Signal Processing Methods for Non-Invasive Respiration Monitoring

Laura Mason

Supervised by Professor Lionel Tarassenko

Trinity College
Michaelmas 2002

This thesis is submitted to the Department of Engineering Science, University of Oxford, in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Declaration

I declare that this thesis is entirely my own work, and except where otherwise stated, describes my own research.

C. L. Mason,
Trinity College

Laura Mason
Trinity College

Doctor of Philosophy
Michaelmas Term, 2002

Signal Processing Methods for Non-Invasive Respiration Monitoring

Abstract

This thesis investigates the feasibility of using a set of non-invasive biomedical signals to monitor respiration. The signals of interest being the electrocardiogram (ECG), photoplethysmography (PPG) and impedance plethysmography (IP) signals.

The work has two main aims; the first being to estimate breathing rates from the signals, the second being to detect apnoeas from the signals.

The fusion of information from different signals is used throughout in developing algorithms that give more accurate respiratory information than that obtained using one signal alone.

Respiratory waveforms are derived from the signals, and the accuracy of detecting individual breaths from the waveforms is assessed and compared objectively. Results from evaluations on two separate databases show there is no waveform that gives sufficient accuracy to consider using alone.

A novel fusion method is developed which uses measurements from all three signals. This fusion method is based on weighting the estimates from each signal, according to the innovation from a Kalman filter model, applied to each respiratory waveform separately. The fused estimates give a higher overall correlation with respect to the reference breathing rate values than any of the breathing estimates derived from a single waveform.

The detection of both central and obstructive sleep apnoea from the signals is investigated. It is shown the accuracy of detecting central apnoeas from the IP signal using a time-domain method, often used in practice, can be improved by combining it with information from the frequency-domain.

When discriminating between obstructive sleep apnoeic and non-apnoeic data it is seen that combining features from two signals results in a superior classification accuracy than using features from just one signal. The proposed classification system using just one of these signals, the ECG, is shown to give a performance accuracy comparable to that found in the literature.

Acknowledgments

I would like to thank my supervisor, Professor Lionel Tarassenko, for his encouragement and advice throughout the course of my D.Phil. Thankyou also to all those who have worked in the research group during my time there, particularly Neil Townsend and Stephen Payne and also James Price for his clinical advice.

Thanks to friends who have made my time at Oxford very enjoyable; especially to Tamsin for believing I could do it and Lex for being able to empathise. Finally thankyou to James, for making me laugh and helping me to keep things in perspective, and to my parents for their continual support and advice, which is greatly appreciated.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Aim

The aim of this thesis is to investigate the feasibility of performing reliable respiratory monitoring from non-invasive sensors. The signals of primary interest are:

- the electrocardiogram (ECG),
- the transthoracic electrical impedance plethysmography signal (IP),
- the photoplethysmography signal (PPG).

These signals are chosen for investigation because they can all be obtained non-invasively, they are known to be influenced by respiration and all are currently measured in clinical environments. Hence a successful solution based on these signals would not require additional hardware or monitoring devices.

1.2 Motivation

Respiration is defined as [5]:

> *"a.The act or an instance of breathing or*
>
> *b. a single inspiration or expiration; a breath".*

There are many clinical reasons for requiring a reliable measure of respiration. Some of these are found in the literature on methods for respiratory monitoring. For example, Moody *et al.* [71] maintain that the clinical significance of certain cardiac arrhythmias can be understood only by reference to respiration. Knowledge of the respiration cycle is also useful for synchronisation and compensation of MRI scans of sequences of the heart and thorax [34]. Lindberg *et al.* [58] quote several examples where respiration information is important. For example, in postoperative care it is extremely important to monitor both the cardiac function and respiration as post-operative patients can be influenced by analgesics that depress respiration.

The detection of sleep apnoea is also important. Apnoea is defined as a temporary cessation of breathing. There are two main types of sleep apnoea, central and obstructive. Central apnoea is very common in neonates, while obstructive sleep apnoea currently affects up to 8% of the male population in the 40-59 year age group [67].

Several methods have been proposed to measure respiration. A description of these is found in Section 1.2.1. Although there are reliable methods which are used in high-dependency or intensive care units, these are either invasive or can cause discomfort. Hence these methods are not used on the general ward. Current methods of detecting apnoeas are given in the chapters which focus on this topic (Chapters 6 and 7).

1.2.1 Current monitoring devices

Several approaches have been designed to measure respiration and they can be categorised into direct and indirect methods.

In direct methods, a sensor is coupled to the airway and measures the movement or other properties of air transported into and out of the lungs. For example nasal thermistors measure the temperature changes in the air and carbon dioxide sensors measure the change in carbon dioxide in inhaled and exhaled air [12]. Respiration can also be monitored indirectly by measuring changes in body volume; examples of such techniques are transthoracic inductance and impedance plethysmographs, strain gauge measurement of thoracic circumferences, pneumatic respiration and whole body plethysmographs [71]. Transthoracic impedance and inductance plethysmography are the indirect methods most commonly employed. In inductance plethysmography, compliant inductance loops are placed around the

chest and abdomen. During inhalation and exhalation the volumes of the chest and abdomen change, and this changes the area of the coils and hence their inductance. Impedance plethysmography measures the change in electrical impedance across the chest and abdomen with respiration.

Each method has advantages and disadvantages. Direct measurements tend to be more accurate but can interfere with normal respiration. Although the whole body plethysmograph can be highly accurate and does not interfere with respiration, it requires immobilising the patient. Transthoracic impedance plethysmography is often combined with heart rate monitoring. However changes in blood flow generate a cardiac-synchronous change in impedance which can be a significant source of error in the generated respiratory signal; aortic blood flow can give rise to a signal of the same order of magnitude as the variations caused by respiration [58]. The impedance plethysmography signal is also very sensitive to body movement [88]. Nevertheless, this method is non-invasive and therefore features in this thesis. It is considered in more detail in Chapter 2.

1.3 Objectives

The objective of this thesis is to use signal processing techniques to improve or provide respiration information from signals routinely obtained in a hospital or other clinical environments. The main objectives of the work are now summarised.

Although there are many published works relating to the task of deriving respiration from signals such as the ECG and PPG including [34], [58], [71], [107] and [117], none of these present objective, reproducible evaluations of the performance of their methods. Such evaluation is necessary to compare the accuracy of one method with another. Therefore the first objective of this thesis is to:

- establish objective and reproducible evaluation procedures for methods of obtaining breath-by-breath information

Methods developed in this work for obtaining respiration are tested and compared with methods described in the literature. A logical question to then ask is whether information from the different signals can be combined (or fused) to obtain an improved measure of respiration. Therefore a second objective of this thesis is to:

- develop a method for combining breathing rate measurements from multiple signals to obtain a more accurate estimate of breathing rate than is available from any individual signal.

Of as much clinical importance is the ability to detect the cessation of respiration. This motivates the final objectives of determining:

- the accuracy of detecting central sleep apnoeas, using information from the non-invasive signals of interest,
- the accuracy of detecting episodes of obstructive sleep apnoea using the non-invasive signals of interest.

1.4 Overview of thesis

This thesis is organised as follows:

Chapter 2 introduces the relevant biomedical signals, namely the ECG, blood pressure, IP, and PPG. An overview of the origin and measurement of these signals is given. The influence of respiration on these signals is also described. The databases used in this thesis are also introduced.

Chapter 3 reviews the signal processing techniques used in this thesis. Classification and Auto-Regressive modelling for frequency-domain estimation are discussed. An overview of data fusion is given, as data fusion is used in a number of applications in this thesis.

Chapter 4 presents a literature review of previous approaches to extracting respiration information from the biomedical signals in question. An evaluation procedure for assessing the performance of methods that automatically detect breaths from the signals is proposed. The different signal processing methods for deriving the breaths are then discussed. Methods suggested in the literature are also implemented for comparison. Finally, results of testing these methods on the databases are presented.

Chapter 5 investigates methods of improving the estimates of breathing rate measurements. The Kalman filter is introduced as an optimal estimator and data fusion methods using Kalman filter theory are presented. Due to the inherent nature of the biomed-

ical signals under investigation, the design of accurate models of measurement noise is a very difficult if not impossible task. Kalman filter methods rely on accurate models of both process and measurement. Therefore a novel method of combining the breathing rate measurements from the different signals is derived.

Chapter 6 investigates methods of detecting central apnoeas using the IP, PPG and ECG signals. It is noted that central apnoea monitors tend to have a very high false alarm rate. A review of studies carried out with existing apnoea monitors which employ non-invasive techniques is given. A time-domain approach based on existing methods is implemented and provides a set of benchmark results against which a novel frequency-domain technique is evaluated. A method that fuses information from the time and frequency domains at a decision level is also assessed.

Chapter 7 investigates methods of detecting obstructive apnoeas using the ECG and blood pressure signals. A classification system is proposed that classifies time windows into the set $\{apnoeic; non - apnoeic\}$. An Auto-Regressive model parameterisation of the signals forms a set of features, used as the input to a statistical classifier. Two approaches to combining information from the ECG and blood pressure are assessed. The classification system is tested on two databases.

Chapter 8 summarises the findings and draws conclusions on obtaining respiration information from the non-invasive signals under consideration in this thesis. Possible avenues of further work are also discussed.

Chapter 2

The Primary Signals

2.1 Introduction

The aim of this thesis is to investigate the feasibility of reliably monitoring respiration using a set of non-invasive biomedical signals. The signals of primary interest are:

- the electrocardiogram (ECG),
- the transthoracic (electrical) impedance plethysmography signal (IP),
- the photoplethysmography signal (PPG),

these signals being chosen because they are all measured non-invasively and generally available in clinical environments. The work is carried out on two databases, both of which include recordings of the ECG, IP, and airflow respiration signals, the latter being used as the reference respiration measure. The database that is part of the Physionet archive [70] however, does not include a PPG signal but an intra-arterial blood pressure signal. Several clinical papers argue that the PPG signal can in fact be used as a non-invasive measure of the blood pressure waveform [36],[45]. This argument is further discussed in Section 2.4.1. In work with the Physionet database, the blood pressure signal is used for "proof of concept", although ultimately all signals would be measured non-invasively.

Figure 2.1 shows a one-minute section of typical signals from one of the databases used in this work. This chapter reviews the physiological origin of the signals, how they are recorded

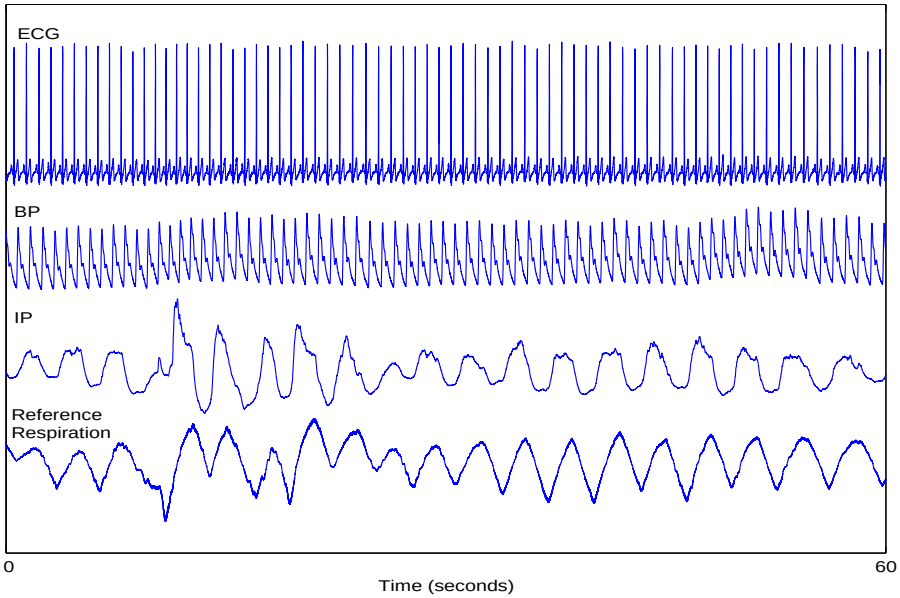

Figure 2.1: A one-minute section of the signals under investigation from the Physionet Polysomnography database. The signals have been scaled to allow visual comparison, hence no vertical scales are provided.

and how they are affected by respiration. The effect of apnoea on the signals is left until later, in the chapters relevant to apnoea detection. Finally Section 2.6 describes the two databases on which the work in this thesis is carried out.

2.2 The electrocardiogram (ECG)

An ECG is a graphic recording of the electrical activity produced by the heart. The heart is a four-chambered pump that provides the driving force for the circulation of blood [100]. With each heart beat the synchronised depolarisation spreading through the heart causes currents in the extracellular fluid that establish field potentials over the whole body. These potential differences can be detected by electrodes placed on the body's surface. The pattern of the ECG varies according to the electrodes' position but certain features are always present. These features were labelled as PQRST by Einthoven [32]. Details of the underlying cardiac events that cause these features can be found in [49]. Figure 2.2 shows the ECG with respect to time obtained during one complete cardiac cycle. This cycle is repeated with every heart beat. It is the QRS complex of the ECG that is of most interest in this work.

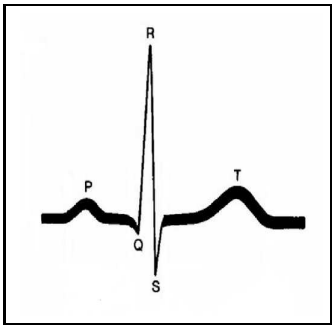

Figure 2.2: A schematic representation of part of an ECG waveform, corresponding to a cardiac cycle. The y-axis corresponds to potential and the x-axis to time. Taken from [4].

Electrical activity radiates from the heart in all directions. The location of recording electrodes influences the view point of the heart recorded. An electrocardiographic lead is a pair of polar terminals connected to electrodes [8].

The electrodes are fixed at points that are both convenient and likely to reveal the pattern of greatest clinical interest. The time-honoured method, originally described by Einthoven [32] and known as Einthoven's Triangle, is to affix them to the right arm, left arm and left leg. The connections between them are known as Lead I (right arm and left arm), Lead II (right arm and left leg) and Lead III (left arm and left leg). In short-term monitoring for diagnostic purposes, a 12-lead ECG is usually employed. In this work, information from a single lead ECG is used; Lead II unless otherwise stated.

2.2.1 The ECG and respiration

It is well recognised that there are breathing-related characteristics in the ECG. The effects on the ECG of heart displacements produced by respiratory movements were first systematically analysed by Einthoven *et al.* in 1913 [32]. Nowadays it is a well known fact that respiratory action produces a rotation of the cardiac vector (see 2. below); as is the effect of the Respiratory Sinus Arrhythmia on the ECG signal. The three dominant effects of respiration on the ECG are listed below:

1. **Respiratory Sinus Arrhythmia** (RSA) refers to the cyclic variation in heart rate that is associated with respiration. Heart rate accelerates during inspiration and slows during expiration [60]. The magnitude of the oscillation is variable and varies from individual to individual. The amount of RSA also depends on breathing frequency,

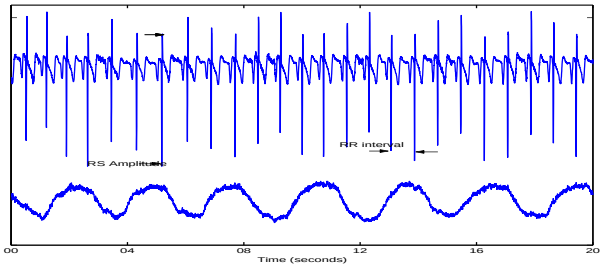

Figure 2.3: A reference respiration signal (lower trace), obtained from a nasal thermistor and an ECG (upper trace). The amplitude of the R-S distance can be seen to be modulated by respiration.

tending to decrease with breathing rate [17].

2. **R-S Amplitude Modulation** During inspiration, the apex of the heart is stretched towards the abdomen because of the filling of the lungs, helped by the shifting down of the diaphragm. During expiration the elevation of the diaphragm, that helps the emptying of the lungs, compresses the apex of the heart towards the breast. Thus respiration changes the angle that the electric cardiac vector makes with a reference vector [107]. These changes modulate the amplitude of the ECG signal. It is noted that the modulation of the QRS amplitude is particularly significant. Appropriate positioning of the ECG electrodes can maximise the respiration-induced modulation; it is suggested that Lead II for example, shows greater modulation than lead I [77]. Figure 2.3 shows a 20-second section of an ECG and a corresponding reference respiration signal obtained from a nasal thermistor. The R-S amplitude of the ECG can be seen to change with respiration.

3. **Baseline Wander** Low frequency wander of the ECG signal can be caused by respiration [8]. The expansion and contraction of the chest that accompanies respiration results in the motion of chest electrodes with respect to the heart. This can cause a baseline wander in the ECG, usually only seen in deep or exaggerated breathing.

2.3 Blood pressure

Blood pressure is the pressure of the blood flowing through the blood vessels against the vessel walls. It depends on the quantity of blood flow and the resistance of the blood vessels

to this flow. Each time the heart beats, a surge of blood is pumped from the heart into the arteries. This increases the pressure in the arteries. In between heart beats the pressure in the arteries decreases. The blood pressure is reported as two numbers, e.g. 120/80. The first, higher number (systolic) is the pressure of the blood against the artery walls when the heart contracts. The second, lower number (diastolic) is the pressure against the artery walls when the heart relaxes between beats.

Blood pressure may be measured using direct or indirect techniques. Direct measurements use catheters for the invasive determination of blood pressure, whereas indirect methods utilise a variety of non-invasive techniques.

Direct measurements employ one of two different types of sensors, in conjunction with a catheter. Extravascular sensors are located outside of the body and use the principle of wave propagation to transmit vascular pressure from the measurement site to the sensor via a fluid-filled catheter. In contrast, intravascular sensors, such as fiber-optic pressure sensors, are positioned on the tip of the catheter and inserted into the artery. Catheter systems are often used for direct, continuous measurement of intra-arterial blood pressure in the aorta, which is considered the “gold standard” for blood pressure measurement.

The non-invasive measurement of blood pressure may be performed in several different ways. One such method employs an air-filled cuff to the upper arm to temporarily occlude blood flow through the brachial artery. The cuff pressure is monitored with an in-line transducer as the cuff is slowly deflated. The systolic and diastolic pressures can be computed from the cardiac-synchronous oscillations in cuff pressure during the process.

2.3.1 Blood pressure and respiration

1. **Pulsus Paradoxus** is the inspiratory decrease in systolic blood pressure which is proportional to changes in intrathoracic pressure during inspiration and expiration [36]. Inspiratory decline of the arterial pulse was first described in asthma sufferers by Floyer in 1717 [51]. The exaggerated decrease in blood pressure was termed pulsus paradoxus by Kussmaul in 1873. Pulsus paradoxus is increased by respiratory diseases, such as asthma, the degree of pulsus paradoxus reflecting the severity of the underlying disorder. A value of greater than 10 $mmHg$ is considered significant. Although strictly speaking, the term pulsus paradoxus refers to a systolic blood pressure

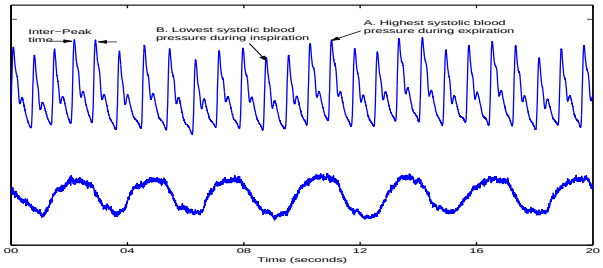

Figure 2.4: The upper trace shows a blood pressure signal obtained from an intra-arterial catheter and the lower trace a reference respiration signal.

variation of more than 10 $mmHg$, it is used in this thesis to refer to any respiratory dependent systolic blood pressure variation.

Figure 2.4 shows a 20-second section of an intra-arterial blood pressure signal and a corresponding respiration signal obtained from a nasal thermistor from the Physionet MIT-BIH Polysomnography Database [3]. **A** marks the highest systolic blood pressure during the respiration cycle (i.e., during expiration) and **B** the lowest.

2. **Blood Pressure Variability** As described in Section 2.3, when the heart contracts, blood is pushed through the aorta into the arteries. This sudden increase in volume of blood causes vasodilatation and the blood vessels expand. After each dilation, the arteries force the blood along in a series of waves. Each wave is known as a pulse and the peak of the pulse corresponds to the systolic blood pressure. Beat-to-beat time intervals of systolic blood pressure have been shown to have a cyclical variation related to respiration [21].

2.4 Photoplethysmography signal (PPG)

Photoplethysmography is the electro-optic technique of measuring the cardiovascular pulse wave found throughout the human body. This pulse wave is caused by the periodic pulsations in arterial blood volume and is measured by the consequential changing optical absorption that this induces. The measurement system consists of a light source (usually infra-red), a detector (positioned in reflection or transmission mode) and a signal recovery/processor/display system. Infra-red light is predominantly used since it is relatively well absorbed in blood and weakly absorbed in tissue; blood volume changes are therefore ob-

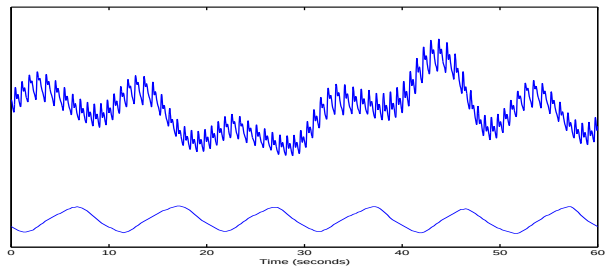

Figure 2.5: A PPG signal (upper trace) and the corresponding reference respiration signal

served with reasonable contrast. The PPG measurement is entirely non-invasive and can be applied to any blood bearing tissue, although the majority of investigations performed to date are of peripheral microcirculation [46].

The use of PPG in pulse oximetry has become a standard means of obtaining valuable blood oxygenation data in a non-invasive and continuous manner [45]. Pulse oximetry uses two wavelengths of light to determine oxygen saturation (SaO_2). Haemoglobin is the carrier of oxygen in the blood. The two common forms of the molecule, oxidised haemoglobin HbO_2 and reduced haemoglobin Hb have very different optical spectra in the wavelength range from 500nm to 1000nm. This wavelength range also corresponds to the range for which there is least attenuation of light by body tissues [106]. Oxygen saturation is defined as the ratio of oxyhaemoglobin (HbO2) to the total concentration of haemoglobin present in the blood.

$$SaO_2 = \frac{HbO_2}{HbO_2 + Hb} \tag{2.1}$$

Figure 2.5 shows a PPG signal from one of the databases used in this work, recorded from a healthy subject over one minute. Pulse rate can be measured by tracking the interval between consecutive peaks on the PPG waveform (about 1 pulse per second in Figure 2.5) and also the amplitude of the waveform may reflect the adequacy of peripheral perfusion. The attenuation of light by the body can be thought of as being due to three independent components; arterial blood flow, venous blood and tissues. The high frequency component of the signal is produced by pulsatile arterial blood volume with each heart beat. The absorption by nonpulsatile arterial blood, venous and capillary blood and tissue absorption gives rise to the steady state signal. However, other slow variations in blood flow also affect the modulation

of light intensity. These slower modulations can be seen in Figure 2.5. One of these is the variation in venous return to the heart, due to changes in intrathoracic pressure caused by respiration [58]. Figure 2.5 clearly shows the modulation due to respiration at a frequency of about $0.16Hz$ (6 breaths in the one-minute segment shown).

2.4.1 Respiration and PPG: relationship between the PPG and blood pressure

PPG measures the pulse wave caused by periodic pulsations in arterial blood volume. These periodic pulsations give rise to the blood pressure waveform. Several studies have attempted to determine the correlation between the respiration induced effects in the blood pressure and PPG signals, the aim of these studies being to investigate the possibility of non-invasive analysis of pulsus paradoxus using PPG. Steele *et al.* [97] report on different methods of continuous non-invasive monitoring of pulsus paradoxus. The reference measurement of pulsus paradoxus is taken as the measurement derived from an arterial catheter — i.e., the difference between the highest (during expiration) and lowest (during inspiration) systolic blood pressure. Pulsus paradoxus determined by a non-invasive continuous blood pressure signal (obtained using a Finapres, see Appendix C) and the PPG signal are compared with the reference measurement. A respiratory cycle is determined using a strain gauge placed around the chest. It is concluded that the PPG pulsus paradoxus measurement, although less accurate than the Finapres measurement shows a strong linear correlation with the reference measurement.

Frey *et al.* [36] carried out a study with the aim of evaluating the relationship between pulsus paradoxus measured intra-arterially and PPG wave changes. Recordings of the PPG waveform, arterial blood pressure and breathing cycle are taken from sixty two nonintubated patients. All the analysed PPG waves and arterial blood pressure waves are in phase, with the lowest values occurring in inspiration and the highest values in expiration. It is concluded that the PPG signal appears to be a rapid and easily performed, non-invasive method for the objective estimation of the degree of pulsus paradoxus.

Hartert *et al.* [45] studied 26 people who had severe breathing difficulties (asthma or emphysema). The change in baseline in the PPG signal between expiration and inspiration was recorded. The magnitude of this change was found to correlate strongly with the strength of

pulsus paradoxus. After treatment and clinical improvement the change in baseline between expiration and inspiration was greatly reduced, and the modulation due to respiration was no longer visible by eye.

These studies show that the effect of respiration can be clearly seen in the PPG of subjects with breathing difficulties such as asthma. More recently it has been shown that the sensors used to measure PPG are sensitive enough to detect the pulsus paradoxus effect in the blood volume change even in healthy subjects [53].

2.5 Impedance plethysmography (IP)

The IP signal is a measure of respiratory effort and is sometimes used to display a respiratory waveform in a clinical environment. A low amplitude current is passed through the subject via electrodes. As the patient inhales and exhales, the electrical impedance of the chest cavity changes as a result of changes in air volume in the lungs and hence the conductive path between the electrodes.

Air in the thoracic cavity undergoes large changes in volume in the lung during normal breathing but the volume of blood also varies over the cardiac cycle due to changes in the amount of blood in the heart and blood vessels [4]. Thus, the electrical impedance of the lungs and heart changes both with change in air and blood volumes. Most monitors utilise two standard ECG electrodes placed on the subject's anterior chest. The impedance across the chest is seen to increase with inspiration due to increased gas volume in the thorax and blood volume during the cardiac cycle. The impedance contributed by muscle and fat remains relatively constant. Usually the variation in thoracic impedance is greater during respiration than during the cardiac cycle. Impedance monitoring has inherent short comings in accurate respiration detection in cases of body movement and postural changes. In cases of shallow breathing, the monitor may detect the cardiac cycle and count them as respirations, causing a false estimation of displayed respiratory rate [94].

2.6 Databases

Descriptions of the two databases on which the work in the thesis is carried out are found in Sections 2.6.1 and 2.6.2 below. The first database, the MIT-BIH Polysomnography database is part of the Physionet archive of biomedical signal databases. The second database was collected during the course of this DPhil. study. It was designed with investigating the feasibility of monitoring respiration from the non-invasive signals discussed in this chapter.

2.6.1 MIT-BIH Polysomnography database

The PhysioNet website contains an archive of well characterised multi-parameter biomedical signals for use by the research community [39]. The use of objective assessment procedures alongside the available databases allows researchers to objectively test and compare algorithms.

The MIT-BIH Polysomnography Database comprises 16 continuous polysomnography recordings from male subjects all of whom suffer from sleep apnoea. The recordings are between 2 and 7 hours in duration and digitised at 250 Hz with 12 bit resolution.

The signals used in this thesis are the ECG, IP and blood pressure signals measured using an intra-arterial catheter. A direct respiratory signal is also available, obtained from a nasal thermistor, i.e., a thermistor placed in the airflow of the nasal passages. This is used as the reference respiration signal. The following information is available for each recording:

- the location of QRS complexes in each ECG record;
- markers for events such as apnoeas, arousals, and movement.

2.6.2 Controlled-breathing database

The second database used in this work was collected in January 2002 by the Signal Processing and Neural Networks research group at Oxford University. It is referred to as the Controlled-breathing database throughout this thesis. For the purpose of the respiration monitoring feasibility study, the protocol included subjects breathing at a range of rates and depths. Central apnoeas were also mimicked by asking the subject to stop breathing for short periods. The full protocol is found in Appendix A. The data was collected from 10 male

subjects, none of whom had known medical conditions. The signals were collected using a multi-parameter patient monitoring system, developed to collect and analyse biomedical signals [102]. The relevant signals are:

- ECG: a three lead ECG was recorded, the signals being digitised with 12 bit accuracy at a sampling rate of $256Hz$.
- IP: a $22KHz$ electrical signal of very low amplitude (μA) is injected into the subject through two of the ECG electrodes. The IP signal is sampled at $68Hz$ with 16-bit precision.
- PPG: this signal is sampled at $81.3Hz$ with an accuracy of 16 bits.
- a direct respiratory signal: an oral airflow trace was obtained by recording the breath-by-breath temperature variations within a tube through which the subject breathed.

2.6.3 Labelling the data

The Polysomnography database [113] is made available with a set of labels, corresponding to each 30-second epoch of data. The labels define events that occur within the 30-second epoch. These events include central and obstructive sleep apnoeas and hypopneas. The Controlled-breathing database was labelled as the data collection was undertaken. Exact times of beginnings and endings of different breathing conditions were noted as the protocol in Appendix A was followed.

Breath-by-breath annotation files corresponding to salient points in the breathing cycle, the start of inhalation and exhalation, were created as part of this work. To construct breath-by-breath reference time-stamps for each database, the peaks of the reference airflow respiratory signals are taken to be the start of each respiratory cycle. Every peak is labelled to define a breath in the reference annotation files. Peak detection was first performed using an automated peak detection algorithm. The results from this algorithm were subsequently edited manually using the Wave Analyzer, Viewer and Editor (WAVE) [113].

Chapter 3

Signal Processing Methods

3.1 Introduction

This chapter gives an overview of some of the signal processing techniques used in this thesis. Feature extraction is introduced and a technique used for feature extraction, Auto-Regressive modelling, are described in Sections 3.2 and 3.3. Section 3.4 discusses classification, with particular reference to neural networks. A discussion of the topic of *data fusion* is presented in Section 3.10.

3.2 Feature extraction

The goal of features is to characterise data by measurements whose values are very similar for objects in the same class, and very different for objects in a different class. As well as providing discriminatory information, one of the most important functions of feature extraction is dimensionality reduction of the data.

In this work, on respiratory signals, it is known that discriminatory information can be found in the frequency-domain. The most well known frequency-domain transformation is the Fast Fourier Transform (FFT). However the FFT transformation results in a set of coefficients of equal number to the original data. The Auto-Regressive (AR) model estimate of the power spectrum allows a low dimensional representation of the data.

3.3 AR spectral estimation

AR spectral estimation often gives a very significant improvement in frequency resolution compared to the traditional periodogram method as implemented by the FFT [79]. The estimated AR spectrum is a continuous function of frequency and can thus be evaluated numerically at any number of frequencies — uniformly spaced or otherwise — in the interval, $0 \leq f \leq 0.5f_s$ where f_s is the sampling frequency. Conversely the periodogram is a discrete spectrum, evaluated only at uniformly spaced frequencies, determined by the sequence length. If the sequence is short the periodogram may fail to resolve spectral peaks that are close together. The sampling frequencies of the signals used in this work range from $68Hz$ to $256Hz$, with breathing frequencies of interest being between $0.1Hz$ and $0.5Hz$. Therefore long sequences are needed to resolve these frequencies, however stationarity cannot be assumed for such long sequences. A second limitation is due to the implicit windowing of the data that occurs when processing the periodogram. Windowing manifests itself as "leakage" in the spectral domain, i.e., energy in the main lobe of a spectral response "leaks" into the sidelobes, obscuring and distorting other spectral responses that are present. In fact, weak signal spectral responses can be masked by higher sidelobes from stronger spectral responses. Skilful selection of tapered data windows can reduce the sidelobe leakage, but always at the expense of reduced resolution. These two performance limitations of the periodogram approach are particularly troublesome when analysing short data records. Short data records occur frequently in practice because many measured processes are brief in duration or have slowly varying time spectra that may be considered stationary only for short record lengths. Another advantage of the AR spectral estimation is that very few cycles or even fractions of a cycle can often be reliably detected [56].

Parametric time series analysis uses mathematical models to estimate the spectral density and covariance function of discrete time signals. An AR model framework assumes that an all-pole linear filter describes the generation of the signal under consideration and that the filter is driven by a white noise signal. The AR model therefore specifies the shape of the signal's spectrum.

In the time-domain an AR model can be viewed as a linear predictor. A discrete time series $x(t)$ may be described as the output of an AR model of order p as follows:

$$x(t) = \sum_{k=1}^{p} a_k x(t-k) + e(t) \tag{3.1}$$

where a_k (*k=1,2...p*) are the coefficients of the AR model, t is the discrete time index and $e(t)$ is the prediction error, i.e., the error between the actual value and the predicted value.

Eq. 3.1 can be expressed in the z-domain

$$X(z) = X(z)(\sum_{k=1}^{p} a_k z^{-k}) + E(z) \tag{3.2}$$

or

$$X(z) = H(z)E(z) \tag{3.3}$$

where $X(z)$ and $E(z)$ are the z transforms of $x(t)$ and $e(t)$ respectively and the model transfer function, $H(z)$ is

$$H(z) = \frac{1}{1 - \sum_{k=1}^{p} a_k z^{-k}} = \frac{1}{\prod_{i=1}^{p}(1 - z_i z^{-1})} \tag{3.4}$$

The z-domain PSD is expressed as

$$P_x(z) = |H(z)|^2 P_e(z) \tag{3.5}$$

The frequency-domain representation of the PSD may be estimated by substitution of

$$z = e^{j2\pi f \Delta t} \tag{3.6}$$

where Δt is the sampling interval and the expression in 3.5 is evaluated over $\frac{-1}{2\Delta t} \leq f \leq \frac{1}{2\Delta t}$. If $e(t)$ is a white noise sequence with zero mean and variance σ^2 then its power spectrum is flat with amplitude $\sigma^2 \Delta t$. The AR frequency power spectral density estimate $P_x(f)$ is given by

$$P_x(f) = \frac{\sigma^2 \Delta t}{\left|1 - \sum_{k=1}^{p} a_k exp^{(-j2\pi f \Delta t)}\right|^2} \tag{3.7}$$

Thus to estimate the PSD one need only estimate $a_1, a_2, a_3..., a_p, \sigma^2$. A number of algorithms exist for estimating the parameters and details and derivations may be found in [56, 79]. In this thesis the Maximum Entropy Method, that is known to produce a minimum bias solution [56], is used. The PSD obtained using the AR model method consists of a number of Gaussian curves, centred at the resonant frequencies. Each resonant frequency f_c is related to a pair of complex conjugate poles. A pth order model has $\frac{p}{2}$ poles or resonant frequencies for an even value of p, and $\frac{p}{2} + 1$ for an odd value of p. To estimate the power at a certain peak or resonant frequency f_c, Eq. 3.7 maybe solved for $f = f_c$. Alternatively, Andersen *et al.* [54] suggest a graphical method, using knowledge of the pole distribution inside the unit circle.

Re-writing Eq. 3.5 as

$$P_x(z) = \frac{\sigma^2 \Delta t}{H(z) H^*(\frac{1}{z^*})} \tag{3.8}$$

and noting from Eq. 3.4

$$H(z) = \frac{1}{\prod_{i=1}^{p}(1 - z_i z^{-1})} = \frac{z^p}{\prod_{i=1}^{p}(z - z_i)} \tag{3.9}$$

Substitution of Eq. 3.9 into Eq. 3.8 gives

$$P_x(z) = \frac{\sigma^2 \Delta t}{\prod_{i=1}^{p} |z - z_p|^2} \tag{3.10}$$

This equation has geometric interpretations in terms of vectors connecting the poles of $P(z)$ to the point on the unit circle with phase of the frequency of interest.

The peak associated with a resonant frequency occurs at the angular location of the pole z_i

$$f_c = 2 \times \pi \times \arctan(Im(z_i)/Re(z_i)) \tag{3.11}$$

These equations give the relationship between the magnitude of the poles in the pole plot in the z-domain and the power at the corresponding frequency. If a pole is close to the unit circle, a well resolved peak appears in the spectrum. Poles distant from the unit circle contribute little to the shaping of the spectrum. Their main effect is to scale the overall spectrum up or down, thus ensuring conservation of the total power.

3.3.1 Reflection coefficients

As well as the coefficients a_k (*k=1,2...p*), the reflection coefficients ρ_k, also provide a parameterisation of the power spectrum of the data x_n.

The reflection coefficients define the reduction in residual signal-model error, E, when the AR model increases its order from $m-1$ to m,

$$E_m = \left(1 - \rho_m^2\right) E_{m-1} \tag{3.12}$$

Reflection coefficients have the advantage that an increase in model order does not effect the coefficients from the previous order and hence there is little cross-correlation between the coefficients, making them more suitable for pattern analysis techniques than the correlated AR coefficients [23]. A further advantage is that reflection coefficients are independent of signal amplitude and always lie between -1 and +1, so they do not always require normalisation when used as features for classification. The AR model or reflection coefficients form a compact representation of the signal power spectrum. In addition the representation is adaptive since the parameters of the estimated spectrum give the location and sharpness of the dominant frequencies wherever they may be in the spectrum (unlike the FFT that provides a fixed frequency representation).

AR models are implemented in Matlab 6.1 using the Burg algorithm [56], which solves the Maximum Entropy Method, wherever they are used in this thesis.

3.4 Classification

A classifier can be viewed as a mapping from a set of input variables (features) $\mathbf{x}$, representing the object or data to be classifed to an output y which represents a class label

$$y = f(\mathbf{x}, \mathbf{w}) \tag{3.13}$$

where $\mathbf{w}$ denotes the vector of parameters representing the mapping. In general it is not possible to determine a suitable form for the required mapping, except with the help of a data set of examples [10]. The process of using data to determine the parameters of a classifier is referred to as training. Many different training techniques exist, using either supervised or unsupervised training methods. Unsupervised techniques take no account of class information and so the mapping that takes place may lead to very poor grouping for a classification problem. Unsupervised techniques are often used for visualisation or cluster analysis in an attempt to discover any underlying structure in the data. Supervised techniques use class information to find an optimal decision boundary to be selected with respect to the target class.

Parametric methods of estimation of the parameter vector $\mathbf{w}$ asssume a specific functional form for the mapping. The data is used to fit the model to the data set. Non-parametric estimation does not assume a particular functional form and allows the mapping to be determined entirely by the data.

The simplest type of decision boundary is linear. A linear discriminant function, which is linear in the components of $\mathbf{x}$, can be written as

$$y(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + w_0 \tag{3.14}$$

For a two class problem, an input vector $\mathbf{x}$ is assigned to class C_1 if $y(\mathbf{x}) \geq 0$ and to class C_2 if $y(\mathbf{x}) \leq 0$ [10].

Linear discriminants can be extended to a k class problem by using a discriminant function $y_k(\mathbf{x})$ for each class C_k. A new input vector $\mathbf{x}$ is assigned to class C_k if $y_k(\mathbf{x}) \geq y_j(\mathbf{x})$ for all $j \neq k$. This leads to a set of decision regions that are always simply connected and convex. However there are many real-world problems for which the optimal decision boundaries are

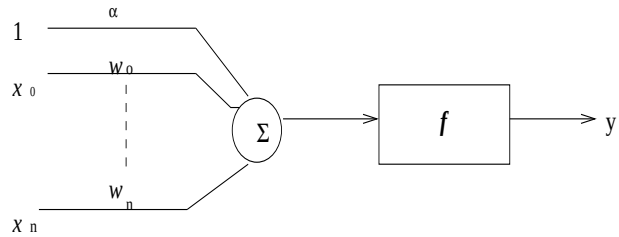

Figure 3.1: Schematic diagram of a perceptron

not linear.

There are many supervised classification techniques that lead to non-linear decision boundaries, including Support Vector Machines (SVM) [92], Gaussian Mixture Models (GMM), Hidden Markov models (HMM) [59], Bayesian discriminants [112], Radial Basis function and Multi-Layer Perceptron (MLP) neural networks [10]. GMMs and HMMs like MLPs (under certain conditions, See Section 3.5) all output normalised probability scores, the posterior probability of belonging to a class. GMMs and HMMs make the assumption that the system or data can be modelled or represented by a weighted sum of Gaussians, but require a lot of data for training.

A neural network is used as a classifier in this work as no explicit analytical model of the data is needed. The section which follows gives an overview of the MLP; this is the neural network chosen to solve the classification problem of Chapter 7.

3.5 Multi-layer perceptron neural networks

The theory of neural networks is well documented in texts such as [10],[29] and [101] a brief overview is given here.

Figure 3.1 shows a schematic diagram of a perceptron. The output y is formed by passing a weighted sum of the input vector $\mathbf{x}$, added to a bias term α, through a non-linearity f. The non-linearity f is known as the activation function and is normally one of a number of standard monotonic functions. In the original work by Rosenblatt [101] the function is the Heaviside function, a hard-limiter which gives an output of 1 for one class and 0 for the complementary class. Another commonly used non-linearity is the logistic sigmoid function.

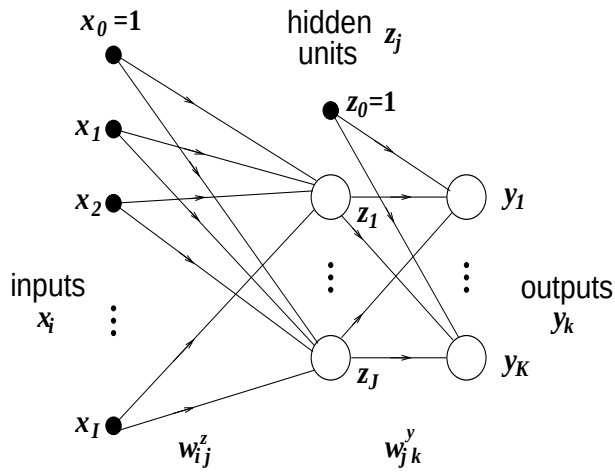

Figure 3.2: An $I - J - K$ MLP

$$f(y) = \frac{1}{1 + exp(-\beta y)} \tag{3.15}$$

The logistic sigmoid activation function has the advantage over the original Heaviside function that allows the outputs of the discriminant to be interpreted as posterior probabilities. This implies that such a discriminant is providing more than simply a classification decision, and is potentially a very powerful result [10].

Figure 3.1 shows a single-layer neural network. A single-layer neural network can be regarded as a generalised linear discriminant function. The input variables are linearly combined and transformed with a non-linear function. Single-layer neural networks have a number of limitations in terms of the range of functions they can represent. It is known however that networks with two layers of weights are capable of approximating any continuous functional mapping [10]. Multi-layered networks having either threshold or sigmoidal activation functions are known as Multi-Layer Perceptrons (MLP). An example of a two-layer MLP is shown in Figure 3.2. The units that are not output units are known as hidden units. MLPs are sometimes referred to as having an $I - J - K$ architecture , where I is the number of input parameters, J the number of hidden units and K the number of outputs.

3.6 Training

Learning or training consists of adjusting the weights $\mathbf{w}$ to minimise the classification error. Training of a neural network consists of gradually adjusting the weight vector $\mathbf{w}$ to minimise a cost function representing the error between the desired and actual outputs when the neural network is presented with a set of input data, known as the training set.

There are various choices for the error function. The simplest is the sum-of-squares error function:

$$E(\mathbf{w}) = \frac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{c}(y_k(\mathbf{x}^n;\mathbf{w}) - t_k^n)^2 \tag{3.16}$$

For a single-layer network this error function is a smooth function of the weights, a quadratic function, and hence its derivatives with respect to the weights are linear functions of the weights.

Training an MLP is a more complex procedure. In the original perceptron model the hard limiting activation function meant that there is no obvious way of attaching target values to the hidden layer units. Therefore it is impossible to determine the contribution of any hidden layer unit to the output error, and hence the weights between the input and hidden layers cannot be adjusted. If the hard-limiting activation function is replaced by a continuously differentiable function, such as the sigmoid function, then the activations of the output units become differentiable functions of both the input variables, and of the weights and biases [10]. If an error function is chosen that is a differentiable function of the network outputs, then the error is itself a differentiable function of the weights. The aim of the training process is to find the minimum point of the error function in weight space. The minimisation of non-linear, multi-dimensional functions is covered in Section 3.7. A computationally efficient algorithm for evaluating the derivatives of the error function is known as the error back-propagation algorithm. This is described fully in [10],(pages 141-148).

In the problems considered in this thesis, there are only two classes of patterns. It is shown for a two class problem [10],(p.230) that appropriate choices of activation and error functions

are the logistic activation function:

$$f(x) = \frac{1}{1 + \exp(-x)}, \tag{3.17}$$

and the cross-entropy error:

$$E = -\sum_n \{t^{(n)} \ln y^{(n)} + (1 - t^{(n)}) \ln(1 - y^{(n)})\}, \tag{3.18}$$

where $t^{(n)}$ is the desired output value for the n-th input vector and $y^{(n)}$ is the network output for that input vector.

3.7 Parameter optimisation techniques

Numerical algorithms for minimising an error function usually consist of iterative processes of the form

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \triangle \mathbf{w}_n \tag{3.19}$$

where $\triangle \mathbf{w}_n$ is the wight update, i.e., the change in weight from the current weight to the next weight. These numerical minimisation techniques face a number of implementation issues. Firstly, the algorithm needs to converge, and secondly an identified minimum may not be a true global minimum, merely one of many local minima.

The simplest minimisation algorithm is that of gradient descent. An initial weight vector $\mathbf{w}_0$ is chosen and the gradient of the error function is calculated with respect to each weight. The next weight vector $\mathbf{w}_1$ is obtained by moving a small distance in the direction of the steepest descent, i.e., along the negative of the gradient. For an individual weight w_i, the weight update $\triangle w_i$ is given by [101]:

$$\triangle wi = -\beta \frac{\delta E}{\delta w_i} \tag{3.20}$$

where β is a small, arbritary fixed parameter that sets the step size. The gradient descent method is in general an inefficient method.

Conjugate gradient methods use a line search strategy. This differs from gradient descent in that rather than moving a fixed step size in the direction of the local negative gradient of the error function, movement takes place along the direction of the local negative gradient to find the point at which the error is minimised. This gives an automatic procedure for setting the step length. At the new point the component in the search direction of the error gradient vanishes. A new search direction is chosen that does not "spoil" the minimisation achieved in the previous direction, i.e., that keeps the projection of the gradient in the previous direction null. Minimisation again occurs in this new direction, and the procedure is repeated iteratively. Scaled conjugate gradient methods include some further developments to the line search algorithm and these are described in [10], pages 282-285.

Single-layer neural networks can also be trained using the algorithms outlined above for MLPs. However, it is also possible to take advantage of the linear structure of the network and use a special-purpose training algorithm known as iterated re-weighted least squares (IRLS) [73].

In Chapter 7 the scaled conjugate gradient algorithm is used when training MLPs and the IRLS algorithm is applied to the training of the single-layer networks. The algorithms are implemented using the Netlab pattern recognition software [72].

3.8 Network architecture

The number of inputs I in an $I - J - K$ MLP is determined by the number of features or input parameters. The number of outputs K is equal to the number of classes (although a two-class problem only requires a single output unit). Too high a value of J forms a decision boundary that is too complex, the network tends to fit the noise in the training data, rather than learning the general mapping. A network with too high a number of hidden units does not generalise well on an unseen test set. Too small a value of J forms a network not capable of separating the different classes.

The choice of J is also influenced by the amount of training data available. The number of weights in a network increases linearly with the number of hidden units, therefore the

amount of training data also increases linearly. Tarassenko [101] recommends that there should be at least as many training examples as weights, and preferably an order of magnitude greater.

Training therefore consists in determining the optimal number J of hidden units. For every value of J considered in this search, at least ten different $I - J - K$ networks should be trained, with a different set of initial weights in each case. The different random initialisations produce different weight sets when training is stopped, each corresponding to a local minimum [101]. The optimal set of weights is the one chosen to give the best classification performance on a validation set.

3.8.1 Overfitting

A network that is over-trained learns the details of the training data rather than the underlying input-output mapping and is therefore likely to perform poorly when given new data that it has not previously seen [101]. "Early-stopping" is one method of preventing over-training. The overall training set is partitioned into a smaller training set and a validation set. After each training epoch is completed the classification error or Mean Square Error (MSE) of the validation set is assessed. When the validation error stops decreasing, or starts increasing training should be stopped.

Regularisation can also be used to prevent overfitting. Regularisation aims to prevent the network from modelling the noise in the training data by limiting the complexity of the decision boundaries. Regularisation involves adding a penalty term to the error function E which is designed to penalize mappings which are not smooth [10]. The new error function, E', is given by

$$E' = E + \nu R, \tag{3.21}$$

where E is the original error function, R is the regularisation term and ν is a parameter that determines the relative importance of the complexity of the boundaries with respect to the network performance on the training set.

One of the most common regularisation terms is known as *weight decay* and is given by

$$R = \frac{1}{2} \sum_i w_i^2. \tag{3.22}$$

This term is motivated by the idea that a complex boundary has a high curvature which, in turn, requires large network weights [101]. The sum-of-squares form of this regulariser encourages the weights to remain small.

The weight update equation is then

$$\nabla E' = \nabla E + \nu \mathbf{w}. \tag{3.23}$$

The regularisation term, ν, described here is applied to all weights individually. Regularisation is not normally used for bias weights so that they remain free to be modified in the face of offsets in input values [10].

The process of regularisation merely prevents the network from over-fitting the training data, it does not provide an indication of when the training process should be stopped. In the work described in Chapter 7, an initial study is carried out in which the misclassification error of the validation set is examined after each training epoch, the number of training epochs required being chosen on the basis of this inspection.

3.9 Data sets

The data available for a neural network study is partitioned into three groups, the training, validation and test sets. The training data is used for the supervised learning of the network, the validation set is used to decide when to stop training ("early-stopping") and to decide upon a suitable network architecture. The optimal network, in terms of performance on the validation set is then evaluated on the test data. Ideally the amount of data available allows a ratio of 1:1:1 between the training, validation and test sets. [101]. The strategy for partitioning the data also needs to be considered. The data could be partitioned randomly or if data is limited a cross-validation procedure may be used [10]. The data is divided into N segments. A network is then trained using data from $N - 1$ of the segments and its performance evaluated using the Nth segment. The process is repeated for each of the N

possible combinations of the data.

3.10 Data fusion

Data fusion is the term used to describe the integration and combination of data/information, either from multiple sources or from the same source over time.

Data fusion is useful in situations where a required parameter cannot be measured directly [96]. Examples where it is useful include remote sensing, target identification, machine condition monitoring and diagnosis, and more relevantly in non-invasive medicine [35],[31],[104]. It is a means of improving the performance of a pattern recognition or measurement system by integrating multiple sensors' data.

In the literature on data fusion, fusion is often described as taking place at one of three different levels of representation, namely [22]:

- **Data level:** combines (unprocessed) sensor data;
- **Feature level:** combines features extracted from different sensor data;
- **Decision level:** combines detections (or detection probabilities) from different sensors.

There are numerous methods that have been developed for all levels of data fusion. A complete review may be found in Chapter 2 of [6]. An overview is given in the following sections with a discussion of methods relevant to the work described in this thesis.

3.10.1 Data level fusion

One means of implementing signal or data fusion is by taking a weighted average of the composite signals or data, with weights based on estimated variances of the signals [6]. Whereas this method allows for real-time processing of dynamic low level data, in most cases, a Kalman filter is preferred because it provides a method that is nearly equal in processing requirements and in contrast to a weighted average, it results in an estimate for the fused data that is optimal in a statistical sense. For systems that have linear dynamics and linear observation models, the standard Kalman filter is the optimal (in the sense of the minimum

mean square error) estimator. It is also possible to utilise multiple models (with different dynamics or process noise models) to improve the reliability of the estimate. These approaches either switch between models (via a decision rule) or combine them via a weighted average based on their *a posteriori* likelihood of being the "correct" model. Kalman-filter-based fusion is considered in Chapter 5 and the underlying theory is presented there.

3.10.2 Feature level fusion

In practice, features are usually fused by combining features from multiple sources or sensors and forming a single input vector to a statistical classifier. Sanderson *et al.* [91] use a combination of features from speech and visual signals as an input to a Gaussian Mixture Model to design a person verification system. In building a landmine detection system Stanley *et al.* [95] combine features from two different types of sensors. This combination of features is used as an input to a neural network. Jimenez *et al.* [52] also use a neural network with a combination of features from different sensors to classify remote images.

3.10.3 Decision level fusion

Decision level fusion can be categorised into hard and soft decision level fusion; hard decisions being all-or-nothing declarations, while soft decisions utilise information about the confidence each sensor places in its individual decisions [43]. Soft decisions may be in the form of, for example, an *a posteriori* probability estimate, or in the form of a score or ranking from the decision maker. Whether the decision is hard or soft defines how the decisions are combined. Soft decisions are usually considered superior to hard decisions as they result in more information being used to make the ultimate decision.

Hard decisions are usually combined using majority voting [57] or logical **AND** or **OR** operators [111]. Soft decisions may be combined in a multitude of ways. At the simplest level, summation, averaging or weighted averages, products and geometric means of the individual decisions have all been suggested as a way of arriving at a resultant single soft decision. If the decisions are interpreted as fuzzy membership values, belief values or evidence fuzzy rules, belief functions and Dempster-Shafer techniques are used [57]. Finally it is possible to treat the decision making as a classification problem and train a classifier using the decisions

as new features.

A number of empirical studies have been carried out to compare different methods of combining soft decisions. For example, Verlinde *et al.* [111] compare the performance of several methods for combining the scores of a number of biometric identity verification systems. A variety of statistical classifiers including a k-Nearest Neighbour classifier, an MLP, a maximum *a posteriori* probability classifier based on Bayes decision theory and linear and quadratic classifiers are used, as well as **AND** and **OR** operators after thresholding the decisions to produce hard decisions. Although a Bayesian approach leads to an optimal classifier in the sense that it implements the lowest Bayes risk, in practice, poor *a priori* knowledge leads to inferior results. It is concluded that, for this application, a logisitic regression classifier gives the best performance with a simple **AND** operator giving comparable results.

Kittler *et al.* [57] suggest combining the decisions using product, sum, minimum and maximum rules. It is shown in two separate experiments, of different applications, that the sum rule outperforms all the individual classifiers and gives the best performance of the combined classifiers.

Cremer *et al.* [22] give a comparison of decision-level sensor-fusion methods for anti-personnel landmine detection. Both non-statistical fusion techniques, including Dempster-Shafer theory, fuzzy probabilities, rule-based threshold functions and voting methods are used as well as an optimal Bayes statistical technique. All the fusion methods perform better than the best single sensor, however no method is shown to be optimal in terms of sensitivity and specificity.

Duc *et al.* [28] compare a Bayesian approach with a geometrical mean method for combining two soft decisions from experts using either audio or video information for person identification. In this work it is shown that the Bayesian approach gives the better performance in terms of the total error rate.

3.11 Relevance to this thesis

Some of the signal processing techniques used throughout this thesis are described in this chapter. Feature extraction using AR models is described in Chapters 6 and 7. In Chapter 7 a neural network classification of these features is presented.

Chapter 5 describes the use of a Kalman filter based fusion technique to combine breathing rate measurements from the different primary signals. As pre-processing of the signals takes place before fusion it is argued this fusion occurs at a feature level.

A description of an approach that fuses hard decisions from two sources of information from the same signal is found in Chapter 6. The fusion of features is described in Chapter 7. Features are fused from two signals by concatenating the feature vectors from both signals before classification. Results of fusion at a soft decision level, combining the outputs of classifiers that use features from a single signal, are also reported.

Chapter 4

Detection of Breath-by-Breath Respiration

4.1 Introduction

This chapter describes an investigation into automated time-domain methods for detecting breath-by-breath respiration from the ECG, blood pressure, PPG and IP signals. The IP signal is sometimes used to display a respiratory waveform in clinical environments. However it is susceptible to artefacts, such as chest distortion resulting from the beating heart, and noise. It is also very sensitive to body movement [88].

The aim of this work is to assess a number of methods in order to determine which gives the most reliable and accurate measurement of breath-by-breath respiration. Ideally such a method is robust to inter and intra patient variability. Respiratory waveforms are derived from the signals and breaths are detected from these waveforms. Following this study, several of the derived respiratory waveforms are used in subsequent chapters of this thesis. The studies are carried out on the two databases described in Section 2.6, the Polysomnography and Controlled-breathing databases.

The chapter is structured as follows. In Sections 4.1.1 and 4.1.2 a review of previous work on extracting respiration from the ECG and PPG signals is given. Section 4.2 describes an objective assessment procedure, which is used to evaluate and compare the different signal

processing methods for deriving respiratory waveforms from the four original signals. These methods are described in detail in Sections 4.4-4.6, methods suggested in the literature are also implemented for comparison. The results of testing these methods on the Polysomnography and Controlled-breathing databases are presented in Section 4.7, and Section 4.8 draws conclusions from these results.

4.1.1 Previous work on extracting respiration information from the ECG

Much of the previous work on extracting respiration information from the ECG follows an approach of using direct measurements of respiratory-induced characteristics of the time-series ECG signal. These measurements are used to derive respiratory waveforms.

In their definitive paper, Moody *et al.* [71] outline a method for generating an ECG Derived Respiratory waveform (EDR) using two ECG leads. After subtracting the baseline, the area of each normal QRS complex in each of the two leads is measured over a fixed window (the width of which is determined during a learning phase to match the interval from the PQ junction to the J-point[1] of a normal QRS). Since the window width is fixed, the area is proportional to the mean amplitude of the signal, hence to the projection of the mean cardiac electrical vector onto the lead axis. Assuming that the leads are orthogonal, the arctangent of the ratio of the areas measured in the two leads gives the angle of the mean axis with respect to one of the axes. The axis direction measurement provides one sample of the EDR. A pseudo-continuous EDR is then obtained by interpolating using a cubic spline. The EDR is compared visually to chest Pneumatic Respiration Transducer (PRT) measurements.

Zhao *et al.* [117] use a similar, if not identical, method to Moody *et al.* [71]. However they use a more analytical method to determine how well the derived waveforms correspond to the recorded respiration signal. The power spectra (over a two minute period of fixed rate breathing) of both the EDR and PRT signals are obtained and compared statistically.

Travaglini *et al.* [107] again use the angle changes in the cardiac vector to derive an EDR. However in this work eight ECG leads are utilised. Their method is based upon the hypothesis that the breath-representative points are laid out around the preferred direction in

[1] the J-point is defined as the end of the S-wave.

an 8-D space. Once this main direction, named the "respiratory direction", is located, the method produces an EDR. A centre of gravity is defined for each lead as the average area of the relevant QRS complex for the first 16 beats. A vector of eight elements for every subsequent heart beat is then determined, the element for each lead being the area of the QRS complex minus the centre of gravity. The projection of this vector onto the respiration direction is then taken as the EDR sample for that point. Results are again assessed visually, by comparing the EDR to a PRT for 10 volunteers who underwent periods of normal breathing interspersed with breath holding.

Felblinger *et al.* [34] use the amplitude of the R wave to obtain an EDR. The height of each R peak is plotted as a step function to represent the respiratory signal. Results are compared visually against the reference respiratory signal, which was obtained from a pneumatic pressure belt.

Earlier work simply examined the power-spectra of the ECG to establish the presence of components at a respiratory frequency. Pallas-Areny *et al.* [78] describe a study aimed at showing that rotation of the cardiac vector introduces respiratory information in the power spectrum of the ECG. An ECG was recorded on a patient wearing a pacemaker (to eliminate the effects of RSA) and also undergoing artificial ventilation. The electrodes for standard leads were placed at extremities so that there were no respiratory-induced movements (baseline wander). The power spectra of leads I, II and III were obtained. Respiration harmonics were found to be present in the spectra. As the cardiac rhythm was constant, the presence of these harmonics was attributed to the displacement of the heart vector.

4.1.2 Previous work on extracting respiration information from the PPG

In comparison with the work done on deriving respiration from the ECG, very little has been done using the PPG. Lindberg *et al.* [58] employ a band-pass filter to eliminate all frequencies of the PPG other than the range in which respiration is likely to be present, the filtered signal being the derived respiratory waveform. The pass-band of the filter is chosen to be from 0.1 to $0.5Hz$. Subjects were kept in a supine position and asked to breathe normally. The number of breaths (peaks) recorded by the reference IP method are compared with the number of breaths (peaks) in the derived respiratory waveform over a ten-minute period.

Nakajima *et al.* [74] also use a band-pass filter approach. In their work the cut-off frequency of the low-pass filter is selected depending on the current heart rate; increasing from $0.3Hz$ to $0.55Hz$, as the heart rate rose. In order to assess the performance of this method, the median value of the breathing rate over five sequential breaths detected from the derived respiratory waveform is compared with the same measure from a reference transthoracic impedance plethysmogram. The study involved the subjects increasing their heart and breathing rates by pedalling on a stationary bicycle. It is found that the largest errors in derived breathing rate are at higher breathing rates, an error of 7 breaths min^{-1}, when the actual rate was 35 breaths min^{-1}. At lower breathing rates the results are more accurate.

4.2 Quantitative assessment of breath-by-breath respiration algorithms

In the previous work (described in Sections 4.1.1 and 4.1.2) that has attempted to derive a respiration measure from other physiological signals, for example the ECG [71],[107], results have been assessed visually by comparing the EDR with a reference respiratory signal (commonly using a pneumatic respiration transducer). The visual assessment is inevitably subjective and not easily reproducible. In their work on extracting an EDR from the MIT-BIH and AHA Arrhythmia Databases [48][62], Moody *et al.* [71] acknowledge that signals *qualitatively* similar to respiratory waveforms could be recovered in all cases. However, since no independent respiration measurements were available, the fidelity of the recovered signals could not be evaluated quantitatively. No quantitative, objective method for evaluating respiratory derivation algorithms has previously been reported.

An evaluation method that is used to assess the performance of different methods for the automatic detection of breath-by-breath respiration is described in the next section.

4.3 Evaluation method

Standardised methods exist for evaluating ECG analysis algorithms [1][2], but at present there are no similar standards for algorithms that derive a respiratory measure from other physiological signals. Using methods and statistics analogous to those published for ECG

analysers, a method for evaluating the performance of respiratory derivation algorithms is now described.

Most ECG analysis algorithms attempt to detect heart beats (i.e., the QRS complex of the ECG) before classifying these beats and performing subsequent analysis. The Association for the Advancement of Medical Instrumentation (AAMI) has published standards for testing beat-by-beat analysis algorithms so that meaningful comparisons between different algorithms may be performed [1],[2]. When detecting beats it is specified that there must be a close temporal match between the beat label in a reference annotation file and that produced by the algorithm under test. This beat matching must occur within a time window of $\pm 150ms$.

In this work, the aim is to detect breaths from a Derived Respiratory waveform (DR). By introducing a labelling scheme whereby a salient point of the respiration cycle, e.g. the start of the cycle, is labelled, a similar approach to the detection of heartbeats in the ECG may be used. The time at which a breath is detected by the algorithm under test is compared with the time of the corresponding 'gold standard' reference breath annotation which is derived from the available respiratory signal. As a breath interval is greater than a heart rate interval, the time window (about the reference time) within which the derived breath should be found is expanded with respect to that used in ECG analysis.

To illustrate this, Figure 4.1 shows a reference respiratory signal (lower trace) and an example EDR (upper trace). The EDR has been cubic-splined for qualitative appearance. The waveforms can be seen to exhibit an antiphase relationship. This is due to the fact that in certain leads the amplitude of the QRS complex increases with inhalation while in others it decreases. For example, in lead I the amplitude decreases with inhalation and in lead III it increases [117]. The times of the peaks of the reference signal serve as the breath labels and in this example the troughs of the EDR are the salient points labelled as breaths by the algorithm under test. The times of the derived breaths can then be compared with the times of the corresponding reference breaths.

As in the standards for ECG analysis the detection statistics employ the mutually exclusive categories of True Positives (TP), False Positives (FP), and False Negatives (FN) that are well-recognised in medical testing [41]. A TP is a derived breath that falls within the time-match window associated with the reference breath label. A FP is a derived breath which

has no corresponding breath in the reference annotation file, and a FN occurs when a breath in the reference annotation file is not matched by a corresponding derived breath.

To evaluate how well an algorithm performs, the results are recorded in terms of sensitivity and positive predictivity. Using positive predictivity ensures that high values of sensitivity caused by high false positive rates are clearly identified.

- *Sensitivity* (Se) is the fraction of real events (breaths) that are correctly detected:

$$Se = \frac{TP}{TP + FN} \times 100\% \tag{4.1}$$

- *Positive Predictivity* (+P) is the fraction of detections that are real events (breaths):

$$+P = \frac{TP}{TP + FP} \times 100\% \tag{4.2}$$

4.3.1 Databases for detection of breath-by-breath respiration

Breath-by-breath respiration is obtained using the methods described in the following sections on a subset of the two databases used in this thesis. The assessment procedure of Section 4.2 is used to evaluate each of the algorithms and to allow for a fair comparison between these different methods.

The presence of apnoeas in the Polysomnography database creates complications in the objective evaluation of the respiratory derivation algorithms. In view of this a database is

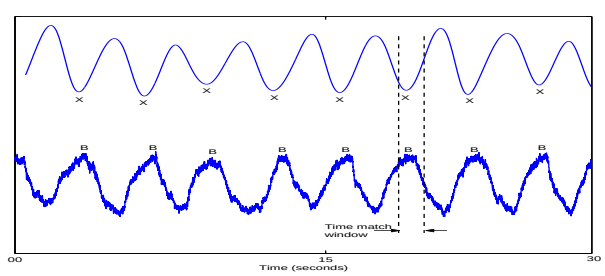

Figure 4.1: A reference respiration signal (lower trace), obtained from a nasal thermistor and an EDR (upper trace), the EDR was cubic-splined. The signals have been manually scaled to permit visual comparison, hence no vertical axis scales are provided. The signals are taken from record slp02b in the Polysomnography database.

Table 4.1: Time durations of the longest sections of non-apnoeic clean data for each record in the Polysomnography database.

Record	Start Sample No	End Sample No	Duration(mins)
slp02b	205853	464176	17.2
slp03	2748292	3472103	48.25
slp04	3832987	4393917	37.40
slp14	569743	998036	28.55
slp16	2908003	3206856	19.92
slp32	1875243	1906295	2.07
slp37	2515411	2574371	3.93
slp41	3498887	4559463	70.71
slp45	5362042	5691310	21.95
slp48	2756692	3595661	55.93
slp59	298131	425570	8.50
slp60	1571615	1888612	21.13
slp66	1400280	1629701	15.29

created consisting of only non-apnoeic data. The data is first selected using the annotation files, that are part of the Polysomnography database, to find the longest sections of data from each record that do not contain apnoeas. This data is further examined visually and any sections of non-apnoeic data that do not contain very clear reference respiration signals are also discarded. This is to ensure that all reference labels are as accurate as possible. Table 4.1 shows the duration and location of the selected data for each record of the Polysomnography database. As explained in Section 2.1 this database does not contain a PPG signal. An intra-arterial blood pressure signal is used in this work for proof of concept.

Appendix A gives the protocol of the data collection procedure for the Controlled-breathing database. For the work in this chapter three separate data sets are extracted from the overall database, corresponding to different data collection conditions. Data Set A includes all data, apart from periods of breath holding and sections for which the subjects were asked to breathe at a rate which was as fast or faster than the cardiac rate. The data in Data Set A contains breathing at different rates, from 6 to 20 breaths per minute, and different depths of inspired air, from 300 ml to 1500ml. It includes a period during which the subjects undertook light exercise, as well as periods during which the subjects breathed normally. Data Set B consists of four minutes of resisted breathing. During these times the subjects were asked to breathe through resistance tubes at normal rates. Finally Data Set C consists of 15-minute periods during which the subjects breathed at their own normal (relaxed) rate and depth.

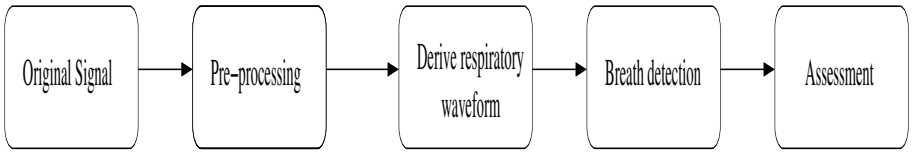

Figure 4.2: Schematic diagram of the method of deriving breath-by-breath respiration from the original signals; ECG, blood pressure and PPG.

4.4 Overview of signal processing and evaluation procedure

The following sections of the chapter describe a number of different methods of deriving breath-by-breath respiration measures from the ECG, blood pressure, PPG and IP signals. Two general approaches are taken. The first approach makes direct measurements of the characteristics in the signals that are known to be modulated by respiration. These respiratory-induced characteristics are described in Sections 2.2.1, 2.3.1 and 2.4.1.

A second, more conventional, signal processing approach is used for comparison. The signals are band-pass filtered, leaving only components in the frequency range in which breathing is normally expected to occur. Both approaches result in derived respiratory waveforms.

For evaluation, time stamps for each breath are then defined from the derived waveforms and compared against the reference breath time stamps, as described in Section 4.3. The generation of the reference breath time stamps is described in Chapter 2. The breath time stamps identified from the IP signal, that is often used to display a respiratory waveform in a clinical environment, are also compared with the reference breaths.

Figure 4.2 summarises the sequence of processes for deriving and assessing breath-by-breath respiration from the original signals (ECG, blood pressure, PPG and IP). After pre-processing methods for deriving a respiratory waveform are carried out on each signal. Once the waveform is derived, breath times are defined by automatically searching for the peaks in the signal. To assess the performance of each algorithm, the breath times are compared with the known reference breath times.

Section 4.5 describes the pre-processing steps. Section 4.6 outlines each of the methods for obtaining respiratory waveforms and Section 4.6.8 presents the generic peak detection algorithm used to define breath times from the waveforms.

4.5 Signal preprocessing

The three signals, ECG, blood pressure and PPG, used to derive respiratory waveforms are all filtered to eliminate any artefact (outside the breathing range) or unwanted components of the signal.

There are a number of noise sources that can cause interference in the ECG. These include mains interference ($50Hz$), muscle artefact and electrode contact noise as well as patient movement [8]. The ECG is band-pass filtered, the high-pass cut-off frequency being chosen as $0.1Hz$ to eliminate baseline wander. This assumes that the lowest possible breathing frequency is $0.1Hz$, 6 breaths minute^{-1}. The low-pass cut-off is $40Hz$ to eliminate mains noise and any other high-frequency noise whilst preserving characteristics of the ECG of interest, for example the QRS amplitude.

Artefacts in the PPG signal can be induced by any phenomenon which causes a transient change in the light received by the PPG sensor. Any variation in the optical coupling between the probe head and the subject or physiological changes which dynamically alter the amount of light transmitted (or received) give rise to what is commonly termed as motion artefact. A subject raising or lowering their hand whilst attached to a finger probe dynamically alters the pressure their finger exerts on the probe which in turn alters the optical coupling, whilst simultaneously causing a change in venous blood flow which also affects light transmission through the tissue [46]. The intra-arterial blood pressure signal is less prone to noise than the PPG, however some low frequency baseline wander is sometimes seen. Both the PPG and blood pressure signals are therefore band-pass filtered, from 0.1 to 40 Hz, as is the ECG. Finite Impulse Response (FIR) filters are used throughout this work, because they can be made symmetrical in form. This produces an ideal linear-phase characteristic, equivalent to a pure time delay of all frequency components passing through the filter, i.e., there is no phase (and hence shape) distortion.

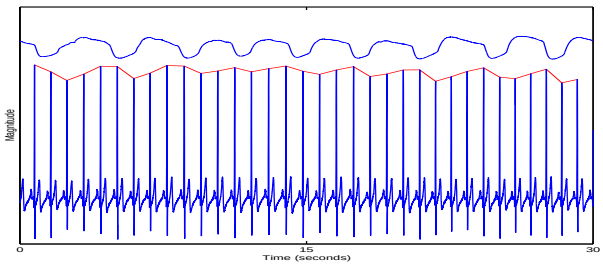

Figure 4.3: The figure shows an ECG, the R-DR (red), formed by tracing the height of the R-wave and the reference nasal thermistor respiratory signal (top trace) for record slp14. The signals have been manually scaled to permit visual comparison, hence no vertical axis scales are provided.

4.6 Deriving respiratory waveforms

4.6.1 Deriving respiration from QRS amplitude modulation of the ECG

As described in Section 2.2.1, the amplitude of the QRS wave is modulated as the heart rotates during the breathing cycle. The method described here makes direct measurements of the QRS amplitude to form an EDR. The position of each QRS complex in the ECG is first located. In the case of the Polysomnography database, the QRS annotations are used and a further scan made in a symmetrical window (of length 1-second) about the annotation to identify the exact position of the R-wave. The QRS complexes of the Controlled-breathing database are found using Engelese and Zeelenberg's 'sqrs' algorithm [33].

The amplitude of the peak of the R-wave is measured both with respect to the baseline (after high-pass filtering to reduce baseline wander) and with respect to the amplitude of the S wave. The latter is found by searching for the minimum value of the ECG in a time window, of length 0.1 seconds, beginning at the peak of the R-wave.

An EDR, referred to as the R-DR, is formed by plotting the difference in amplitude between the peak of the R-wave and the baseline, at the time of the occurrence of the R-peak. A waveform is created by interpolating between these samples. The RS-DR is formed in the same way but the samples of the derived waveform are the differences in magnitude between the R and S waves of the QRS complex. To illustrate the R-DR Figure 4.3 shows an ECG, a reference respiration signal (obtained using a nasal thermistor) and the R-DR.

Once the EDR is formed, the fiducial point corresponding to each breath must be defined. This point is either at the peak or trough of the respiration frequency EDR. Whether the EDR is in or out of phase with the reference respiratory signal depends on the lead used and the placement of the electrodes. For example, lead II is known to be affected by respiration and normally the QRS wave is a positive slope followed by a negative slope with the R-wave being at the most positive potential of the ECG (see Section 2.2). However if the positions of the two lead II electrodes are swapped the R-wave occurs at the lowest potential. The derived waveforms are analysed to determine whether they are in-phase or anti-phase with the reference respiratory signal in order to decide whether the peak or trough should be labelled as the breath time stamp.

4.6.2 Detecting breath-by-breath respiration from Respiratory Sinus Arrhythmia

Respiratory Sinus Arrhythmia (RSA), as described in Section 2.2.1, refers to the cyclic variation in heart rate which is associated with respiration. Heart rate accelerates during inspiration and slows during expiration [60]. The method uses a direct measurement of RSA to form an EDR, the RSA-DR. The times of the R-wave peaks are identified using the method in Section 4.6.1. To form a heart rate sequence the time between successive R-wave peaks (R-R interval) is measured in seconds.

Instantaneous heart rate is then defined as:

$$\text{Heart rate} = \frac{60}{\text{RR interval}} \text{beats min}^{-1} \tag{4.3}$$

There are a number of recognised methods for combining the R-R intervals in order to form a signal that represents heart rate. The method used here plots the value of R-R interval against the time at which the interval ends [26]. This heart rate signal forms the RSA-DR. Figure 4.4 shows an example of an RSA-DR and the corresponding nasal thermistor respiratory signal for a record from the Controlled-breathing database. Heart rate accelerates during inspiration. The times of the troughs of the RSA-DR are therefore defined as the breath times to be compared with the reference annotation labels (which correspond to the start of each respiration cycle).

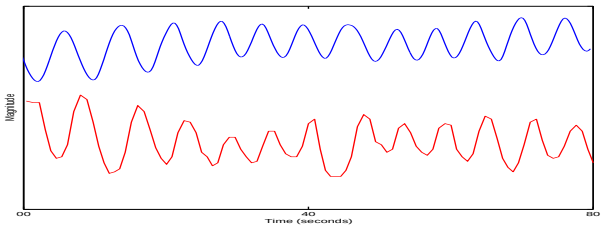

Figure 4.4: RSA-DR (upper trace) and nasal thermistor respiratory signal (lower trace) for a record from the Controlled-breathing database. The waveforms have been manually scaled to permit visual comparison, hence no vertical axis scales are provided.

4.6.3 Detecting breath-by-breath respiration using pulsus paradoxus

Pulsus paradoxus is described in Section 2.3.1 as the inspiratory decrease in systolic blood pressure signal. The following method uses a direct measurement of pulsus paradoxus in the blood pressure to form a derived respiratory waveform. The peaks of the blood pressure signal are first detected using the automated peak detection algorithm described in Section 4.6.8. Due to the presence of a dicrotic notch in the arterial blood pressure signal a magnitude threshold is used to identify the main peaks. An average of the peak-to-trough value in the first five minutes of the blood pressure signals from each patient in the Polysomnography database is found (after high pass filtering as described in Section 4.5 to eliminate baseline wander). After some investigation the amplitude threshold is chosen to be 66% of the peak-to-trough value in order to eliminate the dicrotic notch from the analysis.

The amplitude of each systolic peak provides one sample of a pulsus paradoxus-derived respiratory waveform (PP-DR). An example of a PP-DR and the corresponding respiratory signal are shown in Figure 4.5. As noted in Section 2.3.1, the minimum systolic blood pressure occurs during inspiration, and the maximum during expiration. The reference respiration labels for the Polysomnography database mark the start of the respiration cycle i.e., the start of inspiration. Therefore the times of troughs of the PP-DR are detected and defined as the breath time stamps.

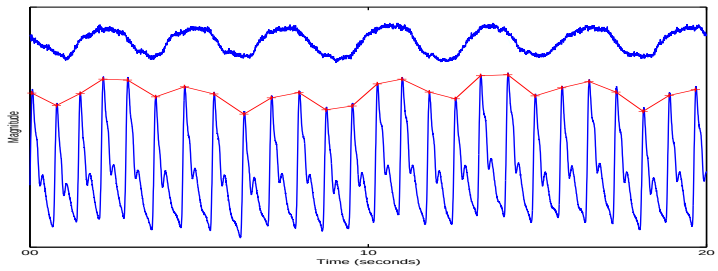

Figure 4.5: A section of blood pressure signal, with the PP-DR (red) and nasal thermistor signal (upper trace) from record slp02b. The waveforms have been manually scaled to permit visual comparison, hence no vertical axis scales are provided.

4.6.4 Detecting breath-by-breath respiration using blood pressure peak variability

The periodic pulses seen in the arterial blood pressure signal are due to the heart pumping blood from the aortic valve. The times at which the pulses arrive at the blood pressure measuring site therefore have a dependence on the heart rate. If the measuring site is a lag distance from the heart, other factors also start to influence the relative timing of the pulses. The stiffness and tension in the arterial walls are the principal factors determining the speed of transmission of the pulse wave, and these change as blood pressure changes [83]. However if the measuring site is close enough to the heart, these factors may be ignored and the times of the pulses assumed to depend only on heart rate. To form a derived waveform, the times of the peaks of the blood pressure signal are identified as in Section 4.6.3. The time between successive peaks is calculated. The peak interval is plotted against the time at which the interval ends. This signal then forms a BPV-DR. The times of the peaks of this waveform are then compared with the reference breath time stamps.

4.6.5 Detecting breath-by-breath respiration using pulsus paradoxus induced effects of the PPG

As discussed in Section 2.4.1, the effects of pulsus paradoxus are also seen in the PPG signal. A derived respiratory waveform is generated from the pulsus paradoxus effect on the PPG signal using the same method as for the blood pressure signal. The peaks of the pulse waves are first identified using the peak detection algorithm and the height of these peaks used as a sample of the derived waveform at the time at which the peak occurs. Figure 4.6 shows a

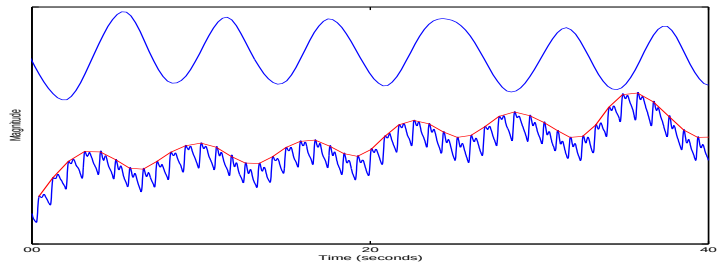

Figure 4.6: A PPG signal (lower trace), showing the pulsus paradoxus derived respiratory waveform and the reference nasal thermistor reference signal (upper trace). The signals have been manually scaled to permit visual comparison, hence no vertical axis scales are provided.

Table 4.2: The characteristics of the filters used in band-pass filtering to obtain respiratory waveforms.

Signal	Filter	Stopband Fc (Hz)	Passband Fc (Hz)	Attenuation (dB)	order
ECG (Polysomnography)	High Pass	$5 * 10^{-7}$	0.1	40	374
	Low Pass	0.5	0.4	40	355
BP (Polysomnography)	High Pass	$5 * 10^{-7}$	0.1	40	374
	Low Pass	0.5	0.4	40	355
PPG(Controlled-breathing database)	High Pass	$5 * 10^{-7}$	0.1	40	122
	Low Pass	0.5	0.4	50	138

section of a PPG signal, a respiratory waveform as described (red) and a reference respiratory signal obtained using a nasal thermistor.

4.6.6 Band-pass filtering

All of the signals considered contain a frequency component at the respiration frequency. Lindberg *et al.* [58] suggest deriving a respiration signal directly by band-pass filtering the PPG signal. The suggested high and low-pass cut-off frequencies are 0.1 Hz and 0.5 Hz respectively. This idea is used here and band-pass filtering is compared, as a means of deriving a respiratory waveform, with the non-linear methods described in the previous sections. Each of the signals used in this work is band-pass filtered at 0.1-0.4Hz. To enable the design of filters with a sharp roll-off in the transition band, the signals are downsampled by a factor of 10. Equiripple FIR filters are used in all cases and Table 4.2 shows the filter characteristics used for each of the signals.

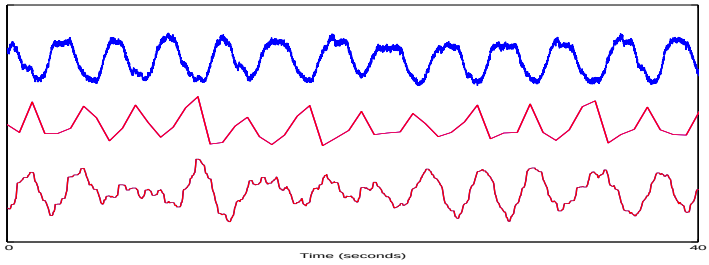

Figure 4.7: A derived respiratory waveform, generated through band-pass filtering (red lower trace), a derived respiratory signal using RS amplitude measurements (red middle) and a reference respiration signal (blue) obtained using a nasal thermistor, from record slp02b of the Polysomnography database. The waveforms have been manually scaled to permit visual comparison, hence no vertical axis scales are provided.

Figure 4.7 shows two ECG derived respiratory waveforms in red, the lower trace derived by band-pass filtering of the ECG and the middle waveform derived from RS amplitude measurements of the same ECG signal. The upper blue signal shows the reference respiration signal for this 40-second segment. Although the respiratory frequency can be seen in the band-pass derived waveform, other frequency components are also present, whereas the RS amplitude derived waveform tends here, to contain only the respiratory frequency.

The peak detection algorithm described in Section 4.6.8 is used to determine the time stamps for each breath from the band-pass filtered signal. Amplitude thresholds are introduced to eliminate the detection of spurious peaks, most of these being due to the fact that the original signals contain a number of components in the frequency range 0.1 Hz to 0.4 Hz other than the respiratory induced component. A good summary of the components that may occur in the PPG and blood pressure signals can be found in Lindberg [58].

4.6.7 Impedance pneumography (IP)

The IP signal was recorded as part of the Controlled-breathing database and is also available with a number of the records of the Polysomnography database. The IP signal is filtered between 0.1 Hz and 0.4Hz and breaths detected using the peak detection algorithm as for the other methods.

4.6.8 Peak detection algorithm

As illustrated in Figure 4.1, peaks of the derived respiratory signals are defined as the breath time stamps for comparison with the reference times.

Peak detection is one of the most commonly encountered problems in the digital processing of biological signals. Many schemes for peak detection can be found in the literature [30]. Those relevant to the detection of peaks in the time-series analysis of respiratory waveforms include [88] and [116]. The main difficulty lies in distinguishing between artefactual peaks (caused for example, by patient movement) and authentic peaks. To overcome this problem it is often suggested that arbitrary thresholds be set to eliminate spurious peaks. These thresholds may be applied to amplitude, slope or duration — or a combination of them [30].

The peak-trough detection algorithm used in this work is a rule-based method, using a gradient approach. There are only a few points for each breathing cycle in every derived respiratory waveform. After detrending, changes in sign of the gradient are searched for and a peak/trough is defined to occur at a given point if:

1. the appropriate change of gradient has occurred (i.e., positive to negative for peak and opposite for trough),
2. the previous extremum labelled is the opposite of that being detected currently,
3. the value of the point in question is above the mean for that section of data (or below for the trough),
4. the time elapsed since a similar extremum detected is greater than 2 seconds.

This simple peak-detection algorithm is found to work well. The waveform from which peaks are to be detected in this work are derived from the respiratory induced characteristics in other signals. The relative magnitude of these characteristics depend on the depth and rate of breathing. For example the modulation of the QRS wave in the ECG reduces with shallow breathing, and RSA diminishes with faster breathing rates. The derived waveforms are also susceptible to artefactual peaks introduced through body movement and contain components other than those due to breathing. These artefactual peaks are often of a comparable size to that of the peaks caused by respiration. An amplitude threshold is therefore very difficult to choose and tends to lead to a decrease in sensitivity and an increase in positive predictivity.

This is confirmed in an earlier study in which peaks are detected from an ECG derived respiratory waveform [63].

Peak detection from all the derived respiratory waveforms follows the algorithm described in this section. Any modifications in threshold levels are described where appropriate.

4.6.9 Choosing a window length

The assessment procedure described in Section 4.3 requires a time-match window around the reference breath within which each test breath must be found. The length of the window is investigated here. If too long a window is used, when shorter breath periods occur the reference windows may be longer than the breath period. This results in the overlap of windows of consecutive reference breaths. To ensure that each test breath is classified just once the assessment procedure is designed so that the detected test breath is classified with respect only to the closest reference breath. Extraneous breaths, whether they occur within the time window or outside the time window are counted as false positives. As all test breaths are classified once with reference to the nearest reference breath increasing the window size does not give rise to the same test breath being classified twice.

Figures 4.8 and 4.9 show the effect of varying the length of the time-match window when testing algorithms on the Polysomnography and Controlled-breathing databases respectively. The top plot of Figure 4.8 shows the effect on sensitivity and positive predictivity when expanding the time-match windows while using an ECG derived respiratory waveform to detect breaths. The lower plot shows the same parameters when using a blood pressure derived respiratory waveform. As can be seen from the plots, there is a significant improvement in performance for both methods when the windows are extended from 1 to 2 seconds (7% or 8% sensitivity and positive predictivity in both cases). However when the window is lengthened from 2 to 3 seconds, the performance improvement is negligible; less than or equal to 1%.

Figure 4.9 shows the effect of changing the window size from 1 second to 4 seconds when testing methods on the Controlled-breathing database. The top plot shows the results from a PPG derived respiratory waveform and the lower plot the results from an ECG derived algorithm. It can be seen from the figure that increasing the window size from 1 second to 2 seconds gives rise to an increase of over 20% in sensitivity and positive predictivity in both the PPG and ECG based methods. The increase from 2 seconds to 3 seconds results

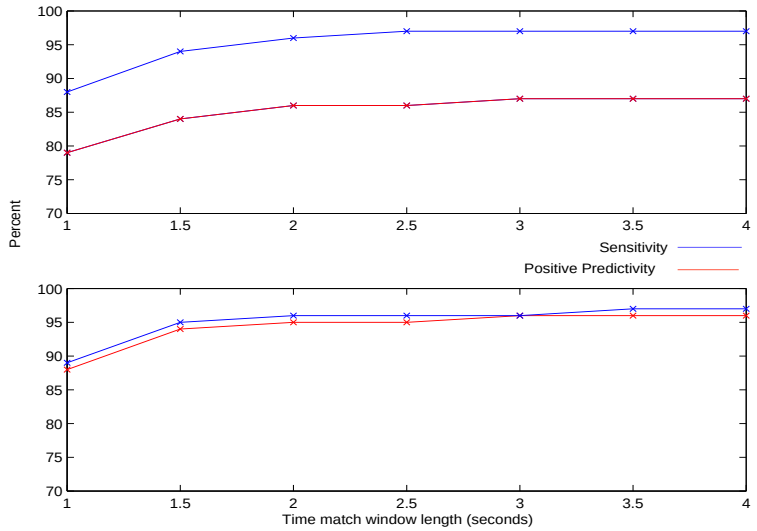

Figure 4.8: Sensitivity and positive predictivity as a function of window length for an ECG derived algorithm (upper trace plot) and a blood pressure derived algorithm (lower trace). The analysis was carried out on the Polysomnography database.

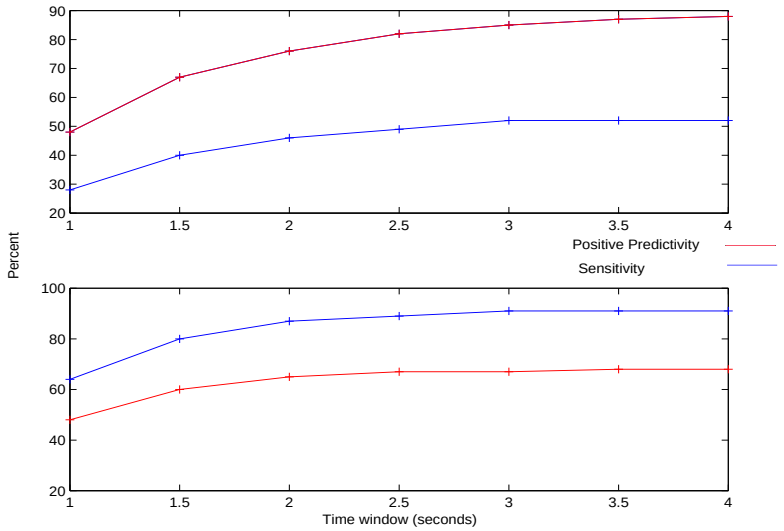

Figure 4.9: Sensitivity and positive predictivity as a function of window length for an ECG derived algorithm (top plot) and a PPG derived algorithm (lower trace). The analysis was carried out on the Controlled-breathing database.

in notable increases in the performance of the PPG method (9% sensitivity and 9% positive predictivity respectively), but smaller increases in the ECG method (3% sensitivity and 3% positive predictivity).

Figure 4.9 suggests that the phase between the respiration signal and the derived waveform is not constant and changes over time, i.e., increasing the window around the reference breath allows more true positive breaths to be detected. The mean breathing periods of the Polysomnography and Controlled-breathing databases are found to be 3.8 and 5.5 seconds respectively. Using this information, together with the results of Figures 4.8 and 4.9 a symmetrical window of 2 seconds is chosen when testing the algorithms on the Polysomnography database, while a 3 second window is chosen for the Controlled-breathing database.

4.7 Results

4.7.1 Polysomnography database

Tables 4.3 and 4.4 show the results of deriving breath-by-breath respiration using the ECG and blood pressure signals respectively. The breath time stamps calculated from the various derived respiratory waveforms are compared with the reference breath times which are obtained from a nasal airflow thermistor. A symmetrical time-match window of 2 seconds length before and after the reference breath is used.

Table 4.3 gives a summary of the results of ECG based methods. The first section shows the results obtained using the RS amplitude measure method (RS-DR), the R amplitude method (R-DR), and the RSA method (RSA-DR).

The second section shows the results of using the MIT EDR (downloaded from [113]) to compare this alternative method of deriving an EDR with the methods proposed in this work. The MIT EDR method is described in Section 4.1.1. It should be noted that the MIT algorithm is used in the exact form in which it is found on the web and that the default value of an 80 $msec$ window over which to take the area of the QRS complex is used (i.e., 40 $msec$ before and after the QRS time). The times at which the QRS complex occur are taken as those available in the annotation files available with the Polysomnography database. The MIT algorithm is used for each subject in its default form; it is expected that making it adap-

Table 4.3: Summary of gross results of different methods of deriving breath-by-breath respiration from the ECG on the Polysomnography database.

Method	TPs	FPs	FNs	Se(%)	+P(%)
RS-DR	5028	906	180	97	85
R-DR	4840	953	369	93	84
RSA-DR	4314	762	939	82	85
MIT-EDR	4656	2124	554	89	69
ECG band-pass	4556	1058	654	87	81

Table 4.4: Summary of gross results of different methods of deriving breath-by-breath respiration from the blood pressure on the Polysomnography database.

Method	TPs	FPs	FNs	Se(%)	+P(%)
PP-DR	4982	179	220	96	97
BPV-DR	4498	650	712	86	88
Blood pressure band-pass	4814	940	396	92	84

tive proffers better results. Moody *et al.* [71] also report that the method works best when two orthogonal ECG leads are available, here it is used on the one available lead.

It can be seen from the table, that of the ECG based methods those which use measures of the QRS amplitude, the RS-DR and R-DR give the best performance in terms of both sensitivity and positive predictivity, with the RS-DR slightly out-performing the R-DR. The other methods show comparable results, with the MIT-EDR showing a lower positive predictivity.

Table 4.4 compares the results for different methods of deriving breath-by-breath respiration from the invasive blood pressure signal. The pulsus-paradoxus method PP-DR is shown as the first result, the peak interval time method BPV-DR as the second and the result of band-pass filtering the signal is shown as the third result. It is seen that the pulsus-paradoxus based method PP-DR performs extremely well, giving a sensitivity of 96% and positive predictivity of 97%. The BPV-DR gives results comparable to the RSA-DR, an ECG based method, which is expected as they are based on the same physiological phenomenon, i.e., Respiratory Sinus Arrhythmia. The band-pass filtered method gives a high sensitivity but a lower positive predictivity due to other frequency components in the filtered range.

Table 4.5 shows the results of detecting breaths from the IP signal and again comparing them with the reference breaths obtained from the nasal thermistor signal. Not all records of the Polysomnography database contain an IP signal. The IP gives results comparable with the better ECG based methods but does not perform as well as the best blood pressure method

Table 4.5: Breath-by-breath detection derived from the band-pass filtered IP signal: Total number of breaths 2648.

Method	TPs	FPs	FNs	Se(%)	+P(%)
IP	2482	313	166	93.73	88.80

(PP-DR).

4.7.2 Controlled-breathing database

Tables 4.6, 4.7 and 4.8 give summaries of the results for detecting breath-by-breath respiration from the Controlled-breathing database. Methods using three non-invasive signals, namely the ECG, PPG and IP signals, are compared.

The window size used is a symmetrical window of 3 seconds either side of the reference breath (in comparison to the window used on the Polysomnography database which consists of a 2 second symmetrical window).

Each of the tables is made up of three sections, corresponding to different data collection conditions, as described in Section 4.3.1. Data Set A is taken from all data, apart from periods of breath holding and sections for which the subjects where asked to breathe at a rate which was as fast or faster than the cardiac rate. Subjects breathed at different rates from 6 to 20 breaths per minute, and different depths of inspired air, from 300 ml to 1500ml. The subjects also underwent light exercise, as well as periods of normal relaxed breathing. Data Set B consists of four minutes of resisted breathing. During these times the subjects were asked to breathe through resistance tubes at normal rates. Finally, Data Set C consists of 15-minute periods during which the subjects breathed at their own relaxed rate and depth. The aim of breaking up the data into sections for analysis is to establish whether different conditions have an effect on the methods. However as can be seen from the tables the results are consistent across the different breathing conditions.

Looking at the results for the ECG derived waveforms in Table 4.6 the methods based on QRS amplitude measurements show greater sensitivity than the method that uses RSA. However the RSA based method consistently shows greater positive predictivity. The ECG band-pass filtering method consistently shows a lower sensitivity than the QRS amplitude based methods with a comparable or lower positive predictivity.

Table 4.6: Comparison of obtaining breath-by-breath respiration from the ECG, using both direct measurements and band-pass filtering methods.

METHOD	TPs	FPs	FNs	Se(%)	+ P(%)
Data Set A: All non-apnoeic data (13907 breaths)					
RS-DR	12610	5791	1297	90.67	67.87
R-DR	12224	4394	1683	87.90	73.56
RSA-DR	11630	2709	2277	83.63	81.11
Band-Pass Filt 0.4-0.5 (50dB atten)	11953	6571	1954	85.95	64.53
Data Set B: Resisted breathing (637 breaths)					
RS-DR	579	196	58	90.88	74.71
R-DR	566	139	81	87.28	80.0
RSA-DR	538	105	99	84.46	83.67
Band-Pass Filt 0.4-0.5 (50dB atten)	545	301	92	85.56	64.42
Data Set C:Normal breathing (2041 breaths)					
RS-DR	1918	852	123	93.97	69.24
R-DR	1884	583	157	92.31	76.37
RSA-DR	1821	313	220	89.22	85.33
Band-Pass Filt 0.4-0.5 (50dB atten)	1818	795	216	89.38	69.58

Table 4.7: Comparison of obtaining breath-by-breath respiration from the PPG, using both amplitude of the PPG pulse measurements and band-pass filtering methods.
[Comparison of obtaining breath-by-breath respiration from the PPG, using both amplitude of the PPG pulse measurements and band-pass filtering methods.

METHOD	TPs	FPs	FNs	Se(%)	+ P(%)
Data Set A: All non-apnoeic data (13907 breaths)					
PP-DR	9164	1559	4745	65.89	85.6
Band-Pass Filt 0.4-0.5 (50dB atten)	9088	2166	4821	65.34	80.75
Data Set B: Resisted breathing (637 breaths)					
PP-DR	456	46	181	71.59	90.84
Band-Pass Filt 0.4-0.5 (50dB atten)	452	88	185	70.96	83.70
Data Set C: Normal breathing (2041 breaths)					
PP-DR	1401	109	633	68.88	92.78
Band-Pass Filt 0.4-0.5 (50dB atten)	1358	266	676	69.80	83.62

]

Table 4.8: Breath-by-breath detection derived from the band-pass filtered IP signal of the Controlled-breathing database.

METHOD	TPs	FPs	FNs	Se(%)	+ P(%)
Data Set A: All non-apnoeic data (13907 breaths)					
IP	11693	3987	1849	86.35	74.57
Data Set B: Resisted breathing (637 breaths)					
IP	550	206	87	86.34	72.75
Data Set C:Normal breathing (2041 breaths)					
IP	1900	558	134	93.41	77.30

From Table 4.7 the two PPG methods, the pulsus-paradoxus method PP-DR and band-pass filtered method show similar sensitivities across all data sections, with the PP-DR giving higher positive predictivities. The PPG methods are most successful (in terms of both sensitivity and positive predictivity) when used on resisted breathing. This is in agreement with the reported clinical papers (see Section 2.4.1) that describe the increase in pulsus paradoxus which occurs in resisted breathing i.e., breathing difficulties such as asthma. The PPG based methods give rise to a much lower sensitivity than the ECG based methods. The PPG methods however, give a very high positive predictivity. This suggests that at times the effect of pulsus paradoxus is not significant enough to modulate the PPG signal. It is also important to note that this work is carried out on healthy subjects none of whom had diagnosed respiratory difficulties. Pulsus paradoxus is therefore not expected to be prominent.

Deriving breaths from the IP signal gives lower sensitivities but higher positive predictivities than the best performing ECG algorithm.

4.8 Discussions and conclusions

The work in this chapter investigates different methods for automatically detecting breath-by-breath respiration from four signals, the ECG, blood pressure, IP and PPG. Two sets of data are used. The Polysomnography database is recorded over-night from subjects with suspected sleep breathing-related problems. The Controlled-breathing database is recorded from healthy subjects. All methods are carried out on the same sections of data, and breaths detected in the same automated way in order to make the assessment procedure as fair and as objective as possible.

Some of the methods use direct measurements of the respiratory induced characteristics in the ECG, blood pressure and PPG signals described in Sections 2.2.1, 2.3.1 and 2.4.1. A simple band-pass filtering method is used as a comparison to these non-linear methods. As a benchmark method, breaths are detected from the IP signal, which is sometimes used to display a respiratory signal in a clinical environment.

Results are compared quantitatively using the assessment method described in Section 4.2. This assessment method is used to make the comparison between respiration derivation algorithms as objective as possible. It is acknowledged that the method of detecting breaths,

i.e., peak detection, is a source of subjectivity. The same algorithm is used for detecting the breaths from each algorithm. The possible use of amplitude thresholds is discussed. Setting amplitude thresholds for the peak detection from these derived waveforms is however problematic. The waveforms are all derived from biomedical signals which are known to contain respiratory induced components. The magnitude of these components show inter-patient variability as well as intra-patient variability changing unpredictably over time due to any of the following: patient movement, sensor movement and breathing rate and depth. An earlier study carried out and published [63] showed that, as suspected, an amplitude threshold results in a trade off between sensitivity and positive predictivity.

4.8.1 Comparison of results on the two databases

Looking first at the ECG derived methods, it is seen that the methods using direct measurements of the QRS amplitude give better results (higher values of both sensitivity and positive predictivity) on the Polysomnography database than on the Controlled-breathing database. The reasons for this may be due to the fact that the position of the ECG electrodes in the Controlled-breathing database is chosen to give an optimum IP signal, also generated from these electrodes rather than in the position (nominally lead II) that sees the biggest effect of the cardiac axis change with breathing.

The second direct measurement method, the RSA method, gives comparable results on both databases. This is despite the databases being collected from different subject groups: one consists of young healthy males, who are expected to exhibit RSA, the second being older males suffering from sleep disorders. RSA is not always found in less healthy individuals.

When deriving respiration from the blood pressure (Polysomnography database) and PPG (Controlled-breathing database) signals, far better results are obtained using the blood pressure signals . The reasons for this are thought to be two-fold,

- the blood pressure signal is intra-arterial, while the PPG signal is a non-invasive, indirect measurement of blood flow. The PPG is prone to more artefact as well as being less sensitive to the sometimes subtle changes in systolic blood pressure due to respiration than the invasive blood pressure signal.
- the Controlled-breathing database (PPG signal) was recorded from healthy subjects

with no known breathing related difficulties, while the Polysomnography database was recorded from subjects with known or suspected breathing difficulties. Therefore pulsus paradoxus is expected to be much more pronounced in subjects from the latter.

In all cases the methods that make direct measurements of the respiratory induced effects outperform the band-pass filtering methods. It is suggested that this is because the band-pass filter used has a fixed pass-band to allow for a range of breathing frequencies. The original signals contain other frequency components in this range. This makes automatically detecting breaths a much harder problem due to other peaks in the resultant signal.

The IP signal, which is sometimes already used to display a respiratory waveform does not show significantly better results than the best performing non-invasive alternative methods.

Exploring further the meaning of the values of sensitivity and positive predictivity in relation to the breathing rate obtained from these methods, the results of the R-amplitude method are used as an example. Results of 88% and 74% of sensitivity and positive predictivity respectively are obtained on the Controlled-breathing database. If the errors in breath detection occurred evenly across the entire database over a 1-minute data section, when a subject is breathing at a normal rate of 12 breaths per minute, there would be on average 12% breaths not detected and about 26% extra breaths i.e., an extra $\sim$1.5 breaths detected. However a visual inspection finds that many of the extra detected breaths occurred together due to noisy sections of the ECG. The results in this chapter give an overall indication of the number of breaths that could be correctly detected. The results give no indication of where the errors occur or whether they tend to occur together in time or manifest as constant off-set errors.

To address these questions and also investigate the possibility of combining information from different methods of obtaining breath-by-breath respiration a further study is carried out. This study is described in Chapter 5.

4.8.2 Conclusion

The aim of the work in this chapter is to investigate how accurately breath-by-breath respiration can be found from one of a number of biomedical signals. The results are compared with those derived from an IP signal. While it is possible to obtain a sensitivity of 96% and

positive predictivity of 97% when using the intra-arterial blood pressure signal ultimately the measurement technique would be non-invasive. Of the non-invasive methods no single technique results in a significant improvement in deriving breaths over using the IP signal.

Chapter 5

Breathing Rate

5.1 Introduction

An evaluation of methods for the automatic detection of respiration, on a breath-by-breath basis, from several sources of information is given in Chapter 4. It is seen that an invasive signal, the arterial blood pressure signal, gives an almost perfect breath-by-breath measure of respiration but outside of the intensive care unit it is not ethical to monitor patients invasively.

The aim of this work is to derive robust and accurate breathing rates from the information available from non-invasive sensors. In this chapter a novel method that uses a Kalman filtering technique is presented for combining breathing rate measurements from multiple sources (derived respiratory waveforms) to obtain a more accurate estimate than is available from any individual source.

The individual accuracy of the methods reviewed in Chapter 4 for tracking breathing rate over time are first considered. The evaluation parameters for assessment and comparison are presented in Section 5.3. A Kalman filter method to estimate the state of a system from a single source of noisy measurements is introduced in Section 5.6. This method is used to derive improved estimates of breathing rate from each of the single source methods of Chapter 4. Section 5.8 presents specific methodology and models for the estimation of breathing rate using Kalman filtering. Section 5.9 then introduces a novel method of fusing

the information from the different methods. Results of the fusion method are presented and discussed in Section 5.10 and compared with traditional fusion methods as well as the estimates from a single source of information. It is seen that the fusion method results in more accurate estimates than any single source of measurements. Finally conclusions and suggestions for further work are presented in Section 5.11.

5.2 Estimating breathing rates

Chapter 4 investigates a number of methods for automatically deriving breath-by-breath respiration from the PPG and ECG in the Controlled-breathing database. These methods use characteristics induced in these signals in order to derive respiratory waveforms. From the derived waveforms breath times are defined. The performance of these methods for deriving breathing are compared with that of detecting breaths from the IP signal, a conventional method for obtaining a respiratory waveform non-invasively. The reference respiration signal was derived from an oral airflow thermistor.

Four of these methods, each based on a different breathing-induced characteristic, are investigated here for tracking breathing rate over time. The methods investigated are those termed the R-DR, RSA-DR, PP-DR and IP methods in Chapter 4. Collectively the methods are referred to as the "four respiratory waveform methods". Individually the methods are referred to as the R-Amplitude (R-DR), RSA (RSA-DR), PPG (PP-DR) and IP methods. In the first three cases this indicates the phenomenon from which the respiratory waveform arises. These methods have been chosen as they each rely on a different source of respiration information.

The aim of this chapter is to develop an algorithm for deriving the most accurate estimate of breathing rate from the respiratory information that can be extracted using the four non-invasive methods.

5.2.1 Instantaneous breathing rate

Once the breath time is calculated from the derived respiratory waveform the instantaneous breathing rate can be obtained. The instantaneous breathing rate is the rate calculated at

the instant at which the current breath is defined from the derived respiratory waveform. The breathing rate in breaths per minute (BPM) is defined as

$$\frac{60}{\text{time interval between current and previous breath (seconds)}} \tag{5.1}$$

The instantaneous breathing rate sequence can be viewed as an unevenly sampled time series. Figure 5.1 shows the instantaneous breathing rate sequences for one of the subjects from the Controlled-breathing database, obtained using each of the four respiratory waveform methods. The breathing rate sequences are superimposed on the reference breathing rate sequence (acquired from the oral air-flow thermistor). Visually the IP method shows the closest agreement with the reference breathing rate. The other three methods show good agreement in some sections but there are also times when the derived and reference rates deviate, in this example tending to over-estimate the breathing rate. This is due to artefactual peaks in the derived respiratory waveform, resulting in false positive breath detections.

Figure 5.2 shows data from a different subject. During the time shown the subject was asked to breathe for five-minute periods at rates of 6, 10 and 20 BPM. All the methods show some sections of reasonable agreement interspersed with noisy estimates. The RSA and R-Amplitude methods show a tendency to over-estimate the breathing rate at lower frequencies (6 BPM). The PPG method seems to be best at tracking this lower frequency. The IP and R-Amplitude methods show good agreement during the 10 and 20 BPM sections. The PPG method appears unable to follow the highest breathing rate.

5.3 Evaluation criteria for comparing methods of tracking breathing rates

In order to evaluate and compare the performance of each of the four respiratory waveform methods when estimating breathing rate, a number of performance parameters are defined. As discussed in Chapter 4, there is always a lead or lag between the reference breath time stamp and the time at which the breath is defined from the derived respiratory waveform. In Chapter 4 the evaluation procedure for methods that obtain breath-by-breath times defines a symmetrical time-match window within which the test method must identify a breath in

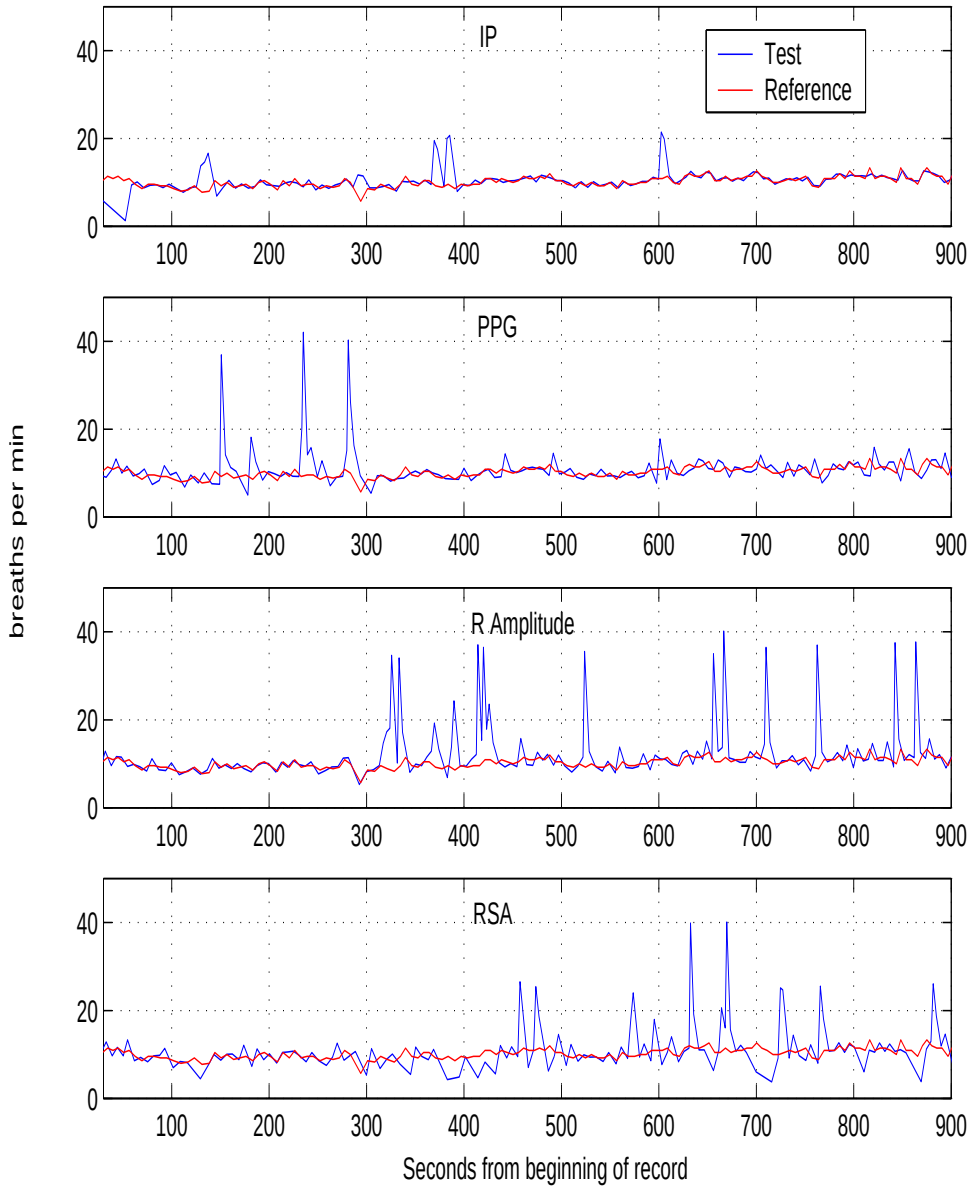

Figure 5.1: Instantaneous breathing rate sequences from each of the four respiratory waveform methods (blue), IP, PPG, R-Amplitude and RSA superimposed on the (red) reference breathing rate (obtained using an oral airflow thermistor) in each case. This data is from a period of time during which the subject was asked to breathe in a relaxed fashion at a rate and depth of their choice.

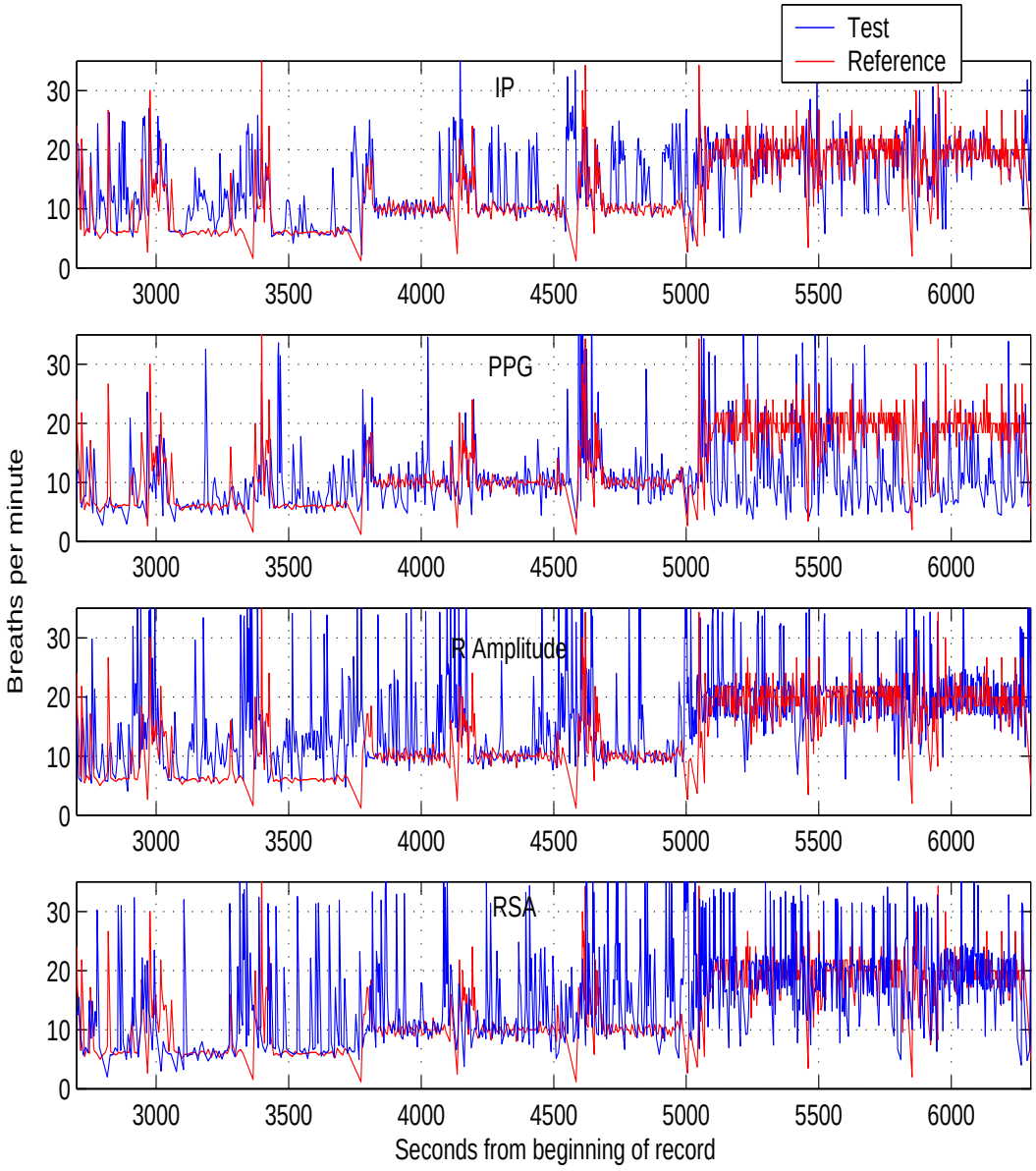

Figure 5.2: Instantaneous breathing rate sequences for record 3010. In each plot the red line corresponds to the reference rate (obtained using an oral airflow thermistor) and blue line is the breathing rate obtained from one of the test methods (labelled at the top of each plot). This data is from a period of time during which the subject was asked to breathe at a rate of 6,10 and 20 BPM.

order for a true positive to be declared. In the general clinical environment an instantaneous breathing rate is not required. Instead, the clinicians require the breathing rate over the last minute or, at best, the last 30 seconds. Hence the rest of this chapter concentrates on computing the average breathing rate over 30 seconds in order to assess the performance of each method. This also means that the time lag between the defined breath times and the reference time stamps is unimportant. The performance criteria that are examined are

- the overall correlation between the average breathing rates estimated by the test methods and the reference signal;
- the percentage error in the estimates.

5.3.1 Correlation between average reference and test breathing rates

Correlation enables us to measure the degree of linear relationship between two variables [18]. Pearson's correlation coefficient ρ is used:

$$\rho = \frac{\text{covariance of } x \text{ and } y}{\text{standard deviation of } x \times \text{ standard deviation of } y} \tag{5.2}$$

where in this case x is the reference breathing rate and y the test rate. Values of ρ lie between -1 and +1 with positive values implying that x and y are positively correlated (increase in x implies increase in y) and negative values implying x and y are negatively correlated (increase in x implies decrease in y). A value of +1 shows the data follows a relationship of the form

$$y = mx + c \tag{5.3}$$

where m and c are real numbers. Ideally in this case the relationship is as near to $y = x$ as possible. Once the correlation between the reference and test has been calculated, the statistical significance of the result is tested. The general procedures for significance tests can be found in [11],[18].

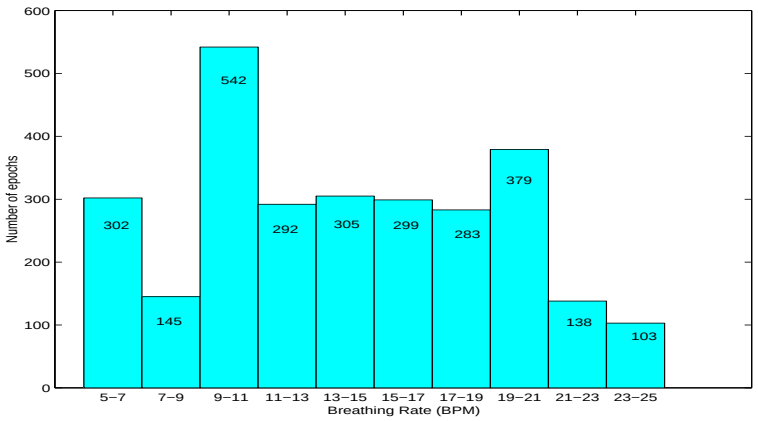

Figure 5.3: Bar chart to show distribution of averaged breathing rates of 30-second reference epochs.

5.3.2 Percentage error curves

Correlation gives an overall measure of how close the test breathing rates are to the reference breathing rates. The percentage error at different breathing rates are examined in order to investigate how well the methods perform at different breathing rates. The errors calculated are *absolute* values. Confidence intervals (95%) are also considered when comparing final results. Ideally percentage errors are as low as possible at all breathing rates.

5.3.3 The data on which the methods are tested

The different methods for estimating breathing rates are tested on the Controlled-breathing database (Section 2.6). All the data from this database is used with the exception of breathing rates known to be above 30 BPM. In evaluating the methods attention is paid to the sections of data corresponding to controlled breathing rates of 6, 10 and 20 BPM. An evaluation of the entire database is also undertaken. In all, 2788 30-second epochs (∼24 hours) of data are used from a total of 10 subjects. Figure 5.3 shows how these epochs are distributed across the breathing rate range.

5.4 Evaluation of the four respiratory waveform methods for tracking breathing rate

The performance of estimating breathing rates from the four respiratory waveform methods is evaluated in terms of the parameters defined in Section 5.3. These parameters are used as benchmark measures for the rest of this chapter.

The left-hand plots in Figure 5.4 are the scatter plots showing the test breathing rates with respect to the reference breathing rates for each of the four test methods. The corresponding right-hand plots show the Pearson correlation coefficient of the test data with the reference data represented as a horizontal black line. The correlation coefficients of the random permutations of the reference and test breathing rates are also given on these plots. These surrogate correlation coefficients are always much lower than the correlation coefficient of interest.

As the reference and test breathing rates are found to have approximate Gaussian distributions, the t significance test is carried out. In all cases the correlations are shown to be very highly significant (99.9%).

Figure 5.5 shows the percentage error curves for each of the four respiratory waveform methods. The mean percentage error (over all the subjects in the database) for a given band of breathing rates is calculated and plotted against the rate.

Table 5.1 shows the correlation coefficients for the four respiratory waveform methods for individual subjects in the database. The table illustrates how the different methods show different degree of success on individual subjects.

5.4.1 Discussion

Figure 5.4 shows that the IP method gives the highest correlation (ρ = 0.64) between the reference and test breathing rates, while the R-Amplitude method shows the weakest correlation (ρ = 0.36). The PPG and RSA methods have similar correlations of ρ = 0.45 and 0.48 respectively.

From Figure 5.5 it is seen that the PPG method gives the flattest percentage error curve, showing that the percentage error is fairly consistent across all breathing rates. The other

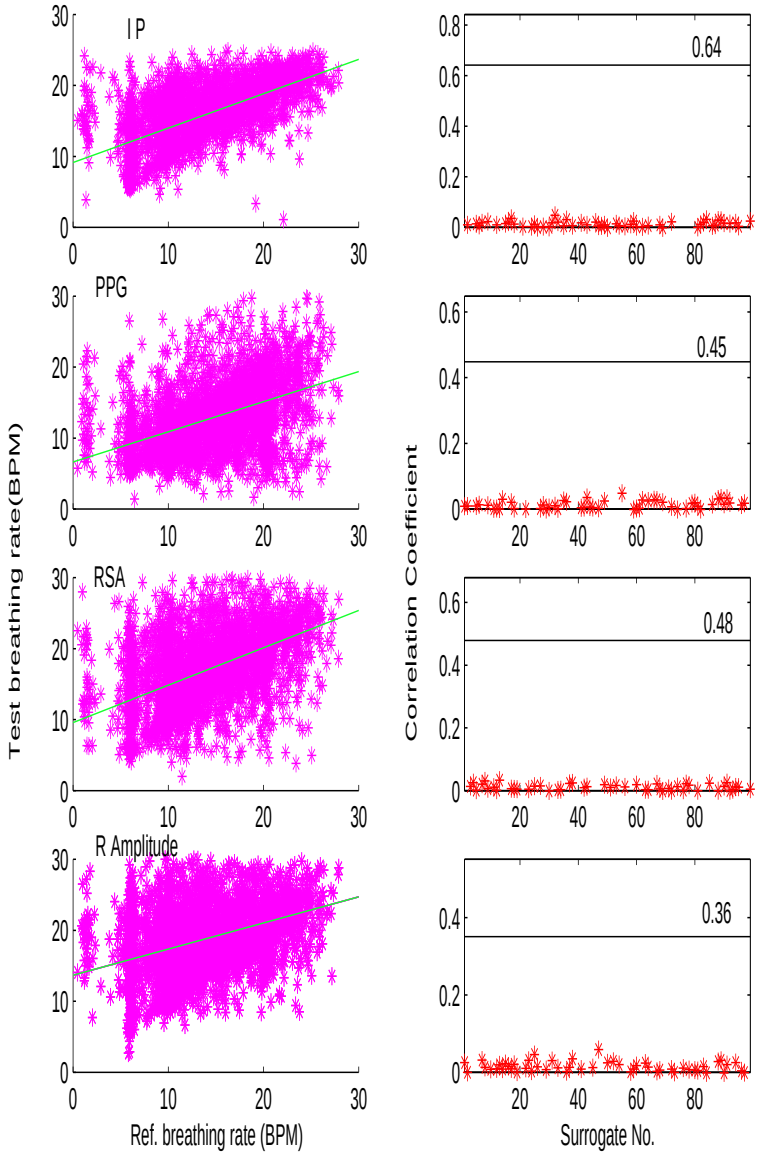

Figure 5.4: The left-hand figures show the scatter plots for the test breathing rates (y-axis) against the reference breathing rates (x-axis). The right-hand plots show the correlation coefficient level (horizontal black line on each plot) the asterisks being the correlation coefficients of the random permutations of the reference and test values.

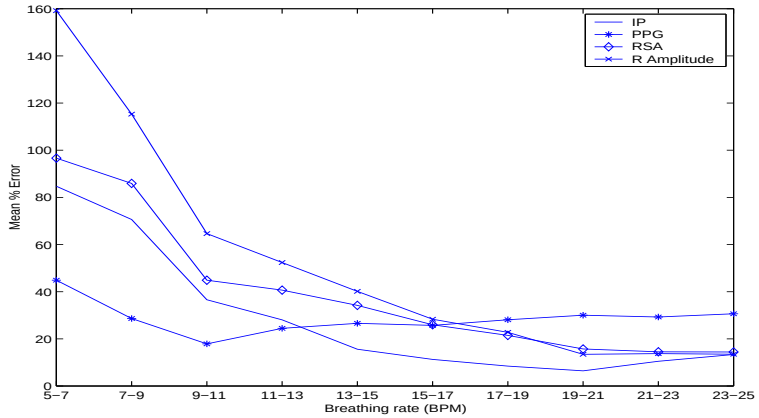

Figure 5.5: Mean percentage errors in breathing rate for each of the four respiratory waveform methods.

Table 5.1: Correlation coefficients of four respiratory waveform methods for individual subjects in the Controlled-breathing database. Each column corresponds to a different method for estimating breathing rate, the rows correspond to different test subjects. It can be seen that no single method performs well on all subjects.

Subject	IP	PPG	RSA	R Amp
3002	0.70	0.51	0.72	0.10
3003	0.80	0.39	0.31	0.31
3004	0.46	0.54	0.41	0.31
3005	0.75	0.51	0.60	0.37
3006	0.56	0.30	0.14	0.06
3007	0.37	0.29	0.57	0.44
3008	0.58	0.49	0.65	0.24
3009	0.55	0.66	0.47	0.51
3010	0.73	0.31	0.59	0.64
3011	0.81	0.29	0.52	0.43
Gross	0.64	0.45	0.48	0.36

methods show a trend of performing better at high breathing rates. The IP method, which gives the best performance overall, has a higher percentage error than the PPG method at rates below 11 BPM but at higher rates the IP method gives smaller errors.

The potential for high percentage errors in the R-Amplitude method is apparent from Figure 5.2. This figure shows the over-estimation of the breathing rate at the lower breathing rate by the R-Amplitude methods and to a lesser extent the RSA method. These results are also confirmed by those of Chapter 4, where it is seen that the R-Amplitude method shows a consistently lower positive predictivity than the other methods. A low positive predictivity occurs when extraneous breaths are defined (false positives), leading to an over-estimation of breathing rate and high percentage errors. From Figure 5.5 it is clear that these false positive detections mostly occur at lower breathing rates.

5.5 Optimal estimation

The results in Section 5.4 show that the four respiratory waveform methods produce noisy estimates of the breathing rate. The rest of this chapter investigates methods to estimate the state of the system (in this case the breathing rate) from these noisy measurements. Estimations from a single respiratory waveform method (source) are first investigated. The sources are all independent to some degree and the causes of artefact and noise do not affect each source in the same way. Therefore fusing information from more than one source to obtain a more accurate estimate is investigated. The derived respiratory waveform methods are referred to as “sources” of information in keeping with the literature on data fusion.

An optimal estimator is a computational algorithm that processes measurements to deduce a minimum error estimate of the state of a system by utilizing: knowledge of system and measurement dynamics, assumed statistics of system noise and measurement errors, and initial condition information[38]. Among the advantages of this type of data processor are that it minimises the estimation error in a well defined statistical sense and that it utilises all measurement data plus prior knowledge about the system. The corresponding potential disadvantages are its sensitivity to erroneous *a priori* models and statistics.

Probably the best known optimal filtering technique is that developed by Kalman [55].

5.6 Kalman filter

The discrete-time Kalman filter is a first-order recursive filter used to estimate the first and second order statistics of a signal in the presence of noise. A Kalman filter combines all available measurement data plus prior knowledge about the system and measuring device to produce an optimal estimate of the desired variables in such a manner that the error is minimised statistically [64].

An estimate, $\hat{x}$, is the computed value of a quantity, x, based upon a set of measurements, z. An *unbiased* estimate is one whose expected value is the same as that of the quantity being estimated. A *minimum variance* estimate has the property that its error variance is less than or equal to that of any other unbiased estimate. A *consistent* estimate is one which converges to the true value of x as the number of measurements increases. The Kalman filter is an unbiased, minimum variance, consistent estimator [38].

5.6.1 Derivation of Kalman filter algorithm

The Kalman filter is formulated using the state-space approach in which the system is described by a set of state variables. The state contains all the information regarding the system at a certain point in time. This information should be the least amount of data one is required to know about the past behaviour of a system in order to predict its future behaviour.

The proof and derivation of the Kalman filter are given in this chapter for completeness. For readers wishing to gain a brief overview, the equations needed to implement the filter are outlined in rectangular boxes. The familiar reader may wish to skip to Section 5.7. The notation and derivation follows that found in [16].

It is assumed the process to be estimated can be modelled in the form

$$\mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{w}_k \tag{5.4}$$

The observation (measurement) of the process is assumed to occur at discrete points in time

in accordance with the linear relationship

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \tag{5.5}$$

where:

$\mathbf{x}_k$ =$(n \times 1)$ process state vector at time t_k.

$\mathbf{A}_k$ =$(n \times n)$ matrix relating $\mathbf{x}_k$ to $\mathbf{x}_{k+1}$ in the absence of a forcing function.

$\mathbf{w}_k$ =$(n \times 1)$ vector — assumed to be a white (uncorrelated) sequence with known covariance structure, $\mathbf{Q}$.

$\mathbf{z}_k$ =$(m \times 1)$ vector measurement at time t_k.

$\mathbf{H}_k$ =$(m \times n)$ matrix giving the ideal (noiseless) connection between the measurement and the state vector at time t_k.

$\mathbf{v}_k$ =$(m \times 1)$ measurement error — assumed to be a white sequence with known covariance, $\mathbf{R}$ and uncorrelated with the $\mathbf{w}_k$ sequence.

The covariance matrices for $\mathbf{w}_k$ and $\mathbf{v}_k$ are given by

$$\begin{aligned} \mathbf{E}[\mathbf{w_k}\mathbf{w_i^T}] &= \mathbf{Q_k} \qquad i = k \\ &= 0 \qquad i \neq k \end{aligned} \tag{5.6}$$

$$\begin{aligned} \mathbf{E}[\mathbf{v_k}\mathbf{v_i^T}] &= \mathbf{R_k} \qquad i = k \\ &= 0 \qquad i \neq k \end{aligned} \tag{5.7}$$

$$\mathbf{E}[\mathbf{w_k}\mathbf{v_i^T}] = 0 \qquad \textit{for all } k \textit{ and } i \tag{5.8}$$

It is assumed at this point that we have an initial estimate of the process at some point in

time t_k, and that this estimate is based on all of our knowledge about the process prior to t_k. This *a priori* estimate is denoted $\hat{\mathbf{x}}^-$ where the circumflex denotes an estimate and the "super minus" is a reminder that this is our best estimate prior to assimilating the measurement at time t_k. It is also assumed that the error covariance matrix associated with $\hat{\mathbf{x}}_\mathbf{k}^-$ is known. That is the estimation error is defined to be

$$\mathbf{e}_\mathbf{k}^- = \mathbf{x}_\mathbf{k} - \mathbf{x}_\mathbf{k}^- \tag{5.9}$$

and the associated error covariance matrix (assuming the estimation error has zero mean) is

$$\mathbf{P}_k^- = \mathbf{E}[\mathbf{e}_\mathbf{k}^- \mathbf{e}_\mathbf{k}^{-\mathbf{T}}] = \mathbf{E}[(\mathbf{x}_\mathbf{k} - \hat{\mathbf{x}}_\mathbf{k}^-)(\mathbf{x}_\mathbf{k} - \hat{\mathbf{x}}_\mathbf{k}^-)^T] \tag{5.10}$$

In many cases, there are no prior measurements. Then, if the process mean is zero, the initial estimate is zero and the associated error covariance matrix is simply the covariance matrix of $\mathbf{x}$ itself. Given the prior estimate of the system state at time t_k, an updated estimate is calculated based on the use of a measurement $\mathbf{z}_\mathbf{k}$. This estimate is linear and recursive in form.

$$\boxed{\hat{\mathbf{x}}_\mathbf{k} = \hat{\mathbf{x}}_\mathbf{k}^- + \mathbf{K}_k(\mathbf{z}_\mathbf{k} - \mathbf{H}_k\hat{\mathbf{x}}_\mathbf{k}^-)}$$

where

$\hat{\mathbf{x}}_\mathbf{k}$ = updated estimate

$\mathbf{K}_k$ = weighting matrix, as yet unspecified.

The problem now is to find the particular weighting matrix $\mathbf{K}_k$ that yields an updated estimate that is optimal in some sense. The minimum mean-square (MMSE) error is used as the performance criterion. The error covariance matrix associated with the updated (*a posteriori*) estimate is

$$\mathbf{P}_\mathbf{k} = \mathbf{E}[\mathbf{e}_\mathbf{k}\mathbf{e}_\mathbf{k}^\mathbf{T}] = \mathbf{E}[(\mathbf{x}_\mathbf{k} - \hat{\mathbf{x}}_\mathbf{k})(\mathbf{x}_\mathbf{k} - \hat{\mathbf{x}}_\mathbf{k})^T] \tag{5.11}$$

By substituting Eq.5.5 into Eq.5.11 and then substituting the resulting expression for $\hat{\mathbf{x}}_\mathbf{k}$ into Eq.5.11 we obtain

$$\mathbf{P_k} = E[(\mathbf{x_k} - \hat{\mathbf{x}}_\mathbf{k}^-) - \mathbf{K}_k(\mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k - \mathbf{H}_k\hat{\mathbf{x}}_k^-)][(\mathbf{x_k} - \hat{\mathbf{x}}_\mathbf{k}^-) - \mathbf{K}_k(\mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k - \mathbf{H}_k\hat{\mathbf{x}}_k^-)]^T \quad (5.12)$$

Performing the indicated expectation and noting that $\mathbf{x}_k - \hat{\mathbf{x}}_k^-$ is the *a priori* estimation error that is uncorrelated with the measurement error $\mathbf{v}_k$ we have

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)_T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \quad (5.13)$$

Eq. 5.13 is a general expression for the updated error covariance matrix and applies for any gain $\mathbf{K}_k$, suboptimal or otherwise.

Returning to the optimisation problem, we wish to find the particular $\mathbf{K}_k$ that minimises the individual terms along the major diagonal of $\mathbf{P}_k$, because these terms represent the estimation error variances for the elements of the state vector being estimated. The cost function for the optimisation is chosen to be

$$\mathbf{J}_k = E[\mathbf{e}_k^T\mathbf{S}\mathbf{e}_k] \quad (5.14)$$

where $\mathbf{S}$ is any positive semi-definite matrix. It can be demonstrated that the optimal estimate is independent of $\mathbf{S}$ [38]; hence we may as well choose $\mathbf{S} = \mathbf{I}$, yielding

$$\mathbf{J}_k = \text{trace}[\mathbf{P}_k] \quad (5.15)$$

This is equivalent to minimising the *a posteriori* error vector. To find the value of $\mathbf{K}_k$ which provides a minimum it is necessary to take the partial derivative of $\mathbf{J}_k$ with respect to $\mathbf{K}_k$ and equate it to zero. Use is made of the relation for the partial derivative of the trace of the

product of two matrices **A** and **B** (when **B** is symmetric [38],

$$\frac{\delta}{\delta A}[\text{trace}(\mathbf{ABA}^T)] = 2\mathbf{AB} \tag{5.16}$$

From Eqs. 5.13 and 5.14 the result is

$$-2(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-\mathbf{H}_k^T + 2\mathbf{K}_k\mathbf{R}_k = 0 \tag{5.17}$$

Solving for $\mathbf{K}_k$,

$$\boxed{\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}_k^T[\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{R}_k]^{-1}}$$

which is referred to as the Kalman gain matrix. The optimised value of the updated estimation error covariance matrix can now be found by substitution of Eq.5.18 into Eq.5.13

$$\begin{aligned}\mathbf{P}_k &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \\ &= \mathbf{P}_k^- - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_k^- - \mathbf{P}_k^-\mathbf{H}_k^T\mathbf{K}_k^T + \mathbf{K}_k(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{R}_k)\mathbf{K}_k^T\end{aligned} \tag{5.18}$$

Substitution of Eq.5.18 leads to

$$\boxed{\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-}$$

We now have a means of assimilating the measurement at t_k by the use of Eq.5.11 with $\mathbf{K}_k$ set equal to the Kalman gain as given by Eq.5.18. Note that we need $\hat{\mathbf{x}}_k^-$ and $\mathbf{P}_k^-$ to accomplish this and we can anticipate a similar need at the next step in order to make optimal use of the measurement $\mathbf{z}_{k+1}$. The updated estimated $\mathbf{x}_k$ is projected ahead via the transition matrix **A**. The contribution of $\mathbf{w}_k$ in Eq. 5.4 can be ignored because it has zero mean and is uncorrelated with the previous **w**'s. Thus we have

$$\boxed{\hat{\mathbf{x}}_{k+1}^- = \mathbf{A}_k\hat{\mathbf{x}}_k^-}$$

The error covariance matrix associated with $\hat{\mathbf{x}}_{k+1}^-$ is obtained by formulating the expression

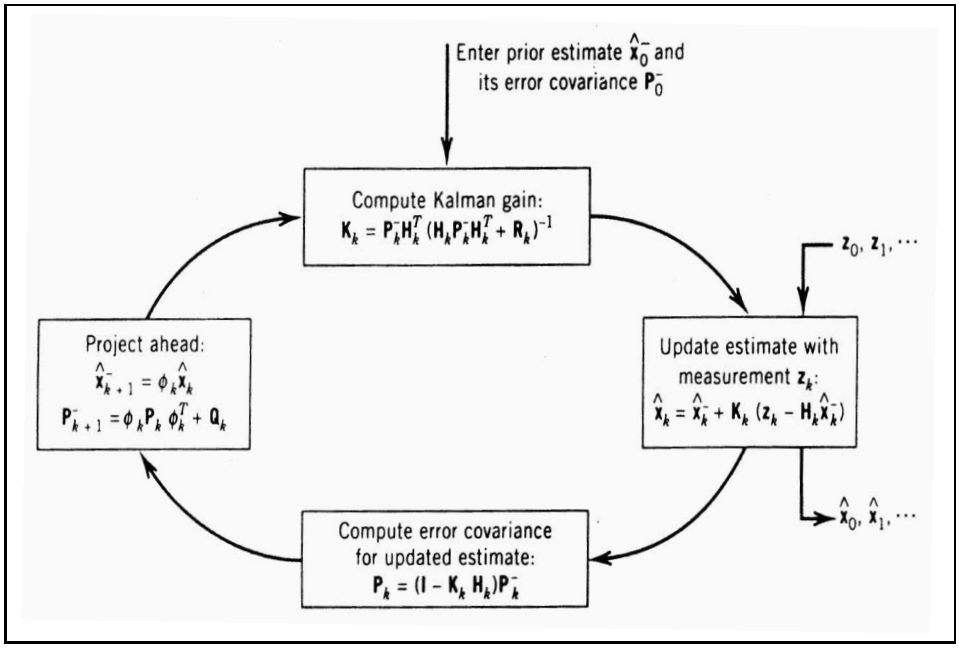

Figure 5.6: Kalman filter loop, taken from [16]. Note that ϕ is used in place of A in this figure.

for the *a priori* error

$$\begin{aligned} \mathbf{e}_{k+1}^- &= \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1}^- \\ &= (\mathbf{A}_k \mathbf{x}_k + \mathbf{w}_k) - \mathbf{A}_k \hat{\mathbf{x}}_k \\ &= \mathbf{A}_k \mathbf{e}_k + \mathbf{w}_k \end{aligned} \tag{5.19}$$

Noting that $\mathbf{w}_k$ and $\mathbf{e}_k$ are uncorrelated the expression for $\mathbf{P}_{k+1}^-$ is

$$\begin{aligned} \mathbf{P}_{k+1}^- &= E[\mathbf{e}_{k+1}^- {\mathbf{e}_{k+1}^-}^T] \\ &= E[(\mathbf{A}_k \mathbf{e}_k + \mathbf{w}_k)(\mathbf{A}_k \mathbf{e}_k + \mathbf{w}_k)^T] \end{aligned} \tag{5.20}$$

$$\boxed{\mathbf{P}_{k+1}^- = \mathbf{A}_k \mathbf{P}_k \mathbf{A}_k^T + \mathbf{Q}_k}$$

We now have the needed quantities at time t_{k+1} and the measurement $\mathbf{z}_{k+1}$ can be assimilated as in the previous step. Eqs.5.11, 5.18, 5.19, 5.19 and 5.21 comprise the Kalman filter recursive equations. These equations and the sequence of computational steps are shown pictorially in Figure 5.6, which summarises what is known as the Kalman filter loop.

5.6.2 Filter parameters Q and R

In the derivation of the Kalman filter it is assumed that process and measurement noise $\mathbf{w}$ and $\mathbf{v}$ are independent of each other and white with Gaussian probability distributions

$$p(\mathbf{w}) \sim N(\mathbf{0}, \mathbf{Q}) \tag{5.21}$$

$$p(\mathbf{v}) \sim N(\mathbf{0}, \mathbf{R}) \tag{5.22}$$

$\mathbf{Q}$ represents the uncertainty or inherent noise in the system model. For example, in a constant velocity model the target may not have constant velocity for all time. There is uncertainty in the target trajectory, the target accelerating or turning at any given time. $\mathbf{R}$ is the amount of error the measurement sensor introduces. In the actual implementation of the filter the measurement noise covariance, $\mathbf{R}$, is usually measured prior to operation of the filter. The determination of $\mathbf{Q}$ is generally more difficult as we do not always have the ability to directly observe or indeed have the prior knowledge to predict the process. If enough knowledge of the process or measurement covariances is known to model either of the variables mathematically, these models can be integrated into the Kalman filter equations; for example, Grewal *et al.* [40] show how to incorporate correlated noise into the equations.

5.7 Kalman filter for data fusion

The Kalman filter lends itself well to the fusion of measurements from different sources or sensors. Two commonly used methods exist for Kalman-filter-based fusion, known as the state vector fusion and measurement fusion methods [37]. State vector fusion methods use a bank of Kalman filters to obtain individual sensor-based estimates which are then fused to obtain an improved joint state estimate. Measurement fusion methods fuse the sensor measurements to obtain a weighted or combined measurement and then use a single Kalman filter to obtain the final state estimation based upon the fused observation. Figure 5.7 shows a diagrammatic representation of these two methods.

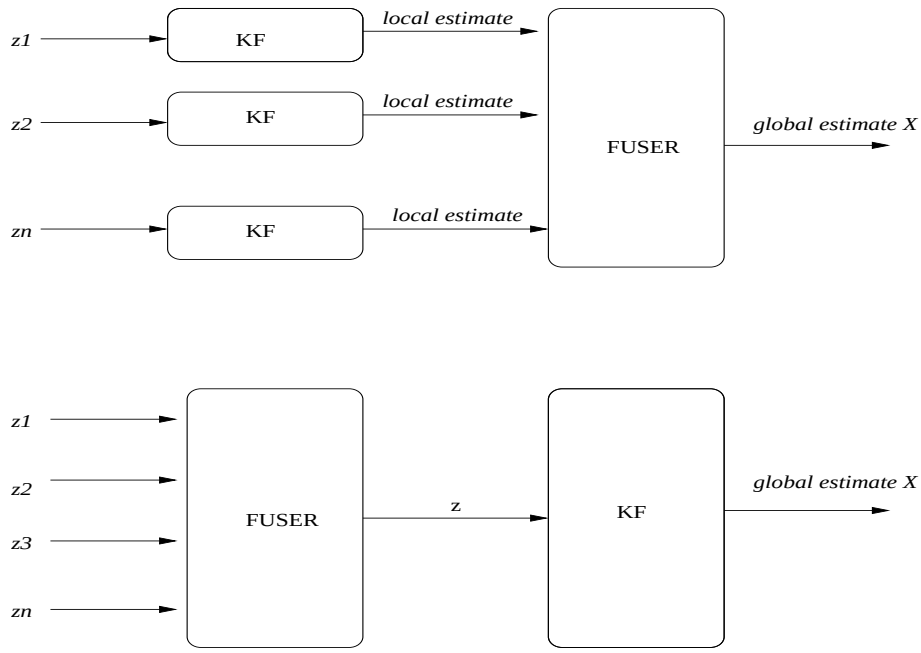

Figure 5.7: The two types of Kalman filter fusion methods. The upper diagram shows the state-vector method, for which estimates ($\hat{x}$) from parallel Kalman filters are combined via Eq.5.23. The lower figure shows the measurement fusion method. Here parallel measurements (z) are combined via Eq. 5.25.

5.7.1 State vector fusion

State vector fusion (an example of decentralised fusion) combines estimates of each Kalman filter in a minimum mean square error (MMSE) sense. The mechanism for data fusion in state vector fusion is

$$\hat{\mathbf{x}}_k = (\sum_{i=1}^{n}(\mathbf{P}_k^i)^{-1})^{-1} \sum_{i=1}^{n}[(\mathbf{P}_k^i)^{-1}\hat{\mathbf{x}}_k^i] \tag{5.23}$$

where $\hat{\mathbf{x}}_k$ is the global estimate of state at each time k and the $(\mathbf{P}_k^i)^{-1}$ are the inverses of the state error covariance matrices for each of the n sensors. $\hat{\mathbf{x}}_k^i$ is the local state vector estimate at each sensor. The global state error covariance matrix is given by

$$\mathbf{P}_k = (\sum_{i=1}^{n}(\mathbf{P}_k^i)^{-1})^{-1} \tag{5.24}$$

It is noted that under conditions when $\mathbf{Q}_k$ and $\mathbf{R}_k$ are constant, both the estimation error covariance $\mathbf{P}_k$ and the Kalman gain $\mathbf{K}_k$ stabilise quickly and remain constant. Therefore the state vector estimates are given the same weighting throughout.

5.7.2 Measurement fusion

The measurement fusion method (an example of centralised fusion) combines the measurements from each information source in a MMSE sense [37]

$$\mathbf{z}_k = (\sum_{i=1}^{n}(\mathbf{R}_k^i)^{-1})^{-1} \sum_{i=1}^{n}[(\mathbf{R}_k^i)^{-1}\mathbf{z}_k^i] \tag{5.25}$$

$$\mathbf{H}_k = (\sum_{i=1}^{n}(\mathbf{R}_k^i)^{-1})^{-1} \sum_{i=1}^{n}[(\mathbf{R}_k^i)^{-1}\mathbf{H}_k^i] \tag{5.26}$$

$$\mathbf{R}_k = (\sum_{i=1}^{n}(\mathbf{R}_k^i)^{-1})^{-1} \tag{5.27}$$

5.7.3 Previous work using the Kalman filter

The Kalman filter has found numerous applications since 1960 when Kalman's seminal paper was published. The filter is used to estimate the state parameters from noisy sensor measurements in navigation, surveying, vehicle tracking (aircraft, spacecraft, missiles), geology, oceanography, fluid dynamics, steel/paper/power industries, and demographic estimation, to mention just a few of the many application areas. The Kalman filter has been used extensively to fuse or integrate the measurements of a number of sensors or sources of information. Much of the published work is applied to the task of multi-sensor tracking of vehicles [40].

New applications for the Kalman filter are continually being found. For example, Hellebrandt *et al.* [47] use a Kalman filter to optimise the estimation of the position and velocity of mobiles in a cellular radio network, using measurements of the signal strength of the

mobile to surrounding base stations. Recently Kalman filters have been used to combine multi-media sources (audio and video) for object localisation. This has applications in video conferencing, automatic scene analysis and security monitoring [99]. Fewer applications of the Kalman filter are currently found in the biomedical signal processing area. Vauhkonen [110] uses a Kalman filter to track fast impedance changes in the body using electrical impedance tomography measurements [1]. Ebrahim *et al.* [31] propose a method for the robust fusion of the estimation of the heart rate from a number of sensor measurements including the ECG, PPG and intra-arterial blood pressure signals. Measurement values of heart rate from each signal are combined in a measurement fusion scheme. However prior to fusion a threshold based decision is made on whether a measurement is nominal, in which case it should be included in the fusion scheme, or artefactual in which case it should not be. This decision is based firstly on whether the individual measurements are physiologically plausible. If they are, further tests are carried out to determine whether the measurements show consensus. These tests are based on finding the maximum likelihood that the sensor measurements fit one of a number of hypothesis models, the hypotheses covering every possible combination of each of the sensors being artefactual or nominal (2^N hypotheses for N sensors).

5.8 Kalman filtering of signals from single sources

The aim of the work described in this chapter is to track the breathing rate from a number of noisy measurements acquired from several sources.

The first step in investigating the use of a Kalman filter to obtain an improved estimate of the measurements, is to determine the process and measurement models.

5.8.1 The process model

In this work the state being estimated is the breathing rate of a human subject. The breathing rate of a resting healthy subject is approximately constant. There are clinical conditions,

[1] in electrical impedance tomography an estimate for the cross-sectional impedance distribution is obtained from the body by injecting a small high-frequency current using surface electrodes and making voltage from the electrodes.

for example Cheyne-Stokes respiration [9], for which the breathing rate varies rapidly. However for the subjects in the Controlled-breathing database, it is reasonable to assume that the breathing rate remains approximately constant, with small changes from breath-to-breath. Thus the process (Eq.5.4) can be modelled as a first-order Markov process

$$x_{k+1} = x_k + w_k \tag{5.28}$$

In this case the state vector (breathing rate) is a scalar and hence matrices in Eqs.5.11, 5.18, 5.19, 5.19 and 5.21 become scalar quantities and we use variances instead of covariance noises. In addition A from Eq. 5.4 is equal to 1, and the variance of the process noise w is now related to the amount by which the breathing rate may be assumed to change from breath to breath.

The measurement is made directly of the state, leading to a measurement model (Eq.5.5)

$$z_k = x_k + v_k \tag{5.29}$$

with H of Eq.5.5 equal to 1.

An approach to choosing the variances of the process and measurement noise (Q and R) is proposed by Mehra [68]. In this method the variances Q and R are adjusted based on a test of the filter optimality. A fundamental property of the Kalman filter is that if the physical system actually evolves according to the state and measurement equations (Eq.s 5.4 and 5.5), the filter generates an innovation sequence [2] that is zero mean and white [108].

For large enough n the autocorrelation sequence (correlogram) of a white noise sequence has distribution

$$r_k \backsim N(0, 1/n) \tag{5.30}$$

values of r_k greater than $2/\sqrt{n}$ can be regarded as significant at about the 5% level. An off-line test may be carried out and Q and R adjusted to make the innovation as white as possible by examining the autocorrelation sequence. As our innovation sequence is unevenly sampled this test is not so easy to carry out. Also it is not practicable to perform "off-line"

[2] the innovation is the difference between the predicted measurement and the observed measurement.

Table 5.2: The statistics of the measurement errors for each of the different sources. These statisitics are available as in this work the actual state values are known.

Method	Expected Error (BPM)	Variance in Error
IP	1.9	17.9
PPG	-1.6	31.5
RSA	2.9	34.5
R amp	3.52	37.7

tests on all subjects to be monitored subsequently.

The chosen value of Q needs to be large enough to track changes in breathing rate, however the larger the value Q the less smooth the tracking. A value of $Q = 2$ is chosen as a reasonable estimate, corresponding to a standard deviation of $\pm\sqrt{2}$ breaths per minute.

The measurement noise, R, is very difficult if not impossible to determine, especially as the estimation method is intended to be generic rather than subject specific. The breathing rate measurements are made from features or characteristics extracted from a number of different biological signals. Variation in the measurement noise both inter-and-intra subject is expected. It is seen in Table 5.1 that there is a subject dependence in the quality of results obtained from the different sources. Also the measurement noise varies intra-subject across time, depending on numerous factors; patient movement, depth and rate of breathing, and the relative position of the sensors on the body.

In this work the actual state, the breathing rate, is directly observed as a reference using an oral airflow thermistor. Therefore the statistics in the measurement errors over the whole database may be obtained by examining the error between the actual state values (the reference values obtained from the thermistor) and the measured values. Table 5.2 shows the expected values and the variances of the measurement errors for each of the methods. The values have been calculated over all subjects in the Controlled-breathing database.

As can be seen from this table the measurement error values are all slightly biased. The bias in the measurement noise may be incorporated into the Kalman filter equations by simply subtracting it from the measurement before using it. Equation 5.5 then becomes

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k - \text{bias} \tag{5.31}$$

However, although there is an *overall* bias in the measurements in this data, the bias is unlikely to be in the form of a constant offset in the measurement, especially across all patients and therefore Eq. 5.31 is not necessarily a more accurate representation.

Although a value of R could be chosen for each subject, intra-subject noise causes unpredictable variations over time. Instead the variances shown in Table 5.2 are chosen as the values for R for each method, as these are our best estimates of a "generic" measurement noise. It is further assumed that the measurement noise is white.

5.8.2 Initialisation

Part of the requirements for a Kalman filter is the specification of initial conditions for the state vector $\mathbf{x}_0$ and the state covariance matrix $\mathbf{P}_0$. An estimate of the values is needed as they are not generally known. A simple way to initialise the state vector is to calculate an average of measurement values obtained during an initialisation period, prior to the start of filtering. The state covariance matrix is usually set to be a multiple (usually 10) of the process noise matrix $\mathbf{Q}$. It can be shown that provided the system is observable the error due to poor initialisation tends to zero as $k \to \infty$. Good initialisation is not essential in linear system models, the estimator merely takes longer to settle down [76].

In the Kalman filters implemented in this work, the initial state, x_0, is taken to be the average of measurements taken over an initialisation period of one-minute for each subject. The value of P_0 is chosen as 20 ($10 \times Q$).

5.8.3 Implementation

Acknowledging that the measurement noise model is non-optimal a Kalman filter is used to estimate the value of breathing rate from each of the four respiratory waveform methods. The Kalman filters for each respiratory waveform have the same process model, but the values of R depend on the particular respiratory waveform. The incorporation of a bias in measurement noise (Eq.5.31) is compared with assuming zero-mean measurement noise (Eq.5.5). Results are given in the following section.

Table 5.3: Pearson correlation coefficients for the unfiltered measurements and the filtered estimates (with and without bias) with respect to reference breathing rate measurements. All correlation coefficients are found to be very highly significant.

Method	Unfiltered	With bias	Without bias
IP	0.64	0.67	0.67
PPG	0.45	0.51	0.51
RSA	0.48	0.54	0.53
R amp	0.36	0.38	0.37

5.8.4 Results of estimating breathing rate using a single source and a Kalman filter

Figure 5.8 shows the percentage error curves for the Kalman filter estimates of average breathing rate. The data for these curves is calculated in the same way as those in Figure 5.5 which show the results for the unfiltered measurements of the four sources (respiratory waveform methods). Thus, the average breathing rate is calculated over a 30-second window, with no overlap between windows. The percentage error for each window is calculated and a mean percentage error for separate breathing rate ranges is found.

For comparison, the error curves obtained using the unfiltered measurements of the single sources are also shown. Table 5.3 shows the correlation of the different methods, using unfiltered measurements and Kalman filter estimates. The Kalman filter is used both with and without incorporating the bias in the measurement model.

In all methods the Kalman filter estimates are seen to give a better correlation with the reference breathing rate, than that obtained using the unfiltered measurement. The effect of incorporating the bias in the measurement error model can be seen in Figure 5.8. The use of the bias in the measurement decreases the percentage error in the low breathing rate range, but has the opposite effect at higher breathing rates, showing that the bias is not consistent across the different breathing rates, and that the constant bias measurement model is not accurate. The difference in overall correlation when using bias and when not using it is negligible. The bias just has the effect of “shifting” the errors in the breathing rate spectrum.

The fact that there is an improvement in correlation with the Kalman filter estimates shows that the filter, as suggested in the literature, still functions even under poor model conditions

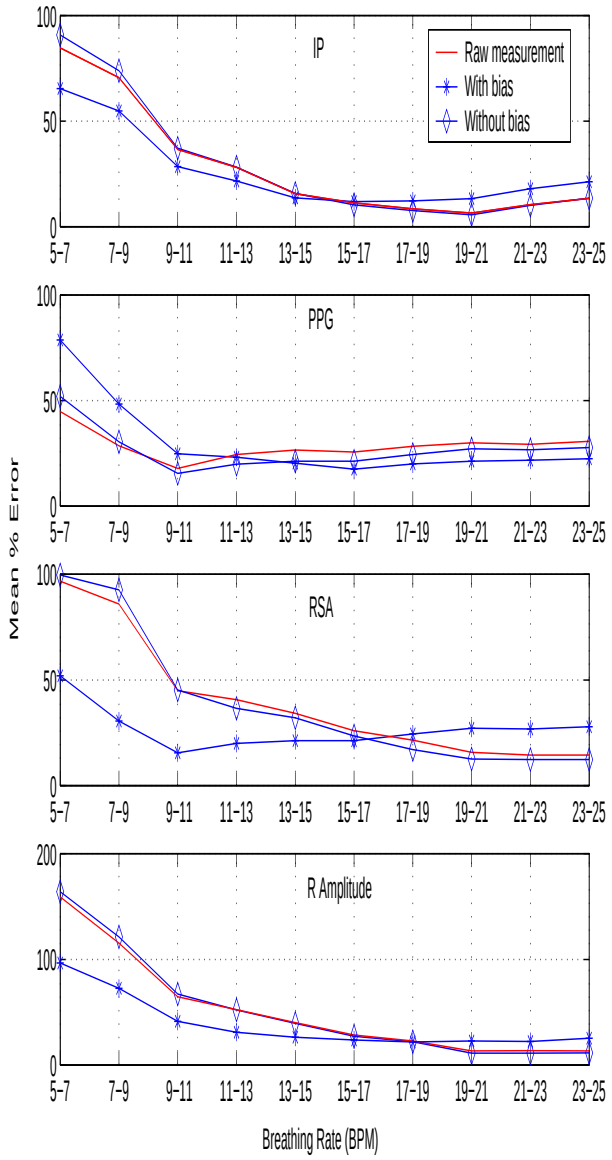

Figure 5.8: Each plot the percentage errors of the estimates from a Kalman filter with and without bias in the measurement model. The percentage errors for the unfiltered measurements (reproduced from Figure 5.5) are shown for comparison. Each plot corresponds to one of the four respiratory waveform methods.

[115].

5.9 Innovation fusion method

As mentioned in Section 5.5 the input sensors, used for acquiring the signals from which the respiratory waveforms are derived, are independent to some degree: the physiological phenomenon giving rise to breathing information in the different sensors are related but not the same and artefact affects each sensor differently. This motivates the investigation into fusing the measurements from each derived respiratory waveform.

Although Section 5.8.4 shows that a Kalman filter applied to noisy measurements of breathing rate can lead to improved estimates, the inability to accurately predict, and hence model both the process noise and the measurement noise leads to sub-optimal estimation. Similarly, the standard methods for data fusion with a Kalman filter discussed in Section 5.7 all rely on accurate models which incorporate well-characterised prior knowledge. Such information is not available for physiological signals which suffer considerable intra- and inter-patient variation.

In this section we therefore suggest an alternative method for fusing the measurements. The innovation σ_k of the scalar Kalman filter operation is defined as (in scalar notation)

$$\sigma_k = z_k - H_k x_k^- \tag{5.32}$$

i.e., the difference between the measurement z_k and the measurement prediction $H_k x_k^-$ at time k. If the physical system evolves according to the state and measurement equations (Eq.s 5.4 and 5.11), the Kalman filter generates a σ_k sequence that is zero mean and white [108]. In two separate works by Tylee [108] and Gustafson *et al.* [44] the innovation is used as a means of assessing information about a measured signal.

Gustafson *et al.* [44] aim to classify R-R intervals into specific arrhythmia patterns. Four Kalman filters are implemented in parallel, each with a different process model. All the Kalman filters operate on the same observed data, the R-R interval sequence. The statistical description of the innovation sequence for a given time epoch from each of the parallel filters

is examined. This information, together with the probabilities of the previous epoch having been generated by one of the models, is used to calculate the probabilities of the current sequence being generated by one of the models.

In the work of Tylee [108] a bank of Kalman filters are used to generate optimal estimates of the state of a nuclear plant. To increase safety, a method for detecting the sensor failure is needed. The innovations from each sensor after Kalman filtering of each reading are compared. If one innovation shows a large discrepancy from the innovations of the other sensors, a potential failure of that sensor is detected. If the innovations of all the sensors are all unpredictably large it is assumed that there is a true change in the estimated state.

In the proposed fusion method the innovation is used as a *confidence measure* in the current measurement; a separate Kalman filter is applied to each of the noisy measurements of breathing rate from the four respiratory methods. The innovations of each Kalman filter output are used as a confidence measure in the current measurement from each source of information. The confidence levels are used to calculate the weighting given to each measurement in the fusion method.

The measurements are combined linearly based on a MMSE criterion. Appendix A shows that n estimates of a value z_p, each with a confidence value σ_p, can be combined to obtain an optimum (in a MMSE sense) estimate $\hat{z}_o$ in a linear fashion

$$\hat{z}_o = \sum_{p=1}^{n} \alpha_p \hat{z}_p \tag{5.33}$$

with a weighting parameter α_p defined as

$$\alpha_p = \frac{1}{\sigma_p^2 \sum_{i=1}^{n} \frac{1}{\sigma_i^2}} \tag{5.34}$$

when n is equal to 2 this gives

$$\alpha_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \tag{5.35}$$

$$\alpha_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \tag{5.36}$$

Consider the following situations with reference to Eq. 5.34:

- Large innovations from all Kalman filters: there is a change in state and the measurements are given similar weightings. The fusion method follows the change with little or no lag, which is not usually the case in conventional Kalman filter methods.
- One innovation much greater than the others: it is assumed that there is no change in state and hence there must be a large error in one of the measurements caused either by the respiration characteristic not being found in the signal or artefact appearing in the derived respiratory waveform. The "rogue" measurement is given a small weighting compared to the others.
- Small innovations from the output of all filters imply that there is no change in the system state; all measurements are considered accurate and given similar weightings.

5.9.1 Implementation of the innovation fusion algorithm

Figure 5.9 represents the fusion process. A Kalman filter is used to obtain individual estimates from each source (derived respiratory waveform). The unfiltered measurements z_i and innovations σ_i are then combined according to Eq.s 5.33 and 5.34 to obtain an improved joint state estimate. This differs from conventional decentralised fusion methods where the state estimate and the error covariance are used to obtain the joint estimate. Fusion occurs whenever a measurement is made. For example when fusing two channels, each measurement from one channel is fused with the two nearest (in time) estimates from the other channel. For each Kalman filter, the process model, measurement model and parameter values Q and R of Section 5.8 are used.

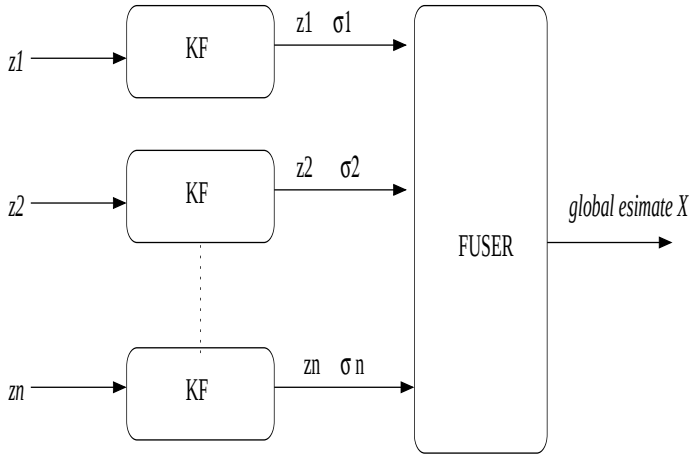

Figure 5.9: The innovation fusion process. z_n is the measurement from each source, σ_n the innovation from the corresponding Kalman filter. The fusing algorithm is described by Eqs. 5.33 and 5.34.

5.9.2 Correlation between measurement noise

The best fused estimates are obtained by combining those sources that do not show a strong correlation in their measurement errors. Table 5.4 shows a matrix of the correlations between the absolute errors in the breathing rates of the four derived respiratory waveforms across the entire Controlled-breathing database.

The RSA and R-Amplitude sources show the highest correlation in errors. This may be due to the fact that these sources both rely on the correct detection of the QRS complex. The PPG source shows a negligible correlation with all other sources.

Figure 5.4 showed that the breathing rate obtained from the IP method gives the best correlation with the reference rate by a substantial margin. Therefore the IP source is first combined with the other 3 sources in turn, to give 3 combinations of two sources. It is expected that the combined PPG and IP estimates would give the best estimate due to the low correlation between the measurement errors from these methods. The fusion of three sources of information is also explored, combining the PPG, IP and RSA sources as well as the PPG, IP and R-Amplitude sources.

Table 5.4: The correlation between the measurement errors of the four respiratory waveform methods. The RSA and R-Amplitude sources show the highest correlation, while the PPG source shows the least correlation with all other sources.

Sources	IP	PPG	RSA	R-Amplitude
IP	1	0.06	0.25	0.32
PPG	0.06	1	0.09	0.04
RSA	0.25	0.09	1	0.35
R-Amplitude	0.32	0.04	0.35	1

5.10 Results of estimating breathing rate by combining sources using innovation fusion

The results are evaluated in terms of the parameters described in Section 5.3 and are compared with the results previously obtained when using the unfiltered measurements from a single source. The percentage error curves are shown in Figure 5.10. For reference, the upper plot shows the percentage errors for each of the breathing rate bands of the unfiltered measurements for each of the four sources. The middle plot shows the percentage errors when two of the sources are combined. These are superimposed on the curves for the unfiltered IP and PPG sources (in blue) which are reproduced from the plot above for comparison. Fusing the PPG and IP sources gives the lowest errors of all methods from 9-17 BPM, with just slightly higher errors at lower rates than the single PPG measurement source. At rates above 17 BPM the combined IP and RSA method shows the best performance. The combined PPG and IP method gives rise to 8 of the 10 breathing rate bands having mean percentage errors of less than 20%, compared to only 6 in the combined PPG and IP.

The lower plot shows the curves obtained when fusing three sources. The combined plots are in black and superimposed are the curves for the combination of two sources, reproduced from the middle plot for comparison. It can be seen that the two combinations of three sources show similar curves. However there is a large deviation at lower breathing rates, where the method that uses RSA rather than the R-Amplitude source gives lower errors.

When the curves for the fused methods are compared with those obtained from measurements of a single source, the characteristics of the single sources are evident in the curves for the fused methods. For instance, the two combinations of three sources in the lower plot are identical at high breathing rates but show deviation in the lower two breathing rate

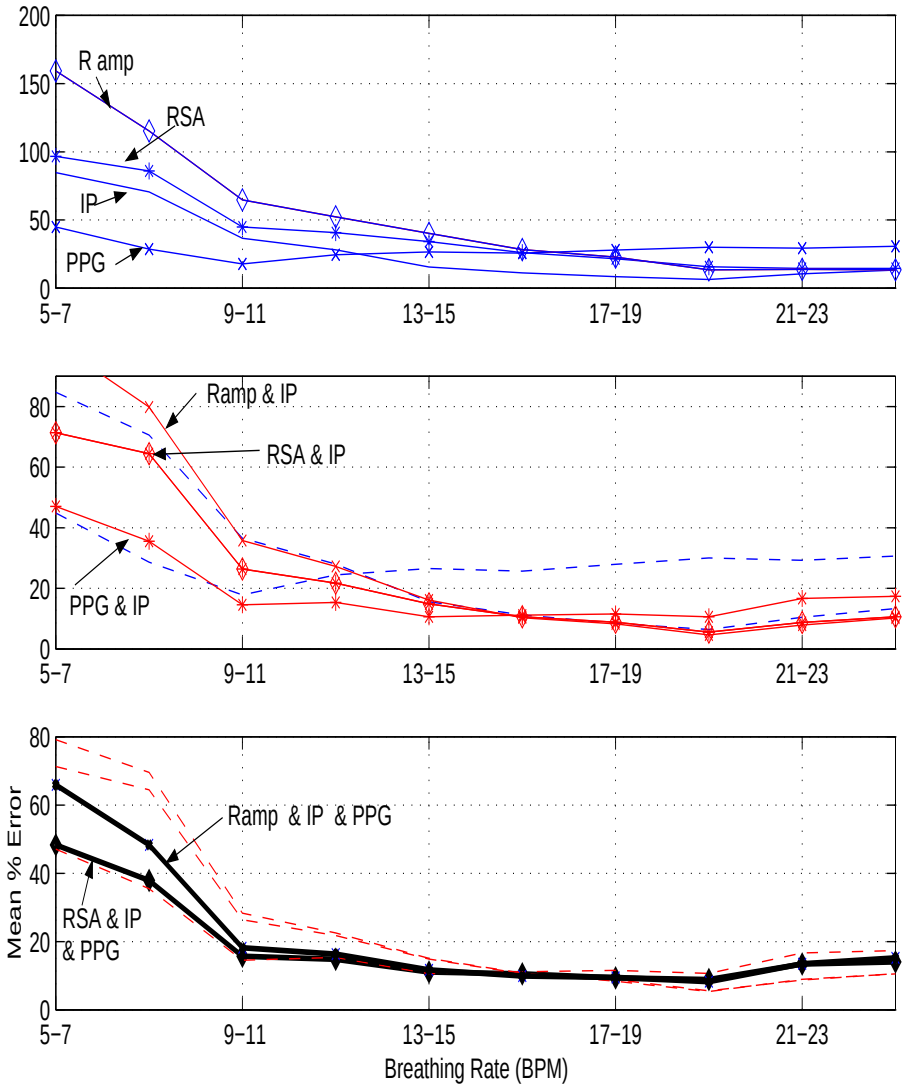

Figure 5.10: The percentage errors curves resulting from fusing more than one source of information. The upper plot shows the unfiltered measurement percentage error curves, as in Figure 5.5 reproduced here for comparison. The middle plot shows the error curves when fusing two sources of information, and the lower plot shows the result of fusing three sources of information. The maximum value of the y-axis decreases as the figures go from top to bottom.

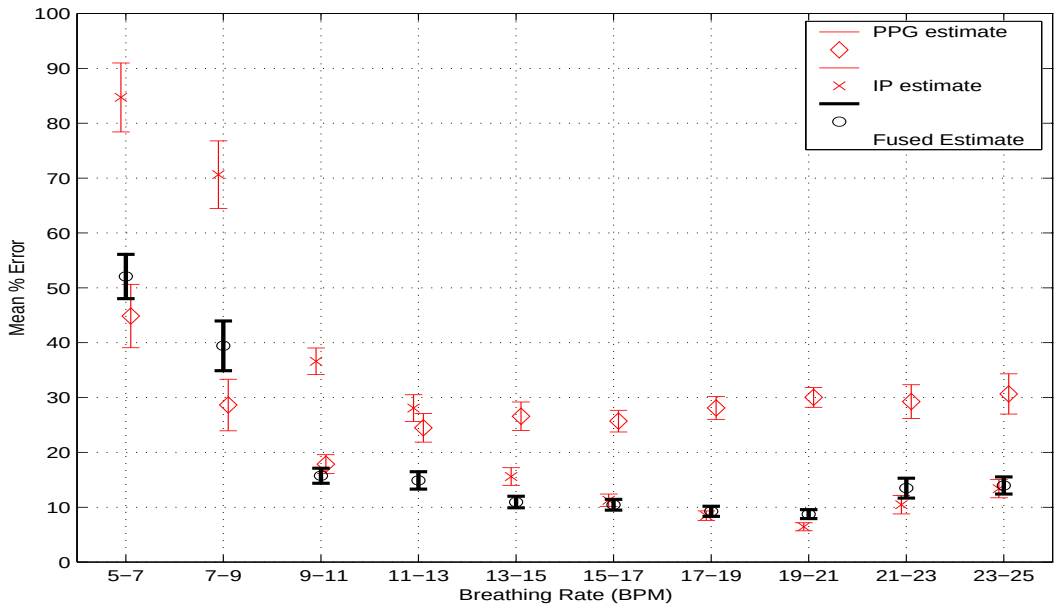

Figure 5.11: The mean and 95% confidence intervals for the fused and best performing single estimates.

bands. The fusion method that uses RSA rather than the R-Amplitude source has a lower error rate in these bands as could be expected from the top plot.

Figure 5.11 shows the averaged means for the best performing single sources, IP and PPG with the 95% confidence intervals drawn in. Also shown (in black) are the corresponding values for the fused estimate that uses the IP, PPG and RSA sources. It can be seen that in 8 of the 10 frequency bands the confidence interval is either lower or overlaps the best performing single source confidence interval.

Table 5.5 shows the overall correlation of the fused estimates of breathing rate with the reference rates. The values in the left-hand column show the correlations for the innovation fusion method, corresponding to the methods represented in Figure 5.10. To compare the performance of the proposed innovation fusion method with the conventional Kalman filter based fusion methods of state vector and measurement fusion (described in Section 5.7), the middle and right-hand columns show the correlations of the estimates obtained when implementing these methods. These implementations used the Kalman filter models of Section 5.8, as does the innovation fusion method. The correlations for these methods are all lower than the values corresponding to the innovation method, showing that the lack of accuracy in the measurement noise model leads to sub-optimal estimation.

Table 5.5: Pearson correlation coefficients for the fusion of different sources. Results for the Innovation fusion method are shown along with the result of using the State Vector and Measurement fusion methods. All correlation coefficients are found to be very highly significant.

	Correlation		
Sources	Innovation Fusion Method	State Vector	Measurement Vector
IP and PPG	0.77	0.66	0.69
IP and R Amp	0.69	0.57	0.67
IP and RSA	0.73	0.64	0.70
IP,RSA and PPG	0.80	0.69	0.72
IP,R Amp and PPG	0.77	0.65	0.69

When the correlation values for the innovation fusion method are considered in more detail the fusion of PPG and IP is seen to give the highest correlation (ρ = 0.77) of any of the two source combinations. This is improved by combining with the RSA source ($\rho = 0.8$), however no improvement is seen when incorporating the R-Amplitude source.

Although the innovation fusion method does not necessarily give a lower mean percentage error at all breathing rates (see Figure 5.10), Table 5.5 demonstrates that overall the estimates from this method are more strongly correlated with the reference breathing rate, than the correlations of any single source estimate (See Table 5.1). This shows that the variance in the errors is lower than that in the single source estimate errors.

To illustrate the improvement in performance in tracking the breathing rate using the innovation fusion method with the best choice of sources (IP, PPG and RSA), Figures 5.12, 5.13, 5.14 and 5.15 show plots of the instantaneous breathing rates over time for four different subjects. In each of the plots the reference breathing rate is shown in red, while the test breathing rate is shown in blue. In the top three plots of each figure, the breathing rates obtained from three unfiltered measurements (the sources labelled IP, PPG and RSA) are displayed, the lowest plot in each case being the breathing rate sequence obtained by fusing the three sources using the innovation fusion method. It can be seen that the fused estimate of breathing rate closely follows the reference breathing rate in all four examples. Figures 5.12 and 5.13 represent times during which the subjects were asked to breath at a relaxed rate and depth.

Figure 5.14 shows a period of time during which a subject was asked to breathe at specified rates of 6, 10 and 20 BPM for 5 minute durations. At the end of each five minute duration

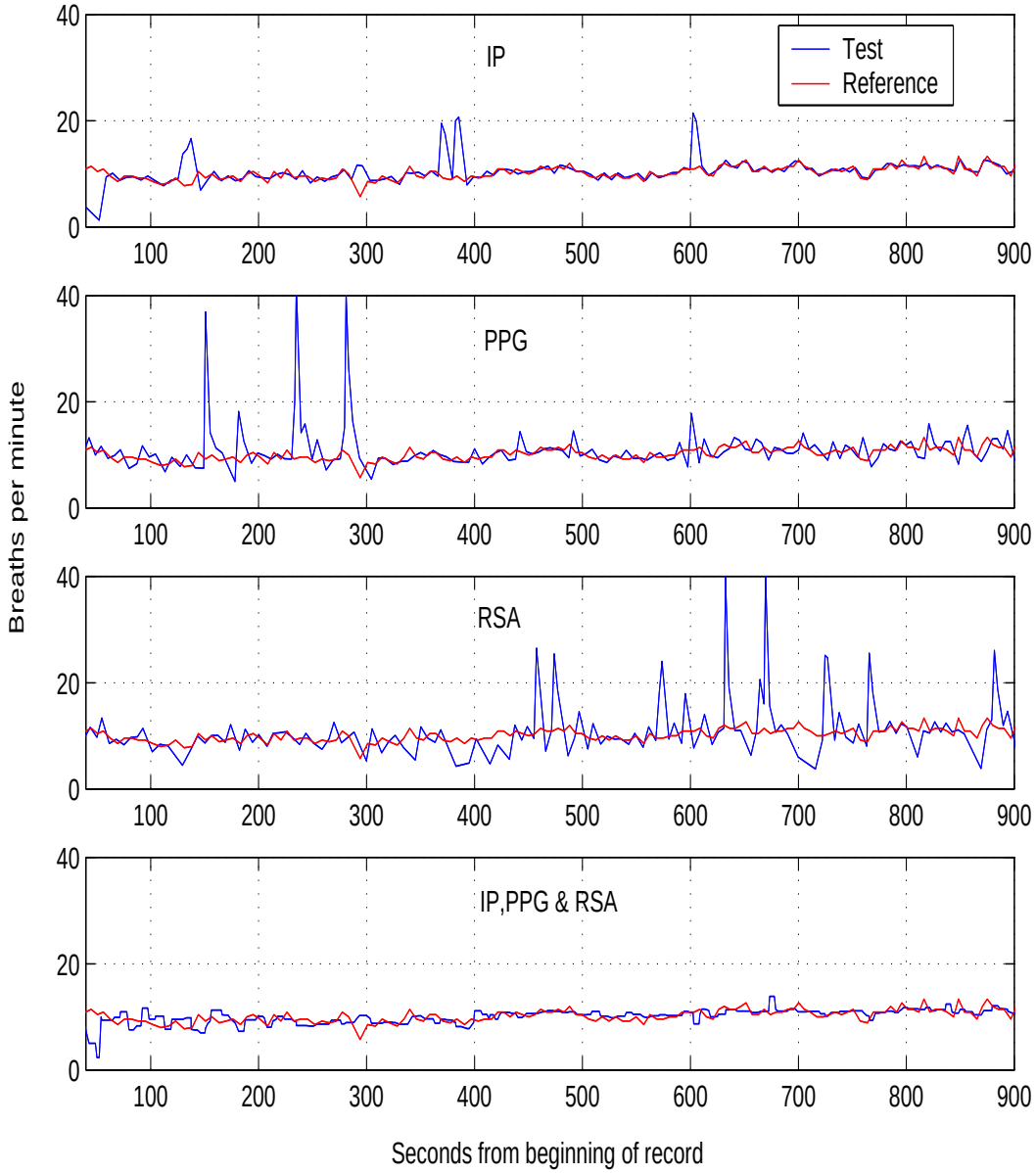

Figure 5.12: Instantaneous breathing rate sequences for record 3011. In each plot the red line corresponds to the reference rate (obtained using an oral airflow thermistor). In the top three plots the blue line is the breathing rate obtained from one of the test methods (labelled at the top of each plot).The fourth plot shows the breathing rate from the IP,PPG and RSA combined using the innovation fusion method. This data is from a period of time during which the subject was asked to breathe in a relaxed fashion at a rate and depth of their choice.

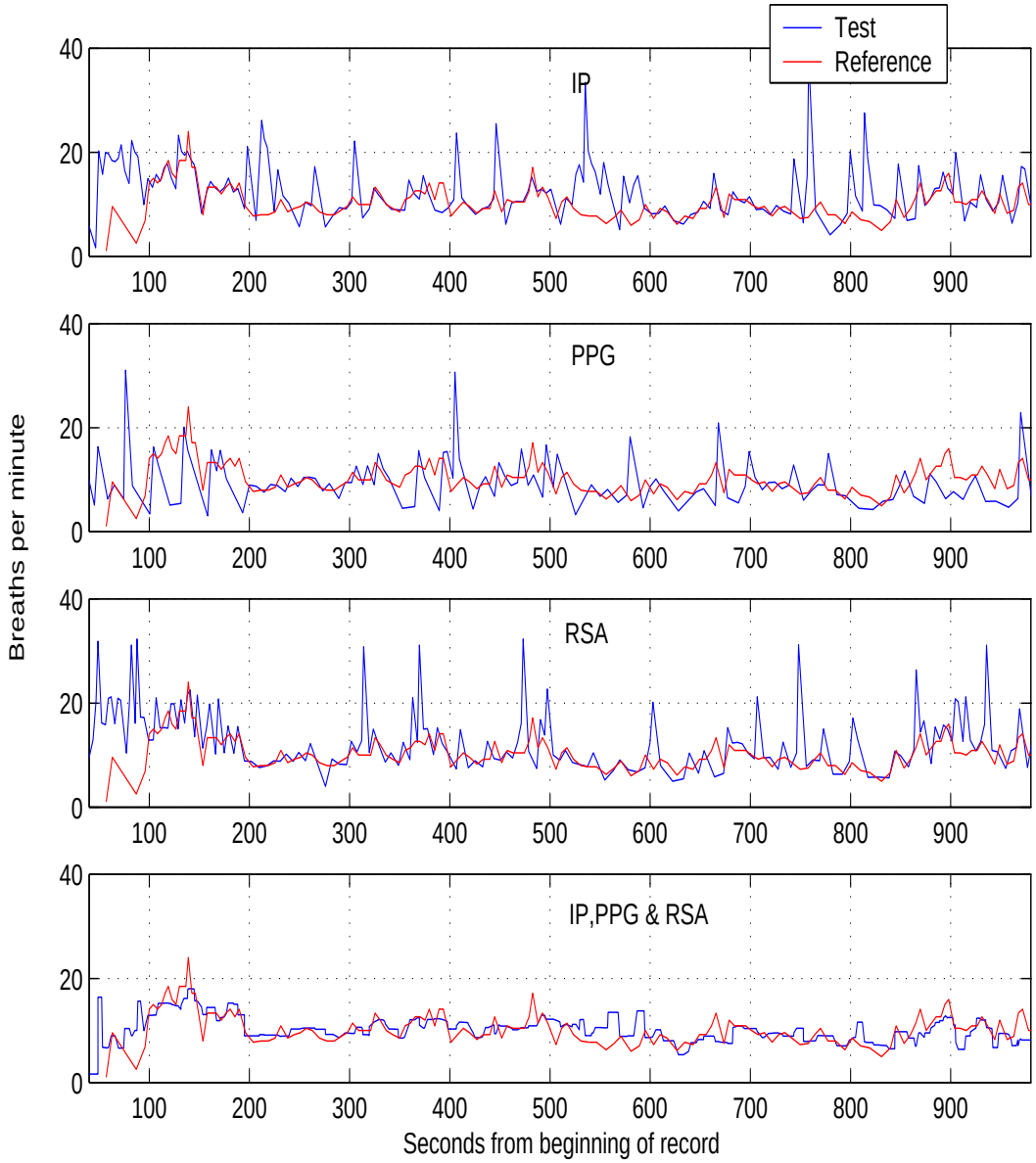

Figure 5.13: Instantaneous breathing rate sequences for record 3002. In each plot the red line corresponds to the reference rate (obtained using a signal from an oral airflow thermistor). In the top three plots the blue line is the breathing rate obtained from one of the test methods (labelled at the top of each plot). The fourth plot shows the breathing rate from the IP,PPG and RSA combined using the innovation fusion method. This data is from a period of time during which the subject was asked to breathe in a relaxed fashion at a rate and depth of their choice.

the subject was asked to hold their breath for 30 seconds. The single IP source estimates the breathing rate accurately at 20 BPM, but is noisier and tends to over-estimate the values of the lower breathing rates. The PPG source gives very good estimates at 10 BPM, and in the last two of the 6 BPM sections. The RSA source shows noisier estimates, at times being very near the reference rate and others showing sharp deviations. The result of fusing these sources using the innovation fusion method is accurate estimates at 10 and 20 BPM, as the steady sources are given higher weighting than the noisier sources. In the first section of 6 BPM breathing, all single sources are noisy and over-estimate the rates. In this situation the fusion method has the effect of smoothing the estimates, but can obviously not track the reference values when all the sources are giving inaccurate measurements. The periods during which the subject held their breath can be seen on the reference rate signal as dips after each constant section. None of the single sources detect this and therefore the fusion method does not track this either.

Figure 5.15 shows similar breathing conditions as those in Figure 5.14 but shows results for another subject (3010).

5.11 Conclusion

Methods for estimating the breathing rate from a number of noisy sources obtained from non-invasive signals are investigated in this chapter. A Kalman filter is used to estimate the rate from the measurements of a single source. The fact that the measurements are made from signals that are prone to unpredictable human inter and intra-variation, artefact such as subject movement and sensor movement means that a robust, accurate generic model of measurement noise is impossible.

This makes this type of Kalman filtering sub-optimal. A fusion method that overcomes the need for an accurate measurement model is presented. This fusion method is able to track the reference breathing rate, and results in better performance than either the use of a single source only or fusing the sources with conventional Kalman filter fusion methods.

The innovation fusion method does not give accurate results when all source measurements are noisy, or when one source gives inaccurate measurements with a constant offset from the true value. If this situation continues for long enough the innovations from this source

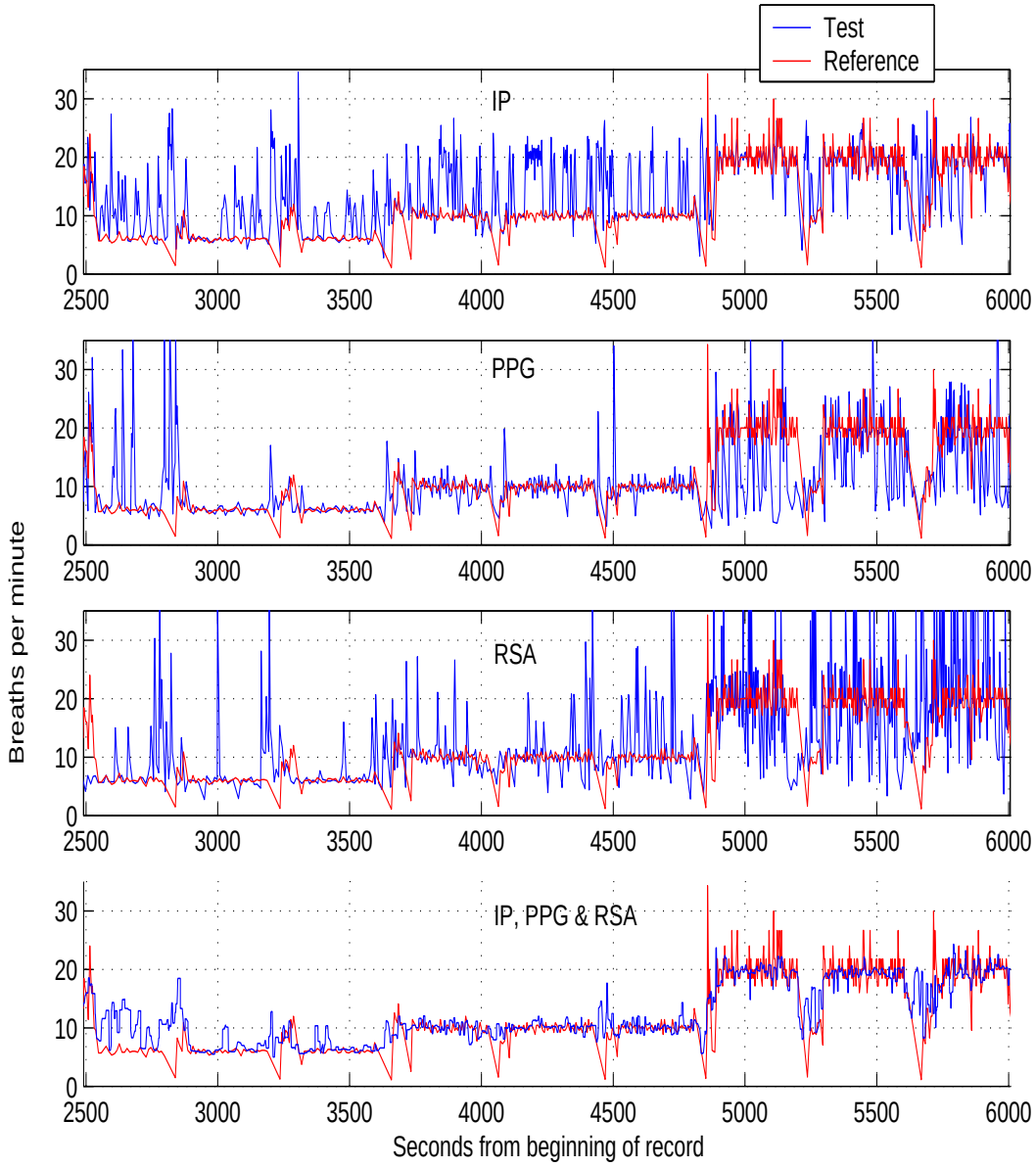

Figure 5.14: Instantaneous breathing rates for record 3005. In each plot the red line corresponds to the reference rate (obtained using a signal from an oral airflow thermistor). In the top three plots the blue line is the breathing rate obtained from one of the test methods (labelled at the top of each plot).The fourth plot shows the breathing rate from the IP,PPG and RSA combined using the innovation fusion method. This data is from a period of time during which the subject was asked to breathe at specific rates of 6 ,10 and 20 BPM.

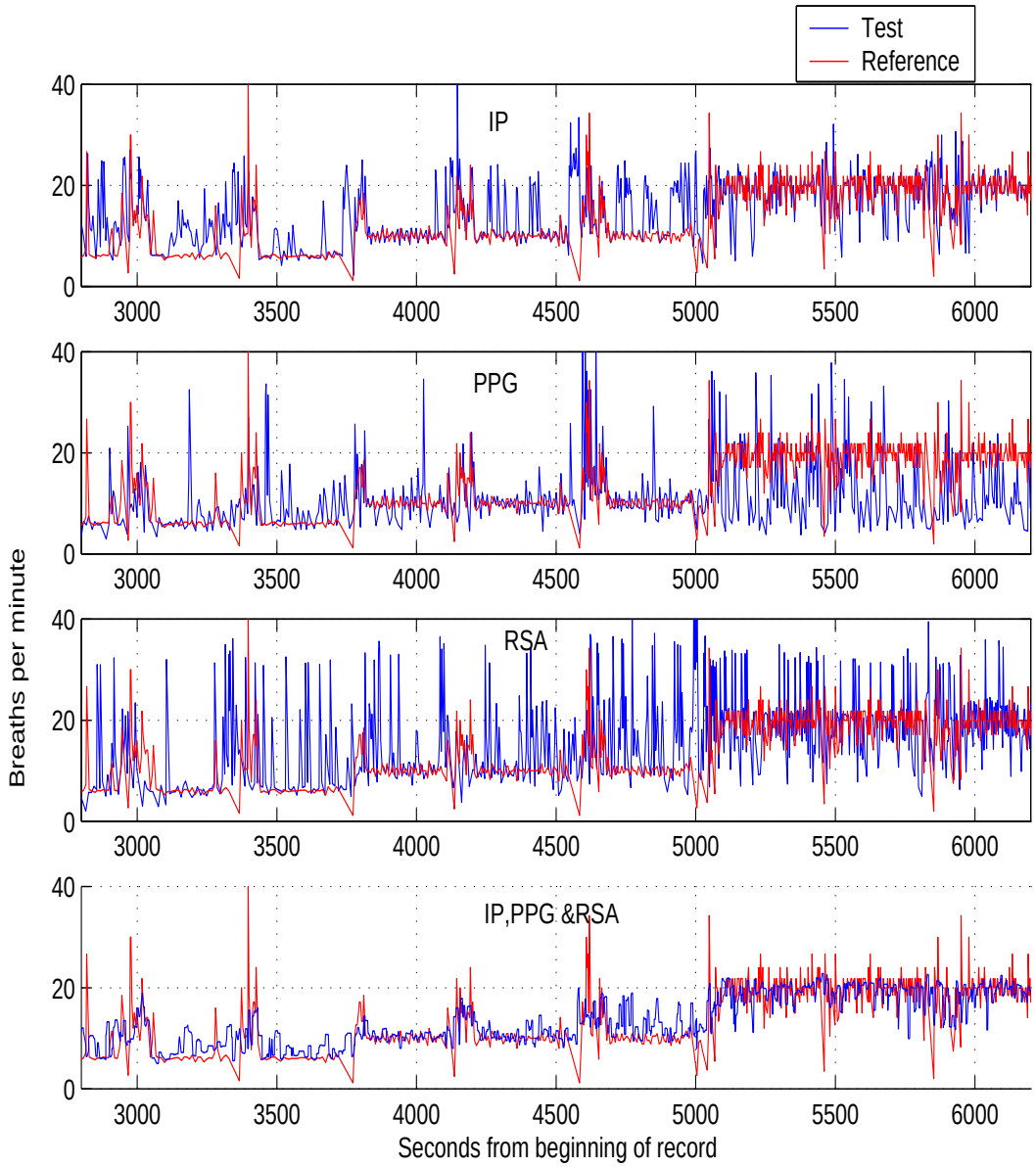

Figure 5.15: Instantaneous breathing rates for record 3010, c.f Figure 5.2. In each plot the red line corresponds to the reference rate (obtained using a signal from an oral airflow thermistor). In the top three plots the blue line is the breathing rate obtained from one of the test methods (labelled at the top of each plot).The fourth plot shows the breathing rate from the IP,PPG and RSA combined using the innovation fusion method. This data is from a period of time during which the subject was asked to breathe at specific rates of 6 ,10 and 20 BPM.

become small and the inaccurate measurements given too large a weighting.

The innovation fusion method works very well on all breathing rates above 9 BPM, the mean percentage error being less than 20% in all ranges. However less accurate tracking is observed at lower breathing rates. It should be noted that these lower breathing rates incorporate periods of time during which the subjects were asked to stop breathing. It is known that there is a tendency [63] for the single source methods considered here to detect breaths during these apnoeic periods.

Chapters 6 and 7 looks at detecting episodes of apnoea as a separate issue.

Chapter 6

Central Apnoea

6.1 Introduction

This chapter looks at how accurately central apnoeas can be detected from the non-invasive signals IP, ECG, and PPG. Central apnoea monitors (usually designed for infants) tend to have a very high false alarm rate. A review of studies carried out with existing apnoea monitors that employ non-invasive techniques is found in Section 6.2. Section 6.3 discusses the performance parameters used to evaluate apnoea detection in this work. Section 6.4 describes a time-domain approach for detecting central apnoeas from non-invasive signals. This approach is based on existing methods and provides a set of benchmark results against which a novel frequency-domain technique is evaluated. The frequency-domain approach is described in Section 6.5. The possibility of fusing information from the time and frequency-domain methods is discussed in Section 6.7 and a simple fusion technique proposed. Results obtained with this technique are presented in Section 6.7.1.

6.2 Central Apnoeas and current detection methods

Apnoea is the absence of ventilation and is abnormally long if it exceeds 15 seconds. Shorter events, especially 10-12 seconds or less in duration, are commonly seen in most infants studied in sleep [9]. There are three different categories of apnoea:

- **Central Apnoea** - absence or suppression of the signal stimulating the inspiratory muscles of respiration. There is a complete cessation of airflow in the absence of respiratory effort.
- **Obstructive Apnoea** - respiratory effort persists, however is ineffective due to the airways being obstructed.
- **Mixed Apnoea** - a single event composed of contiguous central apnoea and intervals of obstruction or ineffective breaths.

This chapter focuses on detecting central apnoeas. Central apnoea monitoring is usually associated with infants.

Respiratory pauses in infants are of clinical significance because infants have a higher rate of oxygen consumption, lower oxygen stores and smaller lung volumes. They are therefore far more sensitive and vulnerable to a lack of oxygen [94]. Home infant apnoea monitoring grew out of the suspected relationship between apnoea and sudden infant death syndrome [86]. Classical infant apnoea monitors tend to record the ECG and an indirect respiration signal, usually obtained using either impedance or inductance plethysmography. These monitors set alarms for apnoea and bradycardia [1] according to a set of pre-determined rules, based on heart rate thresholds and the absence of breath detection from the respiration signal. These monitors however all suffer from an unacceptably high false alarm rate. As described in Section 2.5, impedance plethysmograpy (IP) monitoring has inherent short comings in accurate respiration detection when there is frequent body movement and postural changes, as there are in infants. Cardiac artefact is a major cause of interference and placement of the leads is also critical. Recent "memory" or "documentation" monitor systems, (for example the Arvee 4800, Arvee Medical, Texas, USA and the Nellcor Edentec 336 assurance monitor, Nellcor Puritan Bennett, Minnesota, USA) record ECG trends, respiratory waveforms and pulse oximetry data so that whenever an alarm occurs, the carer of the monitored infant is able to use the recorded data make a decison as to whether or not the alarm indicates a true positive event.

A number of published studies emphasise the problem of false alarm rates for commercial apnoea monitors. Weese-Mayer *et al.* [114] report on a study carried out to assess the alarms triggered by a number of commercial infant apnoea monitors. These monitors use transtho-

[1] a slowing of the heart rate to below 60 beats per minute.

racic IP and ECG to trigger alarms for apnoea and bradycardia. Data was collected from 83 patients. Of the 14,131 alarms recorded in a retrospective visual inspection of the data collected only 8% were classified as a true event, a true event being defined as one where the IP and ECG signals verified apnoea or bradycardia. The automatic detection methods used in the monitors are not described.

Brouillette *et al.* [15] report on a comparison of inductance plethysmography and IP for apnoea monitoring. Apnoeas are automatically detected from the impedance plethysmograph whenever the signal was below a fixed generic threshold (i.e., not subject specific) during a 6-second period. It was reported that all of the 60 central apnoeic events in the study were detected using the processed IP signals, however 14 false positive events were also recorded. The inductance plethysmography method detected apnoeas using a similar threshold based approach; however the breath detection threshold was made subject-specific and set to 25% of the average tidal volume calculated during a 10 minute calibration period. Inductance plethysmography also gave 100% sensitivity, however only 4 false apnoeas were detected.

The Collaborative Home Infant Monitoring Evaluation (CHIME) study is described by Ramanthan *et al.* [86]. ECG and inductance plethysmography signals were recorded from 1079 babies using the CHIME monitor (NonInvasive Monitoring Systems, Miami, USA). The monitor recognises a breath whenever there is an excursion of at least 25% of the peak amplitude in the IP signal determined during a calibration period (the first 5 minutes after the monitor is turned on). An apnoea is defined to occur when there is no breath detected for a time exceeding a specified threshold. The four-year study concluded that less than 10% of the alarmed events corresponded to clinically relevant events.

In view of the high false alarm and low positive predictivity of current methods for detecting central apnoeas this chapter focuses on achieving the high sensitivity required whilst increasing the positive predictivity.

6.3 Detection from non-invasive signals

Current apnoea detection methods tend to take a time-domain approach. When detecting apnoeas, latency is a key issue, as an alarm is needed as soon as possible. If a frequency-domain approach is adopted, the time-frequency resolution is an important consideration.

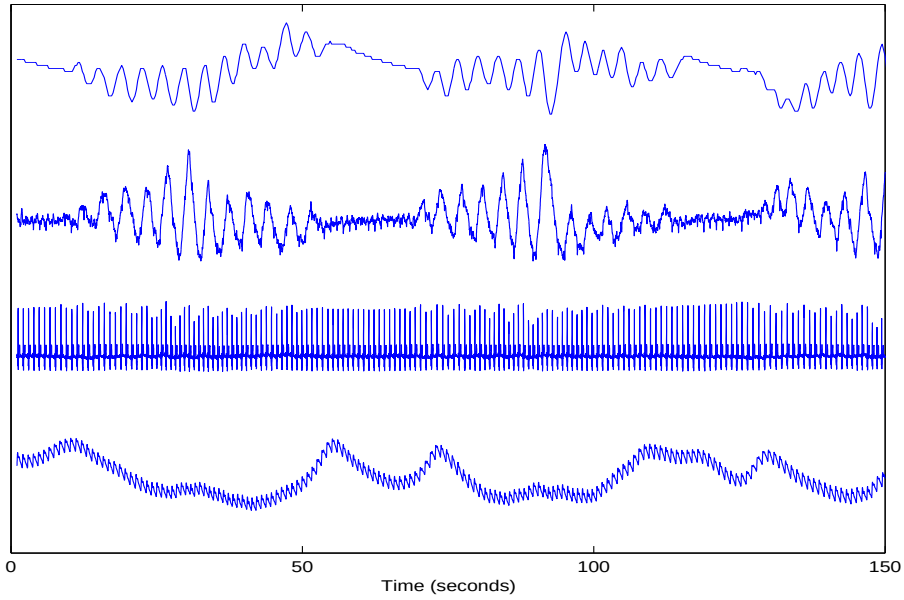

Figure 6.1: Cyclical central apnoeas. The top signal is the reference respiration signal followed by the IP, ECG and PPG signals. Three central apnoeic events occur, most clearly seen in the reference respiration signal.

Any window used needs to be short enough in duration to give a safe time delay between the start of the apnoea and its subsequent detection. On the other hand, accurate spectral estimation needs a long enough window for frequency resolution.

In this chapter, both time and frequency-domain analysis of the ECG, PPG and IP are considered for the detection of central apnoeas. The approaches described in the following sections are applied to the IP signal, respiratory waveforms derived from the ECG (R-DR and RSA-DR, described in Section 4.6.1), and a respiratory waveform derived from the PPG (PP-DR, described in Section 4.6.3).

6.3.1 The central apnoeic data

The Controlled-breathing database includes, in each record, a section mimicking cyclical central apnoeas. The subjects were asked to cease breathing every minute, for a period between 10 and 20 seconds, depending on how comfortable they felt. This cycle lasted for 10 minutes. The times of the start of each apnoea were recorded during the data collection. The evaluation of the automated apnoea detection method is carried out on this data, i.e., in total 100 minutes of data including 100 apnoeic events. Figure 6.1 shows a 150-second section

of the IP, ECG and PPG signal during the cyclical central apnoeas from one of the records of the Controlled-breathing database. The top signal shows the reference respiration signal (recorded with an oral airflow thermistor).

6.3.2 Evaluation procedure

The performance of the automated central apnoea detection methods are assessed using the performance parameters of sensitivity and positive predictivity. These are the same parameters used in the evaluation of automated detection of breaths in Chapter 4. In this case however, a true positive (TP) corresponds to the detection of an apnoea, given that an apnoea has occurred and a false positive (FP) is a detection of an apnoea when an apnoea has *not* occurred. A false negative (FN) is the failure to detect an apnoeic event.

A time window is defined around the *known* start time of the apnoea. To be defined as a true positive the automatic detection system must specify that the start of the apnoea has occurred within this time-match window.

Figure 6.2 shows the time-match window with respect to the beginning (t_b) and end (t_e) of an apnoeic event. The time-match window is declared to end 15 seconds after the start of the apnoea, this time is thought to be an acceptable delay from the start of the apnoea to an alarm being sounded. The value of x is chosen to be half the length of the data window in which analysis takes place.

6.4 Time-domain approach

The time-domain method for detecting apnoeas from non-invasive signals is based on the rule-based, threshold methods found in current apnoea monitors. An apnoea is defined if the amplitude of the respiration waveform is below a pre-determined threshold for a given time period.

In this work the respiratory waveforms investigated are the IP, R-DR, RSA-DR and PP-DR. An average peak-trough amplitude of the respiratory waveform is calculated in a 5-minute calibration period at the beginning of the record for each subject; it is known that the subject was breathing at a relaxed rate and depth during this period.

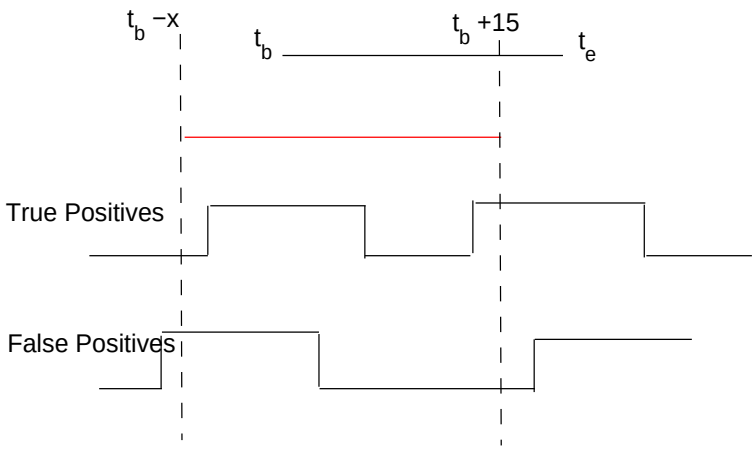

Figure 6.2: The time-match window for apnoea detection. The top line shows the timings of the apnoeic event. The red line shows the time match window and the line below this shows two true positive detections. The lowest line shows two false positive detections, the starts of the detections are outside the time match window. x is taken to be half the window length.

The breath detection threshold, THRESH_{amp} is set to be a percentage of the average peak-trough amplitude. If a breath is not detected for a given time threshold, THRESH_{time}, an apnoea is declared. A range of values of THRESH_{amp} and THRESH_{time} are investigated.

6.5 Frequency-domain approach

Respiration is not a stationary process, indeed it is abrupt changes respiration corresponding to apnoeas that are of interest. Spectral estimation requires stationary stochastic processes. In practice, however the definition of stationarity can be treated loosely [14]. For example, speech can be considered as quasi-stationary over small intervals and these short-time spectra provide useful information.

The approach of segmenting data in order to assume quasi-stationarity and hence compute the spectra (estimated by an AR model) is commonly used in the EEG, to detect for example changes in vigilance [89], or in sleep states [80]. Anderson *et al.* [7] use AR model coefficients computed over 0.25-second windows, to discriminate between various mental tasks from EEG recordings.

Cerutti *et al.* [19] use AR model estimation to detect atrial fibrillation episodes from the ECG.

[2]. If the AR spectrum of the windowed R-R interval sequence does not contain at least one relevant and stable rhythmic component, atrial fibrillation is assumed. The presence of the rhythmic component is assessed by looking at the maximum modulus of the complex poles of the model; if the maximum modulus is below a threshold, atrial fibrillation is defined.

In this work, an AR model is used to detect the presence (or absence) of dominant poles within the physiologically viable breathing frequencies. Since breathing usually occurs at a rate of between 6 and 25 breaths per minute, this range of frequencies is taken to be $0.1Hz$ to $0.42Hz$. The respiration waveform is windowed and the AR spectral estimate for each window calculated. The z-plane representation of the AR model is examined for the presence of a pole within the angular range that corresponds to breathing frequencies. The absence of this pole indicates an apnoea.

A "breathing-pole", z_b therefore has a phase ϕ_b that falls within the corresponding angular range in the unit circle

$$2\pi \times 0.1\Delta t \leq \phi_b \leq 2\pi \times 0.42\Delta t \tag{6.1}$$

The magnitude of the pole is also a decision criterion. As is discussed in Section 3.3 the magnitude of the pole in the z-plane is (non-linearly) related to the power in the estimated spectra. The more dominant a frequency in a signal, the nearer the position of the corresponding pole to the unit circle. A lower bound threshold, T_b is set. For a pole to be classed as a breathing-pole, z_b it must have phase ϕ_b and magnitude

$$T_b \leq |z_b| \leq 1 \tag{6.2}$$

Figure 6.3 shows five 15-second windows of the time-domain IP signal (upper plots) and the corresponding pole plots and estimated power spectrum (using a seventh-order AR model) directly below each of the time windows. Only the upper half of the complex plane is shown, as the pole plot is always symmetric about the real axis. The third time window shows an apnoeic period, while all the other windows show normal breathing. The pole plots show the area in which a "breathing-pole" must be found and is outlined in magenta. The phase

[2]Atrial fibrillation is an abnormal heart rhythm, characterised by an irregular heart rate, which may be very fast.

criterion from Eq. 6.1 is used and in this example the magnitude threshold T_b is set to be 0.8. It can be seen that the AR model estimate results in the presence of a breathing pole in the four windows of normal breathing, whereas in the window corresponding to an apnoea the breathing pole is absent.

6.5.1 Model order

When constructing an AR model it is important to determine the order of the model which best fits the data. The model order can be thought of as the number of past data samples needed to predict the present value of the data and it is therefore is dependent on the data sampling rate.

Since the best choice of model order is not generally known *a priori*, it is usually necessary in practice to postulate several model orders. Based on these, one then computes some error criterion that indicates which model order to choose [56]. Too low a value of model order results in a highly smoothed spectral estimate. Too high an order introduces spurious detail into the spectrum. There are a number of suggested criteria for selecting the optimum model order. As the prediction error of a model decreases monotonically with increasing model order, these criteria combine a goodness-of-fit term with a cost function that penalizes some measure of the model's complexity, i.e., some function of p [79]. Well established criteria include the Akaike information criterion (AIC) and the final prediction error (FPE) criterion [56]. Recent work by Broersen [13] suggests a number of criteria for finite samples. Samples are considered to be finite if the maximum candidate model order for selection is greater than $\frac{N}{10}$, where N denotes the number of observations. In this work, it is expected that the respiration signals have at most two predominant frequencies, one due to the breathing frequency and another due to the cardiac frequency that is often present in the recorded signal. Using this *a priori* knowledge of the signal, a model order of at least 4 seems suitable. Model orders from 3 to 9 are considered here.

Figure 6.4 shows the prediction error curve (E_p) with the corresponding FPE curve with respect to model order. The curves are calculated by averaging the model prediction error of 15-second windows of all the IP data used in this study. The FPE is defined as

$$FPE(p) = \left(\frac{N+p+1}{N-p-1}\right) E_p \tag{6.3}$$

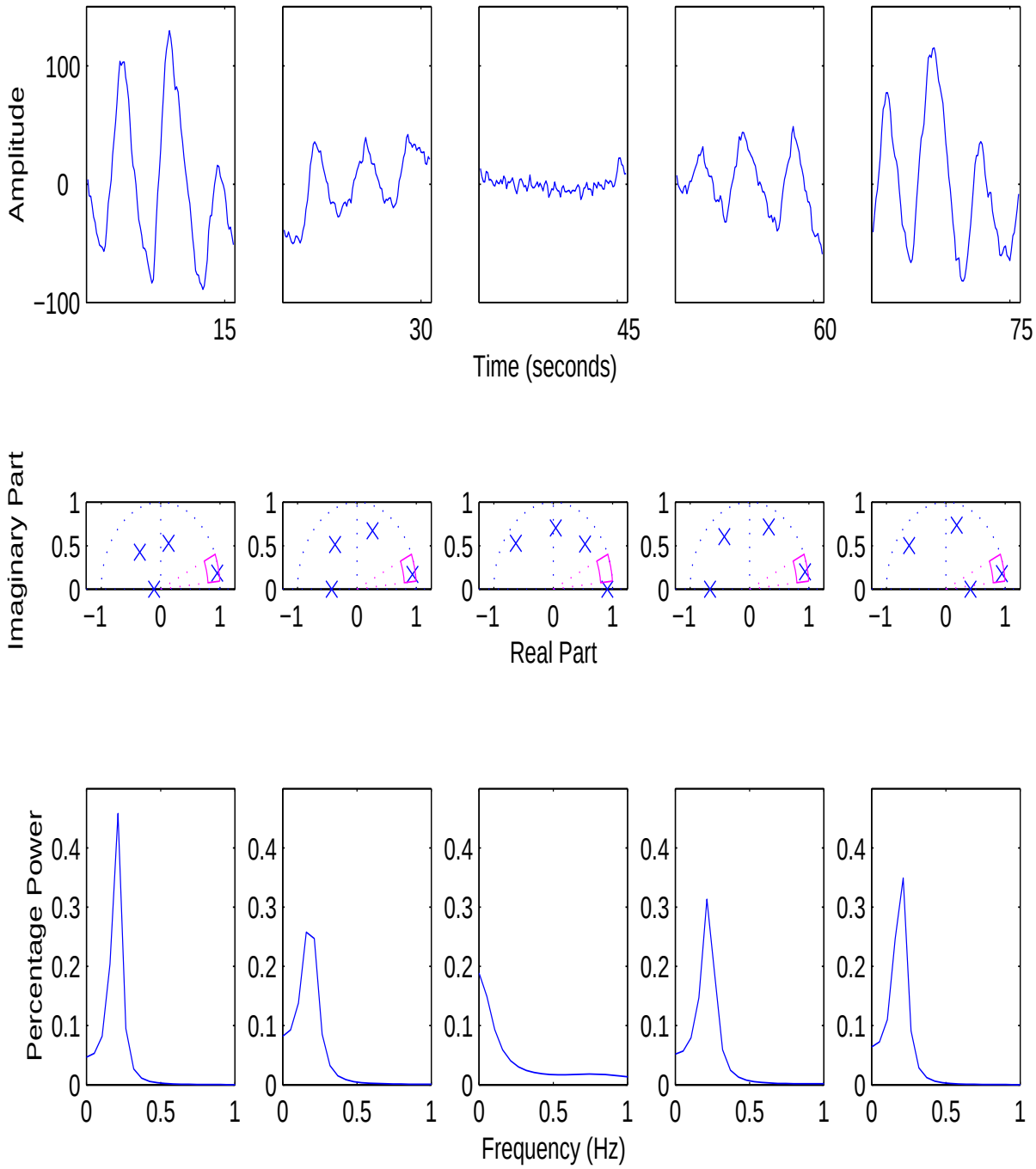

Figure 6.3: The upper row shows a windowed IP signal, with the corresponding pole plots and spectra estimated using a $7th$-order AR model below each time window. The "breathing-pole" region is outlined in magenta. The third column corresponds to an apnoea and it is seen that no breathing-pole is present in the corresponding pole plot. The data is from subject 3011 in the Controlled-breathing database.

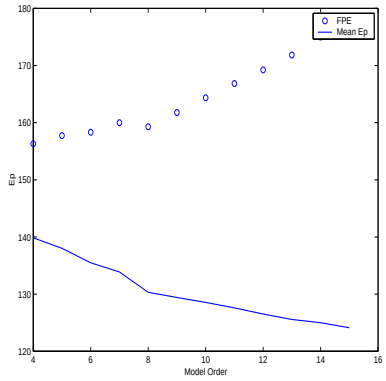

Figure 6.4: Prediction error with respect to model order with corresponding values of the FPE criterion.

The figure shows a minimum in the FPE criterion at a model order of 8, which also corresponds to a 'knee' in the prediction error curve. As a guideline the curves suggest that a model order of 8 represents the data most accurately. However in choosing the final model order, the performance of the different model orders in terms of their success in detecting central apnoeas is considered the most important criterion.

6.5.2 Windowing

As discussed in Section 6.3 when detecting apnoeas, latency is an issue: an alarm should be generated as soon as possible. Any window needs to be short enough in duration to give a safe time delay between the start of and the detection of an apnoea. Apnoeic episodes can be as short as 10 or 12 seconds, giving an idea of a suitable window length.

A window length of 15 seconds is selected, and the evaluation is carried out on both overlapping windows (by 10 seconds) and non-overlapping windows.

6.6 Results

A high sensitivity in the detection of apnoeas is the most important performance criterion, but Section 6.2 illustrates that in practice a high sensitivity is often achieved at the cost of a very low positive predictivity.

Initial analysis of the Controlled-breathing database showed that the application of the time and frequency-domain algorithms to respiratory waveforms derived from the ECG and PPG

results in very poor performance. For example, when applying the time-domain approach to an ECG-derived respiratory waveform, the RSA-DR, results show that a sensitivity of greater than 98% can not be reached without compromising the positive predictivity to a value of less than 10%. There are similar findings for the frequency-domain approach. This result is interesting in view of a recent study by Thayer *et al.* [103]. In this work an AR model estimation of the R-R interval sequence is used to obtain an estimate of the breathing frequency. It is concluded that the estimated frequency, obtained by measuring the phase of the relevant spectral peak is a useful indicator of the known breathing frequency. It should be noted however that the estimations are made over windows of duration of at least 3.5 minutes (c.f. 15 seconds used here) and that the breathing rates are known to remain constant over these windows. It is suggested therefore that 15 seconds is too short a window over which to estimate the breathing frequency using this approach.

The application of both approaches to the PPG-derived respiratory waveform, the PP-DR, result in a similarly low positive predictivity. Therefore the results reported in the rest of the chapter focus entirely on the analysis of the IP waveform.

6.6.1 Time-domain approach

Figure 6.5 shows the sensitivity (upper plot) with respect to the amplitude threshold THRESH_{amp} for different time thresholds THRESH_{time}. As expected, increasing the value of the thresholds results in a higher sensitivity at the cost of a lower positive predictivity. A number of threshold conditions result in very high sensitivities, but it is decided to aim for a combination that results in the highest positive predictivity given a sensitivity of at least 98%.

A time threshold, THRESH_{time} of 5.5 seconds together with an amplitude threshold, THRESH_{amp} of 0.35 result in a sensitivity of 98% and positive predictivity of 43%.

6.6.2 Frequency-domain approach

The use of an AR model with 15 second non-overlapping windows invariably gives a sufficiently high sensitivity (of at least 98% and in some threshold conditions 100%). The introduction of overlap in the windowing simply results in a lower positive predictivity.

Figure 6.6 shows the sensitivity (upper plot) and positive predictivity of the different model

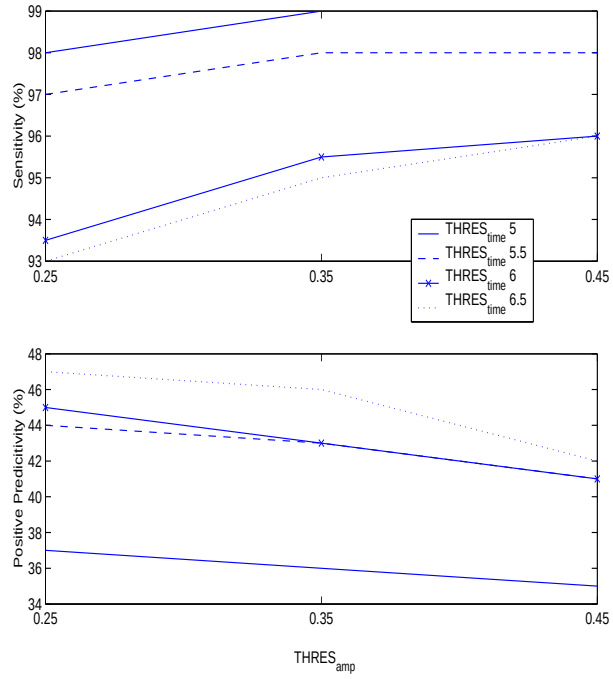

Figure 6.5: Sensitivity (upper plot) and positive predictivity (lower plot) for different values of THRESH$_{time}$ with respect to THRESH$_{amp}$.

orders (from 5 to 9), as the magnitude threshold T_b is increased from 0.4 to 0.9 when applying the pole tracking method to the IP signal.

Compared to the lower order models, the models of higher order (8 and 9) show the lowest sensitivity at all values of T_b, though the difference is more significant at lower values of T_b. This suggests that the model order is too high; the complex poles, in the absence of definite periodic frequencies, are dispersed in phase around the origin and there is a higher likelihood that one of these poles happens to fall into the breathing phase range (Eq.6.1), with the result that the apnoea is not detected. Figure 6.7 shows an example of this phenomenon. The left-hand column shows the windowed IP signal (time going from top to bottom), apnoeic episodes being plotted in red. The middle column shows the corresponding pole plots for the $9th$-order model, and the right-hand column the corresponding pole plots for the $5th$-order model. For the time windows corresponding to the first and third apnoeas, the $9th$-order model disperses the poles approximately equally in phase around the origin, with the moduli of the poles tending to be high (in this case the moduli are greater than 0.8, the value of the magnitude threshold T_b chosen in this example). The model is trying to fit the noise that occurs in the IP signal during the apnoea. The $5th$-order model however gives poles of

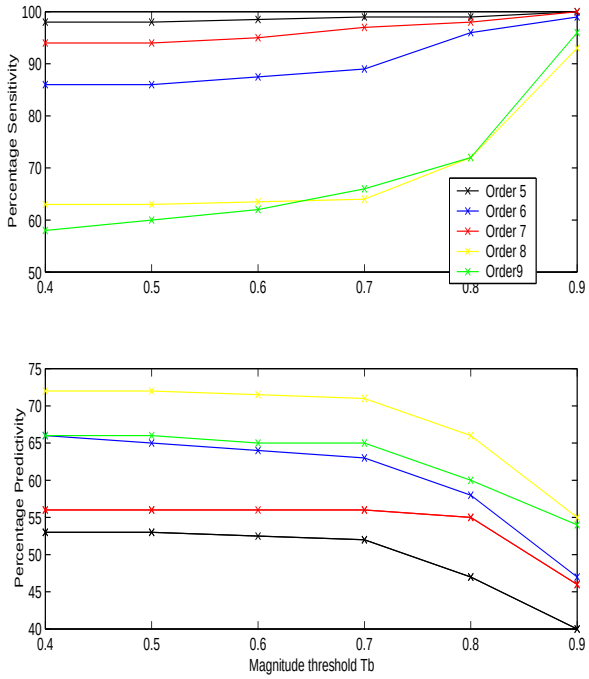

Figure 6.6: Sensitivity (upper plot) and positive predictivity (lower plot) for AR models of orders 5-9 with respect to T_b.

lower modulus resulting in a flat spectrum more representative of the random noise in the IP signal at this time.

Conversely the $5th$-order model gives a very high sensitivity at the cost of a lower positive predictivity for all values of T_b. The high sensitivity and low positive predictivity suggests that at times a model of $5th$ order is not sufficiently complex to estimate the IP signal. Figure 6.8 illustrates this with an example of a noisier IP signal than that shown in Figure 6.7. In the second row of the figure, the $9th$-order model can be seen to have sufficient poles to fit some of the dominant noise as well as the breathing component that can be detected visually. In contrast, there are not enough poles available in the $5th$-order model. This phenomenon also occurs in the seventh and tenth rows of this figure.

The sensitivity and positive predictivity curves for each model order show similar trends in Figure 6.6. In most cases there is a greater gradient in the positive predictivity and sensitivity curves when T_b increases from 0.7 to 0.9 than at lower values of T_b. This shows that there are some poles in the breathing phase range (Eq.6.1) that have magnitudes of approximately 0.7 that correspond to both breathing and apnoea time windows.

If the curves are extrapolated back to lower values of T_b, it can be seen that positive pre-

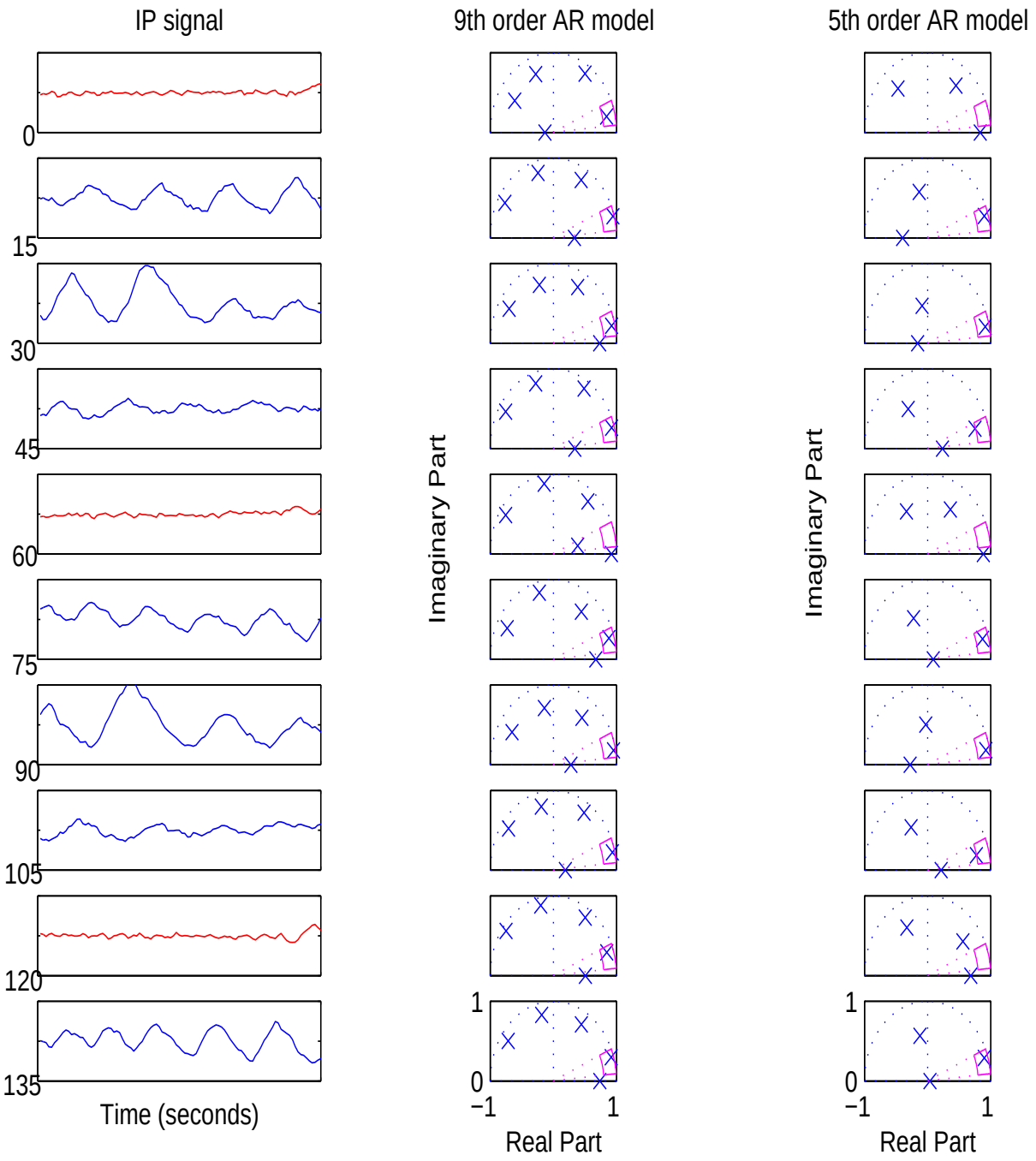

Figure 6.7: The first column shows the windowed IP signal with the corresponding AR models of $5th$ and $9th$ order to the right of the relevant window. The apnoeic events are shown in red in the IP signal. The $9th$ order model places poles in the "breathing pole" area resulting in a low sensitivity in detecting apnoeas. The x-axis scales on the left hand column correspond to the time at the start of the window, the bottom pole plots show axis scales that are the same for all the pole plots.

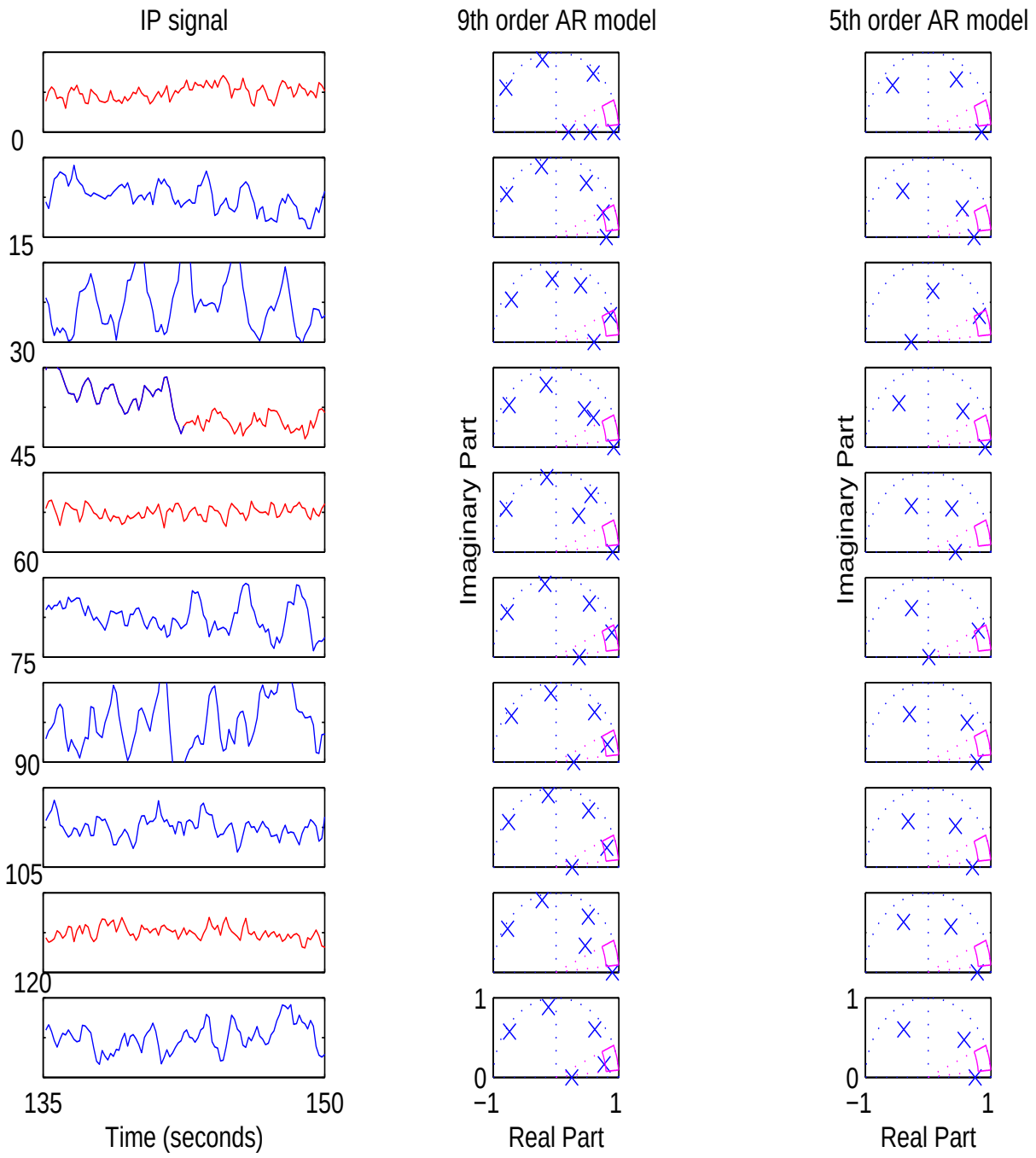

Figure 6.8: The first column shows the windowed IP signal with the corresponding AR models of $5th$ and $9th$ order to the right of the relevant window. The apnoeic events are shown in red in the IP signal. The $5th$ order model does not have enough complex pole pairs to accurately model the noisier IP signal, hence poles are not placed in the "breathing pole" area and apnoeas are falsely detected. The x-axis scales on the left hand column correspond to the time at the start of the window, the bottom pole plots show axis scales that are the same for all the pole plots.

Table 6.1: The values of sensitivity and positive predictivity (%) for varying values of T_b when detecting apnoeas from the IP signal using AR models of $6th$ and $7th$ order.

	Model Order 6		Model Order 7	
T_b	Se	+P	Se	+P
0.4	86	66	94	56
0.5	86	65	94	56
0.6	86	65	95	56
0.7	89	63	97	56
0.8	96	58	98	55
0.9	99	47	100	46

dictivity does not reach 100%, implying that in some cases, a pole does not exist within the breathing phase range, even when breathing is occurring. Thus there are always false positives regardless of the value of T_b.

A $6th$ or $7th$-order model gives the best compromise between sensitivity and positive predictivity, implying that these provide better models for the purpose of detecting apnoeas than either the $5th$ or $9th$ order models. Table 6.1 shows the sensitivity and positive predictivity values for increasing values of T_b for $6th$ and $7th$ order models. When T_b is set to 0.9, the models show similar results with sensitivities of 99 and 100% and positive predictivities of 47 and 46%. A model order of 7 is selected with $T_b = 0.8$ as this satisifies the criteria for selecting the time-domain method thresholds, i.e., the highest positive predicitivity given a sensitivity of at least 98%.

6.7 Fusion

Information fusion is the process of combining evidence from different information sources in order to make a better judgement. It can also be the process of combining information from different algorithms applied to the same sensor output to make a better decision. As discussed in Section 3.10 fusion can take place at different levels of representation, at data, feature or decision level. Decision level fusion combines the decisions from each information source. This process reduces diverse data to a common format (binary decisions or detection probabilities), which are readily combined using conventional statistical techniques. Decision level fusion can be categorised into hard and soft decision level fusion; hard decisions being all-or-nothing declarations, while soft decisions utilise information about the confi-

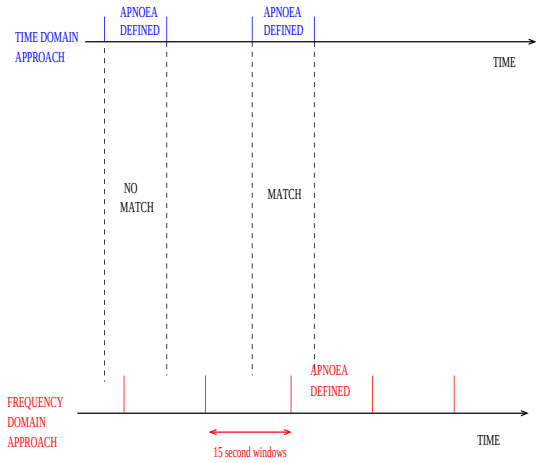

Figure 6.9: Illustration of the definition of concurring apnoeas in the time and frequency-domain approaches.

dence each sensor places in its individual decisions [43].

In the work described in the previous sections, two separate decisions regarding the presence or absence of a central apnoea are made from the same IP signal, one from a time-domain method and the other from a frequency-domain method. The decisions are considered as hard decisions. They are reached using threshold techniques in both the time and frequency-domain methods. The values of these thresholds are selected as they give the best results on the data available.

Both the time and frequency-domain approaches give a very high sensitivity; the frequency-domain approach, 98% and time-domain 98%, with positive predictivity values of 55% and 43% respectively. To improve the accuracy of detection, the number of false positives should decrease without compromising sensitivity. To investigate whether logical operators combining the information from the frequency and time-domain would produce this, the coincidence in the timing of the false positives from the two approaches is investigated.

The timing of the apnoea detections is defined to match if the apnoeic period defined by the time-domain approach overlaps (to any extent) the window over which the frequency-domain approach declares an apnoea. Figure 6.9 illustrates this definition.

On the database studied in this chapter, the frequency and time-domain approaches result in 88 and 139 false positives respectively. Of these 58 are coincident. This suggests that an **AND** voting system would improve the accuracy of detection by lowering the incidence of

false positives, assuming coincidence of timing for true positives.

An **AND** voting decision level fusion approach is therefore investigated as a means of fusing the outputs from the two analysis methods. For each 15-second window, a hard (binary) decision is made by the frequency and time-domain approaches. These decisions are fused using an **AND** voting rule. In light of the results presented in Section 6.6, the AR model of order 7 is chosen for the frequency-domain approach. For the time-domain method, the chosen combination is a THRESH_{amp} of 0.35 and THRESH_{time} of 5.5 seconds, as reported in Section 6.6.

6.7.1 Results

Figure 6.10 shows the performance curves (in red) resulting from the decision level fusion. For comparison the figure shows the performance curves obtained using the frequency-domain approach (with a $7th$ order model) as this gives the better results of the two unfused approaches. The sensitivity (upper) and positive predictivity (lower) plots show the same scales. It can be seen that for a small decrease in sensitivity (∼1%), an increase in positive predictivity of 10% can be gained by using the decison fusion algorithm. Hence the optimal apnoea detector is one in which the outputs from the time-domain (with THRESH_{amp} = 0.35 and $\text{THRESH}_{time} = 5.5$ seconds) and the frequency-domain approach (with AR model of 7 and T_b =0.8) are fused to give a sensitivity of 97% and a positive predictivity of 65%.

6.8 Conclusions

This chapter investigates how accurately central apnoeas can be detected from the non-invasive signals, IP, ECG, and PPG.

Time-domain and frequency-domain approaches are employed. The time-domain method is based on the simple methods for detection of central apnoeas described in the literature. The frequency-domain method estimates the spectra of windowed respiratory signals using an AR model. The identification of apnoeas relies on the positions of the poles in the pole plot. For normal breathing it is assumed that a pole has to occur with a phase within a physiologically plausible range and a magnitude above a certain threshold.

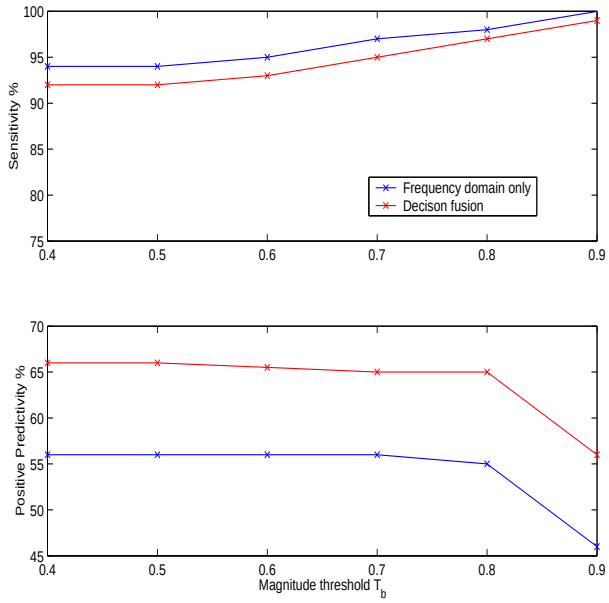

Figure 6.10: Sensitivity (upper plot) and positive predictivity (lower plot) with repect to pole magnitude T_b for the fused estimate.

It is clear that the use of the ECG and PPG signals gives very inaccurate detection of central apnoeas on the Controlled-breathing database. The chapter therefore focused on detection from the IP waveform. With the proviso that sensitivity should be at least 98%, it is found that the time and frequency-domain approaches result in a positive predictivity of 43% and 55% respectively. A further increase of 10% to a positive predictivity of 65% is obtained by fusing the decisions from the two approaches at a cost of only 1% in sensitivity.

As concluded by a number of authors including [15], [12] and [116] the IP signal is inherently noisy, partly due to cardiac artefact. It is demonstrated in Chapter 5 that the accuracy of breath detection from the IP signal could be greatly improved by the fusion of information from other signals, namely the ECG and PPG. However in the detection of central apnoeas, using the methods proposed in this chapter, the ECG and PPG signals do not provide sufficiently accurate information to consider fusing this information with that obtained from the IP signal.

Chapter 7

Obstructive Sleep Apnoea

7.1 Introduction

Detecting obstructive sleep apnoea (OSA) is a very different problem to that of detecting central apnoeas. Obstructive apnoeas occur when the airway becomes obstructed, but respiratory effort still continues. Due to the continuation of respiratory effort the single impedance plethysmography (IP) signal indicates the occurrence of normal breathing.

This chapter investigates the use of a statistical classifier, a neural network, to assign windows of the ECG and blood pressure signal into the set $\{apnoeic, non-apnoeic\}$ (apnoeic referring to OSA). Reflection coefficients of an AR model estimation of derived respiratory waveforms are used as the feature (input) vectors to the classifier. Although various frequency-domain representations of the ECG and blood pressure signals have previously been used as features for classifying OSA, the use of a reflection coefficient parameterisation is novel.

Fusion of information, from the different respiratory waveforms, at both a decision and feature level is considered and the classification performance of the fusion methods compared with that of classifiers that use information from only a single derived respiratory waveform. It is seen that a classifier that fuses at a feature level gives the best classification performance.

Section 7.2 outlines the effect of OSA on the ECG and blood pressure signal and Section 7.3 discusses current methods of detecting OSA. A review of signal processing work that

has been carried out on detecting episodes of OSA from the ECG is given in Section 7.4. Section 7.6 describes the proposed classification system from feature extraction to training and testing of the classifier. The results of the classification of the Polysomnography database are presented in Section 7.7. Fusion techniques are put forward in Section 7.8 and the results of these techniques are given in Section 7.8.3. The classification system is validated on another database, which is used in the previous works described in Section 7.4. Finally a discussion of the results and conclusions are found in Sections 7.10 and 7.11 respectively.

7.2 Effect of OSA on the signals

The physiological phenomena associated with obstructive sleep apnoea are far more complex than those associated with central apnoea. To maintain successful respiration the upper airways have to be kept open by activation of the upper airway muscles. This does not happen during obstructive sleep apnoeas. During each apnoea the upper airways collapse while respiratory efforts persist. The repetitive respiratory effort originating from diaphragm contractions causes effort-related negative intra-thoracic pressure changes. As the apnoea proceeds, the oxygen saturation drops continuously. At the same time the respiratory effort increases in amplitude. Finally the low oxygen levels lead to a central nervous activation, an arousal. An arousal is a shift from the current sleep state to a lighter one or wakefulness and is usually detected by an abrupt shift in the electroencephalogram (EEG). The arousal causes a re-establishing of respiratory intervention and an opening of the upper airways for the next few breaths. The arousal also causes sleep fragmentation, an interruption of the sleep stage, without making the patient aware of this phenomenon [82].

The following sections illustrate the effects of obstructive sleep apnoea on the ECG and blood pressure signals.

7.2.1 Effect of OSA on blood pressure

Systolic blood pressure falls during an OSA and then rises sharply in the period following the apnoea which is coincident with an arousal. The magnitude of the rise in blood pressure following apnoea can be large (more than $10mmHg$ systolic). This blood pressure change occurs reproducibly after every apnoea and hence potentially hundreds of times throughout the

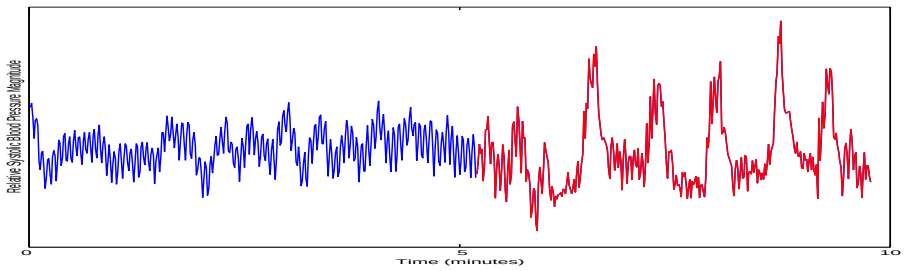

Figure 7.1: Systolic blood pressure (PP-DR) from record slp03 of the Polysomnography database. The first five-minute section (blue) corresponds to a period of quiet sleep, while the second five-minute section (red) shows an obstructive apnoeic period. Five episodes of a decrease followed by a sharp increase in systolic blood pressure can be seen. These episodes correspond to the apnoea labels.

night [24]. Several clinical physiological studies have reported and analysed this response [81]. The size of the inspiratory blood pressure decrease is a semi-quantitative indicator of the degree of effort [25]. Figure 7.1 shows the systolic blood pressure waveform (PP-DR) from a section of the record slp03 of the Polysomnography database. The first 5-minute section corresponds to a period of quiet sleep, while the second section shows a period during which cyclical obstructive apnoeas occur. For each of the five apnoeic episodes, a sharp increase, preceeded by a decrease in systolic blood pressure can be seen. This phenomenon has been considered as a possible indirect marker in the assessment of patients thought to be suffering from obstructive sleep apnoea [25].

7.2.2 Effect of OSA on the ECG

More than 80% of patients with OSA demonstrate a prominent sinus brady-tachyarrhythmia [1] in association with an apnoeic event [81]. Bradycardia during apnoea is reported in some of the earliest studies of patients with OSA [118]. However the relationship between heart rate and OSA is not simple. Recent studies have found that the decrease in heart rate during apnoea varies, in part depending on the current sleep stage. Occasionally, individuals with well characterised OSA do not demonstrate apnoea-related decreases in heart rate. The relationship between heart rate and apnoea depends on the degree of hypoxaemia and oxyhaemoglobin desaturation. In summary, short apnoeas without hypoxaemia do not result in bradycardia, and the combination of apnoea with hypoxaemia appears to be required to

[1] A marked increase in heart rate, followed by a slower than normal heart rate

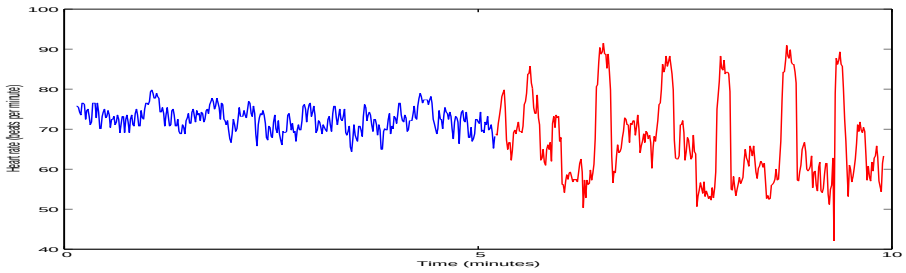

Figure 7.2: R-R interval signal (RSA-DR) from record slp03 of the Polysomnography database. The first five minute section (blue) corresponds to a period of quiet sleep, while the second five minute section (red) shows an obstructive apnoeic period. Five cyclical variations in heart rate can be seen, which correspond to the apnoea labels.

produce significant bradycardia [118].

Guilleminault *et al.* [42] were the first to show that the presence of cyclical variation of heart rate (CVHR) can identify patients with sleep apnoea. Stein *et al.* [98] discovered that there is no common heart rate pattern associated with apnoea, the patterns varying across and within subjects. However they were successful in identifying 29 out of the 30 apnoea patients in the Computers in Cardiology 2000 competition [50] by a visual examination of the heart rate signal.

Figure 7.2 shows the R-R interval waveform (equivalent to the RSA-DR) from two five-minute sections of the record slp03 from the Polysomnography database. The first section (blue) is taken from a period of quiet sleep, while the second section (red) shows a period where cyclical OSAs are taking place. The corresponding PP-DR is shown in Figure 7.1.

Raymond *et al.* [87] note that each apnoea causes a swing in the baseline of the ECG-derived respiratory waveform (EDR). Figure 7.3 shows the R-DR, during the same time period represented by Figures 7.1 and 7.2. The swings in baseline can be seen clearly as each of the five apnoeas occurs.

7.3 Detecting apnoea

OSA is typically treated by surgical intervention or nasal continuous positive airway pressure (CPAP) treatment. Given the prevalence of OSA and the availability of treatment options, it is important that individuals suffering from the disease are identified. The defini-

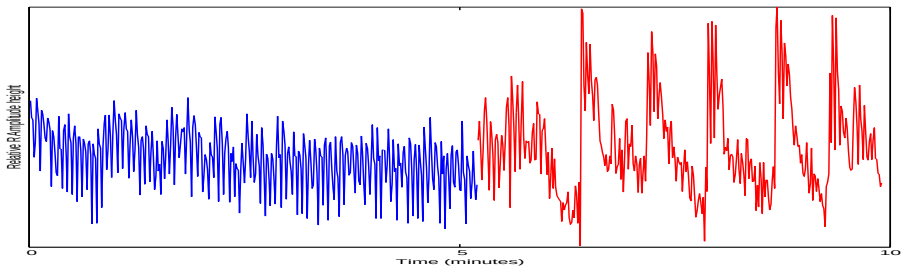

Figure 7.3: R-DR from record slp03 of the Polysomnography database. The first five minute section (blue) corresponds to a period of quiet sleep, while the second five minute section (red) shows an obstructive apnoeic period. Five major oscillations in the signal can be seen and these correspond to the apnoea labels.

tive diagnosis is based on standardised polysomnography (PSG) techniques with overnight recordings of sleep stage (assessed from two channels of EEG and EOG), respiratory effort, oronasal airflow, ECG analysis and oxyhaemoglobin saturation parameters in an attended laboratory setting. This "gold standard" is labour intensive, requires considerable instrumentation, and is expensive to conduct [61]. The raw data from the overnight recordings is screened by clinicians to look for events in the recorded signals that indicate apnoeas or other sleep-related breathing disorders. More recently there have been efforts made at designing portable apnoea monitors. Man *et al.* [61] carried out a study to validate one of these devices. Rather than record the whole set of signals associated with conventional PSG, only oronasal airflow, thorax and abdominal movement, oximetry, ECG and body position were recorded, omitting the EEG. However although the system gave very high sensitivity and specificity compared with a full PSG, it still required experts to analyse the raw data retrospectively, as with conventional PSG analysis.

OSA is associated with arousals. The recognised way to detect an arousal is from the EEG, but this is difficult and time consuming [27]. Pitson and Stradling [84], recognising the need for less labour intensive measures of sleep fragmentation in patients undergoing investigation for OSA, considered a number of non-invasive autonomic markers for the recurrent arousals believed to be stimulated by the increases in inspiratory effort in OSA patients. An automated rule-based method was used to search for transient rises (above certain threshold values) in blood pressure and heart rate. It was found that blood pressure rises were better correlated with the more conventional indices (i.e., EEG micro-arousals and oxygen saturation dips) of sleep fragmentation than heart rate rises. It was concluded that auto-

matic detection of blood pressure rises may provide a useful alternative to manual scoring of EEG micro-arousals (the accepted gold standard) according to the American Sleep Disorders Association (ASDA) criteria.

7.4 Review of methods for detecting OSA from the ECG

Until recently there has been very little signal processing or pattern recognition work carried out on the detection of OSA events from signals recorded routinely in clinical environments. In 2000, the organisers of the IEEE Computers in Cardiology conference [50], made available a database consisting of a single channel of ECG recorded from subjects suffering from OSA. Competitors were invited to either:

a. classify the test recordings as being from subjects who did or did not suffer apnoeas, or

b. generate a minute-by-minute annotation file for each recording specifying when apnoeas occurred.

The goal of the competition was to demonstrate the efficacy of ECG-based methods for detecting apnoea using a large, well-characterised, and representative set of data [50]. Details of the most successful methods are published in the Proceedings of the Computers in Cardiology conference (2000). Both rule-based and learning algorithms were applied to the classification task. The majority of the successful methods used information from the heart rate (R-R interval) signal, computing the short-term magnitude frequency spectrum, and calculating the energy in the low-frequency band (about 0.01 Hz — 0.07 Hz) which is increased during OSA as apnoeas tend to occur periodically (with a period of the order of 15-100 seconds, this periodicity is seen in Figures 7.1, 7.2 and 7.3). Other authors concentrated on the tachy-bradycardia time-domain patterns of heart rate that tend to occur with OSA. The most successful approaches to *automatically* classifying minute-by-minute result in accuracies ranging from 63% to 85%.

Raymond *et al.* [87] report on a mixture model classifier using one feature vector (of only four elements) per minute of recording. Two of these features consist of two discrete-harmonic-wavelet coefficients representing frequency spectra measures of an EDR. The EDR in this work is generated from measurements of the T-wave amplitude. Apnoea causes a swing in the baseline of the EDR. Therefore the wavelet coefficients used as features correspond

to the frequency of the periodic apnoea episode repetition. The arousals that terminate an obstructive sleep apnoea event are marked by a tachy/bradycardia cycle in heart rate. The slope and magnitude of the cycle make up the remaining two features of the mixture model. This method resulted in an accuracy of 81%.

Garcia *et al.* [75] use a wavelet coefficient representation of the ECG to construct a Bayesian hierarchical model. The posterior probability that the test coefficients fit into the class model is calculated and if it exceeds a pre-defined threshold the signal is accepted as belonging to that model. An accuracy of 63% is reported.

McNames *et al.* [66] use various features of the ECG such as the heart rate, S-wave amplitude and an amplitude measurement of the QRS complex to train a hidden Markov model for classification (the exact features used are not specified). Using this approach 78% of the minute epochs are correctly classified.

Shinar *et al.* [93] employ a rule-based approach. A number of parameters of the R-R interval sequence are examined including the power in the frequency range 0.01-0.04 Hz, as well as transient decreases in the time-domain signal. If these parameters do not meet set thresholds an apnoea is defined in the segment of data being examined. This method gives an accuracy of 85%.

Mietus *et al.* [69] apply a Hilbert transform to the R-R interval sequence to derive instantaneous amplitudes and frequencies of the sequence. After calculating their averages and standard deviations over a 5-second moving window, a thresholding technique is used and windows within the threshold limits are detected as apnoeas. 84.5 % of windows are identified correctly.

7.5 Detection of obstructive sleep apnoea from the Polysomnography database

As discussed in Section 7.4 the work published on the detection of obstructive sleep apnoea from the ECG uses a variety of measures (features) of the signal for classification, these features being derived from both the time and frequency-domains. The features are either extracted directly from the ECG signal or from waveforms derived from the signal. In the

rest of this chapter the use of features extracted from respiratory waveforms (derived from the ECG and blood pressure) to classify 30-second epochs of data into the set $\{apnoeic; non-apnoeic\}$ is investigated. The respiratory waveforms are those referred to as R-DR, RSA-DR and PP-DR in Chapter 4.

The use of information extracted from each individual waveform is first considered. Information from the different waveforms is then combined using fusion either at a feature or decision level to investigate whether this improves the classification performance.

To compare the results with those published in the literature, the proposed methods are also applied to the Computers in Cardiology database referred to in Section 7.4. Since this database consists only of the ECG signal, it only allows a direct comparison of classification based on features derived solely from the ECG.

7.5.1 The data set

The Polysomnography database is made available with a set of labels, corresponding to each 30-second epoch of data. The labels define events that occur within the 30-second epoch. The aim here is to distinguish between normal and obstructive apnoeic epochs and so all epochs that are labelled as either obstructive sleep apnoea or normal sleep are selected to construct the data set. Any epochs labelled as hypopnea or central apnoea are discarded.

7.5.2 Evaluation parameters

For a two-class problem such as this the evaluation parameters used are sensitivity (defined in Chapter 4), specificity and accuracy. Specificity is the fraction of normals (in this problem an apnoea is defined as a true event, therefore a normal is a negative event) correctly detected:

$$Specificity = \frac{TN}{TN + FP} \times 100\% \tag{7.1}$$

Accuracy is defined as the percentage of the correct decisions for both the normal and apnoea classes:

$$Accuracy = \frac{TP + TN}{TN + TP + FN + FP} \times 100\% \tag{7.2}$$

The Receiver Operating Characteristic (ROC) of the classification systems is also examined. An ROC curve is a plot of *Sensitivity* (y-axis) with respect to *1-Specificity* (x-axis) as a threshold varies. The curves illustrate two main points:

- the trade-off between sensitivity and specificity (any increase in sensitivity is accompanied by a decrease in specificity).
- the area under the curve is a measure of classifier accuracy.

7.6 The classification system

7.6.1 Features

It can be seen from Section 7.4 that the majority of previous studies on the detection of obstructive apnoeic episodes from the ECG use frequency spectra measures of the R-R interval sequence. However none of them use an AR model estimation, which is an efficient way of representing the frequency-spectra (See Chapter 3). In this work three derived respiratory waveforms (R-DR, RSA-DR and PP-DR) are parameterised using the reflection coefficients of an AR model. As is described in Section 3.3.1, the reflection coefficients define the reduction in residual signal-model error, E, when the AR model increases its order from $m-1$ to m,

$$E_m = \left(1 - \rho_m^2\right) E_{m-1} \tag{7.3}$$

Reflection coefficients have the advantage that an increase in model order does not effect the coefficients from the previous order. Hence there is little cross-correlation between the coefficients, making them more suitable for pattern analysis techniques than the correlated AR coefficients [23]. A further advantage is that reflection coefficients always lie between -1 and +1, hence each feature has a similar dynamic range.

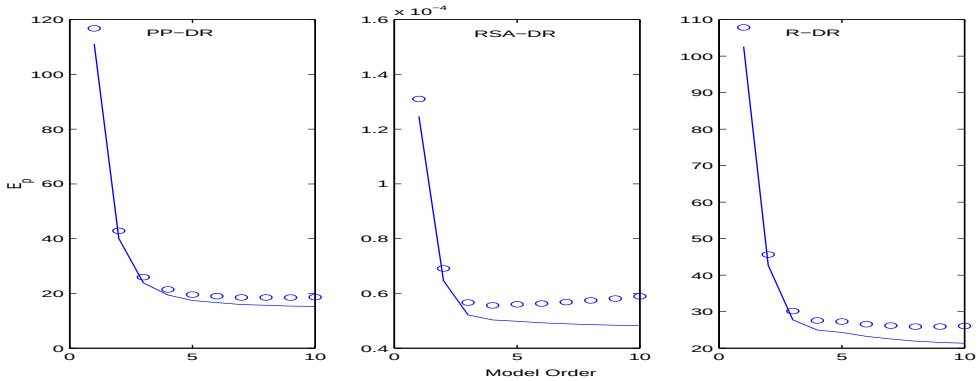

Figure 7.4: Prediction error (E_p) and FPE curves with respect to model order for each of the three derived respiratory waveforms, PP-DR (left), RSA-DR (middle) and R-DR (right). The continuous line corresponds to the E_p curve and the circles to the FPE curve.

All three of the derived respiratory waveforms are unevenly sampled, as they are derived from direct measurements of the source signals. Cubic-spline interpolation and resampling at $4Hz$ are therefore used to generate an evenly-sampled time series.

As in Chapter 6 the model order is chosen taking into consideration both prior knowledge of the signal and the prediction error curves. During normal breathing, the R-DR and PP-DR are expected to exhibit one dominant frequency, while the RSA-DR (R-R interval sequence) is know to contain two dominant frequencies. During OSA the derived signals exhibit a change in characteristic (this is shown in Figures 7.1, 7.2 and 7.3), with a greater proportion of the power in the frequency spectra being at lower frequencies. Figure 7.4 shows the prediction error (E_p) curve and the corresponding FPE curve with respect to model order for each of the derived respiratory waveforms. The curves are calculated by averaging the AR model prediction error for all the 30-second windows in the data set.

The “knee” in the prediction error curves occurs around a value of 3 or 4 for all the derived waveforms. There are no distinct minima in the FPE curves though in all cases the curves either flatten or show a slight increase at a model of order of 5. In view of this a model order of 5 is selected. Figure 7.5 shows the power spectra obtained using a $5th$ order AR model for epochs of each of the derived respiratory waveforms for a subject from the Polysomnography database. The red curves show the power spectra from apnoeic epochs and the blue curves show the power spectra estimates from non-apnoeic epochs. It can be seen that the curves for the two classes are different; hence the AR reflection coefficients are expected to show discriminatory information.

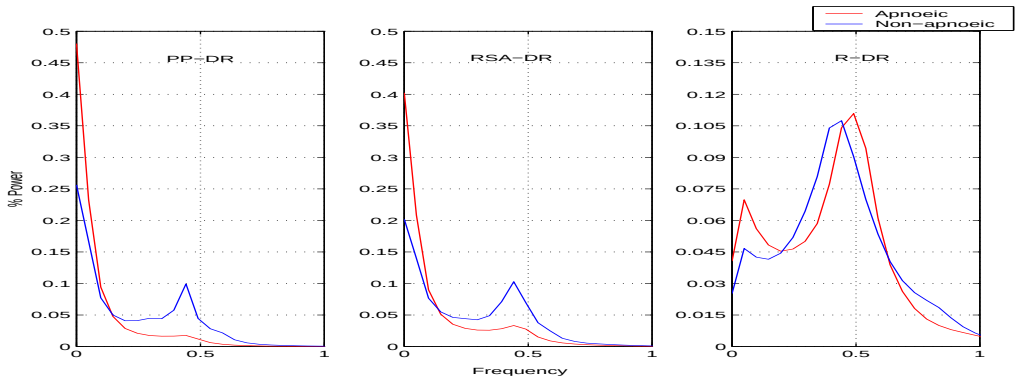

Figure 7.5: Averaged PSDs for the three derived respiratory waveforms, PP-DR (left), RSA-DR (middle) and R-DR (right) for a subject from the Polysomnography database (slp59). The PSDs are estimated with a $5th$ order model and the resultant curves are obtained by averaging the PSD corresponding to every epoch of each class.

7.6.2 Training procedure

In this work the limited amount of available data in the Polysomnography database does not allow the data to be partitioned into equal training and test data sets. Therefore a cross-validation procedure is used [10]. The data is divided into N segments. In this work each of the N segments corresponds to data from each of the N subjects. A network is then trained using data from $N-1$ of the segments and its performance evaluated using the Nth subject. The process is repeated for each of the N possible combinations of the data. This method is sometimes referred to as the "leave-one-out" method.

A balanced database is constructed by selecting an equal number of apnoea and normal epochs from each subject. The number selected is the minimum number of apnoea or normal epochs available for that subject. Table 7.1 shows the partitioning of data for the training and testing of N networks.

7.6.3 Parameter optimisation

As is discussed in Section 3.7 various minimisation algorithms exist for training a neural network.

A scaled conjugate gradient method is used for training the MLPs with a regularisation parameter to ensure that the networks are not over-fitted. The regularisation parameter is chosen to be 0.1, and a total of 1000 training epochs are carried out. On checking the training

Table 7.1: The $N = 10$ different data partitions used in the "leave-one-out" training and testing procedure. The total number of epochs is 2348.

Partition No.	Training data subjects	No. of epochs	Test data subjects	No. of epochs
1	slp03,slp04,slp14,slp16,slp48,slp59,slp60,slp61,slp66	2262	slp02b	86
2	slp02b,slp04,slp14,slp16,slp48,slp59,slp60,slp61,slp66	2278	slp03	70
3	slp02b,slp03,slp14,slp16,slp48,slp59,slp60,slp61,slp66	2076	slp04	272
4	slp02b,slp03,slp04,slp16,slp48,slp59,slp60,slp61,slp66	2154	slp14	194
5	slp02b,slp03,slp04,slp14,slp48,slp59,slp60,slp61,slp66	1934	slp16	414
6	slp02b,slp03,slp04,slp14,slp16,slp59,slp60,slp61,slp66	2064	slp48	284
7	slp02b,slp03,slp04,slp14,slp16,slp48,slp60,slp61,slp66	2186	slp59	162
8	slp02b,slp03,slp04,slp14,slp16,slp48,slp59,slp61,slp66	1930	slp60	418
9	slp02b,slp03,slp04,slp14,slp16,slp48,slp59,slp60,slp66	1928	slp61	420
10	slp02b,slp03,slp04,slp14,slp16,slp48,slp59,slp60,slp61	2320	slp66	28

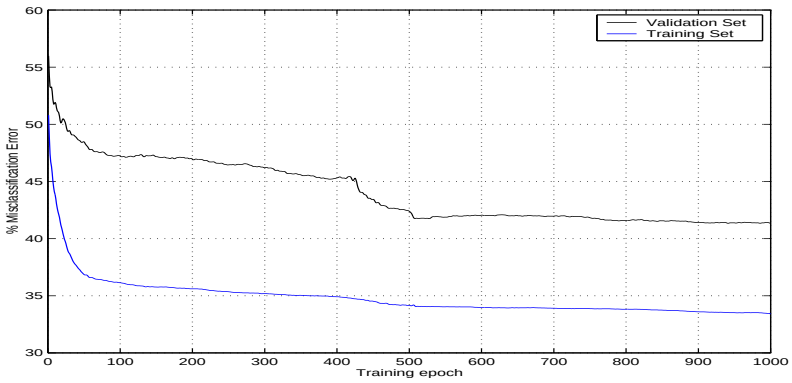

Figure 7.6: Misclassification error (%) with respect to the number of training epochs of a 5-5-1 MLP. The MLP is trained on the feature set from the R-DR and the errors are averaged over the N data partitions.

set error term during training it is seen that this value ensures that under-training does not occur. Figure 7.6 shows the misclassification error with respect to the number of training epochs for a 5-5-1 MLP trained on the R-DR feature set. It can be seen that any change in misclassification error is negligible at 1000 training epochs. Similar curves are obtained for the PP-DR and RSA-DR feature sets.

7.6.4 Network architecture optimisation

The optimum network architecture for the chosen features ($5th$ order reflection coefficients) is investigated using a validation set. This validation set comprises 20% of the epochs (from each class) randomly selected from the training sets in each of the N partitions of the data,

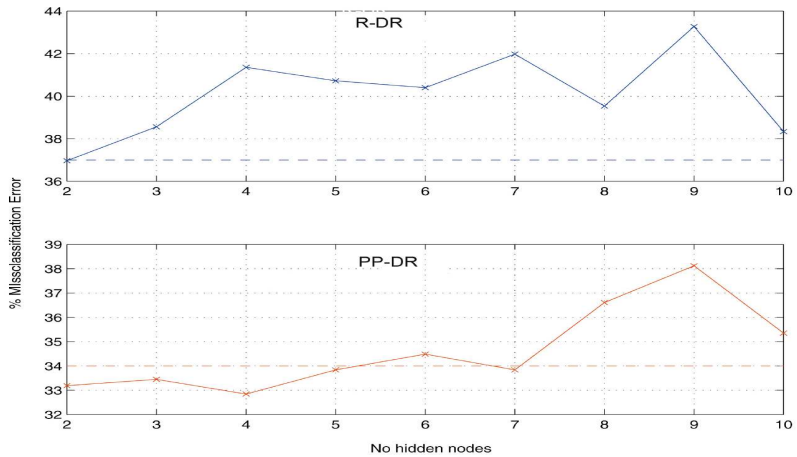

Figure 7.7: Misclassification error (%) for 5-J-1 MLP as J is varied from 2 to 9. The dashed line in each case shows the result for a single-layer neural network. The results in the top plot are obtained using R-DR reflection coefficients, the lower plot uses PP-DR reflection coefficients.

described in Table 7.1. A network is trained for each of the $N = 10$ data partitions, and tested on the validation set. For each data partition 10 networks are trained with different weight initialisations. The error of the best performing network is taken as the error value for that data partition.

Figure 7.7 shows the effect on the percentage misclassification error of changing the number of hidden nodes from 2 to 9. The misclassification value is calculated as the average of the misclassification errors of the validation sets from each of the (best of 10) N networks. Changing the number of hidden nodes has little effect on classification accuracy. The MSE and misclassification rate should fall as the number of hidden nodes is increased beyond 2 if the optimal decision plane is non-linear. A single-layer neural network is therefore trained and the corresponding misclassification errors calculated. The errors are plotted on Figure 7.7 and it can be seen that a single-layer neural network gives similar (or better) accuracy than the MLPs, implying that for the chosen feature sets the optimal decision boundaries are linear.

Table 7.2: Classification accuracy of single-layer network using the features from only a single derived respiratory waveform. The accuracy is averaged over all $N = 10$ test subjects.

Waveforms used	No. of inputs	% Accuracy
R-DR	5	63
RSA-DR	5	58
PP-DR	5	66

7.7 Results of single-layer network classification, using features from one waveform

In view of the results of Section 7.6.4, the rest of this chapter concentrates on the use of a linear discriminant function, in the form of a single-layer neural network, for classification of the reflection coefficient feature set.

Table 7.2 shows the classification accuracy for each of the single source networks. It can be seen that using the reflection coefficients of the PP-DR as features gives the best performance resulting in an accuracy of 66%.

7.8 Fusion of information from more than one derived respiratory waveforms

7.8.1 Feature level fusion

Feature-level fusion is performed by concatenating the feature vectors from combinations of the different derived waveforms. The concatenated vector is used as the input to a single-layer neural network which is trained as in Section 7.6.4. There are three single sources (R-DR, PP-DR, RSA-DR) and therefore three different combinations of two sources, and a single combination of all three. Table 7.3 shows the results in terms of the classification accuracy. In all cases there is at least a marginal improvement with respect to using a single set of features. The biggest improvement is seen when using a combination of features from all three derived respiratory waveforms. However, this is negligible in comparison to fusing only the RSA-DR and PP-DR reflection coefficients, in this case the accuracy increases to 72%, an

Table 7.3: Results of a feature-level fusion using a single-layer neural network. In the first three examples features of two derived respiratory waveforms are combined, whilst in the last example features from all three are combined. It can be seen that combining the features from the RSA-DR and PP-DR gives the best accuracy, with only a slight improvement when using features from all three.

Waveforms used	No. of inputs	% Accuracy
R+PP	10	68
RSA+PP	10	72
RSA+RS	10	65
RSA+R+PP	15	73

improvement of 6% with respect to the best performing single source classifier (which uses the PP-DR feature set).

7.8.2 Decision level fusion

To compare the performance of decision level fusion with that of feature level fusion for this problem, a decision level fusion scheme is proposed. The outputs of individual classifiers which use features from different derived respiratory waveforms are combined. The outputs are in the form of estimates of *a posteriori* probabilities for each class, leading to soft decision level fusion. As discussed in Chapter 3, there are a number of ways of combining soft decisions. The empirical studies by Kittler [57] and Verlinde [111], discussed in Chapter 3, show that non-statistical techniques such as logical operators or sum combinations often result in similar or better results than statistical techniques.

In view of these results, a simple non-statistical fusion technique is applied to the decision outputs of the single source neural networks that give the best feature fusion results, i.e., those using PP and RSA-DR features. A sum rule is used: the outputs of the two networks are averaged and a decision made on the averaged value (a threshold of 0.5 is used initially).

This decision fusion method results in an accuracy of 70%, slightly less than that found using feature fusion (72%).

7.8.3 Comparison of results

The classification schemes proposed in this chapter all output *a posteriori* probabilities for either class. The results reported are all obtained using a decision threshold of 0.5, i.e., a positive classification is made if the probability is greater than 0.5.

Figure 7.8 shows the ROC curve (left) and corresponding accuracy curves with respect to threshold (right). The curves for the performance of the best classifier that uses features from only one derived respiratory waveform (PP-DR) are shown in red. The values of the decision thresholds are varied from 0.1 to 0.9. It is seen that the highest accuracy always occurs at a threshold of 0.5. However the relative values of sensitivity and specificity may be adjusted, depending on the importance of detecting all events versus the number of false positives. The points on the three curves that correspond to a threshold of 0.5 are marked with a square. It can be seen that for the feature fusion case, the 0.5 threshold corresponds to a sensitivity of 71% and a specificity of 72% so the two values are balanced. If a higher sensitivity is required a threshold of 0.4 results in a sensitivity of 77% with a specificity of only 61%, in which case more false positives occur.

Both curves show that the feature fusion classification method outperforms the best classifier that uses features from only a single derived respiratory waveform. The ROC curve for the feature fusion lies nearer the (0,1) point (the point at which 100% sensitivity and specificity are achieved) of the ROC plot at all times, apart from at thresholds of 0.8 or above. This characteristic is also seen in the accuracy curves.

7.9 Validation of method on Computers in Cardiology apnoea database

As there are currently no published results on the use of the ECG and blood pressure signal to classify the 30-second epochs of the Polysomnography database, the classification method proposed in this chapter is also applied to the Computers in Cardiology database [50]. As discussed in Section 7.4 there are a number of published results with which to compare the accuracy of the proposed classification method on this database.

There are several differences between the Polysomnography database and the Computers in

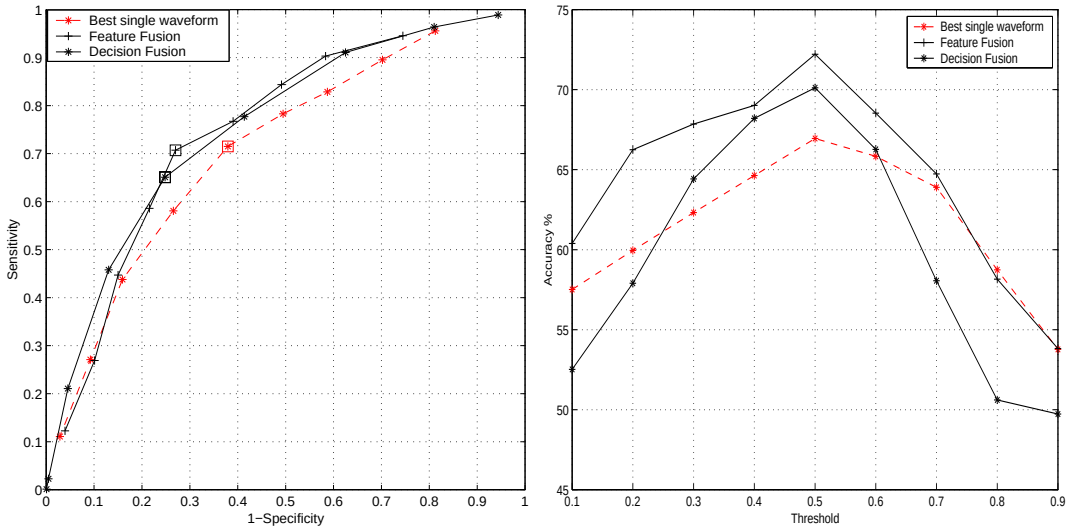

Figure 7.8: The ROC curves (left) and corresponding accuracy curves (right) for the classifiers using feature and decision fusion (of the PP-DR and RSA-DR). The best performing classifier uses only features from a single respiratory waveform (the PP-DR). It can be seen that the feature fusion classifier outperforms the single waveform classifier at all thresholds except high thresholds (greater than 0.8) where it gives a similar performance.

Cardiology apnoea database:

- the Computers in Cardiology database comprises only of a single lead ECG, therefore information is available from only one signal.
- the labels in the Computers in Cardiology database correspond to 60-second epochs, rather than 30-second epochs in the Polysomnography database.
- the Computers in Cardiology database contains a total of 34208 epochs while the corresponding value for the Polysomnography database is only 2348.
- the sampling rate of the Computers in Cardiology database is $100Hz$ in comparison to a rate of $250Hz$ in the Polysomnography database.

7.9.1 Methodology

The Computers in Cardiology database consists of 70 night-time ECG recordings each of a duration of between 7 and 10 hours. The data is partitioned into training and test data sets of 35 recordings each, exactly as in the published work that uses this database.

As the epochs being classified are twice the length of those in the Polysomnography database,

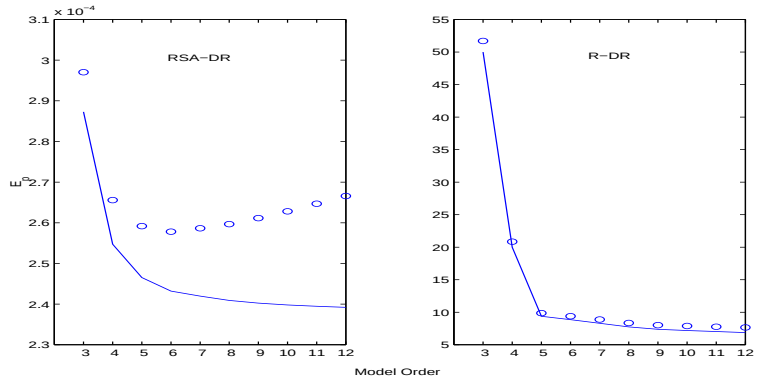

Figure 7.9: Prediction error E_p and FPE curves with respect to model order for the RSA-DR (left) and R-DR (right).The continuous line corresponds to the E_p curve and the circles to the FPE curve.

the AR model order is reconsidered. Figure 7.9 shows the prediction error curve (E_p) curve and the corresponding FPE curve with respect to model order for the two ECG derived respiratory waveforms. The curves are calculated by averaging the model prediction error over all 60-second epochs in the training set. Both prediction error curves show a "knee" at a model order of 5. The FPE curve of the RSA-DR, shows a minimum at 6, while the corresponding curve for the R-DR does not show a minimum. A number of studies indicate that the FPE criterion tends to under-estimate the optimal model order [85], therefore a model order of 7 is chosen for the AR model order for both waveforms.

Two individual single-layer neural network are trained using the reflection coefficients (of a $7th$ order AR model) for 60-second epochs of the R-DR and RSA-DR waveforms. To combine the information from the two respiratory waveforms, a single-layer neural network is also trained that uses a feature vector of 14 elements; this feature vector is formed by concatenating the two $7th$ order vectors corresponding to the R-DR and RSA-DR.

7.9.2 Results of classification of Computers in Cardiology database

Table 7.4 shows the performance accuracy of the single-layer neural networks, using reflection coefficients as features, for classification of the Computers in Cardiology database. The networks that use features from a single derived respiratory waveform give accuracies of 73% (R-DR) and 70% (RSA-DR). Combining information at feature level results in an accuracy of 78%.

Table 7.4: Results of single-layer network using reflection coefficients from two ECG derived respiratory waveforms. The misclassification error is over the whole test set of 17268 epochs. It can be seen that combining information from the two waveforms at a feature level results in an accuracy of 78%. This is an improvement of 5% from using the features from the best single respiratory waveform.

Waveforms used	No. of inputs	% Accuracy
R-DR	7	73
RSA-DR	7	70
RSA+R-DR	14	78

7.10 Discussion and comparison of results

Discussion of results obtained for classification of epochs of the Polysomnography database

It is seen that when using features from only a single derived respiratory waveform, the PP-DR gives the best results (an accuracy of 66%). An improvement in accuracy of 7% can be obtained by fusing the features of all three derived respiratory waveforms (R-DR, RSA-DR and PP-DR). However when fusing just the features from the RSA-DR and PP-DR waveforms an improvement of 6% is still seen. Decision fusion, using the outputs from two single-layer neural networks that use features from the RSA-DR and PP-DR waveforms, results in an accuracy of 70%, not quite as an improvement as that seen when using fusion at a feature level.

Comparison of results of the proposed classification system on the two databases

Only classification methods that use features from the ECG can be compared (as the Computers in Cardiology database only comprises a single channel of ECG). It can be seen from Tables 7.2, 7.3 and 7.4 the classifiers that use features from only one ECG derived respiratory waveform (either R-DR or RSA-DR) give accuracies that are at least 10% better on the Computers in Cardiology data set. The classifier that uses features from both of the ECG-derived respiratory waveforms gives an accuracy of 78% on the Computers in Cardiology database compared with 65% on the Polysomnography database. The reasons for this are thought to be two-fold:

- the Computers in Cardiology training set contains approximately 7 times more training

data than the Polysomngraphy training set.

- the Computers in Cardiology database classifies 60-second epochs (in comparison with 30-second epochs in the Polysomnography database).

It is well known that the more training data available to train a statistical learning classifier, the better the likely performance of the system. Secondly the 60-second epochs are more likely to contain the whole apnoeic event, whereas in 30-second labelling the apnoeic event may span across consecutive epochs.

Comparison of results of classification of Computers in Cardiology database with those in the literature

The performance of the proposed classification method and the performance of methods described in the literature can be compared directly using the Computers in Cardiology database. It is seen that a single-layer network classifier using reflection coefficients of two ECG-derived respiratory waveforms as input features gives an accuracy (78%), comparable to those methods that give the best performance in the Computers in Cardiology competition (the referenced works give a range of accuracies from 63% to 85%).

7.11 Conclusion

It can be first concluded that the classification system proposed in this chapter gives comparable performance to those reported in the literature. This classification system uses a reflection coefficient parameterisation of derived respiratory waveforms as a set of features for input to a linear classifier (a single-layer neural network). The method results in an accuracy of 78%, while classification accuracies of up to 85% are reported ([69],[93]) for methods which use a combination of features from the time and frequency-domains. It is suggested that the addition of time-domain features to the reflection coefficients features could improve performance. A formal feature selection process [2], could be carried out to optimise a final set of features.

It may also be concluded that fusing information from the ECG and blood pressure signals

[2]Numerous feature selection processes exist, a review may be found in Chapter 8 of Webb [112]

leads to improved classification performance. Combining the information at feature level gives slightly better results than combining at a soft decision level. This can be expected as feature level fusion leads to more information being used to make the classification decision.

Chapter 8

Conclusions

The aim of the work described in this thesis is to investigate the feasibility of performing reliable respiratory monitoring from non-invasive biomedical signals that do not measure airflow directly. There are two main objectives, the first being to estimate the breathing rates from the signals, the second being to detect the occurrence of apnoeas. Section 8.1 gives an overview of the research and draws conclusions from the results. Section 8.2 suggests some areas for further work.

8.1 Overview and conclusions

Methods for obtaining an accurate measure of breathing rate from non-invasive signals are described in Chapters 4 and 5. Various methods for deriving respiratory waveforms from the ECG, blood pressure, PPG and IP signals are presented in Chapter 4. An objective evaluation procedure for comparing and evaluating these methods is proposed and implemented.

One of the main aims is to establish whether a method using one of these signals alone can provide reliable breath-by-breath information. Results from evaluations on two separate databases show that no method, based on the analysis of a single non-invasive signal, consistently out-performs the others in terms of sensitivity and positive predictivity. In light of these results a new method is developed for combining breathing rate measurements using information from the three non-invasive signals. This fusion method is based on weighting the estimates from each derived respiratory waveform according to the innovation from a

Kalman filter model applied to each waveform separately.

The fused estimates give a higher overall correlation with respect to the reference breathing rate values than any of the breathing estimates derived from a single waveform. The best performing single signal, the IP signal gives a correlation coefficient of 0.64, while a fused estimate using three of the signals gives a correlation of 0.80.

The fusion method is assessed over a range of breathing rates. The mean percentage error and the corresponding 95% confidence interval are calculated for 10 breathing rate ranges. The method achieves a percentage error of 16% or less at all breathing rates above 9 breaths per minute.

It may be concluded that:

- none of the methods proposed for deriving breath-by-breath respiration are considered sufficiently accurate to use alone.
- a fusion algorithm that fuses breathing rate estimates from the ECG, PPG and IP according to the innovations from Kalman filter models gives a significant improvement in the accuracy of the estimated breathing rates.
- the best fused estimates are obtained when the subject is breathing at a rate above 9 breaths per minute. None of the proposed methods for detecting breath-by-breath respiration perform well at lower rates; this is reflected in the fused estimates.

The problem of detecting apnoeas from the signals is addressed in Chapters 6 and 7. In Chapter 6 methods for detecting central apnoeas are investigated. Current monitors, that analyse non-invasive respiratory signals such as the IP, tend to have a very high false alarm rate, for example Weese-Mayer *et al.* [114] report a positive predictivity of less than 8%. When analysing the IP signal for apnoea detection, two approaches are taken. The first is a time-domain, threshold-based technique based upon published work on central apnoea detection. The second is a frequency-domain approach that uses an AR model estimation of the windowed IP signal. The magnitude of poles in a physiologically plausible breathing frequency range are used to indicate the presence or absence of breathing. Both approaches result in very high sensitivities (greater than 98%), with the frequency-domain approach giving a higher positive predictivity (55%), and hence a lower false alarm rate, than the time domain approach (positive predictivity of 43%). Fusing information from the time and

frequency-domains at the decision level results in a further 10% improvement in positive predictivity without compromising the sensitivity. When evaluating the two methods using respiratory waveforms derived from the ECG and PPG, very low positive predictivites are seen (typically less than 15%). Therefore these signals are not considered for use when detecting central apnoeas.

Chapter 7 describes investigated methods for classifying obstructive sleep apnoea from the ECG and blood pressure signals. A classification system is proposed and tested on two databases, the first being the Polysomnography database [113]. The second, referred to as the Computers in Cardiology database [50], has been used extensively for investigating the detection of obstructive sleep apnoea from the ECG; using the latter database allows a direct comparison with published results.

It is clear that obstructive sleep apnoea provokes changes in the frequency spectra of derived respiratory waveforms. Therefore a feature set is required that gives a compact representation of short sections of the frequency spectra of these waveforms. A feature set comprising reflection coefficients is investigated. These features are used as the input to a single-layer neural network, after concluding that the optimal decision boundary in the feature space is linear.

The classification system is tested on the Computers in Cardiology database using information from the ECG alone. This classifier gives an accuracy of 78%, comparable to the figures reported in the literature, which range from 63% to 85%. It is further shown that the classification accuracy can be improved by combining information from the blood pressure signal and the ECG. The accuracy of classifiers that use information from either only one signal or the two signals combined are compared. Feature level fusion, prior to classification is found to give the best results, with a 10% increase in classification accuracy from the best performance achieved using only a single signal.

It may be concluded that:

- using the proposed methods for detecting central apnoea from the ECG and PPG results in very low positive predictivity. The signals are therefore not considered for use in the detection of central apnoeas.
- when identifying central apnoea events from the IP signal, the detection accuracy can

be improved if the output of algorithms applied to time and frequency-domain features are fused.

- when detecting obstructive sleep apnoeas, a reflection coefficient parameterisation of short sections of derived respiratory waveforms produces a feature set that provides discriminatory information. The optimal decision boundary in the feature space is shown to be linear.
- fusing information, in the form of a reflection coefficient parameterisation of ECG and blood pressure derived respiratory waveforms, at a feature level improves the accuracy of classifying obstructive sleep apnoea.

8.2 Further work

This thesis investigates the feasibility of performing respiratory monitoring from a set of non-invasive biomedical signals. Although it is shown that data fusion of the signals improves the accuracy of respiratory information with respect to that obtained using any signal alone, possible techniques for deriving and fusing this information have not been exhausted.

Two papers published at the time of writing this thesis propose alternative methods for deriving breath-by-breath respiration. Respiratory signals are used as the direct input to neural networks. The signals are different to those used in this work, one using the signal obtained from an invasive nasal airflow thermistor, the second using inductance plethysmography. However the analysis methods could be applied to the non-invasive signals used in this work.

- Varady *et al.* [109] use windows (16 seconds in length) of both the IP and nasal airflow respiration signals as the inputs to four separate neural networks. The neural networks are trained to classify the time windows into the classes of hypopneas, apnoeas, normal or unknown. Each network has slightly different inputs. The best performing network, not surprisingly, uses the nasal airflow signal only. Sensitivities and positive predictivities of greater than 90% are achieved for both normal breathing and apnoea detection. This work was carried out on a subset of the Polysomnography database [113].

- Sa *et al.* [90] use a neural network technique to recognise 2-second windows of an inductance plethysmography respiratory signal that contain the onset of inspiration or expiration. Post-processing of the signal within the detected window is subsequently carried out to search for minimum and maximum values corresponding to the exact start of expiration or inspiration. The respiratory signal used is the sum of the signals obtained from two inductance plethysmography bands (one placed around the ribcage, the other around the abdomen). The neural networks are tested on a large data set ($\sim$ 30,000 breaths), resulting in very high values of sensitivity and positive predictivity ($\geq$ 97%) for the detection of the windows. A further study is carried out on a small subset of the data (162 breaths). The study compares the neural network window detection followed by the post-processing search algorithm with two more conventional algorithms designed to detect the onset of inspiration or expiration from the respiratory signal. The neural network algorithm performs well on this small dataset, giving values of 96.7% and 93.8% for sensitivity and positive predictivity when detecting expiration. It is also noted that on this small dataset one of the conventional algorithms (which involved a search for maxima and minima in the signal, similar to that used in this work) also performed well with a sensitivity of 94.1% and positive predictivity of 94.6%. This indicates the quality of the respiratory information in the inductance plethysmography signals.

An AR model is used to estimate frequency-domain information throughout this thesis. AR model estimation assumes that the signal being parameterised is stationary, or at least that it can be segmented into windows within which it can be assumed quasi-stationary. The use of the wavelet transform may be an alternative for extracting respiratory information from the signals. The basic idea underlying wavelet analysis is that a signal may be expressed as a linear combination of a particular set of basis functions, obtained by shifting and dilating a single function sometimes called a mother wavelet [20]. The wider or more dilated the function the poorer the time location and the lower the frequency component being represented. High frequency components are represented by narrower mother wavelets with a sharper time resolution. Wavelets have advantages for signal processing when signals contain discontinuities or sharp spikes. One of the aims of this work is to detect apnoeas. Apnoeas, both central and obstructive introduce abrupt changes in the signals.

Prior to any further research work it would be useful to extend the scope of the data available;

in size to ensure statistically meaningful results and to also be representative of any targeted subject groups. Notable additions would include:

- the inclusion of a PPG signal in all databases (in this thesis some of the work, notably that in Chapter 7, is carried out using the available invasive blood pressure signal for proof of concept).
- data from a larger cross-section of subjects, the data used in this work is all collected from adult male subjects. A particular subject group of interest are neonates who are known to suffer respiratory difficulties.
- including more data at lower breathing rates, as it is shown that the methods presented in this work give less accurate results at lower breathing rates.

When collecting the data it is recommended that a human observer is present. The labelling of the data may then be carried out with knowledge of breathing and subject conditions during the data collection period. The data in the Polysomnography database is labelled by a retrospective visual examination of the data, with no such prior knowledge. This leads to uncertainties in the labelling. It is also noted that the positioning of sensors on the body, in particular ECG electrodes has an effect on the quality of the respiratory information.

Appendix A

The Controlled-Breathing database data collection protocol

Table A.1: Protocol of data collection for Controlled-Breathing database

Duration (minutes)	Description
15	Normal relaxed breathing
0.5	Isometric handgrip
1.5	Relaxation
0.5	Isometric handgrip
1.5	Relaxation
5	6rpm, 300ml volume, no resistance
≤ 0.5	Breath hold
1	Rest
5	6rpm, 500ml volume, no resistance
≤ 0.5	Breath hold
1	Rest
5	6 rpm, 1000ml volume, no resistance

continued from previous page

Duration (minutes)	Description
≤ 0.5	Breath hold
5	Rest
5	10rpm, 300ml volume, no resistance
≤ 0.5	Breath hold
1	Rest
5	10rpm, 500ml volume, no resistance
≤ 0.5	Breath hold
1	Rest
5	10rpm, 1000ml volume, no resistance
≤ 0.5	Breath hold
1	Rest
5	20 rpm, 300ml volume, no resistance
≤ 0.5	Breath hold
1	Rest
5	20 rpm, 500ml volume, no resistance
≤ 0.5	Breath hold
1	Rest
5	20 rpm, 1000ml volume, no resistance
≤ 0.5	Breath hold
1	Rest
1	Breath 60 rpm
2	Rest- 'normal' breathing
1	Breath 120rpm
2	Breathing normally
10	Mimic central apnoeas, 60-second cycle time, apnoeic period 15 seconds -apnoea 22.5 $\rightarrow$37.5 every min
2	Rest
2	'Normal' rate, resistance: 20mmHg H_2O
1	Rest

continued from previous page

Duration (minutes)	Description
2	'Normal' rate, resistance: 50mmHg H_2O
2	Rest
5	Cycling
5	Rest

Appendix B

MMSE combination of estimates [105]

Given n estimates of the value $\hat{x}$:

$$x_i = \hat{x} + \delta_i \tag{B.1}$$

where $i = 1...n$. And δ_i is drawn from a Gaussian distribution of mean 0, variance σ_i^2 and covariances γ_{ij}:

$$\mathcal{E}\left[\delta_i\right] = 0 \tag{B.2}$$

$$\mathcal{E}\left[\delta_i^2\right] = \sigma_i^2 \tag{B.3}$$

$$\mathcal{E}\left[\delta_i \delta_j\right] = \gamma_{ij} \tag{B.4}$$

Note that $\gamma_{ii} = \sigma_i^2$. Let x_O be an the 'optimum' estimate of $\hat{x}$ using a linear combination of the other estimates:

$$x_O = \sum_{i=1}^{n} \alpha_i x_i \tag{B.5}$$

$$\sum_{i=1}^{n} \alpha_i = 1 \tag{B.6}$$

To minimise the error in this, use equation B.1:

$$\hat{x} + \delta_O = \sum_{i=1}^{n} \alpha_i (\hat{x} + \delta_i) \tag{B.7}$$

$$\Rightarrow \delta_O = \sum_{i=1}^{n} \alpha_i \delta_i \tag{B.8}$$

We wish to minimise:

$$\sigma_O^2 = \mathcal{E}\left[\delta_O^2\right] \tag{B.9}$$

$$= \mathcal{E}\left[\sum_{j=1}^{n} \alpha_j \delta_j \sum_{i=1}^{n} \alpha_i \delta_i\right] \tag{B.10}$$

$$= \mathcal{E}\left[\sum_{j=1}^{n} \sum_{i=1}^{n} \alpha_i \alpha_j \delta_i \delta_j\right] \tag{B.11}$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} \alpha_i \alpha_j \mathcal{E}\left[\delta_i \delta_j\right] \tag{B.12}$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} \alpha_i \alpha_j \gamma_{ij} \tag{B.13}$$

From equation B.6, a Lagrange multiplier can be defined:

$$\lambda \left(\sum_{i=1}^{n} \alpha_i - 1 \right) = 0 \tag{B.14}$$

which is added to equation B.13:

$$\sigma_O^2 = \sum_{j=1}^{n}\sum_{i=1}^{n}\alpha_i\alpha_j\gamma_{ij} + \lambda\left(\sum_{i=1}^{n}\alpha_i - 1\right) \tag{B.15}$$

To optimise the values of each α this equation is then differentiated with respect to each α_p (where p is between 1 and n inclusive) and the differential is set to zero:

$$\frac{\partial\sigma_O^2}{\partial\alpha_p} = 2\sum_{j=1}^{n}\alpha_j\gamma_{jp} + \lambda = 0 \tag{B.16}$$

$$\Rightarrow \lambda = -2\sum_{j=1}^{n}\alpha_j\gamma_{jp} \tag{B.17}$$

From this it follows that λ can be determined by differentiating with respect to any α_p and that differentiating against, say, α_p and α_i will give the same result for λ at the optimum point. Therefore:

$$\sum_{j=1}^{n}\alpha_j\gamma_{jp} = \sum_{k=1}^{n}\alpha_k\gamma_{ki} \tag{B.18}$$

If it is assumed that the sources are independant, *ie* $\gamma_{ij,i\neq j} = 0$, then, recalling equation B.4, this equation reduces to:

$$\alpha_p\sigma_p^2 = \alpha_i\sigma_i^2 \tag{B.19}$$

$$\Rightarrow \alpha_i = \frac{\sigma_p^2}{\sigma_i^2}\alpha_p \tag{B.20}$$

Combining this equation with equation B.6 gives:

$$\sum_{i=1}^{n} \frac{\sigma_p^2}{\sigma_i^2} \alpha_p = 1 \tag{B.21}$$

$$\alpha_p \sigma_p^2 \sum_{i=1}^{n} \frac{1}{\sigma_i^2} = 1 \tag{B.22}$$

$$\Rightarrow \alpha_p = \frac{1}{\sigma_p^2 \sum_{i=1}^{n} \frac{1}{\sigma_i^2}} \tag{B.23}$$

$$= \frac{1}{\sigma_p^2 \frac{\sum_{i=1}^{n} \prod_{j=1, j \neq i}^{n} \sigma_j^2}{\prod_{i=1}^{n} \sigma_i^2}} \tag{B.24}$$

$$= \frac{\prod_{i=1}^{n} \sigma_i^2}{\sigma_p^2 \sum_{i=1}^{n} \prod_{j=1, j \neq i}^{n} \sigma_j^2} \tag{B.25}$$

Appendix C

Glossary and Abbreviations

Apnoea — A temporary cessation in breathing.

Arousal — A change in sleep state from the current state to a lighter state or wakefulness.

Atrial Fibrillation — The extremely rapid and disorganised pattern of the depolarisation of the atria, leading to a rapid and disorganised heart rate.

Bradycardia — An abnormal heart rate of 60 or fewer bpm.

Finapres — Finapres is the acronym for FINger Arterial PRESsure, the quantity that is measured continuously by the instrument. This is a non-invasive measure of blood pressure.

Heart-rate — The number of beats per unit time; usually measured as beats per minute.

Heart rate variability (HRV) — The normal variability of the heart rate.

Hypopnea — Abnormally slow and shallow breathing.

Hypoxemia — Deficiency of oxygen in the blood.

Pulsus Paradoxus — Inspiratory fall in systolic arterial blood pressure.

QRS wave — Also the QRS complex; the signal on the electrocardiogram that repsresents the ventricular depolarisation and atrial repolarisation of the heart.

Respiratory Sinus Arrhythmia (RSA) — RSA is the variation in heart rate that is seen during inspiration and expiration. Heart rate tends to increase during inspiration and decrease during expiration.

Tachycardia — A rapid heart rate; often used to classify heart rates of above 100 bpm.

EDR — ECG-derived respiratory waveform.

IP — impedance plethysmography.

PP-DR — respiratory waveform derived from pulsus paradoxus measurements of the blood pressure.

R-DR — respiratory waveform derived from R amplitude measurements of the ECG.

RSA-DR — respiratory waveform derived from RSA (R-R interval or heart rate) measurements of the ECG.

Bibliography

[1] ANSI/AAMI EC38:1998 Ambulatory Electrocardiographs.

[2] ANSI/AAMI EC57:1998 Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms.

[3] *The MIT-BIH Polysomnography Database CD-Rom*. Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, USA, August 1992.

[4] *Lecture Notes on Human Physiology*. Blackwell Science Ltd, 1994.

[5] *Oxford English Reference Dictionary*. Oxford University Press, 1995.

[6] M Abidi and R Gonzalez, editors. *Data Fusion in Robotics and Machine Intelligence*. Academic Press Inc., 1992.

[7] C W Anderson, E A Stolz, and S Shamsunder. Multivariate Autoregressive Models for Classification of Spontaneous Electroencephalographic Signals During Mental Tasks. *IEEE Transactions on Biomedical Engineering*, 45(3), March 1998.

[8] S T Anderson, W G Downs &A P Lander, D M Mirvis, C Rizo-Patron, R L Burr, C Jacobson, G A Massumi, and J C Perry. *Advanced Electrocardiography*. SpaceLabs Medical Inc., 1995.

[9] P Berghuis, N Cohen, M Decker, A Gettinger, K Myrabo, J Nilsestuen, K Strohi, and J Yount. *Respiration*. SpaceLabs Medical Inc., 1994.

[10] C. M. Bishop. *Neural networks for pattern recognition*. Oxford University Press, 1995.

[11] M Bland. *An introduction to medical statistics*. Oxford University Press, 2000.

[12] Y Brans and W Hay, editors. *Physiological monitoring and instrument diagnosis in perinatal and neonatal medicine*. Cambridge University Press, 1995.

[13] P M T Broersen. Facts and Fiction in Spectral Analysis. *IEEE Transactions on Instrumentation and Measurement*, 49(4):766–772, August 2000.

[14] P M T Broersen. Automatic Spectral Analysis with Time Series Models. *IEEE Transactions on Instrumentation and Measurement*, 51(2):211–216, April 2002.

[15] R T Brouillette, A S Morrow, D E Weese-Mayer, and C E Hunt. Comparison of respiratory inductive plethysmography and thoracic impedance for apnea monitoring. *Journal of Pediatrics*, 111:377–83, 1987.

[16] R G Brown. *Introduction to Random Signal Analysis and Kalman Filtering*. Wiley, 1983.

[17] E G Caiani, A Porta, M Terrani, S Guzzetti, A Malliani, and S Cerutti. Minimal Adaptive Notch Filter for Respiratory Frequency Tracking. In *Computers in Cardiology Proceedings*, volume 25, pages 511–514, 1999.

[18] F Caswell. *Success in Statistics*. John Murray Ltd, 1995.

[19] S Cerutti, L T Mainardi, A Porta, and A M Bianchi. Analysis of the Dynamics of RR Interval Series for the Detection of Atrial Fibrillation Episodes. In *Computers in Cardiology Proceedings*, pages 77–80, 1997.

[20] S Conforto, T D'Alessio, and S Pignatelli. Optimal rejection of movement artefacts from myoelectric signals by means of a wavelet filtering procedure. *Journal of Electromyography and Kinesiology*, pages 47–57, 1999.

[21] I Constant, D Laude, I Murat, and J L A D Elghozi. Pulse rate variability is not a surrogate for heart rate variability. *Clinical Science*, 97(4):391–7, October 1999.

[22] F Cremer, K Schutte, J G M Schavemaker, and E den Breejen. A comparison of decision-level sensor-fusion methods for anti-personnel landmine detection. *Information Fusion*, 2:187–208, 2001.

[23] E. Curran, P. Sykacek, M. Stokes, S. Roberts, W. Penny, I. Johnsrude, and A. Owen. Cognitive Tasks for driving a Brain Computer Interfacing System: a pilot study. *Technical Report PARG-01-06, March 2001. Submitted to IEEE Transactions on Rehabilitation Engineering*.

[24] R J O Davies. Cardiovascular aspects of obstructive sleep apnoea and their relevance to the assessment of the efficacy of nasal continuous poitive airway pressure therapy. *Thorax*, 53:416–418, 1998.

[25] R J O Davies, K Vardi-Visy, M Clarke, and J R Stradling. Identification of sleep disruption and sleep disordered breathing from the systolic blood pressure profile. *Thorax*, 48:1242–1247, 1993.

[26] R W DeBoer, J M Karemaker, and J Strackee. Comparing Spectra of a series of point events particularly for heart rate variability data. *IEEE Trans. Biomed. Eng*, 31:384–387, 1984.

[27] M J Drinnan, J Allen, P Langley, and A Murray. Detection of Sleep Apnoea from Frequency Analysis of Heart Rate Variability. In *Computers in Cardiology Proceedings*, pages 259–262, 2000.

[28] B Duc, E Bigun, J Bigun, G Maitre, and S Fischer. Fusion of audio and video information for multi-modal person authentication. *Pattern Recognition Letters*, 18:835–843, 1997.

[29] R Duda and P Hart. Pattern Classification and Scene Analysis, 1973.

[30] S R Dumpala, S N Reddy, and S K Sarna. An algorithm for the detection of peaks in biological signals. *Computer Programs in Biomedicine*, 14(3):249–256, June 1982.

[31] M H Ebrahim, J M Feldman, and I Bar-Kana. A Robust Sensor Fusion Method for Heart Rate Estimation. *The Journal of Clinical Monitoring*, 13(6):385–393, November 1997.

[32] W Einthoven, G Fahr, and A de Waart. On the direction and manifest size of the variations of the potential in the human heart on the from of the electrocardiogram. *Pfluger's Arch. f. d. ges. Physiol.*, 50:275–315,1913., English Translation by H. E. Hoff and P. Sekelj, Amer. Heart J., vol. 40 pp.163-193 1950.

[33] W A H Engelese and C Zeelenberg. A single scan algorithm for QRS-detection and feature extraction. *Computers in Cardiology*, 6:37–42, 1979.

[34] J Felblinger and C Boesch. Amplitude demodulation of the electrocardiogram signal (ECG) for respiration monitoring and compensation during MR examinations. *Magn-Reson-Med*, 38(1):129–36, July 1997.

[35] J M Feldman, M H Ebrahim, and I Bar-Kana. Robust Sensor Fusion Improves Heart Rate Estimation: Clinical Evaluation. *The Journal of Clinical Monitoring*, 13(6):385–393, November 1997.

[36] B Frey and W Butt. Pulse Oximetry for assessment of pulsus paradoxus: a clinical study in children. *Intensive Care Med*, 24:242–246, 1998.

[37] Q Gan and C J Harris. Comparison of Two Measurement Fusion Methods for Kalman-Filter-Based Multisensor Data Fusion. *IEEE Transactions on Aerospace and Electronic Systems*, 37(1):273–280, January 2001.

[38] A Gelb. *Applied Optimal Estimation*. The MIT Press, 1996.

[39] A L Goldberger, L A N Amaral, L Glass, J M Hausdorff, P Ch Ivanov, R G Mark RG, J E Mietus, G B Moody, C K Peng, and H E Stanley. PhysioBank, PhysioToolkit, and Physionet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 101(23), June 2000.

[40] M S Grewal, L R Weill, and A P Andrews. *Global Positioning Systems, Inertial Navigation and Integration*. Wiley, 2001.

[41] P F Griner, R J Mayewski, A Mushlin, and P Greenland. Selection and interpretation of diagnostic tests and procedures. Principles and applications. *Annals of Internal Medicine*, 94(4):557–600, 1981.

[42] C Guilleminault, S J Connolly, and R A Winkle. Cyclical variation of the heart rate in sleep apnoea syndrome. *Lancet*, 1:126–131, 1984.

[43] A H Gunatilaka and B A Baertlein. Feature-Level and Decision-Level Fusion of Noncoincidently Sampled Sensors for Land Mine Detection. *IEEE Transactions on Pattern Analysis and Machine Intellegence*, 23(6):577–589, June 2001.

[44] D E Gustafson, A S Willsky, J Wang, M C Lancaster, and J H Triebwasser. ECG/VCG Rhythm Diagnosis Using Statistical Signal Analysis - I. Identification of Persistent Rhythms. *IEEE Transactions on Biomedical Engineering*, 25(4), July 1978.

[45] T V Hartert, A P Wheeler, and J R Sheller. Use of Pulse Oximetry to Recognize Severity of Airflow Obstruction in Obstructive Airway Disease. *Chest*, 115:475–481, 1999.

[46] M J Hayes, P R Smith, and D M Barnett. Quantitative Investigation of Artefact in Photoplethysmography and Pulse Oximetry for Respiratory Exercise Testing. In *Proceedings of the 7th International Symposium on Computer-aided Noninvasive Vascular Diagnostics*, 1997.

[47] M Hellebrandt and R Mathar. Location Tracking of Mobiles in Cellular Radio Networks. *IEEE Transaction on Vehicular Technology*, 48(5), September 1999.

[48] R Hermes and G Oliver. Use of the American Heart Association database. *Amulatory Electrocardiographic Recording*, pages 165–181, 1980.

[49] A Houghton and D Gray. *Making Sense of the ECG*. Oxford University Press, 1997.

[50] Computers in Cardiology apnoea competition website. http://www.physionet.org/cinc-challenge-2000.shtml.

[51] G D Jay, K Onuma, R Davis, M Chen, A Mansell, and D Steele. Analysis of Physician Ability in the Measurement of Pulsus Paradoxus by Sphygmomanometry. *Chest*, 118:348–352, 2000.

[52] L O Jimenez, A Morales-Morell, and A Creus. Classification of Hyperdimensional Data Based on Feature and Decision Fusion Approaches Using Projection Pursuit, Majority

Voting and Neural Networks. *IEEE Transactions of Geoscience and Remote Sensing*, 37(3):1360–1366, May 1999.

[53] A Johansson and P A Oberg. Estimation of respiratory volumes from the photoplethysmographic signal. Part 1:experimental results. *Medical Biological Engineering Computing*, 37:42–47, 1999.

[54] S J Johnsen and N Andersen. On Power Estimation in Maximum Entropy Spectral Analysis. *Geophysics*, 43(4):681–690, June 1978.

[55] R E Kalman. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME— Journal of Basic Engineering*, pages 35–45, March 1960.

[56] S M Kay and S L Marple. Spectrum Analysis — A modern perspective. In *Proceedings of the IEEE*, volume 69, pages 1380–1419, 1981.

[57] J Kittler, M Hatef, R P W Duin, and J Matas. On Combining Classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3), March 1998.

[58] L G Lindberg, H Ugnell, and P A Oberg. Monitoring of respiratory and heart rates using a fibre-optic sensor. *Medical Biological Engineering Computing*, 30:533–537, 1992.

[59] I L MacDonald. *Hidden Markov and Other Models for Discrete-Valued Time*. CRC Press, January 1997.

[60] M Malik and A J Camm, editors. *Heart Rate Variability*. Futura Publishing Company, 1995.

[61] G C W Man and B V Kang. Validation of a Portable Sleep Apnea Monitoring Device. *Chest*, 108:388–93, 1995.

[62] R Mark, P Schluter, G Moody, P Devlin, and D Chernoff. An annotated database for evaluating arrythmia detectors. *Frontiers of Engineering in Health Care, Proc 4th Annual Conf, IEEE Engineering in Medicine and Biology Society*, pages 205–210, 1982.

[63] C L Mason and L Tarassenko. Quantitative Assessment of Respiratory Derivation Algorithms. In *Proceedings of the 23rd Annual International Conference of the IEEE EMBS*, volume 2, pages 1998–2001, October 2001.

[64] P S Maybeck. *Stochastic Models, Estimation and Control, Volume 1*. Academic Press Inc, 1979.

[65] N McGrogan. Neural Network Detection of Epileptic Seizures in the Electroencephalogram. *PhD thesis, University of Oxford*.

[66] J N McNames and A M Fraser. Obstructive Sleep Apnea Classification Based on Spectrogram Patterns in the Electrocardiogram. In *Computers in Cardiology Proceedings*, 2000.

[67] W T McNicholas. Sleep apnoea and driving risk, European respiratory society task force on public health and medicolegal implications of sleep apnoea [editorial]. *European Journal of Respiration*, 13(6):1225–7, June 1999.

[68] R K Mehra. On the identification of variances and adaptive Kalman filtering. *IEEE Transactions on Automatic Control*, AC-15:175–184, 1970.

[69] J E Mietus, C K Peng, P C Ivanov, and A L Goldberger. Detection of Obstructive Sleep Apnea from Cardiac Interbeat Interval Times Se ries. In *Computers in Cardiology Proceedings*, pages 753–756, 2000.

[70] G B Moody, R G Mark, and A L Goldberger. Physionet: A Web-Based Resource for the Study of Physiologic Signals. *IEEE Engineering in Medicine and Biology Magazine*, pages 70–75, May 2001.

[71] G B Moody, R G Mark, A Zoccula, and S Mantero. Derivation of Respiratory Signals from Multi-Lead ECGs. In *Computers in Cardiology Proceedings*, volume 12, pages 113–116, 1985.

[72] I Nabney and C Bishop. http://www.ncrg.aston.ac.uk/netlab/. *Netlab neural network software*.

[73] I T Nabney. *Netlab:Algorithms for Pattern Recognition*. Springer, 2001.

[74] K Nakajima, T Tamura, and H Miike. Monitoring of heart and respiratory rates by photoplethysmography using a digital filtering technique. *Medical Engineering Physics*, 18(5):365–372, 1996.

[75] F Ng, I Garcia, P Gomis, A La Cruz, G Passariello, and F Mora. Bayesian Hierarchial Model with Wavelet Transform Coefficients of the ECG in Obstructive Apnea Screening. In *Computers in Cardiology Proceedings*, pages 275–278, 2000.

[76] A Noble. Estimation II - The Kalman Filter. *Internal Lecture Notes, Department of Engineering Science, Oxford University*, 1999.

[77] R Pallas-Areny and F Riera Canals. Recovering the Respiratory Rhythm out of the ECG. *Medical Biological Engineering Computing*, 23(Supplement, Part1):338–339, 1985.

[78] R Pallas-Areny, J Colominas-Balague, and F J Rosell. The Effect of Respiration-Induced Heart Movements on the ECG. *IEEE Transactions on Biomedical Engineering*, 36(6):585–590, June 1989.

[79] J Pardey, S Roberts, and L Tarassenko. A review of parametric modelling techniques for EEG analysis. *Medical Engineering and Physics*, 18:2–11, January 1996.

[80] J Pardey, S J Roberts, L Tarassenko, and J Stradling. A new approach to the analysis of the human sleep-wakefulness continuum. *Journal of Sleep Research*, pages 201–210, 1986.

[81] J M Parish and J W Shepard. Cardiovascular Effects of Sleep Disorders. *Chest*, 95(7), May 1990.

[82] T Penzel, A Bunde, J Heitmann, J W Kantelhardt, J H Peter, and K Voigt. Sleep Stage-Dependent Heart Rate Variability in Patients with Obstructive Sleep Apnoea. In *Computers in Cardiology Proceedings*, volume 26, 249-252 1999.

[83] D J Pitson, A Sandell, R van-den Hout, and J R Stradling. Use of pulse transit time as a measure of inspiratory effort in patients with obstructive sleep apnoea. *Eur-Respir-J*, 8(10):1669–74, October 1995.

[84] D J Pitson and J R Stradling. Autonomic Markings of arousal during sleep in patients undergoing investigation for obstructive sleep apnoea, their relationship to EEG arousals,respiratory events amd subjective sleepiness. *Journal of Sleep Research*, 7:53–59, 1998.

[85] J G Proakis and D G Manolakis. *Digital Signal Processing: principles, algorithms and applications*. Prentice Hall, 3rd edition 1996.

[86] R Ramanathan and M J Corwin. Cardiorespiratory Events Recorded on Home Monitors. *Journal of the American Medical Association*, 285(17):2199–2207, May 2001.

[87] B Raymond, R M Cayton, R A Bates, and M J Chappell. Screening for Sleep Apnoea Based on the Electrocardiogram - The Computers in Cardiology Challenge. In *Computers in Cardiology Proceedings*, pages 267–270, 2000.

[88] M D Revow, S J England, and H O'Beirne. Robust computer algorithm for detecting breaths in noisy ventilatory waveforms from infants. *Medical Biological Engineering Computing*, 24(6):609–615, November 1986.

[89] S Roberts, I Rezek, R Everson, H Stone, S Wilson, and C Alford. Automated assessment of vigilance using committees of Radial Basis Function Analysers. In *IEE Proceedings Science, Technology & Measurement*, volume 147(6), pages 333–338.

[90] R C Sa and Y Verbandt. Automated Breath Detection on Long-Duration Signals Using Feedforward Backpropagation Artificial Neural Networks. *IEEE Transactions on Biomedical Engineering*, 49(10):1130–1141, October 2002.

[91] C Sanderson and K K Paliwal. Noise compensation in a person verification system using face and multiple speech features. *Pattern Recognition*, 36:293–302, 2001.

[92] D Scholkopf and A J Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond*. MIT Press, December 2001.

[93] Z Shinar, A Bahrav, and S Askelrod. Obstructive Sleep Apnea Detection based on Electrocardiogram Analysis. In *Computers in Cardiology Proceedings*, pages 757–760, 2000.

[94] S E Sittig. Transitional Technology from NICU to Home. *American Association for Respiratiory Care Times*, September 2000.

[95] R J Stanley, P D Gader, and K C Ho. Feature and Decision level sensor fusion of electromagnetic induction and ground penetrating radar sensors for landmine detection with hand-held units. *Information fusion*, 3:215–223, 2002.

[96] A Starr, M Desforges, and J Esteban. Strategies in Data Fusion - A Condition Monitoring Approach. *ncaf forum,Manchester School of Engineering*, July 2000.

[97] D W Steele, R O Wright, C M Lee, and G D Jay. Continuous Noninvasive Determination of Pulsus Paradoxus: A Pilot Study. *Academic Emergency Medicine*, 2(10):894–900, October 1995.

[98] P K Stein and P P Domitrovich. Detecting OSAHS from Patterns seen on Heart-Rate Tachograms. In *Computers in Cardiology Proceedings*, pages 271–274, 2000.

[99] N Strobel, S Spors, and R Rabenstein. Joint Audio-Video Object Localization and Tracking. *IEEE Signal Processing Magazine*, pages 22–31, January 2001.

[100] L Tarassenko. *Lecture Notes for Medical Electronics*. Engineering Science Department, Oxford University, Michaelmas 1995.

[101] L Tarassenko. *A Guide to Neural Computing Applications*. Arnold, 1998.

[102] L Tarassenko, N Townsend, G Clifford, L Mason, J Burton, and J Price. Medical Signal Processing using the Software Monitor. In *DERA-IEE Workshop*, 2001.

[103] J F Thayer, J J Sollers, E Ruiz-Padail, and J Vila. Estimating Respiratory Frequency from Autoregressive Spectral Analysis of Heart Period. *IEEE Engineering in Medicine and Biology Magazine*, pages 41–45, July 2002.

[104] L Thorval, G Carrult, J M Schleich, R Summers, M Van de Velde, and J Diaz. Data Fusion of Electrophysiological and Haemodynamic Signals for Ventricular Rhythm Tracking. *IEEE Engineering in Medicine and Biology*, pages 48–55, November 1997.

[105] N Townsend. Internal Report. *Department of Engineering Scince, Oxford University*, 2001.

[106] N Townsend. Biomedical Engineering Course Lecture Notes. *Oxford University*, 2002.

[107] A Travaglini, C Lamberti, J DeBie, and M Ferri. Respiratory Signal Derived from Eight-lead ECG. In *Computers in Cardiology Proceedings*, volume 25, pages 65–68, 1998.

[108] J L Tylee. On-line Failure Detection in Nuclear Power Plant Instrumentation. *IEEE Transactions on Automatic Control*, AC-23(3):406–415, March 1983.

[109] P Varady, T Micsik, S Benedek, and Z Benyo. A Novel Method for the Detection of Apnea and Hypopnea Events in Respiration Signals. *IEEE Transactions on Biomedical Engineering*, 49(9):936–942, September 2002.

[110] M Vauhkonen, P A Karjalainen, and J P Kaipio. A Kalman Filter Approach to Track Fast Impedance Changes in Electrical Impedance Tomography. *IEEE Transactions on Biomedical Engineering*, 45(4), April 1998.

[111] P Verlinde, G Chollet, and M Acheroy. Multi-modal identity verification using expert fusion. *Information fusion*, 1:17–33, November 1999.

[112] Andrew Webb. *Statistical Pattern Recognition*. Arnold, 1999.

[113] Physionet website. *http://www.physionet.org/*.

[114] D E Weese-Mayer, R T Brouillette, A S Morrow, L P Conway, L M Klemka-Walden, and C E Hunt. Assessing validity of infant monitor alarms with event recording. *Journal of Pediatrics*, 115(5):702–708, November 1989.

[115] G Welch. SCAAT: Incremental Tracking with Incomplete Information. *PhD Thesis, University of North Carolina*, 1996.

[116] A J Wilson, C I Franks, and I L Freeston. Algorithms for the detection of breaths from respiratory waveform recordings in infants. *Medical Biological Engineering Computing*, 20(3):286–292., May 1982.

[117] L Zhao, S Reisman, and T Findley. Derivation of Respiration from Electrocardiogram during Heart Rate Variability Studies. In *Computers in Cardiology Proceedings*, pages 53–56, 1994.

[118] C W Zwillich. Sleep apnoea and autonomic function. *Thorax*, 53(Suppl 3):S20–24, 1998.