each part of the system has been estimated. The frequency response of the air-coupled sensors has been studied in detail. In particular, a method using a laser Doppler vibrometer (LDV) to study a diaphragm type chestpiece (air-coupler) in contact use has been proposed. To the author's knowledge, no such work has been previously reported. It has been shown from the experiments that there are mutual effects of contact. One effect is on the measured signal, that is, the mass loading of the sensor distorts the 'true' signal. The other effect is on the chestpiece property. The initial air volume of the chestpiece in a contact situation is different from that in the free field, thus the transfer function in a contact situation may differ from that of non-contact for a diaphragm type air-coupler. These mutual effects have been isolated, and thus the transfer function for contact use has been obtained for certain types of surface.

The calibrated spectrum of a normal subject's tracheal sounds using estimated contact transfer function compares well with that of the LDV in the frequency range

100-500 Hz. The differences at other frequency ranges may be explained by the mass of the sensor changing the sound pressure beneath it.

Two possible new methods of measuring breath sounds have been investigated. It has been shown that a LDV can be used to measure breath sounds. However, it needs patients to restrict their body movement and the LDV is expensive. For these reasons, the LDV is not likely to be a common choice of transducer for this application. Nevertheless, as the LDV is capable of non-contact measurement, it could be used to monitor a patient's cardiorespiratory system in special situations, such as during anaesthesia.

Some mobile phones can be used as self or remote breath sounds monitoring tools. Others may need modifications to guarantee sound quality. This is because the breath sounds are either treated as a signal or as a background noise. For the latter one, the received signals lose the temporal quality. To avoid the breath sounds being treated as background noises, a simple practical method has been proposed; that is, playing a purposely-designed high frequency narrow-band background noise during the breath sounds measurement. For the mobile phone and telephones tested, this strategy was very successful and breath sounds could be extracted from the wave files. However this method needs to be validated on more mobile phones. For the future it may be possible to make and provide to patients a small portable device that can play the purposely-designed background noise, so that patients play the noise to accompany the recording of the tracheal breath sounds. At the processing side, pre-processing of the received signals by digital filtering can extract the measured breath sounds.

Factors, such as flow rate, pressure and measurement positions, which can influence measurement repeatability, have also been studied. Repeatable measurements can be obtained only when measurement conditions are kept as constant as possible. This implies that for self or remote monitoring of breath sounds, necessary training should be provided to patients so that measured signals are minimally influenced by these factors.

Heart sounds reduction using wavelet decomposition and selective synthesis by hard thresholding has been attempted in chapter 4. It has been shown that heart sounds reduction is achieved at the price of the loss of the breath sounds quality. The heart sounds reduction is necessary if parametric indices are based on a breath sound spectrum. It is not necessary if a time-frequency representation is chosen as heart sounds have a fixed pattern in the time-frequency plane. So if spectrum based indices are chosen, the measurement sites should be selected carefully to avoid strong heart sounds. Wherever possible time-frequency analysis is recommended.

Techniques for simulation of cardiorespiratory sounds and wheeze audibility tests have been developed in chapter 5. A graphical user interface for the simulation has been developed. A variety of breath sounds can be easily obtained by changing adjustable simulation parameters. Simulated breath sounds can be used as a training tool, as well as an evaluation method. Wheeze audibility has been tested. Under certain circumstances, when simulated wheezes are of small amplitude and are audible on their own, they are inaudible when they are embedded in simulated apparent normal breath sounds. This could happen in real

measurements, where wheezes exist but are inaudible. Simulated wheezy sounds with predefined characteristics have been used as an objective method to evaluate the new automatic wheeze detection algorithm.

Chapter 6 describes the successful development of an automatic wheeze detection algorithm based on auditory modelling. The idea is to detect the tonal signals, i.e. wheezes, in a masking noise, i.e. the apparent normal breath sounds. A signal is detectable when the signal's energy is above the masking threshold, which is frequency dependent. The human auditory system is modelled as a bank of band pass filters, in which the bandwidths are also frequency dependent. This algorithm has been validated using simulated and real data. It is superior to previous algorithms, being more reliable in detecting wheezes and less prone to mistakes. However, the performance of the algorithm reduces when the wheezes are faint.

A graphical user interface that incorporates the automatic wheeze detection algorithm has been developed to facilitate operations. Detected wheezes are marked in the spectrogram, while the wheeze parameters, such as mean frequency and percentage occupation of each respiratory phase, can be easily saved in a database to facilitate patient management and monitoring.

Using the new wheeze detection algorithm, the relationship between wheeze and peak expiratory flow rate (PEFR) for long term monitoring has been investigated. This preliminary work showed that in one example of a data set containing imminent wheezes, wheezes only appeared when peak expiratory flow rate

(PEFR) was below the average value. In another example of a data set containing persistent wheezes, wheeze percentage occupation was not found to correlate with PEFR. Due to the lack of sufficient relevant data, no conclusion can be made at this stage. The relationship between wheeze occupation and PEFR is worth of further investigation.

Overall, the work done in this thesis promises a bright future for breath sounds monitoring.

7.2 Future Work

Some improvements can be achieved based on the current work. The contact effects between the diaphragm chestpiece and skin-and-flesh-like material deserve study. Even in-vivo studies could be attempted. It is also worth investigating possible methods of studying the frequency response of diaphragm type air-coupled transducers which cannot be disassembled. The performance of the cardiorespiratory sound simulator needs to be improved if the simulator is expected to be a commercial product. More combinations of adjustable parameters should be provided.

By using the flexibility of mobile phones to capture breath sounds on the measurement side, and using the wheeze detection algorithm to quantify wheezes on the analysis side, it is possible to monitor occupational asthma in the work place. This would be a very useful application, to examine the occupational hazards, and is well worth developing. Another very useful application, which

should be pursued, is evaluation of bronchodilator therapy for asthmatics. This could be particularly useful in monitoring young asthmatics with parents' help.

Wheezes are mainly studied in asthma patients, but other diseases can cause wheezes too. It is worth investigating whether different diseases can be distinguished by wheeze characteristics.

References

Abella, M., Formolo, J., and Penney, D. G., (1992). Comparison of the acoustic properties of six popular stethoscopes, *J Acoust Soc Am*, vol. 91, no. 4 Pt 1, pp. 2224-2228.

Ademovic, E., Pesquet, J. C., and Charbonneau, G., (1998a). Time-scale segmentation of respiratory sounds, *Technol Health Care*, vol. 6, no. 1, pp. 53-63.

Ademovic, E., Pesquet, J. C., and Charbonneau, G., (1998b). Wheezing Lung Sounds Analysis with adaptive local trigonometric transform, *Technol Health Care*, vol. 6, no. 1, pp. 41-51.

al Jarad, N., Strickland, B., Bothamley, G., Lock, S., Logan-Sinclair, R., and Rudd, R. M., (1993). Diagnosis of asbestosis by a time expanded wave form analysis, auscultation and high resolution computed tomography: a comparative study, *Thorax*, vol. 48, no. 4, pp. 347-353.

Anderson, K., Luk, A., McLeod, C., and Moran, F., (1986). The application of pattern-recognition and signal-processing techniques in the diagnosis of asbestosis, *Thorax*, vol. 41, no. 9, p. 715.

Anderson, K., Aitken, S., Carter, R., MacLeod, J. E., and Moran, F., (1990). Variation of breath sound and airway caliber induced by histamine challenge, *Am Rev Respir Dis*, vol. 141, no. 5 Pt 1, pp. 1147-1150.

Anderson, K., Qiu, Y., Whittaker, A. R., and Lucas, M., (2001). Breath sounds, asthma, and the mobile phone, *Lancet*, vol. 358, no. 9290, pp. 1343-1344.

Austrheim, O. and Kraman, S. S., (1985). The effect of low density gas breathing on vesicular lung sounds, *Respir Physiol*, vol. 60, no. 2, pp. 145-155.

Avital, A., Bar-Yishay, E., Springer, C., and Godfrey, S., (1988). Bronchial provocation tests in young children using tracheal auscultation, *J Pediatr*, vol. 112, pp. 591-594.

Bahoura, M., Hubin, M., and Ketata, M., (1998). Respiratory sounds denoising using wavelet packets, in *Proc 2nd Int Conf on Bioelectromagnetism*, IEEE, New York, pp. 11-12.

Banaszak, E. F., Kory, R. C., and Snider, G. L., (1973). Phonopneumography, *Am Rev Respir Dis*, vol. 107, no. 3, pp. 449-455.

Banham, S. W., Urquhart, R. B., MacLeod, J. E., and Moran, F., (1984). Alteration in the low frequency lung sounds in respiratory disorders associated with crackles, *Eur J Respir Dis*, vol. 65, no. 1, pp. 58-63.

Baraniuk, R. G. and Jones, D. L., (1993). Signal-dependent time-frequency analysis using a radially Gaussian kernel, *Sigal Process*, vol. 32, pp. 263-284.

Baughman, R. P. and Loudon, R. G., (1984). Quantitation of wheezing in acute asthma, *Chest*, vol. 86, no. 5, pp. 718-722.

Baughman, R. P. and Loudon, R. G., (1985). Lung sound analysis for continuous evaluation of airflow obstruction in asthma, *Chest*, vol. 88, no. 3, pp. 364-368.

Baughman, R. P. and Loudon, R. G., (1989). Stridor: differentiation from asthma or upper airway noise, *Am Rev Respir Dis*, vol. 139, no. 6, pp. 1407-1409.

Baughman, R. P., Shipley, R. T., Loudon, R. G., and Lower, E. E., (1991). Crackles in interstitial lung disease. Comparison of sarcoidosis and fibrosing alveolitis, *Chest*, vol. 100, no. 1, pp. 96-101.

Beck, R. and Gavriely, N., (1990). The reproducibility of forced expiratory wheezes, *Am Rev Respir Dis*, vol. 141, no. 6, pp. 1418-1422.

Beck, R., Dickson, U., Montgomery, M. D., and Mitchell, I., (1992). Histamine challenge in young children using computerized lung sounds analysis, *Chest*, vol. 102, no. 3, pp. 759-763.

Benedetto, G., Dalmasso, F., Guarene, M. M., Righini, G., and Spagnolo, R., (1983). A method for the acoustical analysis of respiratory crackles in cryptogenic fibrosing alveolitis, *IEEE T Bio-Med Eng*, vol. 30, no. 9, pp. 620-623.

Bergstresser, T., Ofengeim, D., Vyshedskiy, A., Shane, J., and Murphy, R., (2002). Sound transmission in the lung as a function of lung volume, *J Appl Physiol*, vol. 93, no. 2, pp. 667-674.

Bhat, K., Scott, P. H., Eigen, H., and Boggs, P. B., (1985). Evaluation of lung sounds by telephone, *Ann Allergy*, vol. 54, no. 2, pp. 109-111.

Bohadana, A. B., Kopferschmitt-Kubler, M. C., and Pauli, G., (1994a). Breath sound intensity in patients with airway provocation challenge test positive by spirometry but negative for wheezing: a preliminary report, *Respiration*, vol. 61, no. 5, pp. 274-279.

Bohadana, A. B., Massin, N., Teculescu, D., and Peslin, R., (1994b). Tracheal wheezes during methacholine airway challenge (MAC) in workers exposed to occupational hazards, *Respir Med*, vol. 88, no. 8, pp. 581-587.

Bohadana, A. B., Peslin, R., Uffholtz, H., and Pauli, G., (1995). Potential for lung sound monitoring during bronchial provocation testing, *Thorax*, vol. 50, no. 9, pp. 955-961.

Broersen, P. M. T. and de Waele, S., (2000). Detection of methacholine with time series models of lung sounds, *IEEE T Instrum Meas*, vol. 49, no. 3, pp. 517-523.

BTS, (2003). British Guideline on the Management of Asthma, *Thorax*, vol. 58, no. Suppl 1, p. i1-i94.

Bullar, J., (1884). Experiments to determine the origin of the respiratory sounds, *Proc R Soc Lond*, vol. 37, pp. 411-422.

Cabot, R. C. and Dodge, H. F., (1925). Frequency characteristics of heart and lung sounds, *JAMA-J AM MED ASSOC*, vol. 84, pp. 1793-1795.

Cardionics Inc., (2002). PneumoSim Digital Breath Sound Simulator.

Charbonneau, G., Racineux, J. L., Sudraud, M., and Tuchais, E., (1983). An accurate recording system and its use in breath sounds spectral analysis, *J Appl Physiol*, vol. 55, no. 4, pp. 1120-1127.

Charbonneau, G., Sudraud, M., Racineux, J. L., Meslier, N., and Tuchais, E., (1987a). Forced expirations in normal subjects. Is the shape of the flow rate curve related to existence of a wheeze?, *Chest*, vol. 92, no. 5, pp. 825-831.

Charbonneau, G., Sudraud, M., and Soufflet, G., (1987b). Method for the evaluation of flow rate from pulmonary sounds, *Bull Eur Physiopathol Respir*, vol. 23, no. 3, pp. 265-270.

Charbonneau, G., Ademovic, E., Cheetham, B. M., Malmberg, L. P., and Vanderschoot, J., (2000). Basic techniques for respiratory sound analysis, *Eur Respir Rev,* vol. 10, no. 77, pp. 625-635.

Charleston, S. and Azimi-Sadjadi, M. R., (1996). Reduced order Kalman filtering for the enhancement of respiratory sounds, *IEEE T Bio-Med Eng*, vol. 43, no. 4, pp. 421-424.

Charleston, S., Azimi-Sadjadi, M. R., and Gonzalez-Camarena, R., (1997). Interference cancellation in respiratory sounds via a multiresolution joint time-delay and signal-estimation scheme, *IEEE T Bio-Med Eng*, vol. 44, no. 10, pp. 1006-1019.

Chatfield, C., (1989). *The analysis of time series: an introduction*, 4th Edition, Chapamn and Hall.

Cohen, A. and Landsberg, D., (1984). Analysis and automatic classification of breath sounds, *IEEE T Bio-Med Eng*, vol. 31, no. 9, pp. 585-590.

Cohen, A., (1990). Singal processing methods for upper airway and pulmonary dysfunction diagnosis, *IEEE Eng Med Bio*, vol. 9, pp. 72-75.

Cugell, D. W., (1971). Use of tape recordings of respiratory sound and breathing pattern for instruction in pulmonary auscultation, *Am Rev Respir Dis*, vol. 104, no. 6, pp. 948-950.

Cummiskey, J., Williams, T. C., Krumpe, P. E., and Guilleminault, C., (1982). The detection and quantification of sleep apnea by tracheal sound recordings, *Am Rev Respir Dis*, vol. 126, no. 2, pp. 221-224.

Dalmasso, F., Guarene, M. M., Spagnolo, R., Benedetto, G., and Righini, G., (1984). A computer system for timing and acoustical analysis of crackles: a study in cryptogenic fibrosing alveolitis, *Bull Eur Physiopathol Respir*, vol. 20, no. 2, pp. 139-144.

De Oliveira, L. P. L., Roque, W. L., and Custodio, R. F., (1998). Lung sound analysis with time-dependent fractal dimentions, *Chaos, Solitons & Fractals*, vol. 10, no. 9, pp. 1419-1423.

Dekker, E., (1961). Transition between laminar and turbulent flow in human trachea, *J Appl Physiol*, vol. 16, p. 1060-1064.

Doherty, M. J., Spence, D. P., Graham, D., Cheetham, B. M., Sun, X. Q., and Earis, J. E., (1998). A vibrating trachea, *Thorax*, vol. 53, no. 3, pp. 230-231.

Dosani, R. and Kraman, S. S., (1983). Lung sound intensity variability in normal men. A contour phonopneumographic study, *Chest*, vol. 83, no. 4, pp. 628-631.

Druzgalski, C. K., Donnerberg, R. L., and Campbell, R. M., (1980). Techniques of recording respiratory sounds, *J Clin Eng*, vol. 5, pp. 321-330.

East, K. A. and East, T. D., (1985). Computerized acoustic detection of obstructive apnea, *Comput Methods Programs Biomed*, vol. 21, no. 3, pp. 213-220.

Epler, G. R., Carrington, C. B., and Gaensler, E. A., (1978). Crackles (rales) in the interstitial pulmonary diseases, *Chest*, vol. 73, no. 3, pp. 333-339.

Ertel, P. Y., Lawrence, M., Brown, R. K., and Stern, A. M., (1966). Stethoscope acoustics I. The doctor and his stethoscope, *Circulation*, vol. XXXIV, pp. 889-898.

Ertel, P. Y., Lawrence, M., Brown, R. K., and Stern, A. M., (1966). Stethoscope acoustics II. transmission and filtration patterns, *Circulation*, vol. XXXIV, pp. 899-909.

Ertel, P. Y., Lawrence, M., and Song, W., (1969). How to test stethoscopes, *Med Res Eng*, vol. 8, pp. 7-17.

ETSI, (2000). GSM 06.10.

ETSI, (2000). GSM 06.12.

ETSI, (2000). GSM 06.31.

ETSI, (2000). GSM 06.32.

Fahr, G., (1927). Acoustics of bronchial breath sounds, *Arch Intern Med*, vol. 39, pp. 286-302.

Fenton, T. R., Pasterkamp, H., TAL, A., and Chernick, V., (1985). Automated spectral characterization of wheezing in asthmatic children, *IEEE T Bio-Med Eng*, vol. 32, no. 1, pp. 50-55.

Fiz, J. A., Jane, R., Salvatella, D., Izquierdo, J., Lores, L., Caminal, P., and Morera, J., (1999). Analysis of tracheal sounds during forced exhalation in asthma patients and normal subjects: bronchodilator response effect, *Chest*, vol. 116, no. 3, pp. 633-638.

Fiz, J. A., Jane, R., Homs, A., Izquierdo, J., Garcia, M. A., and Morera, J., (2002). Detection of wheezing during maximal forced exhalation in patients with obstructed airways, *Chest*, vol. 122, no. 1, pp. 186-191.

Forgacs, P., (1967). Crackles and wheezes, *Lancet*, vol. 2, no. 7508, pp. 203-205.

Forgacs, P., (1969). Lung sounds, *Br J Dis Chest*, vol. 63, no. 1, pp. 1-12.

Forgacs, P., Nathoo, A. R., and Richardson, H. D., (1971). Breath sounds, *Thorax*, vol. 26, no. 3, pp. 288-295.

Forgacs, P., (1978). *Lung sounds*, Tindall, London.

Fredberg, J. J. and Holford, S. K., (1983). Discrete lung sounds: crackles (rales) as stress-relaxation quadrupoles, *J Acoust Soc Am*, vol. 73, no. 3, pp. 1036-1046.

Gavriely, N., Palti, Y., and Alroy, G., (1981). Spectral characteristics of normal breath sounds, *J Appl Physiol*, vol. 50, no. 2, pp. 307-314.

Gavriely, N., (1983). Generation and transimission of lung sounds. A study in dried,inflated, dog lung model., *Am Rev Respir Dis*, vol. 127 (Suppl), p. 258.

Gavriely, N., (1984). Measurement of tracheal lung sounds, *J Appl Physiol*, vol. 56, no. 3, pp. 817-818.

Gavriely, N., Loring, S. H., and Kelly, K. B., (1984a). Forced expiratory wheezes are a manifestation of flow limitation in normal subjects, *Am Rev Respir Dis*, vol. 129, p. 266.

Gavriely, N., Palti, Y., Alroy, G., and Grotberg, J. B., (1984b). Measurement and theory of wheezing breath sounds, *J Appl Physiol*, vol. 57, no. 2, pp. 481-492.

Gavriely, N., Kelly, K. B., Grotberg, J. B., and Loring, S. H., (1987). Forced expiratory wheezes are a manifestation of airway flow limitation, *J Appl Physiol*, vol. 62, no. 6, pp. 2398-2403.

Gavriely, N., Shee, T. R., Cugell, D. W., and Grotberg, J. B., (1989). Flutter in flow-limited collapsible tubes: a mechanism for generation of wheezes, *J Appl Physiol*, vol. 66, no. 5, pp. 2251-2261.

Gavriely, N. and Herzberg, M., (1992). Parametric representation of normal breath sounds, *J Appl Physiol*, vol. 73, no. 5, pp. 1776-1784.

Gavriely, N. and Cugell, D. W., (1995). *Breath sounds methodology*, Boca Raton : CRC Press.

Gavriely, N., Nissan, M., Rubin, A. H., and Cugell, D. W., (1995). Spectral characteristics of chest wall breath sounds in normal subjects, *Thorax*, vol. 50, no. 12, pp. 1292-1300.

Gavriely, N. and Cugell, D. W., (1996). Airflow effects on amplitude and spectral content of normal breath sounds, *J Appl Physiol*, vol. 80, no. 1, pp. 5-13.

Gross, V., Dittmar, A., Penzel, T., Schuttler, F., and von Wichert, P., (2000). The relationship between normal lung sounds, age, and gender, *Am J Respir Crit Care Med*, vol. 162, no. 3 Pt 1, pp. 905-909.

Grotberg, J. B. and Davis, S. H., (1980). Fluid-dynamic flapping of a collapsed channel:sound generation and flow limitation, *J Biomech*, vol. 13, pp. 219-230.

Grotberg, J. B. and Gavriely, N., (1989). Flutter in collapsible tubes: a theoretical model of wheezes, *J Appl Physiol*, vol. 66, no. 5, pp. 2262-2273.

Hadjileontiadis, L. J. and Panas, S. M., (1997a). Adaptive reduction of heart sounds from lung sounds using fourth-order statistics, *IEEE T Bio-Med Eng*, vol. 44, no. 7, pp. 642-648.

Hadjileontiadis, L. J. and Panas, S. M., (1997b). Higher-order statistics: a robust vehicle for diagnostic assessment and characterisation of lung sounds, *Technol Health Care*, vol. 5, no. 5, pp. 359-374.

Hadjileontiadis, L. J. and Panas, S. M., (1997c). Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter, *IEEE T Bio-Med Eng*, vol. 44, no. 12, pp. 1269-1281.

Hadjileontiadis, L. J. and Panas, S. M., (1998). A wavelet-based reduction of heart sound noise from lung sounds, *Int J Med Inf*, vol. 52, no. 1-3, pp. 183-190.

Hall, L. T., Maple, J. L., Agzarian, J., and Abbott, D., (2000). Sensor system for heart sound biomonitor, *Microelectronics Journal*, vol. 31, no. 7, pp. 583-592.

Hannon, R. R. and Lyman, R. S., (1929). Studies on pulmonary acoustics. The transmisson of tracheal sounds through freshly exenterated sheep's lungs, *Am Rev Tuberc*, vol. 19, pp. 360-375.

Hardin, J. C. and Patterson, J. L., Jr., (1979). Monitoring the state of the human airways by analysis of respiratory sound, *Acta Astronaut.*, vol. 6, no. 9, pp. 1137-1151.

Harper, P., Kraman, S. S., Pasterkamp, H., and Wodicka, G. R., (2001). An acoustic model of the respiratory tract, *IEEE T Bio-Med Eng*, vol. 48, no. 5, pp. 543-550.

Harper, V. P., Pasterkamp, H., Kiyokawa, H., and Wodicka, G. R., (2003). Modeling and measurement of flow effects on tracheal sounds, *IEEE T Bio-Med Eng*, vol. 50, no. 1, pp. 1-10.

Harris, F. J., (1978). On the use of windows for harmonic analysis with the discrete Fourier transform, *Proc IEEE*, vol. 66, no. 1, pp. 51-83.

Hewlett Packard, (1999). Users guide for Stethos electronic stethoscope.

Hetherington, T. E., (2001). Laser vibrometric studies of sound-induced motion of the body walls and lungs of salamanders and lizards: implications for lung-based hearing, *J Comp Physiol [A]*, vol. 187, no. 7, pp. 499-507.

Hidalgo, H. A., Wegmann, M. J., and Waring, W. W., (1991). Frequency spectra of normal breath sounds in childhood, *Chest*, vol. 100, no. 4, pp. 999-1002.

Hiew, Y. H., Smith, J. A., Earis, J. E., Cheetham, B. M., and Woodcock, A. A., (2002). DSP algorithm for cough identification and counting, in *2002 IEEE Int Conf on ASSP*, IEEE, New York, pp. 3888-3891.

Homs-Corbera, A., Jane, R., Fiz, J. A., and Morera, J., (2000). Algorithm for time-frequency detection and analysis of wheezes, in *Proc 22nd Ann Int Conf IEEE EMBS*, IEEE, New York, pp. 2977-2980.

Howard, D. M. and Angus, J., (2001). *Acoustics and psychoacoustics (2nd edition)*, Focal Press.

Howell, W. L. and Aldridge, C. F., (1965). The effect of stethoscope-applied pressure in auscultation. A new instrument for improving discrimination, *Circulation*, vol. 32, no. 3, pp. 430-434.

Ishizaka, K., Matsudair, M., and Kaneko, T., (1976). Input acoustic-impedance measurement of subglotal system, *J Acoust Soc Am*, vol. 60, no. 1, pp. 190-197.

Iyer, V. K., Ramamoorthy, P. A., Fan, H., and Ploysongsang, Y., (1986). Reduction of heart sounds from lung sounds by adaptive filtering, *IEEE T Bio-Med Eng*, vol. 33, no. 12, pp. 1141-1148.

Iyer, V. K., Ramamoorthy, P. A., and Ploysongsang, Y., (1989). Autoregressive modeling of lung sounds: characterization of source and transmission, *IEEE T Bio-Med Eng*, vol. 36, no. 11, pp. 1133-1137.

Jackson, A. C., Giurdanella, C. A., and Dorkin, H. L., (1989). Density dependence of respiratory system impedances between 5 and 320 Hz in humans, *J Appl Physiol*, vol. 67, no. 6, pp. 2323-2330.

Kahya, Y. P., Guler, E. C., and Sankur, B., (1999). Statistical analysis of lung sound data, in *Proc 1st joint BMES/EMBS Conf*, IEEE, New York.

Kanga, J. F. and Kraman, S. S., (1986). Comparison of the lung sound frequency spectra of infants and adults, *Pediatr Pulmonol*, vol. 2, no. 5, pp. 292-295.

Katila, T., Piirila, P., Kallio, K., Paajanen, E., Rosqvist, T., and Sovijarvi, A. R., (1991). Original waveform of lung sound crackles: a case study of the effect of high-pass filtration, *J Appl Physiol*, vol. 71, no. 6, pp. 2173-2177.

Kiyokawa, H., Greenberg, M., Shirota, K., and Pasterkamp, H., (2001). Auditory detection of simulated crackles in breath sounds, *Chest*, vol. 119, no. 6, pp. 1886-1892.

Kiyokawa, H. and Pasterkamp, H., (2002). Volume-dependent variations of regional lung sound, amplitude, and phase, *J Appl Physiol*, vol. 93, no. 3, pp. 1030-1038.

Kompis, M. and Russi, E., (1992). Adaptive heart-noise reduction of lung sounds recorded by a single microphone, in *Proc 14th Annu Int Conf IEEE EMBS*, pp. 691-692.

Kompis, M. and Russi, E. W., (1997). Computer-based lung sound simulation, *Med Biol Eng Comput*, vol. 35, no. 3, pp. 231-238.

Kompis, M., Pasterkamp, H., and Wodicka, G. R., (2001). Acoustic imaging of the human chest, *Chest*, vol. 120, no. 4, pp. 1309-1321.

Kraman, S. S., (1980). Determination of the site of production of respiratory sounds by subtraction phonopneumography, *Am Rev Respir Dis*, vol. 122, no. 2, pp. 303-309.

Kraman, S. S., (1981). Does laryngeal noise contribute to the vesicular lung sound?, *Am Rev Respir Dis*, vol. 124, no. 3, pp. 292-294.

Kraman, S. S., (1983a). Does the vesicular lung sound come only from the lungs?, *Am Rev Respir Dis*, vol. 128, no. 4, pp. 622-626.

Kraman, S. S., (1983b). Speed of low-frequency sound through lungs of normal men, *J Appl Physiol*, vol. 55, no. 6, pp. 1862-1867.

Kraman, S. S. and Austrheim, O., (1983). Comparison of lung sound and transmitted sound amplitude in normal men, *Am Rev Respir Dis*, vol. 128, no. 3, pp. 451-454.

Kraman, S. S., (1984). The relationship between airflow and lung sound amplitude in normal subjects, *Chest*, vol. 86, no. 2, pp. 225-229.

Kraman, S. S., (1985a). New Tools in lung sound research, *Seminars in respiratory medicine*, vol. 6, no. 3, pp. 220-228.

Kraman, S. S., (1985b). Vesicular (Normal) lung sounds: how are they made, where do they come from, and what do they mean?, *Seminars in respiratory medicine*, vol. 6, no. 3, pp. 183-191.

Kraman, S. S., (1986). Effects of lung volume and airflow on the frequency spectrum of vesicular lung sounds, *Respir Physiol* , vol. 66, no. 1, pp. 1-9.

Kraman, S. S., Wodicka, G. R., Oh, Y., and Pasterkamp, H., (1995). Measurement of respiratory acoustic signals. Effect of microphone air cavity width, shape, and venting, *Chest*, vol. 108, no. 4, pp. 1004-1008.

Kraman, S. S., Pasterkamp, H., Kompis, M., Takase, M., and Wodicka, G. R., (1998). Effects of breathing pathways on tracheal sound spectral features, *Respir Physiol*, vol. 111, no. 3, pp. 295-300.

Leblanc, P., Macklem, P. T., and Ross, W. R., (1970). Breath sounds and distribution of pulmonary ventilation, *Am Rev Respir Dis*, vol. 102, no. 1, pp. 10-16.

Lenclud, C., Cuttitta, G., Van Gansbeke, D., Visconti, A., Van Muylem, A., Bellia, V., and Yernault, J. C., (1996). Evaluation of nocturnal bronchoconstriction by all night tracheal sound monitoring, *Thorax*, vol. 51, no. 7, pp. 694-698.

Lessard, C. S. and Wong, W. C., (1986). Correlation of constant flow rate with frequency spectrum of respiratory sounds when measured at the trachea, *IEEE T Bio-Med Eng*, vol. 33, pp. 461-463.

Loudon, R. and Murphy, R. L., (1984). Lung sounds, *Am Rev Respir Dis*, vol. 130, no. 4, pp. 663-673.

Loudon, R. G., (1985). History of lung sounds, *Seminars in respiratory medicine*, vol. 6, no. 3, pp. 157-165.

Mahagnah, M. and Gavriely, N., (1994). Repeatability of measurements of normal lung sounds, *Am J Respir Crit Care Med.*, vol. 149, no. 2 Pt 1, pp. 477-481

Malmberg, L. P., Sovijarvi, A. R., Paajanen, E., Piirila, P., Haahtela, T., and Katila, T., (1994). Changes in frequency spectra of breath sounds during histamine challenge test in adult asthmatics and healthy control subjects, *Chest*, vol. 105, no. 1, pp. 122-131.

Malmberg, L. P., Pesu, L., and Sovijarvi, A. R., (1995). Significant differences in flow standardised breath sound spectra in patients with chronic obstructive pulmonary disease, stable asthma, and healthy lungs, *Thorax*, vol. 50, no. 12, pp. 1285-1291.

Malmberg, L. P., Kallio, K., Haltsonen, S., Katila, T., and Sovijarvi, A. R., (1996). Classification of lung sounds in patients with asthma, emphysema, fibrosing alveolitis and healthy lungs by using self-organizing maps, *Clin Physiol*, vol. 16, no. 2, pp. 115-129.

Marini, J. J., Pierson, D. J., Hudson, L. D., and Lakshminarayan, S., (1979). The significance of wheezing in chronic airflow obstruction, *Am Rev Respir Dis*, vol. 120, no. 5, pp. 1069-1072.

Martini, P. and Meuller, H., (1923). Studien uber das bronchialatmen, *Dtsch Arch F Klin Med*, vol. 143, pp. 159-172.

Mastorocostas, P. A., Tolias, Y. A., Theocharis, J. B., Hadjileontiadis, L. J., and Panas, S. M., (2000). An orthogonal least squares-based fuzzy filter for real-time analysis of lung sounds, *IEEE T Bio-Med Eng*, vol. 47, no. 9, pp. 1165-1176.

McFadden, E. R., Kiser, R., and de Groot, W. J., (1973). Acute bronchial asthma: relations between clinical and physiologic manifestations, *N Engl J Med*, vol. 288, pp. 221-224.

McKusic, V. A., Jenkis, J. T., and Webb, J. N., (1955). The acoustic basis of the chest examination:studies by means of sound spectrography, *Am Rev Turberc*, vol. 72, pp. 12-34.

Meslier, N., Charbonneau, G., and Racineux, J. L., (1995). Wheezes, *Eur Respir J*, vol. 8, pp. 1942-1948.

Moore, B. C. J., (1997). *An introduction to the psychology of hearing (4th edition)*, Academic Press.

Mori, M., Kinoshita, K., Morinari, H., Shiraishi, T., Koike, S., and Murao, S., (1980). Waveform and spectral analysis of crackles, *Thorax*, vol. 35, no. 11, pp. 843-850.

Munakata, M., Homma, Y., Matsuzaki, M., Ogasawara, H., Tanimura, K., Kusaka, H., and Kawakami, Y., (1986). Production mechanism of crackles in excised normal canine lungs, *J Appl Physiol*, vol. 61, no. 3, pp. 1120-1125.

Munakata, M., Ukita, H., Doi, I., Ohtsuka, Y., Masaki, Y., Homma, Y., and Kawakami, Y., (1991). Spectral and waveform characteristics of fine and coarse crackles, *Thorax*, vol. 46, no. 9, pp. 651-657.

Murphy, R. L., Jr., Holford, S. K., and Knowler, W. C., (1977). Visual lung-sound characterization by time-expanded wave-form analysis, *N Engl J Med*, vol. 296, no. 17, pp. 968-971.

Murphy, R. L., Jr., Gaensler, E. A., Holford, S. K., Del Bono, E. A., and Epler, G., (1984). Crackles in the early detection of asbestosis, *Am Rev Respir Dis*, vol. 129, no. 3, pp. 375-379.

Mussell, M. J., Nakazono, Y., and Miyamoto, Y., (1990). Effect of air flow and flow transducer on tracheal breath sounds, *Med Biol Eng Comput*, vol. 28, no. 6, pp. 550-554.

Mussell, M. J., (1992). The need for standards in recording and analysing respiratory sounds, *Med Biol Eng Comput*, vol. 30, no. 2, pp. 129-139.

Nath, A. R. and Capel, L. H., (1974a). Inspiratory crackles and mechanical events of breathing, *Thorax*, vol. 29, pp. 695-698.

Nath, A. R. and Capel, L. H., (1974b). Inspiratory crackles-early and late, *Thorax*, vol. 29, pp. 223-227.

National Instruments, (1998). DAQPad-6200E user manual.

Novak, P. and Novak, V., (1993). Time/frequency mapping of the heart rate, blood pressure and respiratory signals, *Med Biol Eng Comput*, vol. 31, pp. 103-110.

Noviski, N., Cohen, L., Springer, C., Bar-Yishay, E., Avital, A., and Godfrey, S., (1991). Bronchial provocation determined by breath sounds compared with lung function, *Arch Dis Child*, vol. 66, no. 8, pp. 952-955.

O'Donnell, D. M. and Kraman, S. S., (1982). Vesicular lung sound amplitude mapping by automated flow-gated phonopneumography, *J Appl Physiol*, vol. 53, no. 3, pp. 603-609.

Ogden, R. T., (1997). *Essential wavelets for statistical applications and data analysis*, Birkhäuser.

Olson, D. E., Sudlow, M. F., Horsfield, K., and Filley, G. F., (1973). The convective patterns of flow during inspiration in the upper and central airways, *Arch Intern Med*, vol. 131, p. 51-57.

Olson, D. E., Bogyi, M., Schwartz, D. B., and Hammersley, J. R., (1984). Relationship of tracheal breath sounds to airflow, *Am Rev Respir Dis*, vol. 129, p. A256.

Olson, D. E. and Hammersley, R., (1985). Mechanisms of lung sound generation, *Seminars in respiratory medicine*, vol. 6, no. 3, pp. 171-179.

Ono, M., Arakawa, K., Mori, M., Sugimoto, T., and Harashima, H., (1989). Separation of fine crackles from vesicular sounds by a nonlinear digital filter, *IEEE T Bio-Med Eng*, vol. 36, no. 2, pp. 286-291.

Oppenheim, A. V., (1989). *Discrete-time signal processing*, Printice-Hall International Inc..

Oud, M., Dooijes, E. H., and van der Zee, J. S., (2000). Asthmatic airways obstruction assessment based on detailed analysis of respiratory sound spectra, *IEEE T Bio-Med Eng*, vol. 47, no. 11, pp. 1450-1455.

Pasika, H. and Pengelly, D., (1994). Lung sound crackle analysis using generalised time-frequency representations, *Med Biol Eng Comput*, vol. 32, no. 6, pp. 688-690.

Pasterkamp, H., Fenton, R., TAL, A., and Chernick, V., (1984). Tracheal vs lung sounds in acute asthma, *Am Rev Respir Dis*, vol. 129, p. A256.

Pasterkamp, H., TAL, A., Leahy, F., Fenton, R., and Chernick, V., (1985). The effect of anticholinergic treatment on postexertional wheezing in asthma studied by phonopneumography and spirometry, *Am Rev Respir Dis*, vol. 132, no. 1, pp. 16-21.

Pasterkamp, H., Wiebicke, W., and Fenton, R., (1987). Subjective assessment vs computer analysis of wheezing in asthma, *Chest*, vol. 91, no. 3, pp. 376-381.

Pasterkamp, H., Carson, C., Daien, D., and Oh, Y., (1989). Digital respirosonography. New images of lung sounds, *Chest*, vol. 96, no. 6, pp. 1405-1412.

Pasterkamp, H. and Sanchez, I., (1992). Tracheal sounds in upper airway obstruction, *Chest*, vol. 102, no. 3, pp. 963-965.

Pasterkamp, H., Kraman, S. S., DeFrain, P. D., and Wodicka, G. R., (1993). Measurement of respiratory acoustical signals. Comparison of sensors, *Chest*, vol. 104, no. 5, pp. 1518-1525.

Pasterkamp, H., Powell, R. E., and Sanchez, I., (1996a). Lung sound spectra at standardized air flow in normal infants, children, and adults, *Am J Respir Crit Care Med*, vol. 154, no. 2 Pt 1, pp. 424-430.

Pasterkamp, H., Schafer, J., and Wodicka, G. R., (1996b). Posture-dependent change of tracheal sounds at standardized flows in patients with obstructive sleep apnea, *Chest*, vol. 110, no. 6, pp. 1493-1498.

Pasterkamp, H. and Sanchez, I., (1996). Effect of gas density on respiratory sounds, *Am J Respir Crit Care Med.*, vol. 153, no. 3, pp. 1087-1092.

Pasterkamp, H., Kraman, S. S., and Wodicka, G. R., (1997). Respiratory sounds. Advances beyond the stethoscope, *Am J Respir Crit Care Med*, vol. 156, no. 3 Pt 1, pp. 974-987.

Patel, S. B., Callahan, T. F., Callahan, M. G., Jones, J. T., Graber, G. P., Foster, K. S., Glifort, K., and Wodicka, G. R., (1998). An adaptive noise reduction stethoscope for auscultation in high noise environments, *J Acoust Soc Am*, vol. 103, no. 5 Pt 1, pp. 2483-2491.

Pesu, L., Ademovic, E., Pesquet, J. C., and Helisto, P., (1996). Wavelet packet based respiratory sound classification, in *Proc IEEE-SP Int Symp TFTS*, IEEE, New York, pp. 377-380.

Pesu, L., Helisto, P., Ademovic, E., Pesquet, J. C., Saarinen, A., and Sovijarvi, A. R., (1998). Classification of respiratory sounds based on wavelet packet decomposition and learning vector quantization, *Technol Health Care*, vol. 6, no. 1, pp. 65-74.

Piirila, P., Sovijarvi, A. R., Kaisla, T., Rajala, H. M., and Katila, T., (1991). Crackles in patients with fibrosing alveolitis, bronchiectasis, COPD, and heart failure, *Chest*, vol. 99, no. 5, pp. 1076-1083.

Piirila, P., (1992). Changes in crackle characteristics during the clinical course of pneumonia, *Chest*, vol. 102, no. 1, pp. 176-183.

PixSoft Inc. and Medi-wave Inc., (2001). RALE lung sounds.

Plante, F., Kessler, H., Sun, X. Q., Cheetham, B. M., and Earis, J. E., (1998). Inverse filtering applied to upper airway sounds, *Technol Health Care*, vol. 6, no. 1, pp. 23-32.

Ploy-Song-Sang, Y., Martin, R. R., Ross, W. R., Loudon, R. G., and Macklem, P. T., (1977). Breath sounds and regional ventilation, *Am Rev Respir Dis*, vol. 116, no. 2, pp. 187-199.

Ploy-Song-Sang, Y., Macklem, P. T., and Ross, W. R., (1978). Distribution of regional ventilation measured by breath sounds, *Am Rev Respir Dis*, vol. 117, no. 4, pp. 657-664.

Ploy-Song-Sang, Y., Dosman, J., and Macklem, P. T., (1979). Demonstration of regional phase differences in ventilation by breath sounds, *J Appl Physiol*, vol. 46, no. 2, pp. 361-368.

Ploysongsang, Y., (1983). The lung sounds phase angle test for detection of small airway disease, *Respir Physiol*, vol. 53, no. 2, pp. 203-214.

Ploysongsang, Y., Baughman, R. P., Loudon, R. G., and Rashkin, M. C., (1988). Factors influencing the production of wheezes during expiratory maneuvers in normal subjects, *Respiration*, vol. 54, no. 1, pp. 50-60.

Qian, S. and Chen, D., (1996). *Joint time-frequency analysis : methods and applications*, Upper Saddle River, N.J. : PTR Prentice Hall.

Raymond, L. H. and Murphy, Jr., (1985). Discontinuous adventitious lung sounds, *Seminars in respiratory medicine*, vol. 6, no. 3, pp. 210-219.

Reed, C. M. and Bilger, R. C., (1973). A comparative study of S/N_o and E/N_o, *J Acoust Soc Am*, vol. 53, no. 4, pp. 1039-1044.

Rice, D. A., (1980). Sound speed in the upper airways, *J Appl Physiol*, vol. 49, no. 2, pp. 326-336.

Rice, D. A., (1983). Sound speed in pulmonary parenchyma, *J Appl Physiol*, vol. 54, no. 1, pp. 304-308.

Rice, D. A., (1985). Transmission of lung sounds, *Seminars in respiratory medicine*, vol. 6, no. 3, pp. 166-170.

Rietveld, S., Dooijes, E. H., Rijssenbeek-Nouwens, L. H., Smit, F., Prins, P. J., Kolk, A. M., and Everaerd, W. A., (1995). Characteristics of wheeze during histamine-induced airways obstruction in children with asthma, *Thorax*, vol. 50, no. 2, pp. 143-148.

Rietveld, S. and Dooijes, E. H., (1996). Characteristics and diagnostic significance of wheezes during exercise- induced airway obstruction in children with asthma, *Chest*, vol. 110, no. 3, pp. 624-631.

Rietveld, S., Oud, M., and Dooijes, E. H., (1999). Classification of asthmatic breath sounds: preliminary results of the classifying capacity of human examiners versus artificial neural networks, *Comput Biomed Res*, vol. 32, no. 5, pp. 440-448.

Rossi, M., Sovijarvi, A. R. A., Piirila, P., Vannuccini, L., Dalmasso, F., and Vanderschoot, J., (2000). Environmental and subject conditions and breathing manoeuvres for respiratory sound recordings, *Eur Respir Rev*, vol. 10, pp. 611-615.

Sanchez, I. and Pasterkamp, H., (1993). Tracheal sound spectra depend on body height, *Am Rev Respir Dis*, vol. 148, no. 4 Pt 1, pp. 1083-1087.

Sankur, B., Kahya, Y. P., Guler, E. C., and Engin, T., (1994). Comparison of AR-based algorithms for respiratory sounds classification, *Comput Biol Med*, vol. 24, no. 1, pp. 67-76.

Sankur, B., Cagatay, G. E., and Kahya, Y. P., (1996). Multiresolution biological transient extraction applied to respiratory crackles, *Comput Biol Med*, vol. 26, no. 1, pp. 25-39.

Scanlon, M. V., (1999). Acoustically monitor physiology during sleep and activity, in *Proc 1st Joint BMES/EMBS Conf*, IEEE, New York, p. 787.

Schreur, H. J., Sterk, P. J., Vanderschoot, J., van Klink, H. C., van Vollenhoven, E., and Dijkman, J. H., (1992). Lung sound intensity in patients with emphysema and in normal subjects at standardised airflows, *Thorax*, vol. 47, no. 9, pp. 674-679.

Schreur, H. J., Vanderschoot, J., Zwinderman, A. H., Dijkman, J. H., and Sterk, P. J., (1994). Abnormal lung sounds in patients with asthma during episodes with normal lung function, *Chest*, vol. 106, no. 1, pp. 91-99.

Seul, M., O'Gorman, L., and Sammon, M. J., (2000). *Practical algorithms for image analysis : description, examples, and code* , Cambridge University Press.

Shabtai-Musih, Y., Grotberg, J. B., and Gavriely, N., (1992). Spectral content of forced expiratory wheezes during air, He, and SF6 breathing in normal humans, *J Appl Physiol*, vol. 72, no. 2, pp. 629-635.

Shirai, F., Kudoh, S., Shibuya, A., Sada, K., and Mikami, R., (1981). Crackles in asbestos workers: auscultation and lung sound analysis, *Br J Dis Chest*, vol. 75, no. 4, pp. 386-396.

Shykoff, B. E., Ploysongsang, Y., and Chang, H. K., (1988). Airflow and normal lung sounds, *Am Rev Respir Dis*, vol. 137, no. 4, pp. 872-876.

Soufflet, G., Charbonneau, G., Polit, M., Attal, P., Denjean, A., Escourrou, P., and Gaultier, C., (1990). Interaction between tracheal sound and flow rate: a comparison of some different flow evaluations from lung sounds, *IEEE T Bio-Med Eng*, vol. 37, no. 4, pp. 384-391.

Sovijarvi, A. R., Malmberg, L. P., Paajanen, E., Piirila, P., Kallio, K., and Katila, T., (1996). Averaged and time-gated spectral analysis of respiratory sounds. Repeatability of spectral parameters in healthy men and in patients with fibrosing alveolitis, *Chest*, vol. 109, no. 5, pp. 1283-1290.

Sovijarvi, A., Malmberg, L. P., Charbonneau, G., Vanderschoot, J., Dalmasso, F., Sacco, C., Rossi, M., and Earis, J. E., (2000a). Characteristics of breath sounds and adventitious respiratory sounds, *Eur J Respir Dis*, vol. 10, pp. 591-596.

Sovijarvi, A. R. A., Dalmasso, F., Vanderschoot, J., Malmberg, L. P., and Righini, G., (2000b). Definition of terms for application of respiratory sounds, *Eur Respir Rev*, vol. 10, no. 77, pp. 597-610.

Spence, D. P., Bentley, S., Evans, D. H., and Morgan, M. D., (1992). Effect of methacholine induced bronchoconstriction on the spectral characteristics of breath sounds in asthma, *Thorax*, vol. 47, no. 9, pp. 680-683.

Spence, D. P., Graham, D. R., Jamieson, G., Cheetham, B. M., Calverley, P. M., and Earis, J. E., (1996). The relationship between wheezing and lung mechanics during methacholine-induced bronchoconstriction in asthmatic subjects, *Am J Respir Crit Care Med*, vol. 154, no. 2 Pt 1, pp. 290-294.

Sprikkelman, A. B., Grol, M. H., Lourens, M. S., Gerritsen, J., Heymans, H. S., and van Aalderen, W. M., (1996). Use of tracheal auscultation for the assessment of bronchial responsiveness in asthmatic children, *Thorax*, vol. 51, no. 3, pp. 317-319.

Springer, C., Godfrey, S., Picard, E., Uwyyed, K., Rotschild, M., Hananya, S., Noviski, N., and Avital, A., (2000). Efficacy and safety of methacholine bronchial challenge performed by auscultation in young asthmatic children, *Am J Respir Crit Care Med*, vol. 162, no. 3 Pt 1, pp. 857-860.

Sun, X. Q., Cheetham, B. M., Evans, K. G., and Earis, J. E., (1998). Estimation of analogue pre-filtering characteristics for CORSA standardisation, *Technol Health Care*, vol. 6, no. 4, pp. 275-283.

Suzuki, A., Sumi, C., Nakayama, K., and Mori, M., (1995). Real-time adaptive cancelling of ambient noise in lung sound measurement, *Med Biol Eng Comput*, vol. 33, no. 5, pp. 704-708.

Thacker, R. E. and Kraman, S. S., (1982). The prevalence of auscultatory crackles in subjects without lung disease, *Chest*, vol. 81, no. 6, pp. 672-674.

The Mathworks, I., (1999). *Signal processing toolbox user's guide*.

Tolias, Y. A., Hadjileontiadis, L. J., and Panas, S. M., (1997). A fuzzy rule-based system for real-time separation of crackles from vesicular sounds, in *Proc 19th Int Conf IEEE EMBS*, IEEE, New York, pp. 1115-1118.

Tolias, Y. A., Hadjileontiadis, L. J., and Panas, S. M., (1998). Real-time separation of discontinuous adventitious sounds from vesicular sounds using a fuzzy rule-based filter, *IEEE T Inf Technol B*, vol. 2, no. 3, pp. 204-215.

Tran, T., Jones, N. B., and Fothergill, J. C., (1995). Heart sound simulator, *Med Biol.Eng Comput*, vol. 33, pp. 357-359.

Turner, J. D. and Pretlove, A. J., (1991). *Acoustics for engineers*. Basingstoke: Macmillan .

Turtle Beach Systems, (1995). Turtle beach systems TBS-2000 User's guide.

Urquhart, R. B., McGhee, J., MacLeod, J. E., Banham, S. W., and Moran, F., (1981). The diagnostic value of pulmonary sounds: a preliminary study by computer-aided analysis, *Comput Biol Med*, vol. 11, no. 3, pp. 129-139.

Urquhart, R. B., (1983). *Some new techniques for pattern recognition research and lung sound signal analysis,* PhD, University of Glasgow.

Vanderschoot, J. and Schreur, H. J., (1991). Flow and volume related AR-modelling of lung sounds, in *Proc 13th Ann Int Con IEEE EMBS*, IEEE, New York, pp. 385-386.

Vanderschoot, J., Klein, R. C., and Schreur, H. J., (1992). Polynomial regression of lung sound AR parameters on flow and volume, in *Proc 14th Ann Int Con IEEE EMBS*, IEEE, pp. 2578-2579

Vanderschoot, J. and Schreur, H. J., (1994). AR (q, v) modeling of normal lung sounds, *Methods Inf Med*, vol. 33, no. 1, pp. 24-27.

Vanderschoot, J., Helisto, P., Lipponen, P., Piirila, P., and Sovijarvi, A. R., (1998). Distribution of crackles on the flow-volume plane in different pulmonary diseases, *Technol Health Care*, vol. 6, no. 1, pp. 81-89.

Vannuccini, L., Rossi, M., and Pasquali, G., (1998). A new method to detect crackles in respiratory sounds, *Technol Health Care*, vol. 6, no. 1, pp. 75-79.

Vetterli, M. and Herley, C., (1992). Wavelets and filter banks:theory and design, *IEEE T Sigal Proces*, vol. 40, no. 9, pp. 2207-2232.

Vidakovic, B., (1999). *Statistical modelling by wavelets*, John WIley & Sons, Inc.

Vovk , I. V., Zaluskii , K. E., and Krasnyi , L. G., (1994). Acoustic model of the human respiratory system, *Acoustical Physics*, vol. 40, no. 5, pp. 676-680.

Vovk, I. V., Grinchenko, V. T., Krasnyi, L. G., and Makarenkov, A. P., (1994). breath sounds:recording and classification problems, *Acoustical Physics*, vol. 40, no. 1, pp. 43-48.

Vovk, I. V., Grinchenko, V. T., and Oleinik, V. N., (1995). Modelling the acoustic properties of the chest and measuring breath sounds, *Acoustical Physics*, vol. 41, no. 5, pp. 667-676.

Waitman, L. R., Clarkson, K. P., Barwise, J. A., and King, P. H., (2000). Representation and classification of breath sounds recorded in an intensive care setting using neural networks, *Journal of Clinical Monitoring and Computing*, vol. 16, no. 2, pp. 95-105.

Waring, W. M., Beckerman, R. C., and Hopkins, R. L., (1985). Continous adventitios lung sounds: site and method of production and significance, *Seminars in respiratory medicine*, vol. 6, no. 3, pp. 201-209.

Waris, M., Helisto, P., Haltsonen, S., Saarinen, A., and Sovijarvi, A. R., (1998). A new method for automatic wheeze detection, *Technol Health Care*, vol. 6, no. 1, pp. 33-40.

Weiss, E. B. and Carlson, C. J., (1972). Recording of breath sounds, *Am Rev Respir Dis*, vol. 105, no. 5, pp. 835-839.

Weitz, H. H. and Mangione, S., (2000). In defense of the stethoscope and the bedside [editorial; comment], *Am J Med*, vol. 108, no. 8, pp. 669-671.

Welch, P. D., (1967). The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, *IEEE T Audio Electro*, vol. 15, no. 2, pp. 70-73.

West, J. B. and Hugh-Jones, P., (1959). Patterns of gas flow in the upper bronchial tree, *J Appl Physiol*, vol. 14, pp. 753-759.

Whittaker, A. R., Lucas, M., Carter, R., and Anderson, K., (2000). Limitations in the use of median frequency for lung sound analysis, *P I Mech Eng H*, vol. 214, no. 3, pp. 265-275.

Wodicka, G. R., Stevens, K. N., Golub, H. L., Cravalho, E. G., and Shannon, D. C., (1989). A model of acoustic transmission in the respiratory system, *IEEE T Bio-Med Eng*, vol. 36, no. 9, pp. 925-934.

Wodicka, G. R. and Shannon, D. C., (1990). Transfer function of sound transmission in subglottal human respiratory system at low frequencies, *J Appl Physiol*, vol. 69, no. 6, pp. 2126-2130.

Wodicka, G. R., Kraman, S. S., Zenk, G. M., and Pasterkamp, H., (1994). Measurement of respiratory acoustic signals. Effect of microphone air cavity depth, *Chest*, vol. 106, no. 4, pp. 1140-1144.

Woodward, B., Istepanian, R. S. H., and Richards, C. I., (2001). Design of a telemedicine system using a mobile telephone, *IEEE T Inf Technol B*, vol. 5, no. 1, pp. 13-15.

Workum, P., Holford, S. K., DelBono, E. A., and Murphy, R. L., (1982). The prevalence and character of crackles (rales) in young women without significant lung disease, *Am Rev Respir Dis*, vol. 126, no. 5, pp. 921-923.

Workum, P., DelBono, E. A., Holford, S. K., and Murphy, R. L., Jr., (1986). Observer agreement, chest auscultation, and crackles in asbestos- exposed workers, *Chest*, vol. 89, no. 1, pp. 27-29.

Yonemaru, M., Kikuchi, K., Mori, M., Kawai, A., Abe, T., Kawashiro, T., Ishihara, T., and Yokoyama, T., (1993). Detection of tracheal stenosis by frequency analysis of tracheal sounds, *J Appl Physiol*, vol. 75, no. 2, pp. 605-612.

Yong, S. C., Smith, C. M., Wach, R., Kurian, M., and Primhak, R. A., (1999). Methacholine challenge in preschool children: methacholine-induced wheeze versus transcutaneous oximetry, *Eur Respir J*, vol. 14, no. 5, pp. 1175-1178.

Yost, W. A., (2000). *Fundamentals of hearing-an introduction (4th edition)*, Academic Press.

Young, M., Sparrow, D., Gottlieb, D., Selim, A., and Friedman, R., (2001). A telephone-linked computer system for COPD care, *Chest*, vol. 119, no. 5, pp. 1565-1575.

Appendix A

A.1 Frequency Characteristics of Data Acquisition Device

The DAQPad-6020E is a USB-compatible multifunction analogue, digital, and timing I/O device for USB-compatible computers (National Instruments 1998). This product features a 12-bit ADC with eight differential/16 single-ended channels, which function was used here.

A.1.1 Materials and methods

Manually swept sinusoidal signals with amplitude of 1v from an analogue signal generator (Servomex) were input to one of the DAQPad's channels. The frequency range of the swept Input signals was from 1Hz to 2000Hz (due to signal generator limit). The step was 1Hz from 1 to 10Hz, 10Hz from 10 to 100Hz, and 100Hz from 100 to 2000Hz.

During the tests an oscilloscope (Trio CS-1352) was used to monitor the signals. An existing Labview program was used to sample the signals; sampling rate and gain were selectable. The data were written in text files.

The signals were sampled at 11025Hz. Each signal was recorded for about 5 seconds and the middle section from 1 sec to 4 sec was used for the calculation. DC-offset was removed from data before calculation. The magnitude response at each frequency was calculated by the ratio of the root-mean-squared (RMS) value of measured data to the RMS of input data on a dB scale.

Then 100mv, 10mv and 2mv signals at frequencies 1000Hz, 100Hz, 10Hz and 1Hz were tested under suitable software selectable gains of 20, 100 and 100 respectively.

A.1.2 Results

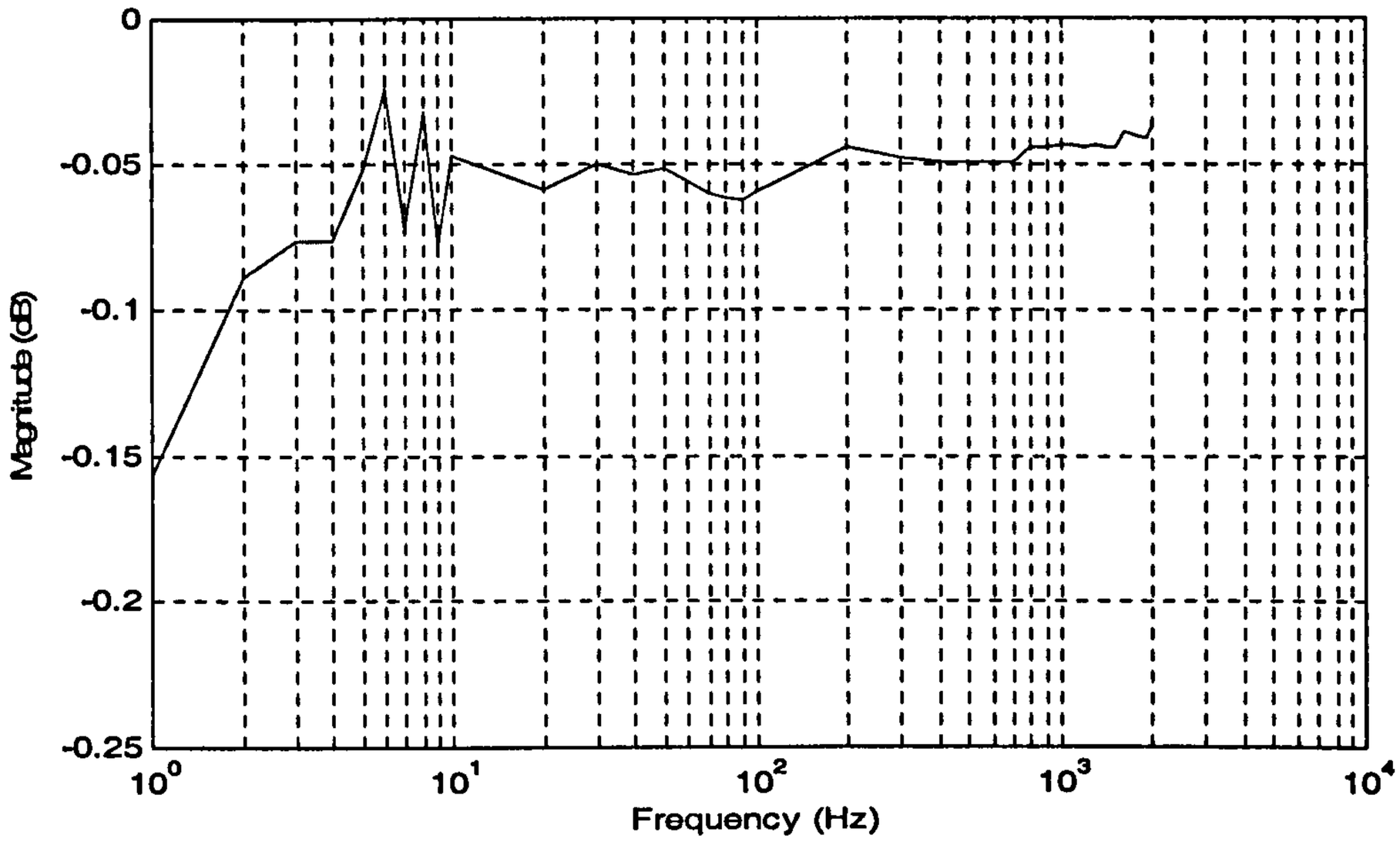

Figure A.1 Magnitude response of DAQPad-6200E

It can be seen from figure A.1 that the DAQPad 6200E had a flat gain from 1Hz to 2000Hz with ± 0.1dB. Figure A.2 shows that at different software selectable gains the measurements are linear, which implies the device-integrated amplifier is linear.

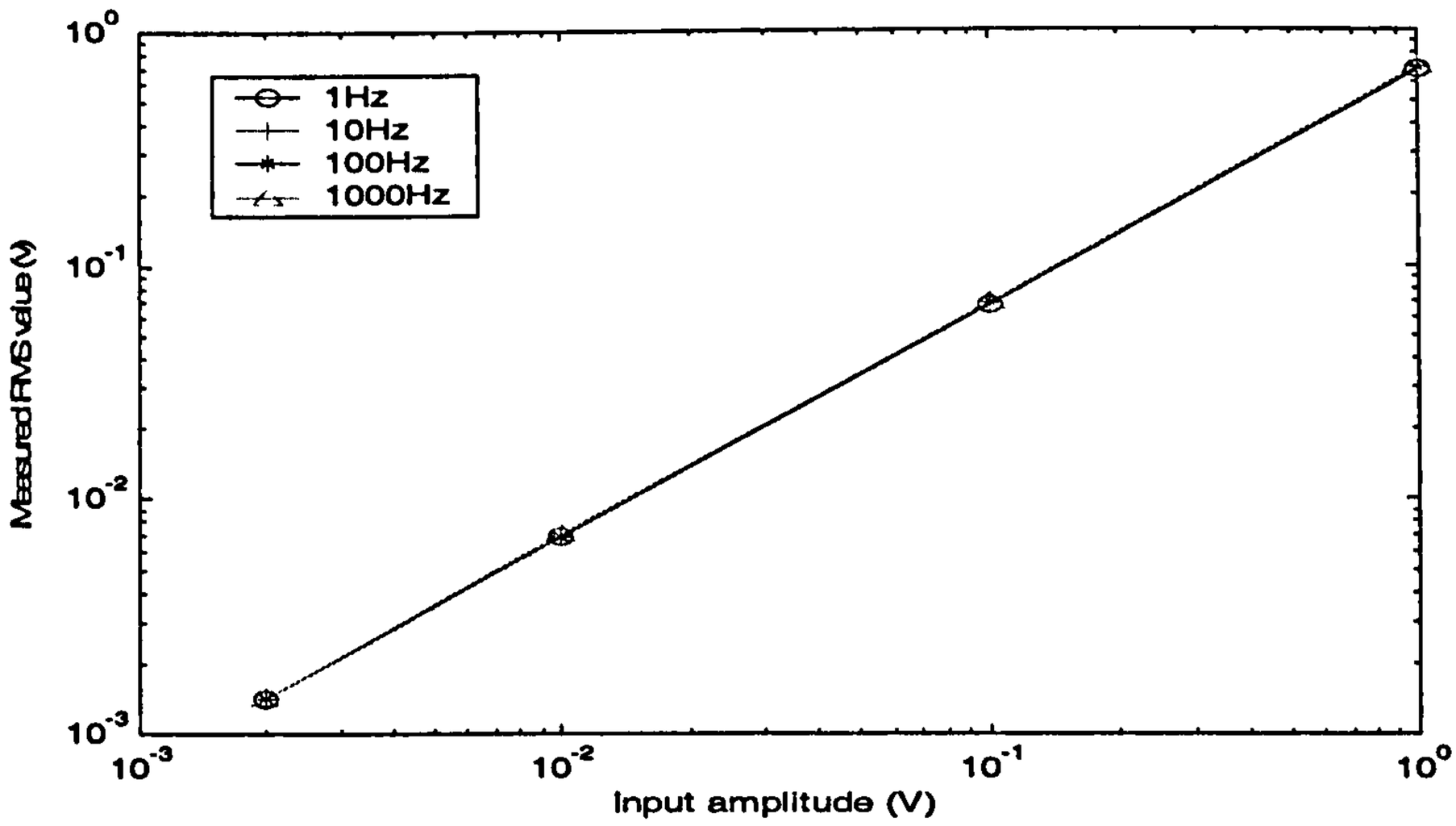

Figure A.2 Linear relationship between input and output signals.

A.2 Sound Card as an Audio Output

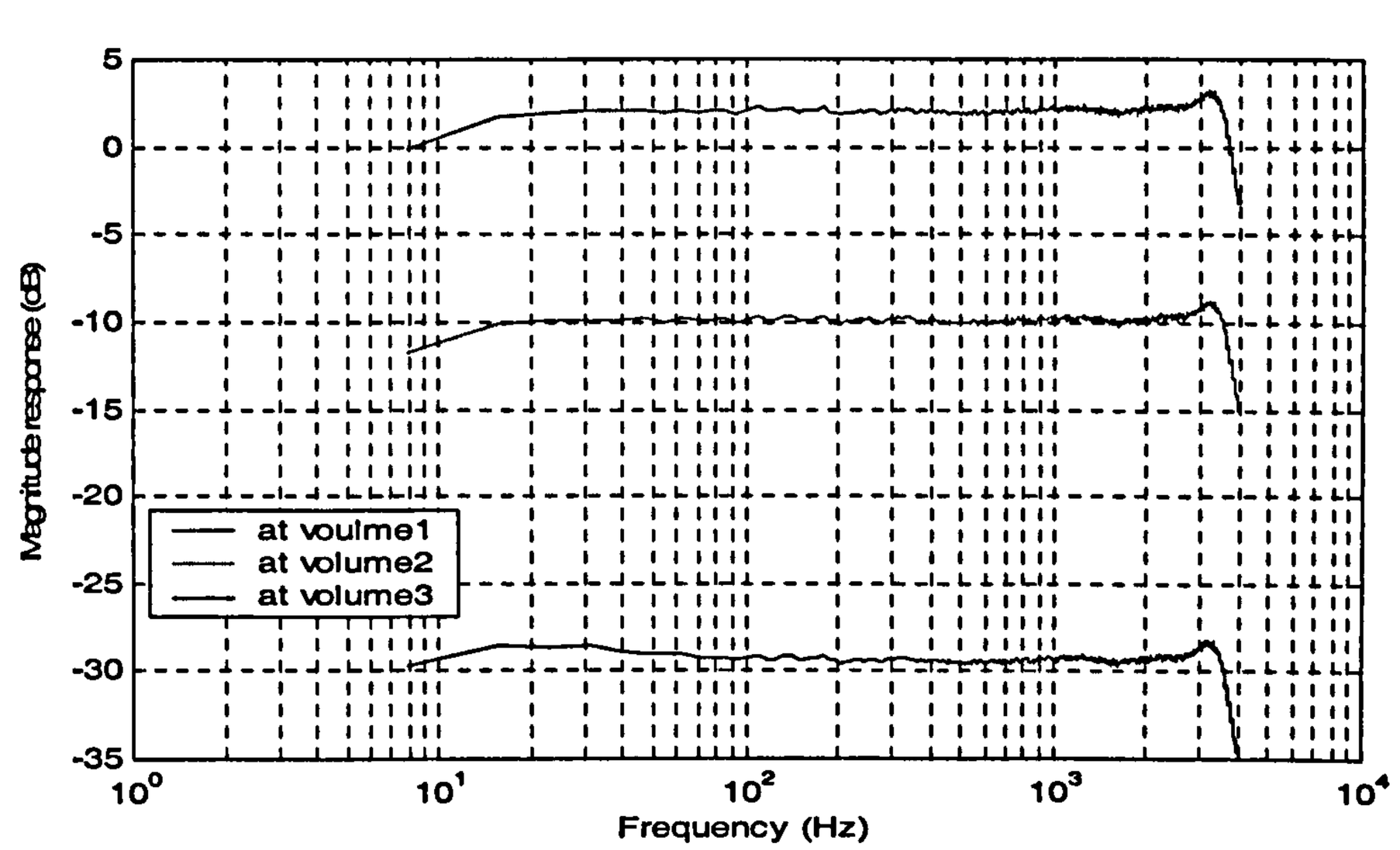

Figure A. 3 Magnitude response of TBS-2000 playback.

Figure A.3 shows that the TBS-2000 has a flat gain from20Hz to Nyquist frequency with ±3dB. The volume control acts like a linear amplifier.

A.3 Frequency Characteristics of Filters

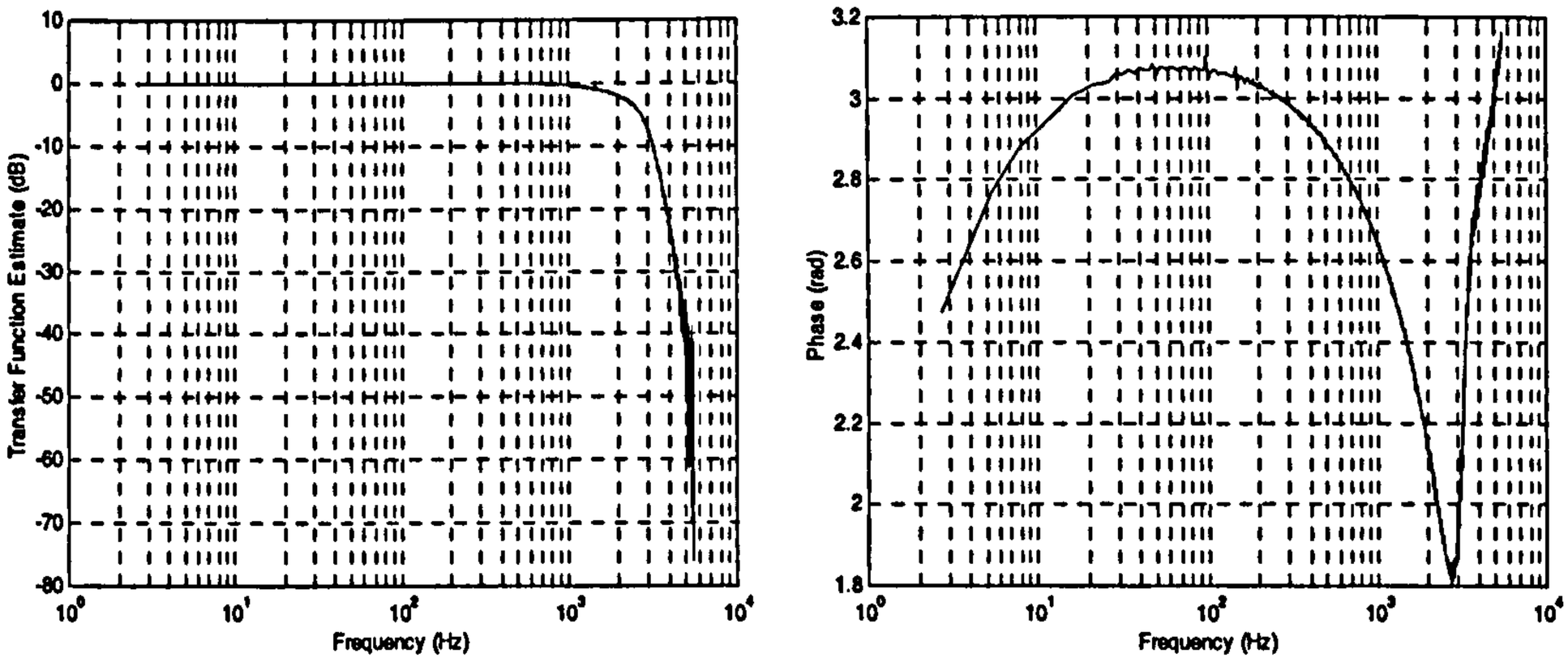

Figure A. 4 Frequency response of low pass filter with cut-off frequency 3kHz.

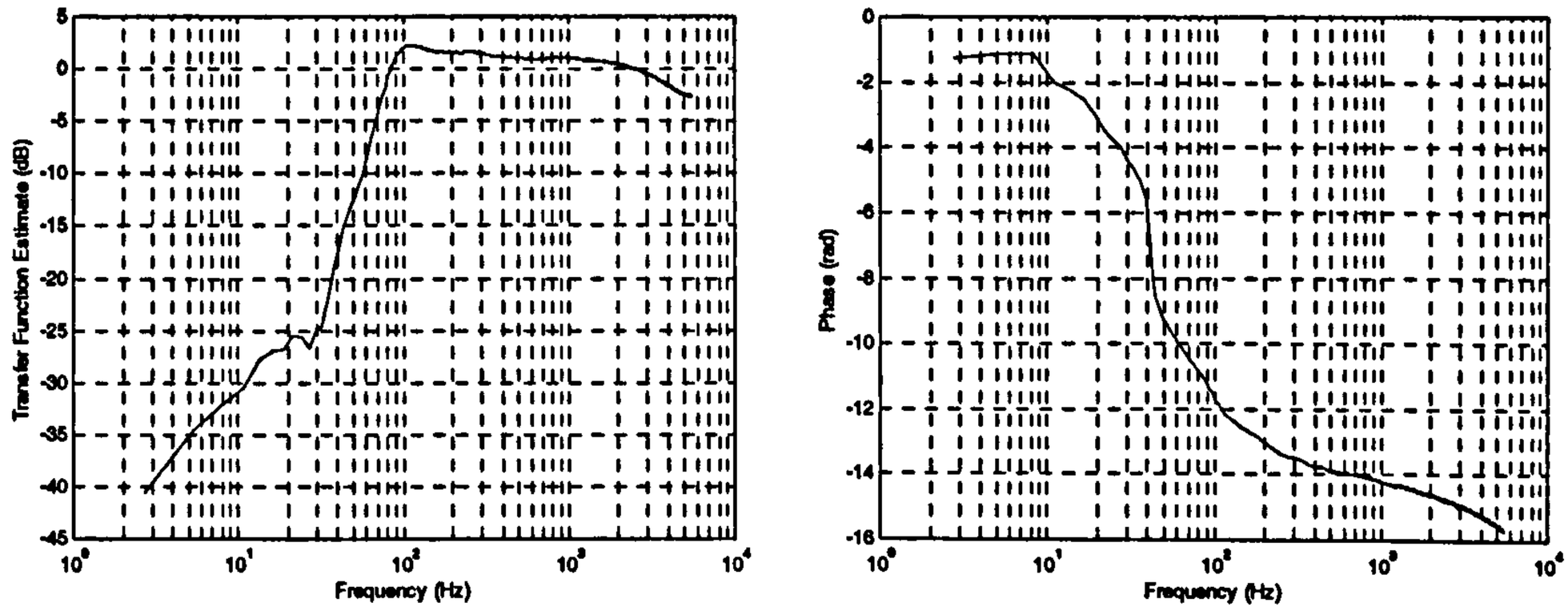

Figure A. 5 Frequency response of high pass filter with cut-off frequency 100Hz.

An example of low pass filter with cut-off frequency 3000Hz is shown in the figure A.4. An example of high pass filter with cut-off frequency 100Hz is shown in figure A.5. In pass band the each filter has an almost flat gain, but phase response is not linear.

A.4 linearity Test Between Vibration Speed and Sound Card Volume

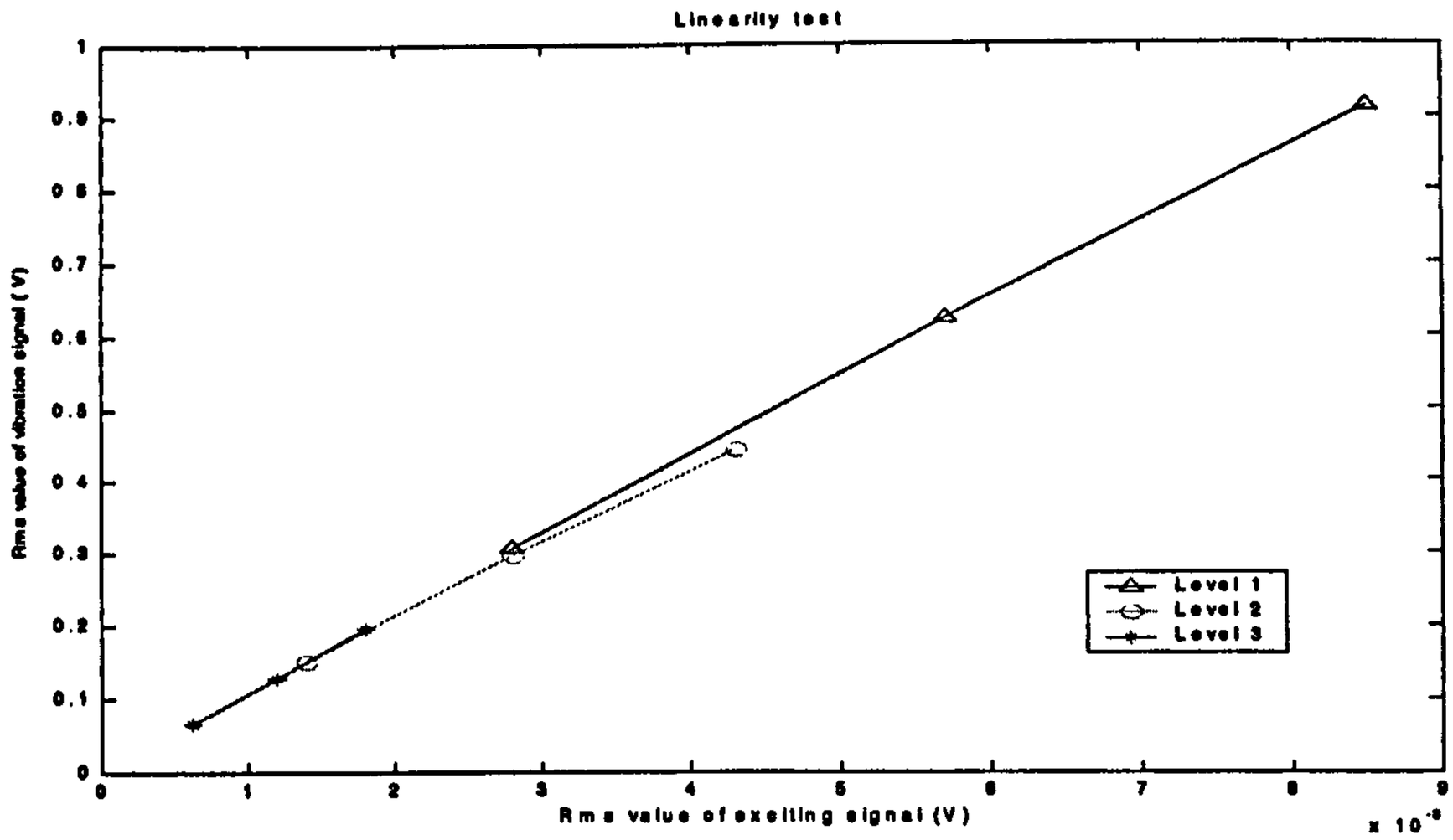

Figure A. 6 Relationship between exciting signal and vibration signal.

A.5 Transfer functions of the Escope

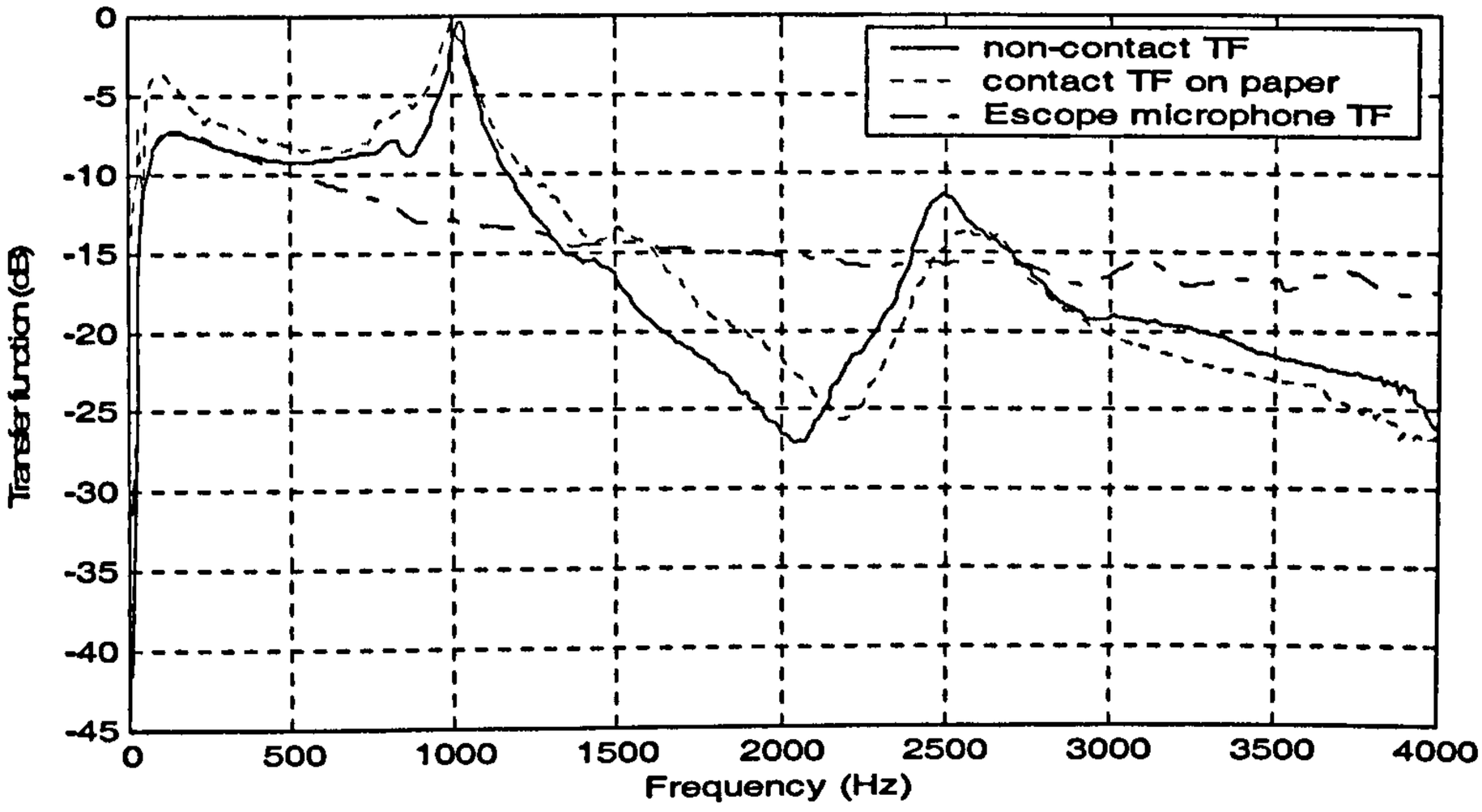

Figure A. 7 Escope non-contact TF; contact TF on paper and its microphone TF.

Figure A.7 shows that the Escope's contact TF with paper was almost the same as that of non-contact. This indicates that the paper was 'acousticlly transparent'.

A.6 LDV Measurements of Centre Velocities

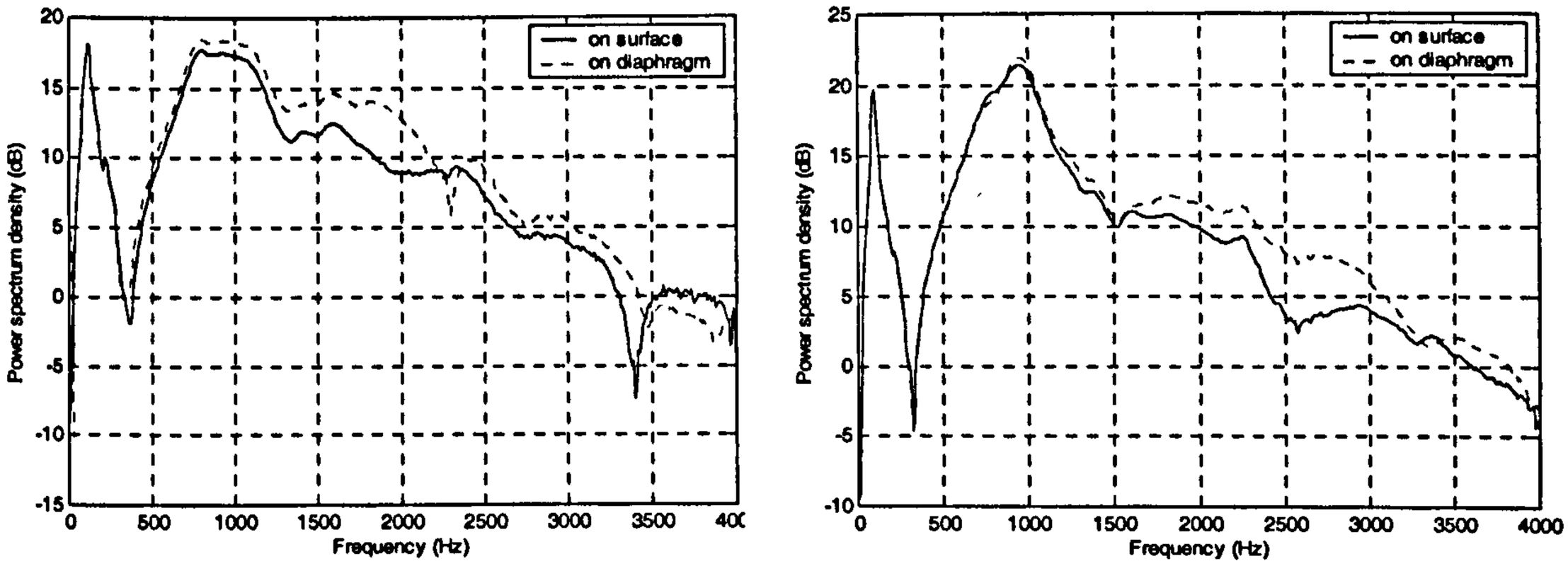

Figure A. 8 PSDs of velocity on surface centre and diaphragm centre respectively. Left-on thick surface; right-on thin surface.

Figure A.8 shows that the diaphragm could almost follow the vibration of the surface which it is contacted.

A.7 Repeatability tests

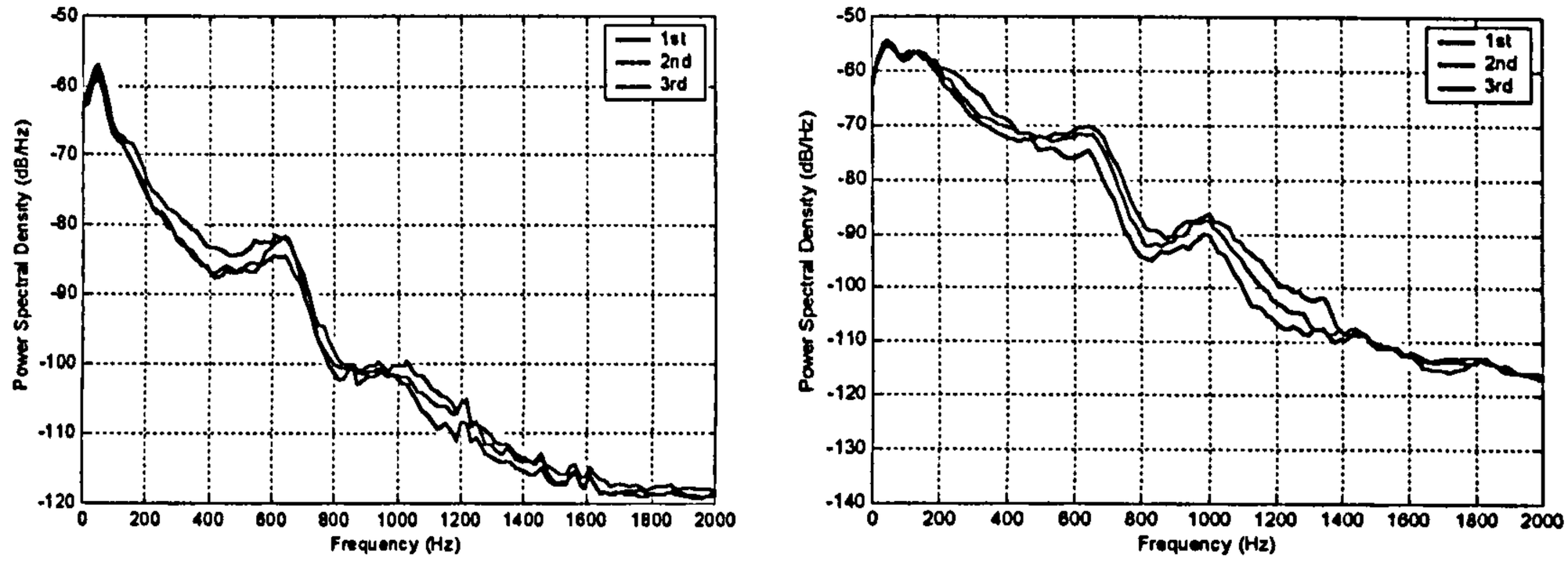

Figure A. 9 Repeatability tests of effects of flow. Left-quiet breath; right-deep breath.

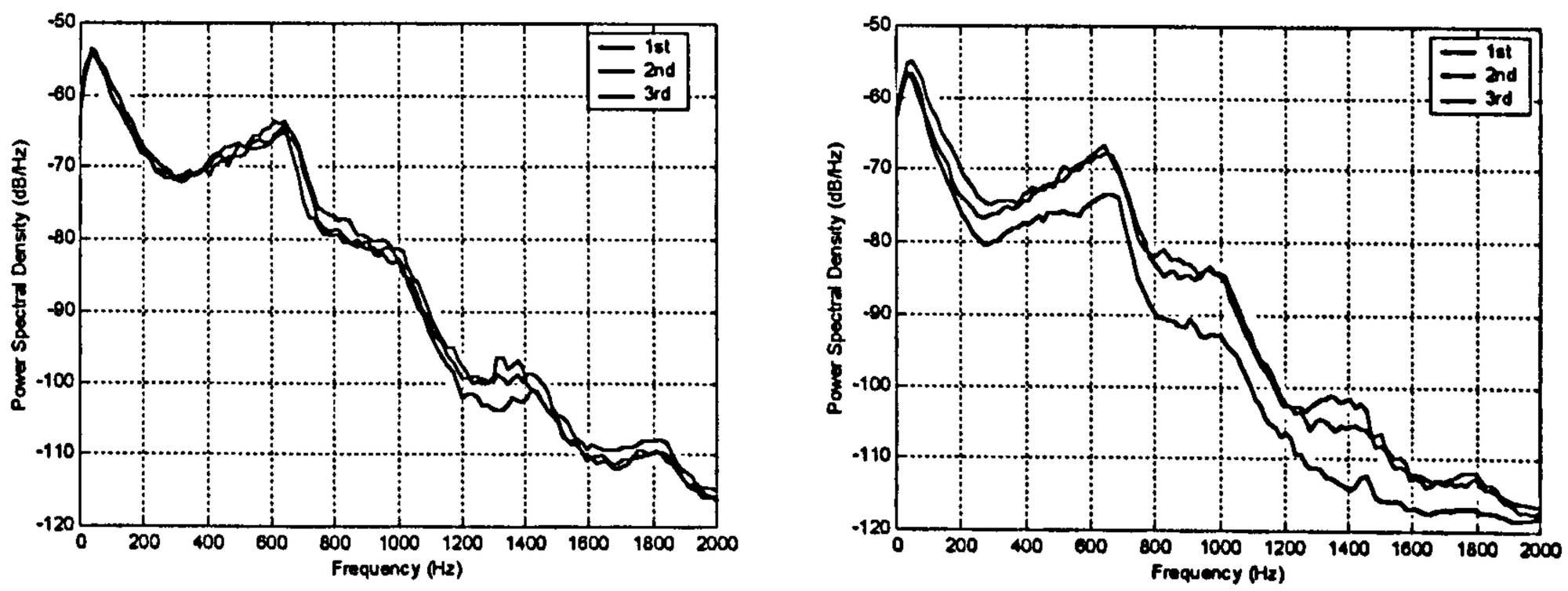

Figure A.10 Repeatability tests of effects of pressure. Left-heavy pressure; right-middle pressure.

Figures A.9 and A.10 show that under same conditions, the measurements are repeatable.

A.8 Two weeks monitoring using mobile phones

Table A.1 Two weeks monitoring using patients' own mobile phones

Subjects No.	Measure-ments	Mention peak flow value	Heard wheezes	Detected wheezes?	Sound quality (clear?)
1	30	yes (missed once)			yes
2	12	yes (second time)			audible; not normal
2	26	no (first time)			audible; not normal
3	27	yes(missed once)	yes	yes	yes
4	28	yes (missed twice)			audible; not normal
5	9	yes			audible; not normal
6	10	yes (missed twice)			audible; not normal
7	20	yes			yes
8	4	yes (missed once)			audible; not normal
9	26	yes			yes
10	26	yes	yes	yes	yes
11	20	yes (missed five)			ripples
12	28	yes	yes	yes	yes
13	23	yes (missed twice)			nothing
14	26	yes (missed four)	yes (once in one cycle)	yes(once in one cycle)	yes
15	32	yes			yes
16	8	no			yes

Appendix B

1. Auditory modeling based automatic wheeze detection program in Matlab

```
%wezedetect.m
%using threshold according to pyscoacoustic principles
%plot spectrogram with wheeze marked & wheeze contour
%threshold using 3.4766*log10(f0)-0.5--regression from Reed's result
%estimated inspiration and expiration duration
%calculate wheeze frequency(mean) and duration(percent)

clear;
f=input('Which .wav file should be processed?--','s');
[xin,Fs,Bits]=wavread(f);
X=xin(:,1); %first channel data
X=detrend(X); %remove DC component

nfft=512;
olaps=256;%about 32ms
[B,F,T]=specgram(X,nfft,Fs,hanning(nfft),olaps); %get STFT of X
% plot 0 to approximately 2kHz
n1=round(0*nfft/Fs)+1;%n1 for start frequency
hl=100;%low frequency range, changable; mobile frequency 300~3400Hz
hf=3400;%high frequency range,changable
n2=round(hf*nfft/Fs); %n2 for end frequency;
C=abs(B(n1:n2,:)).^2; %power
mineng=min(min(C))/10;% minimum energy
meaneng=10*log10(mean(mean(C))); %mean energy in dB

k=length(T); %columns of T and C
l=length(F(n1:n2)); %rows of F and C
D(l,k)=0;%initial D and E, D for audible peaks
E(l,k)=0;% E for peaks
G(l,k)=0;%G for labeled wheezes
H=C; %H for connected wheezes

%estimate inspiration and expiration time duration
%base on sum power <300Hz
m=fix(300/F(2));
pbase=sum(C); %sum of power
pref=5*min(pbase); %reference

i=1;
while i<k & pbase(i)<=pref
   i=i+1;
end
tin1=i; %start point of inspiration
while i<k & pbase(i)>pref
   i=i+1;
end
tin2=i; % end point of inspiration
tinl=(tin2-tin1+1)*T(2);%length of inspiration

while i<k & pbase(i)<=pref
```

```
   i=i+1;
end
tout1=i; %start point of expiration
while i<k & pbase(i)>pref
   i=i+1;
end
tout2=i; %end point of expiration
tout1=(tout2-tout1+1)*T(2);%length of expiration


npeaks=5; %how many peaks, can be changed
[loc,val]=picksharppeak(C,npeaks,2); %find sharp peaks along coloum
%loc puts frequency location, and val puts peak value

for i=1:k %along coloums
   I=find(loc(:,i)>0); %rows
   E(loc(I,i),i)=1; %binary values, 1 for peaks and 0 for others
end


%audible peak detection
for i=1:k  %how many columns
   for j=1:npeaks
      P1=loc(j,i); %position in a coloum
      F0=(P1-1)*F(2); %tonal frequency; centre frequency;
                      %F(2)is frequency resolution
      A1=0; %initial
      if F0>=hl & F0<hf % peaks between hl to hf;
      A1=val(j,i); %signal power
      ERB=24.7*(4.37*F0/1000+1);% Critical bandwidth in Hz
      num=ceil(ERB/F(2)); %how many noise band power to be count
      if mod(num,2)==1 %odd
         nl=round(num/2); %how many bands left to the central
                          %frequency, more effect on masking
         nr=fix(num/2);%how many bands right to central frequency,
                       %less masking effect
      else
         nl=num/2;
         nr=nl;
      end

%threshold of (Energy of signal)/(Energy of noise)in dB
%A2 for band power, bandwidth ERB Hz
      A2=0;%noise power
      if P1-nl>1 %within index>=1
         for nn=P1-nl:P1-1
            A2=A2+C(nn,i);
      end
   else %form index 1
      for nn=1:P1-1
         A2=A2+C(nn,i);
      end
   end

   if P1+nr<1
      for nn=P1+1:P1+nr
         A2=A2+C(nn,i);
```

```
    end
else
    for nn=P1+1:l
        A2=A2+C(nn,i);
    end
end

A1=10*log10(A1);%in dB;
A2=10*log10(A2);%in dB;

tmp=A1-A2;
          %threshold for 100ms
          if tmp>=3.4776*log10(F0)+9.5-10*log10(ERB)
             D(P1,i)=1; %label for value above thresholded
             %power over threshold
             ED(P1,i)=tmp-3.4776*log10(F0)-9.5+10*log10(ERB);
          end
       end
    end
end


%labelling
%add timing factor to elimiate short tones less than 32ms

[wf,wt]=find(D==1); %wf for wheeze frequency location and wt for time value
lwt=length(wt);

tag=1; %to label wheeze sections
if ~isempty (wf)
   G(wf(1),wt(1))=tag;
   for i=2:lwt
   %CONTINESOU IN TIME AND FREQUENCY
     if wt(i)>1 & wf(i)>=2 & wf(i)+2<=l % k for time coloums;
                                        %l for frequency columns
        switch D(wf(i),wt(i))
        case D(wf(i),wt(i)-1) %has same frequency with left cell
           label=G(wf(i),wt(i)-1);%get the label
           G(wf(i),wt(i))=label; %assigne the label to neighbour cell
        case D(wf(i)+1,wt(i)-1)%left cell has higher frequency
           label=G(wf(i)+1,wt(i)-1);
           G(wf(i),wt(i))=label;
        case D(wf(i)-1,wt(i)-1)%left cell has lower frequency
           label=G(wf(i)-1,wt(i)-1);
           G(wf(i),wt(i))=label;
        otherwise %no peaks at left neighbours
           tag=tag+1;
           G(wf(i),wt(i))=tag;
        end
     else
        tag=tag+1;
        G(wf(i),wt(i))=tag;
     end
   end
end
```

```
nw=max(max(G)); %number of wheezes sections
for i=1:nw
    [wf,wt]=find(G==i);
    JT(i)=wt(1); %JT for start time points
    KT(i)=wt(length(wt));%KT for end time points
    JF(i)=wf(1);
    KF(i)=wf(length(wf));

    if KT(i)-JT(i)==0 %wheezes last less than 32ms ,discard
       %no 4.8dB higher power than threshold of 100ms
       if  ED(JF(i),JT(i))<4.8
          G(wf,wt)=0;% reset to 0
          JT(i)=0;
          KT(i)=0;
          JF(i)=0;%JF for start frequency points
          KF(i)=0;%KF for end frequency points
       end
    elseif KT(i)-JT(i)==1 %wheezes last less than 70ms
       if ED(JF(i),JT(i))<3%no 3dB higher power than threshold
          G(wf,wt)=0;% reset to 0
          JT(i)=0;
          KT(i)=0;
          JF(i)=0;%JF for start frequency points
          KF(i)=0;%KF for end frequency points
       end
    end
 end
%wf for wheeze frequency location and wt for time value
for i=1:nw %connect same wheezes-right side
  f1=KF(i);
  t1=KT(i);
  if t1>=1 & f1>1
     while t1<k & f1<l & G(f1,t1) > G(f1,t1+1) ...
           & G(f1,t1) > G(f1-1,t1+1) & G(f1,t1) > G(f1+1,t1+1)
        if E(f1,t1+1)==1%a peak
           G(f1,t1+1)=G(f1,t1);
           t1=t1+1;
        elseif E(f1-1,t1+1)==1
           G(f1-1,t1+1)=G(f1,t1);
           f1=f1-1;
           t1=t1+1;
        elseif E(f1+1,t1+1)==1
           G(f1+1,t1+1)=G(f1,t1);
           f1=f1+1;
           t1=t1+1;
        else
           break;
        end %end if
     end %end while
   end%end if
end%end for


for i=1:nw %connect same wheezes-left side
   f1=JF(i);
   t1=JT(i);
```

```
while t1>1 & f1>2 & f1+1<1 & G(f1,t1) > G(f1,t1-1)
   if E(f1,t1-1)==1%a peak, same frequency
      G(f1,t1-1)=G(f1,t1);
      t1=t1-1;
   elseif E(f1-1,t1-1)==1 %lower frequency
      G(f1-1,t1-1)=G(f1,t1);
      f1=f1-1;
      t1=t1-1;
   elseif E(f1+1,t1-1)==1 %higher frequency
      G(f1+1,t1-1)=G(f1,t1);
      f1=f1+1;
      t1=t1-1;
     else
      break;
     end %end if
    end %end while
end %end for


%for spectrogram display, mark the wheezes
[wf,wt]=find(G>=1);
lwt=length(wt);
for i=1:lwt
   H(wf(i),wt(i))=mineng;
end

%parameters calculation
j=1;
AF=[];
for i=1:nw
   [wf,wt]=find(G==i);
   if ~isempty(wf) %real wheezes
      AF(j)=(mean(wf)-1)*F(2);%mean frequency
      TS(j)=wt(1);
      TE(j)=wt(length(wt));
      AT(j)=(TE(j)-TS(j)+1)*T(2); %duration of time
      if TS(j)>tout1
         OC(j)=AT(j)*100/toutl; %expiration wheeze ocupation percent
      else
         OC(j)=AT(j)*100/tinl;%inspiration percentage
      end
     j=j+1;
   end
end

tmp=0;
for i=tin1:tin2
   if max(G(:,i))>1
      tmp=tmp+1;
   end
end
toci=tmp*T(2)*100/tinl%total inspiration ocupation

tmp=0;
for i=tout1:tout2
   if max(G(:,i))>1
      tmp=tmp+1;
```

```
    end
end
toce=tmp*T(2)*100/toutl % total expiration ocupation

if ~isempty(AF)
    [SF,SI]=sort(AF); %sort frequency-from low to high
    ST=AT(SI);%coresponding time to SF
    SOC=OC(SI);
end

%spectrogram of peaks
figure(1);
imagesc(T,F(n1:n2),E),axis xy;colorbar

%contour
GG=G;
[I,J]=find(G>1);
l=length(I);
for i=1:l
    GG(I(i),J(i))=1;
end

figure(2);
contour(T,F(n1:n2),GG);
colormap(gray);
h=findobj('type','patch');
set(h,'Linewidth',2); %increase line width to 2 points
axis([0 k*T(2) 0 2000]);
xlabel('Time(sec)');
ylabel('Frequency (Hz)');


%spectrogram with wheezes marked
figure(3);
imagesc(T,F(n1:n2),10*log10(H+eps)),axis xy;colorbar;
```

2. Other automatic wheeze detection programs in Matlab

```
%othermethod.m-using other thresholds
%1-Fenton's method
%2-Baughman's method
%3-Homs-Corbera's

function othermethod(method)

error(nargchk(1,1,nargin));

if method>3
   error('input number should be between 1 to 3');
end

s=input('Please input file name--','s');
[x,fs]=wavread(s);

switch method
   case 1
      %using Fenton's threshold
      nfft=fix(0.1*fs);%100ms
      [B,F,T]=specgram(x,nfft,fs,hanning(nfft),0);
      n1=fix(110/F(2));
      n2=fix(1200/F(2));
      B=abs(B).^2;
      C=B(n1:n2,:); %power spectrum in 110 to 1200 Hz
      a=15;
      mineng=min(min(C))/10;
      D=C; %initial D

      [loc,val]=pickpeak(C,5,3);
      [l,m]=size(C); %how many rows & columns
      G(l,m)=0; %binray matix to keep peaks

      for i=1:m
         p=mean(C(:,i));%mena power at a given time
         for j=1:5
            if loc(j,i)>200/F(2) &val(j,i)>a*p
               D(loc(j,i),i)=mineng;
               G(loc(j,i),i)=1;
            end
         end
      end
figure(1);
imagesc(T,F(n1:n2),10*log10(C+eps)),axis xy;colorbar;
figure(2);
imagesc(T,F(n1:n2),10*log10(D+eps)),axis xy;colorbar;

figure(3);
colormap(gray);
contour(T,F(n1:n2),G);
h=findobj('type','patch');
set(h,'Linewidth',2); %increase line width to 2 points
axis([0 m*T(2) 0 2000]);
xlabel('Time(sec)');
```

```
ylabel('Frequency (Hz)');

case 2
   %Baughman's threshold
   nfft=fix(0.25*fs);%250ms

   [B,F,T]=specgram(x,nfft,fs,hanning(nfft),fix(nfft/2));
   n1=fix(150/F(2));
   n2=fix(2000/F(2));
   B=abs(B).^2;
   C=B(n1:n2,:); %power spectrum in 150 to 2000 Hz
   a=3;
   mineng=min(min(C))/10;
   D=C; %initial D

   [loc,val]=pickpeak(C,3,5);
   [l,m]=size(C); %how many rows & columns
   G(l,m)=0; %binary matrix to keep peaks

   sp=mean(C');
   plot(sp);
   s=input('define the baseline power: ','s'); %baseline power
   p=str2num(s);

   for i=1:m
      for j=1:3
         if val(j,i)>a*p
            D(loc(j,i),i)=mineng;
            G(loc(j,i),i)=1;
         end
      end
   end
   figure(1);
   imagesc(T,F(n1:n2),10*log10(C)),axis xy;colorbar;
   figure(2);
   imagesc(T,F(n1:n2),10*log10(D)),axis xy;colorbar;

   figure(3);
   colormap(gray);
   contour(T,F(n1:n2),G);
   h=findobj('type','patch');
   set(h,'Linewidth',2);
   axis([0 m*T(2) 0 2000]);
   xlabel('Time(sec)');
   ylabel('Frequency (Hz)');

case 3
   %using Homs-Corbera's threshold
   nfft=fix(0.0512*fs); %51.2ms;
   [B,F,T]=specgram(x,nfft,fs,hanning(nfft),fix(nfft/2));
   B=abs(B).^2;
   mineng=min(min(B))/10;
   D=B;
   [l,m]=size(B);
   C(l,m)=0;
   G(l,m)=0; %binary matrix for wheezing peaks
   H(l,m)=0; %binary matrix for all peaks
```

```
E=mean(B'); %spectrum
EE=zscore(E); % normalized global spectrum

n1=fix(100/F(2)); %how many rows for 100Hz
for i=1:m
   for j=1:n1:1-n1
      % C is normalized power spectrum in 100Hz band
      C(j:j+n1,i)=zscore(B(j:j+n1,i));
   end
end
```

```
      end


      [loc,val]=pickpeak(C,5,2);
      for i=1:m
         for j=1:5
            H(loc(j,i),i)=1;
         end
      end


%scoring
      for i=1:m
         for j=1:5
            tag=0; %initial score
            tmp=EE(loc(j,i));
            if val(j,i)>0.25*tmp & val(j,i)<1.35*tmp
               if loc(j,i)-3>=1 & ...
               val(j,i)>0.1*mean(C(loc(j,i)-3:loc(j,i)-1,i)) %rule 1
               tag=tag+1;
            end
            if loc(j,i)+3 <=1 & ...
               val(j,i)>0.1*mean(C(loc(j,i)+1:loc(j,i)+3,i)) %rule2
               tag=tag+1;
            end
            if loc(j,i)-1>=1 & loc(j,i)+1 <=1 & ...
               val(j,i)>C(loc(j,i)-1,i)+0.05 & ...
               val(j,i)>C(loc(j,i)+1,i)+0.05 %rule 3
               tag=tag+1;
            end
            if loc(j,i)-2>=1 & loc(j,i)+2 <=1 & ...
               val(j,i)>C(loc(j,i)-2,i)+0.2 & ...
               val(j,i)>C(loc(j,i)+2,i)+0.2 %rule 4
               tag=tag+1;
            end
            if loc(j,i)-2>=1 & loc(j,i)+2 <=1
               if C(loc(j,i)-1,i)< C(loc(j,i)-2,i) | ...
                  C(loc(j,i)+1,i)<C(loc(j,i)+2,i)%rule 5
                  tag=tag-1;
               end
            end
            if loc(j,i)-2>=1 & loc(j,i)+2 <=1 & ...
               4*val(j,i)-C(loc(j,i)-1,i)-C(loc(j,i)-2,i)- ...
               C(loc(j,i)+1,i)-C(loc(j,i)+2,i)>0.8 %rule 6
               tag=tag+1;
            end
            if tag>3
```

```
                G(loc(j,i),i)=1;
            end
        end
    end
  end

  %grouping twice
  n1=fix(50/F(2));
  n2=fix(65/F(2));

for k=1:2
  for i=1:m-2
     I=find(G(:,i)==1); %check for true peaks
     tmp=length(I);
     for j=1:tmp
       if I(j)-n1>=1 & I(j)+n1<1 & ...
          G(I(j),i)==max(G(I(j)-n1:I(j)+n1,i+1))
          %time distance 25ms&frq dis 50Hz
          [va,lo]=max(G(I(j)-n1:I(j)+n1,i+1));
          if lo==5 %upper neighbour
             G(lo+I(j)-n1-2,i)=1;
             %include the middle one between two wheeze peaks
          elseif lo==1 %lower neighbour
             G(lo+I(j)-n1,i)=1;
          end
        elseif I(j)-n2>=1 & I(j)+n2<1 & ...
               G(I(j),i)==max(G(I(j)-n2:I(j)+n2,i+2))
          %time distance 51ms&frq dis 65Hz
          [va,lo]=max(G(I(j)-n2:I(j)+n2,i+2));
          if lo>n2
             G(lo+I(j)-n2-2,i+1)=1;
             %include the middle one between two wheeze peaks
          else
             G(lo+I(j)-n2,i+1)=1;
          end
        end
      end
    end
  end


%elimiate short segments;twice
 for k=1:3
  for i=3:m-2
     I=find(G(:,i)==1); %check for true peaks
tmp=length(I);
for j=1:tmp
   if I(j)-n2>=1 & I(j)+n2<1
      if G(I(j),i)>max(G(I(j)-n2:I(j)+n2,i-2:i-1))... %lonely peek
         & G(I(j),i)>max(G(I(j)-n2:I(j)+n2,i+1:i+2))
         G(I(j),i)=0;
      elseif  G(I(j),i)==max(G(I(j)-n1:I(j)+n1,i-1))... %segs<80ms
            & G(I(j),i)>max(G(I(j)-n2:I(j)+n2,i-2))...
            & G(I(j),i)>max(G(I(j)-n2:I(j)+n2,i+1:i+2))
         G(I(j),i)=0;
      end
```

```
        end
      end
    end
end

 for i=3:m-2
    I=find(G(:,i)==1); %check for true peaks
    tmp=length(I);
    for j=1:tmp
       D(I(j),i)=mineng;
    end
end


 figure(1);
 imagesc(T,F,G),axis xy;colorbar;
 figure(2);
 imagesc(T,F,10*log10(D+eps)),axis xy;colorbar;

 figure(3);
 colormap(gray);
 contour(T,F,G);


 h=findobj('type','patch');
 set(h,'Linewidth',2);
 axis([0 m*T(2) 0 2000]);
 xlabel('Time(sec)');
    ylabel('Frequency (Hz)');

end
```