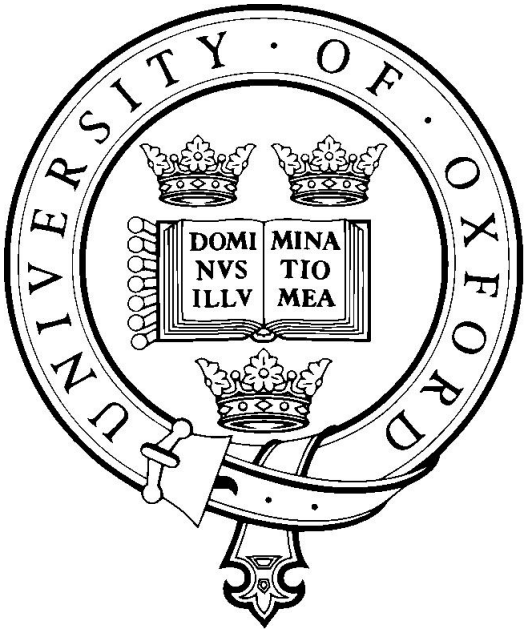
Vital Sign Monitoring and Data Fusion for Paediatric Triage



Syed Ahmar Shah

Linacre College

Supervised by: Professor Lionel Tarassenko

Trinity Term, 2012

This DPhil Thesis is submitted to the

Department of Engineering Science, University of Oxford

Vital Sign Monitoring and Data Fusion for Paediatric Triage

Syed Ahmar Shah

Linacre College

DPhil in Engineering Science

Trinity Term, 2012

Abstract

Accurate assessment of a child's health is critical for appropriate allocation of medical resources and timely delivery of healthcare in both primary care (GP consultations) and secondary care (ED consultations). Serious illnesses such as meningitis and pneumonia account for 20% of deaths in childhood and require early recognition and treatment in order to maximize the chances of survival of affected children. Due to time constraints, poorly defined normal ranges, difficulty in achieving accurate readings and the difficulties faced by clinicians in interpreting combinations of vital signs, vital signs are rarely measured in primary care and their utility is limited in emergency departments.

This thesis aims to develop a monitoring and data fusion system, to be used in both primary care and emergency department settings during the initial assessment of children suspected of having a serious infection. The proposed system relies on the photoplethysmogram (PPG) which is routinely recorded in different clinical settings with a pulse oximeter using a small finger probe. The most difficult vital sign to measure accurately is respiratory rate which has been found to be predictive of serious infection. An automated method is developed to estimate the respiratory rate from the PPG waveform using both the amplitude modulation caused by changes in thoracic pressure during the respiratory cycle and the phenomenon of respiratory sinus arrhythmia, the heart rate variability associated with respiration. The performance of such automated methods deteriorates when monitoring children as a result of frequent motion artefact. A method is developed that automatically identifies high-quality PPG segments mitigating the effects of motion on the estimation of respiratory rate.

In the final part of the thesis, the four vital signs (heart rate, temperature, oxygen saturation and respiratory rate) are combined using a probabilistic framework to provide a novelty score for ranking various diagnostic groups, and predicting the severity of infection in two independent data sets from two different clinical settings.

Acknowledgments

Foremost of all, I would like to thank my supervisor, Lionel Tarassenko, for his invaluable support, guidance and wisdom. I consider myself very lucky to be given the opportunity to work under his supervision. He is a master of what he does, thank you very much.

I would also like to thank Susannah Fleming who provided me with valuable support during the initial days of my DPhil when I joined the group and for her contributions to the setting up of the 8-month clinical trial which forms the core of this thesis.

I am indebted to our clinical collaborators, Matthew Thompson, David Mant, Ann Van den Bruel and Richard Stevens from the "Department of Public Health and Primary Care, University of Oxford", who provided regular feedback and valuable insight. Acknowledgments are also extended to Monica Lakhanpaul and Roderick Macfaul for providing me access to their datasets used in this thesis.

I am extremely grateful to all the staff members and the paediatric nurses at the Paediatric ED, John Radcliffe Hospital, Oxford who contributed to the ED study. I would especially like to thank Sally Beer, Karen Warnes, and Soubera Yousefi for their efforts and patience during the 8-month clinical trial.

I am grateful to all the members of the BSP group, especially David Clifton, Christina Orphanidou, Oliver Gibson, David Meredith, Sara Khalid, Yasmina Borhani, Dave Wong, Marco Pimental, Busi Vilakazi and Mauro Dsantos for their company, friendship, guidance, ideas, and help over the course of this thesis. I am very grateful to Qiao Li and Gari Clifford for their input on the signal quality assessment work.

Although not directly involved in the work, this thesis would still have been a distant dream had it not been for the many friends and family who helped me to keep my head high which keept

me going over the course of many years. It would be too big a task to mention everyone here, I am forever indebted.

I would like to thank Ahmet Hamdi Furat, Said Al Sarmi, Faisal Khan, Zabair Asghar, Habib Baluwala, Wioleta Dudka and so many people who have been a part of my life in Oxford over the years.

I would reserve the final few words for my family, my brothers Athar and Azhar who had profound influence on the person that I am today, my fiance and a constant friend, Saman, who has always been around, and my parents, Papa and Ammi, without their support I would not have been anywhere close to the person I am today.

Contents

Chapter 1

Introduction

Infection is one of the leading causes of death in children in both the developing and the developed worlds [18]. A recent study investigating child deaths in England, Wales, and Northern Ireland revealed that infection accounted for 20% of all child deaths [107]. Amongst children with acute medical illness, respiratory difficulties and infectious illnesses account for 40-50% of all children presenting to GPs or hospital emergency departments [121, 131].

About 25% of all General Practitioner (GP) consultations [121] and Emergency Department visits [107] are concerned with childhood illnesses. A recent confidential enquiry in the UK found that 23% of childhood deaths amongst those deaths in which a GP was directly involved in dealing with the fatal condition were avoidable [107]. Similarly, 21% cases of in-hospital childhood death were avoidable [107]. In both primary and secondary care, an important factor that contributed to these avoidable deaths was the failure to recognise the severity of illness [107].

Early recognition of serious infection is difficult as many mild, self-limiting infections exhibit the same clinical features as serious infection [16]. In addition, the prevalence of serious infection is low in children [149]. In a study in the UK involving 887 paediatric admissions to the Emergency Department, it was found that while the majority of children had mild, self-limiting illnesses, serious illness was found in 13% of the cases, and that this could not have been predicted from the presenting problems [131]. In another study involving 324 children with meningococcal disease presenting in primary care, only 51% of the children were recognised as having a serious infection and were subsequently sent to hospital after their first contact [142].

Changes in vital signs are known to be associated with severe illness [84, 108, 20, 140, 149]. The National Institute for Health and Clinical Excellence (NICE) recommends that the measurement of temperature, heart rate, respiratory rate and capillary refill time be made part of the routine assessment of children with fever [115]. However, the measurement of vital signs in general practice is rare, partly due to time constraints, poorly defined normal ranges, difficulty in achieving accurate readings, and the difficulties faced by clinicians in interpreting the combinations of vital signs [141, 140].

Respiratory rate in particular is recognised as an important vital sign since it has been found to be predictive of lower respiratory infections [118], pneumonia [84], admission to hospitals [20] and death [108]. However, it is measured by only 17% of general practitioners during the initial assessment of children (under 5 years of age). This may be due to the inaccuracy of manual counting [128, 75, 62]. The accuracy of manual counting depends on the counting period and it is desirable to count chest wall movement for at least 60 seconds to obtain reasonably accurate results[128]. The use of currently available technology for estimating respiratory rate during initial assessment is not a feasible option at the moment, as it involves attaching electrodes to the chest which is difficult and inconvenient outside the critical care setting.

1.1 Assessing the severity of illness using vital signs

Primary care, located in the community, is usually the first point of contact for a patient to access essential healthcare. Examples of primary care settings include healthcare providers such as GPs, pharmacists, optometrists, and dentists. Secondary care provides more specialist services in a hospital setting. Typically patients in secondary care are referred from primary care, except for children presenting to an ED which includes both self[1] and primary care referrals. Figure 1.1 shows an overview of primary care and secondary care clinical settings[2]. Various methods are currently employed for the assessment of the severity of illness in children, depending on the clinical setting.

[1] not in the strict sense of the word, as the referral is usually through a parent or carer

[2] It is possible that some patients might not fit the flow given in the figure (e.g. critically ill patients, especially those with trauma, might be brought in by ambulance and taken directly to ICU), however the diagram represents the journey for the majority of patients.

Typically, the time period to assess the severity of illness of a child is limited in both the primary care and the emergency department at the first point of contact because of the available consultation time and because of the inability of an unwell child to tolerate being monitored for too long. The requirement to use only a minimum number of non-invasive monitoring devices, if any, imposes additional constraints in these settings. Section 1.1.1 will describe how patients are currently assessed within such limitations in the context of the ED. A subset of unwell patients presenting to the ED are admitted to hospital (general ward or an intensive care unit). The current methods to assess patients during hospital care are then described in section 1.1.2.

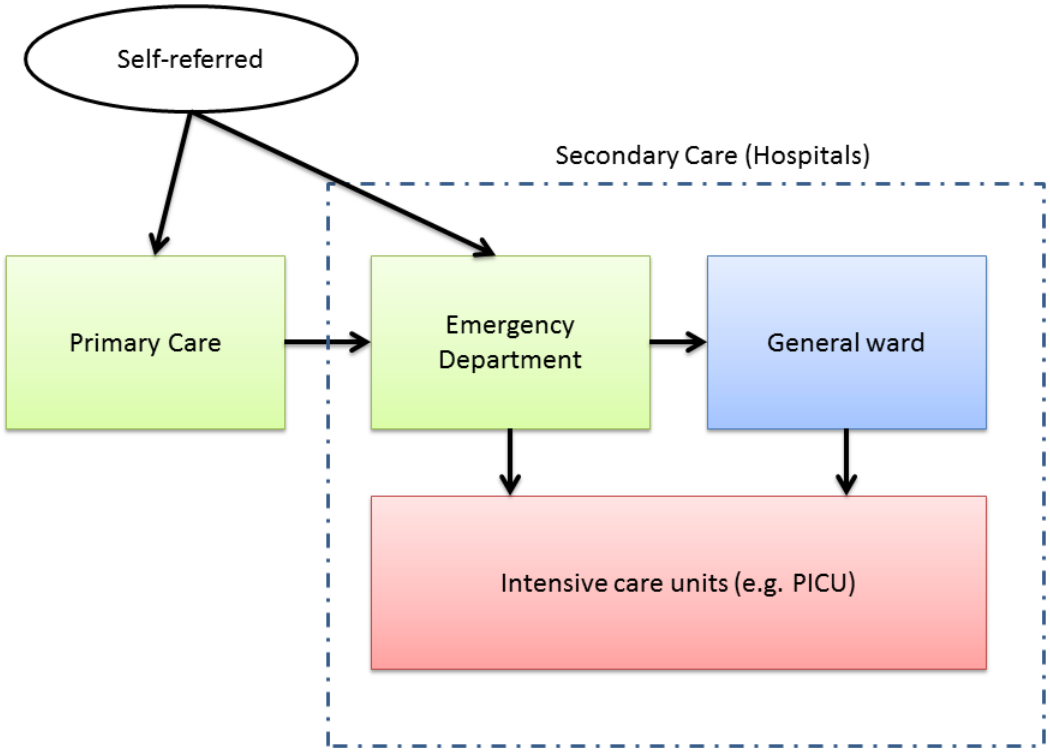


Figure 1.1: Different primary and secondary care clinical settings (PICU=Paediatric Intensive Care Unit)

1.1.1 Assessing patients in emergency departments

The process of assessing the severity of illness in an emergency department is called "triage". This term was first used by the American Army Medical Corps during the First World War to describe

the process of sorting casualties for treatment priority [100]. Triage algorithms are designed to assign transport and treatment priorities to individuals to ensure efficient use of limited medical resources and are critical for the timely delivery of health care and the appropriate allocation of medical resources [100, 83] .

The triage process starts as soon as the patient presents to the emergency department. Existing triage systems use between two and seven priority levels, with the priority decision based on symptoms, risk factors, vital signs, test results and the nursing workload requirements [7]. Ideally, a triage system should improve individual care, should have a simple structure and be easily understood and it should also be able to predict outcomes such as morbidity, mortality, hospitalisation and resource utilization [7, 100]. Lastly, it should be reliable so that the same patient examined by more than one observer would be assigned the same triage category.

Regardless of the triage system used, it is inevitable that some patients will be assigned an inappropriately high or low priority level than that required by the severity of their condition [34]. Undertriage, where a patient is assigned an appropriately lower priority level, is a problem since it puts the patients at risk as they may suffer adverse consequences due to a delay in receiving appropriate treatment. The opposite case, overtriage, where a patient is assigned an inappropriately higher priority level, is also a problem since, although it will not put the overtriaged patient at risk, it will divert medical resources unnecessarily and this may have an adverse effect on the patients that need more urgent medical attention [81, 20].

Current methods of paediatric triage in the Emergency Department

The triage process is usually performed by triage nurses who are trained in a recognised triage system [26]. The literature reports a number of systematic methods of triage practised in different emergency departments.

The Australasian Triage Scale (ATS), Canadian Triage and Acuity Scale (CTAS), and the Manchester Triage Category (MTS) are all based on similar principles and all of them use a five-level classification system [26]. The methodology involves an initial physiological screening in order to identify the most urgent patients followed by problem-specific flow charts. Adjustments

to the triage algorithms to include paediatric patients can be by child-specific flow charts as used in the MTS [123] or by a paediatric scale derived from the adult version, as used in the CTAS i.e. Paed-CTAS [45].

Studies investigating the performance of these scales generally show an adequate to good performance. However, Gouin et al. [45] in their study comparing the performance of Paed-CTAS with a previously used unspecified algorithm, reported that Paed-CTAS is less able to predict cases of admission, requirement for a complete blood count (CBC) and delivery of bronchodilator inhalations. Roukema et al. [116] compared the performance of the MTS in paediatric emergency care with a predefined reference classification based on vital signs, resource utilisation and follow-up and reported a correlation of MTS triage category with both resource utilisation and hospital admissions. The study, however, reported an overtriage in 40% of the patients, most of whom were patients with a fever. Of greater concern was undertriage occurring in 15% of the patients although 96% of them were one urgency category lower than the reference classification.

The Soterion Rapid Triage System (SRTS) is a computerised, five-level triage system developed in the United States [83]. The SRTS begins with an initial screening algorithm to identify potentially life-threatening conditions which, if present, will lead to either Level 1 (immediate) or Level 2 (emergent) triage level. For children who clear the initial screening test, the heart rate, respiration rate and oxygen saturation are measured. If these vital signs are also normal, then one of the 47 paediatric specific algorithms is chosen depending on the chief complaint of the patient [82]. In a study by Maningas et al. [82] to investigate the validity of the SRTS on children attending a mixed emergency department, the triage system was found to be strongly predictive of hospital admission, length of stay and hospital charges. Furthermore, the SRTS was found to be highly reliable, with a 92% exact agreement of triage level between nurses triaging the same patient.

1.1.2 Assessing patients during hospital care

A number of early warning systems have been developed and refined to recognise and respond appropriately to clinical deterioration in hospitalized children to avoid life-threatening events [106, 3]. There is evidence to suggest that serious adverse events (including cardiopulmonary arrest)

in children are preceded by physiological instability, as with adults [22]. With an early warning system in place, an appropriate response can be initiated (e.g. calling a medical emergency team for help [17]).

Table 1.1 provides a summary of nine early warning systems designed to be used to provide early warning of clinical deterioration in hospitalised children. The early warning scores are designed to be used as bedside evaluation tools (usually by the nurses) and typically use a combination of vital signs and additional variables that can readily be observed (e.g. the level of consciousness, respiratory effort at the bedside, need for oxygen, vomiting). Figure 1.2 shows an example of an early warning scoring chart using five vital signs (heart rate, temperature, oxygen saturation, respiratory rate and capillary refill time) and two additional variables obtained through observation (respiratory effort and level of consciousness). Each of these variables is assigned a score from 0 to 3 and the final score is then obtained by adding the scores assigned to each of these individual variables. Finally, a threshold is applied to identify patients who are likely to be at greater risk of clinical deterioration.

Score/Method	Variables	Vital signs				
		HR	Temp	SpO_2	SBP	RR
MEWS [132]	5	●	●	●	●	●
Brighton PEWS [91]	7	●				●
Bristol PEWS [47]	14	●		●		●
SICK score [8]	7	●	●	●	●	●
MET calling criteria [144, 143]	9	●		●	●	●
Toronto PEWS [30]	16	●	●	●	●	●
PAWS [32]	7	●	●	●		●
Cardiff & Vale PEWS [31]	8	●			●	●
Bedside PEWS [105]	7	●		●	●	●

Table 1.1: Summary of design of eight paediatric early warning scores for managing unwell children in a hospital. Abbreviations used: HR (heart rate); Temp (temperature); SBP (systolic blood pressure); RR (respiratory rate); MEWS (modified early warning score); PEWS (paediatric early warning score); SICK (signs of inflammation that can kill); MET (medical emergency team); PAWS (paediatric advanced warning score).

The nine scores summarised in Table 1.1 differ greatly in how they were designed (e.g. prospective [31], retrospective [47]), the parameters used and the outcomes predicted (e.g. cardiac arrest [17], mode of death [47], admission to a Paediatric Intensive Care Unit (PICU) [105]), making

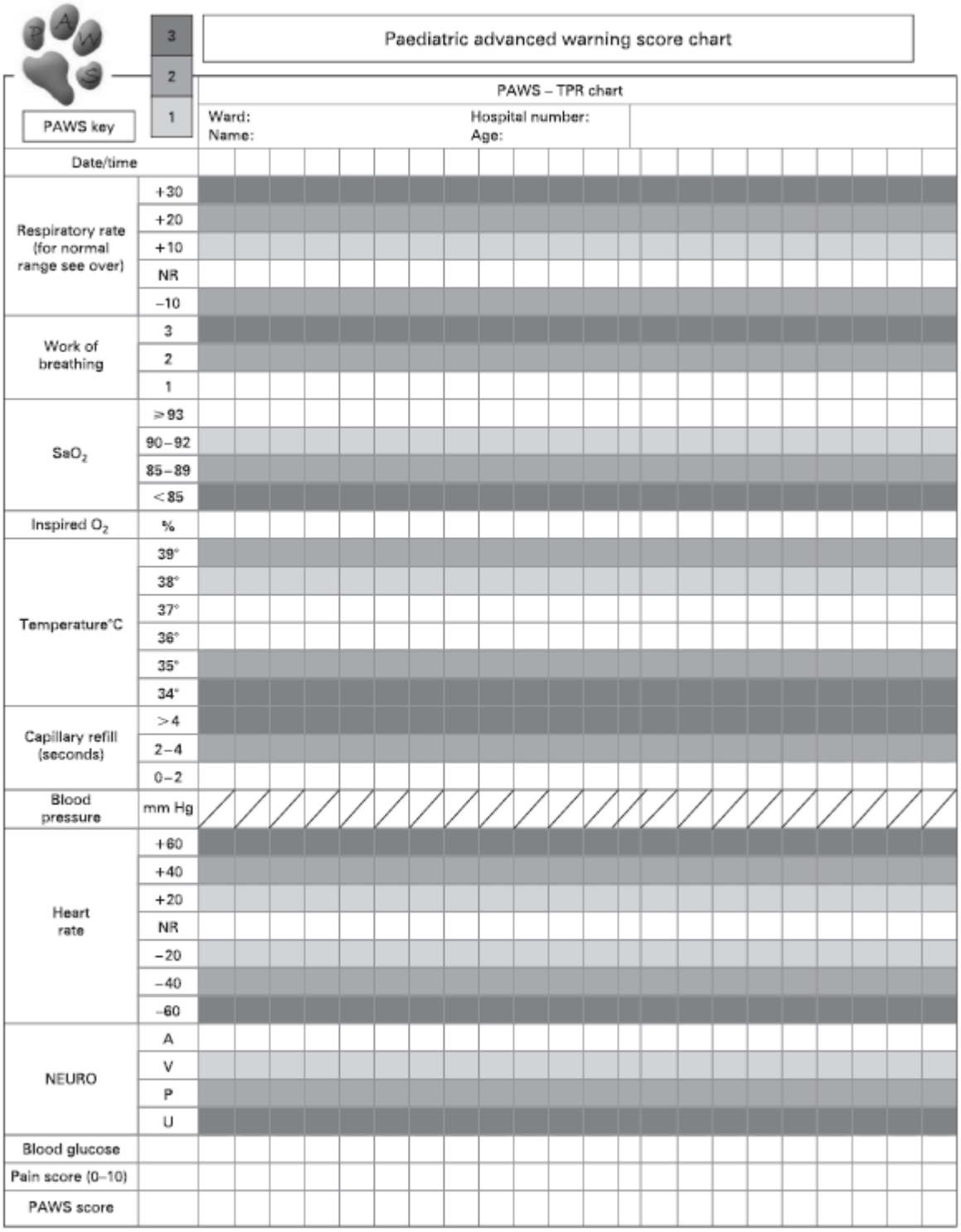


Figure 1.2: Paediatric Advanced Warning Score (PAWS) chart. Abbreviations used: AVPU (alert/responds to voice/responds to pain/unresponsive); NR (normal range); SaO_2 (oxygen saturation). Reproduced from [32] with permission from BMJ Publishing Group Ltd.

it difficult to compare the results reported. Typical results however, include a sensitivity of 89% with a specificity of 64% for adverse outcomes (includes respiratory arrest, admission to PICU and death) on children in general wards for the Cardiff & Vale PEWS [31]; or a sensitivity of 70% with a specificity of 90% for admission of children to a PICU from the general ward for PAWS [32].

The paediatric triage systems (typically employed at presentation to the ED) and the paediatric early warning scores (typically used during monitoring of hospitalised children on general wards) are the most relevant to the application considered in this thesis; development of a robust and reliable monitoring and data fusion system using vital signs acquired non-invasively from children to assess severity of illness. A number of scores have also been developed for assessing patients in intensive care (e.g. PSI [156], PRISM III-APS [108]). These scores typically include a larger number of variables, requiring invasive measurements and blood tests. The outcome measures predicted by such scores in these highly specialised units include mortality [108] and severity of illness [156].

1.2 Monitoring of paediatric vital signs

As described in the previous section, vital sign measurement is an integral part of all paediatric triage and early warning systems. Since the application to be considered in this thesis is the routine screening of sick children, it is essential that only non-invasive monitors are used for the measurement of vital signs. Furthermore, it is desirable to use the minimum number of such sensors, without compromising the utility of the measurements acquired.

As already stated, the National Institute for Health and Clinical Excellence (NICE) recommends that the routine assessment of children with fever should include the measurement of temperature, heart rate, and respiratory rate [115]. Table 1.1 shows that these three vital signs, together with systolic blood pressure and oxygen saturation, feature in most early warning scoring systems. Systolic blood pressure requires the application of a cuff pressure above systolic pressure, usually around the upper arm and this is usually inconvenient and distressing for an unwell child during a routine assessment. The remaining four vital signs (heart rate, temperature, oxygen saturation, and respiratory rate) can potentially be acquired in a short period of time without causing inconvenience or

distress to the child being monitored. They are therefore considered in more detail in the sections that follow.

1.2.1 Heart rate

Heart rate, also referred to as the pulse rate[3], is a measure of the rate at which the heart beats and it is measured in beats per minute. The normal resting heart rate varies with age [36], decreasing during childhood from around 110-160 beats per minute at birth to 60-100 beats per minute by late adolescence [78]. The variation with age of heart rate and respiratory rate will be considered in more detail in this thesis in section 7.2.2 of chapter 7.

Heart rate is included in the design of all nine early warning scores summarised in Table 1.1, and is thus accepted as an important vital sign by the medical community. A raised heart rate can be an indication of a serious medical condition such as septic shock[4] [115].

The gold standard method for measuring heart rate requires the electrocardiogram (ECG) to be recorded by placing electrodes at various points on the thorax. The ECG is a time-varying signal that measures the electrical activity of the heart. It is composed of a set of waveform features which correspond to the various stages that make up a heartbeat. Typical configurations for recording the ECG are either 3 electrodes placed on the chest or 10 electrodes (6 chest and 4 limb), depending on the clinical information desired [4].

Although the ECG is the gold standard for the measurement of heart rate, it is not appropriate to use this for the routine assessment of unwell children in primary care or ED settings for a number of reasons. Firstly, using the ECG for heart rate measurement will require some undressing of the child to allow the attachement of chest electrodes. This will be time consuming and may cause undue discomfort to the unwell child. This is further compounded by the need to ensure accurate placement and attachment of the electrodes to record a good-quality ECG. Finally, ECG electrodes are typically designed to be adhesive and their removal can cause additional pain to the unwell child.

[3] pulse rate is typically used when the rate of heart beating is measured by manual palpation

[4] a serious complication which can lead to multiple organ failure due to inadequate blood supply to various body tissues

The manual measurement of heart rate is common in clinical practice during initial assessment, either by using a stethoscope, or by palpating the wrist (radial artery) [34] or the neck (carotid artery). It is also possible to use a pulse oximeter (explained in detail in section 1.2.3 below) for the measurement of heart rate. The latter has been used in both the primary care [141] and emergency department settings (e.g. [16]; the Paediatric ED, JR Hospital, Oxford, UK; Pinderfields dataset described in chapter 8).

1.2.2 Respiratory rate

Also referred to as the breathing rate[5], respiratory rate is measured in breaths per minute (bpm). As with heart rate, normal respiratory rates vary with age [36], with 30-40 breaths per minute at birth reducing to 15-20 breaths per minute by late adolescence [78].

Respiratory rate in particular is recognised as an important vital sign since it has been found to be predictive of lower respiratory infections [118], pneumonia [84], admission to hospitals [20] and death [108]. All nine early warning scores in Table 1.1 include respiratory rate.

Respiratory rate is typically measured by counting chest wall movement in both the primary care and ED settings, during the initial assessment of a child. Although the method is simple and does not interfere with the breathing mechanism, it has been found to be inaccurate [128, 75] with errors up to 34% reported in the literature [62]. Furthermore, the accuracy of manual counting depends on the counting period and it is desirable to count chest wall movement for at least 60 seconds to obtain reasonably accurate results [128]. However, counting periods of just 15 seconds or at most 30 seconds are far more common in the estimation of respiratory rate (see figure 3.6 in chapter 3). There is also evidence that not all nurses are adequately trained to measure respiratory rate accurately, with one study reporting that only 67% of nurses were able to measure respiratory rate to within 10 breaths per minute during an investigation in a simulated environment [106].

In addition to manual counting, there are a number of other methods that can be used to estimate respiratory rate. These include attaching sensors to nasal/oral region to monitor changes

[5]Breathing rate is the mechanical process of taking in oxygen and expelling carbon dioxide while respiration is the chemical process whereby oxygen is absorbed into the bloodstream from the lungs. This distinction will not be made in this thesis and both will be used interchangeably.

in temperature, humidity, carbon dioxide or to detect sound changes with respiration. A previous study involving healthy children has shown that not all of the children are able to tolerate being monitored with sensors attached to the nasal/oral region [34] suggesting that is not an appropriate option for routine monitoring of children presenting to primary care or ED.

Other non-invasive methods rely on the detection of the changes in tissue composition, volume or movement that occur during respiration [38]. In this category, impedance plethysmography is the most common method in hospital settings exploiting the changes in electrical impedance of the thorax as a result of the variation in air volume within the lungs during the respiratory cycle. A possible reason for its popularity in hospital settings is that it can be measured simultaneously with the ECG, using the same electrodes and thereby eliminating the need to attach extra hardware [34]. Other methods for respiratory rate estimation include measuring the changes in the abdomen and thoracic circumference (using strain-gauge transducers), changes in thorax volume (using capacitance measurement) and changes in muscle activity using electromyography [38].

It is not feasible to use impedance plethysmography in the intended application for the reason that was given in section 1.2.1 (the need to use chest electrodes). This also eliminates the possibility of using electromyography, which also requires electrodes attached to the skin. The use of elasticated bands (for measuring inductance) or of sensors that measure thoracic volume changes is also not feasible because of the inconvenience for children, and additionally the interference with the normal breathing pattern of a child.

1.2.3 Arterial oxygen saturation (SpO_2)

Arterial oxygen saturation (SpO_2) gives a measure of the amount of oxygen carried by the arterial blood [89]. It is given by the percentage of haemoglobin that is attached to oxygen as given in equation 1.1

$$SpO_2 = \frac{[HbO_2]}{[HbO_2] + [Hb]} \times 100\% \tag{1.1}$$

where $[HbO_2]$ is the concentration of haemoglobin attached to oxygen (O_2), and $[Hb]$ is the concentration of reduced haemoglobin. Unlike heart rate and respiratory rate, SpO_2 does not vary

with age and a value of 95-100% is considered normal in both children and adults [78].

The value of SpO_2 is included in seven of the nine early warning scores shown in Table 1.1. There is evidence that reduction in SpO_2 is associated with serious infection [140], especially pneumonia [115].

SpO_2 is typically measured by using a pulse oximeter, which consists of circuits for processing the light transmitted through a small, non-invasive finger probe. The latter includes a light source (an LED[6]) on one side of the finger to which it is attached and a sensor (a photodiode) on the opposite side of the finger. The light received by the photodiode is converted into an electrical signal, commonly termed as the photoplethysmogram (PPG), as shown in figure 1.3. The common measurement sites for pulse oximetry in clinical settings are the finger, toe and ear lobe.

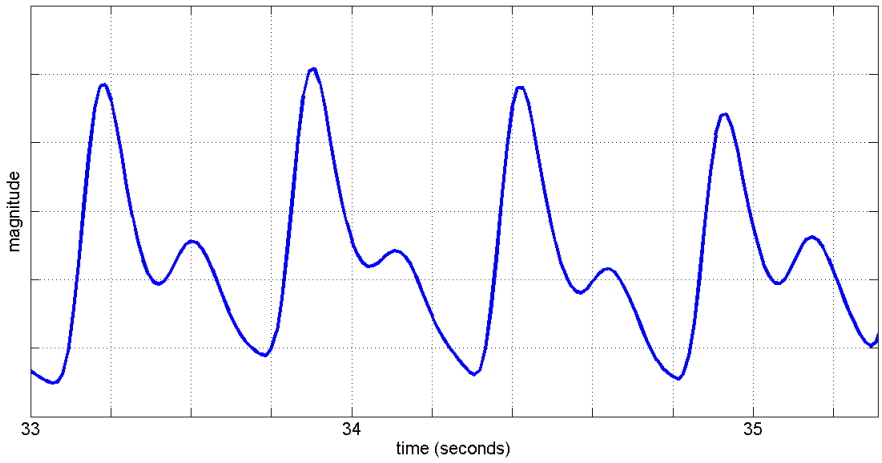


Figure 1.3: A typical PPG waveform

The fundamental principle behind oximetry is the Beer-Lambert law given by equation 1.2

$$I_t = I_{in} e^{-DC\varepsilon} \tag{1.2}$$

where I_t is the intensity of the transmitted light (received by the photodiode), I_{in} is the intensity of the incident light (emitted by the LED), D is the distance travelled through the blood (transmission path length), C is the concentration of the solute ($[Hb]$ or $[HbO_2]$) and ε is the extinction coefficient for the solute (Hb or HbO_2) at the LED wavelength. The values of ε for different wavelengths

[6]light emitting diode

for both oxyhaemoglobin and reduced haemoglobin are given in figure 1.4, with the wavelengths corresponding to the Red and Infrared wavelengths marked. These are the two wavelengths that are typically used in pulse oximeters. The values of extinction coefficient for oxyhemoglobin and hemoglobin differ at these wavelengths, as can be seen in figure 1.4.

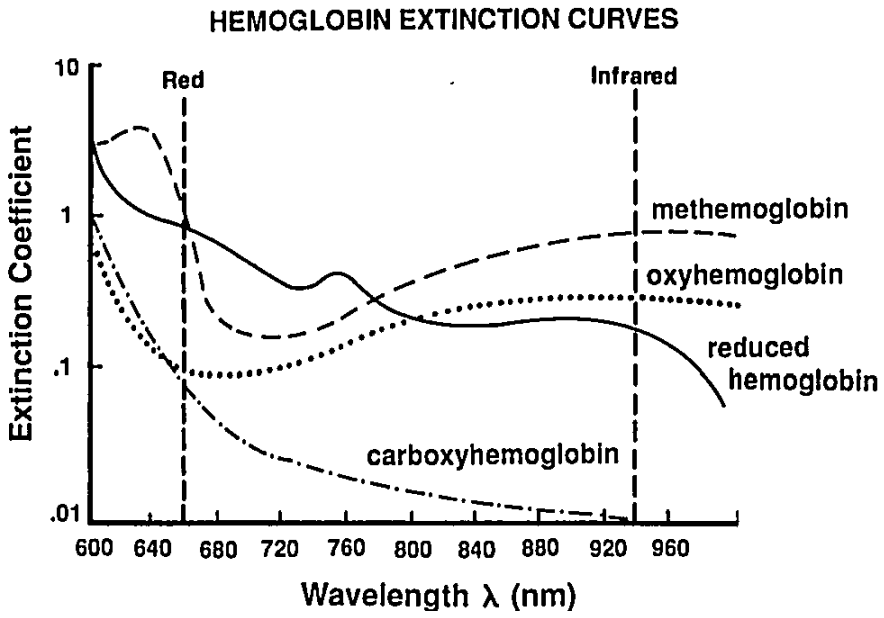


Figure 1.4: Absorption spectra for the different haemoglobin species. Reproduced from [145], Figure 2 with the permission of the copyright owner, Lippincott Williams and Wilkins.

Oximetry exploits the difference in light absorption by the Hb and HbO_2 molecules in the arterial blood at the two different wavelengths to provide a measure of oxygen saturation. Apart from the haemoglobin molecules in the arterial blood, light is also absorbed by the capillary and venous blood, the skin pigment, soft tissue, and the bones through which it passes [89]. Early oximeters were calibrated by finding an estimate of the bloodless tissue light absorption after compression to prevent blood flow [145]. These oximeters also required heating to minimise the effect of venous and capillary blood during oxygen saturation measurement [145].

"Pulse" oximetry relies on separating the two distinct components of light absorption: a pulsatile component (AC) due to the flow of arterial blood (in synchronisation with the heart beat), and a baseline component (DC) corresponding to the light absorption by the non-pulsatile part of the arterial, capillary and venous blood and skin (see figure 1.5 for illustration). From the values of the AC and DC components at the two different wavelengths, the ratio given by equation 1.3 can

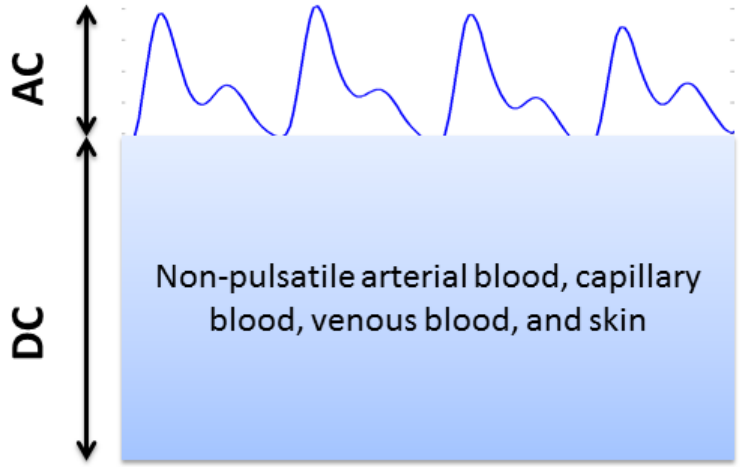


Figure 1.5: Pulsatile (AC) and non-pulsatile (DC) components of light absorption

be determined.

$$R = \frac{AC_{\lambda 1}/DC_{\lambda 1}}{AC_{\lambda 2}/DC_{\lambda 2}} \tag{1.3}$$

where $AC_{\lambda 1}$ and $DC_{\lambda 1}$ are the AC and DC components of light of wavelength λ_1 received by the photodiode, and $AC_{\lambda 2}$ and $DC_{\lambda 2}$ are the AC and DC components of light of wavelength λ_2 received by the photodiode. Ideally, equation 1.2 would be used to find the concentration of oxyhaemoglobin and reduced haemoglobin. However, to account for the effects caused by multiple scattering of the light by red blood cells, calibration curves experimentally derived from healthy young volunteers (by measuring oxygen saturation in parallel using blood gas analysis) are used instead to relate R to the value of oxygen saturation [129].

Pulse oximeters are small, cheap, portable, non-invasive, do not require any prior training for their use and are thus ideally suited to the initial assessment of a child in primary care or emergency departments.

1.2.4 Temperature

The measurement of body temperature allows the identification of patients with fever [33].

The core (internal) body temperature varies with location but core temperature is commonly defined as the temperature measured within the pulmonary artery [33]. Since internal body sites

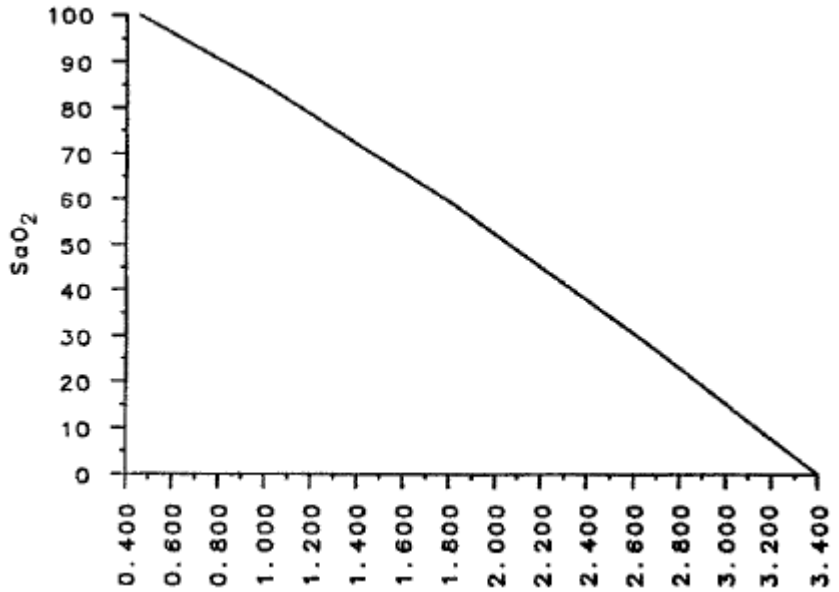


Figure 1.6: Relationship between the numeric value of R and the arterial oxygen saturation. Reproduced from [154], figure 5, with kind permission from Springer Science and Business Media.

cannot be directly accessed, various surrogate sites have been used over the years to provide a reasonable estimate of the core body temperature. These include the rectum, the axilla (armpit), and beneath the tongue. The types of thermometers that can be used to provide temperature measurements at these sites include electronic, and mercury-in-glass thermometers. In addition, there are plastic-encased thermophototropic liquid crystals based thermometers (typically applied on the forehead) that change colour as the temperature rises and infrared based thermometers that measure the infrared radiation near the tympanic membrane and provide a surrogate for the core body temperature [33].

The ideal thermometer should provide rapid and accurate reflection of the core body temperature, be non-invasive, easy to use, and with minimal, if any, risk of cross-infection.

The rectal route is no longer the preferred anymore due to a number of disadvantages including a high chance of cross infection, and rectal perforation [33]. Thermometers placed beneath the tongue are also not preferred for a variety of reasons including the possibility of cross-infection, oral laceration, susceptibility to error (e.g. due to hot/cold drinks, mouth breathing, placement of thermometer bulb) [33]. While disposable plastic-encased thermometers do not pose any risk of cross-infection, they have been found to be inaccurate and not suitable for routine measurement in a clinical setting. Mercury-in-glass thermometers are also avoided due to the risk of mercury

poisoning whenever the glass breaks.

Electronic thermometers measuring tympanic or axillary temperatures are ideally suited to the routine assessment of children in clinical settings since the measurement is quick, reasonably accurate, clean and there is a low [93, 61], low risk of cross-infection.

1.3 Overview of thesis

About 40-50% of all children presenting to GPs or hospital emergency departments (EDs) suffer from respiratory difficulties and/or infectious illnesses [121, 131]. A number of childhood deaths, in both primary care (GP consultations) and secondary care (ED consultations), occur due to failure to recognise the severity of illness [107]. Due to time constraints, poorly defined normal ranges, difficuly in achieving accurate readings and the difficulties faced by clinicians in interpreting combinations of vital signs, vital signs are rarely measured in primary care [141, 140] and their utility is limited in emergency departments [116, 140, 151, 45].

This thesis aims to develop a robust and reliable monitoring and data fusion system, to be used in both primary care and emergency department settings during initial assessment, for the identification of children with serious infection. As discussed in section 1.2, it is not practical to use either invasive techniques or a large number of non-invasive sensors during the assessment of unwell children. The proposed system, therefore relies solely on a pulse oximeter and a thermometer probe to measure heart rate, oxygen saturation, temperature, peripheral perfusion and respiratory rate. The work to be described in the thesis can be divided into two main parts, as shown in figure 1.7: (i) monitoring of vital signs; and (ii) vital sign data fusion.

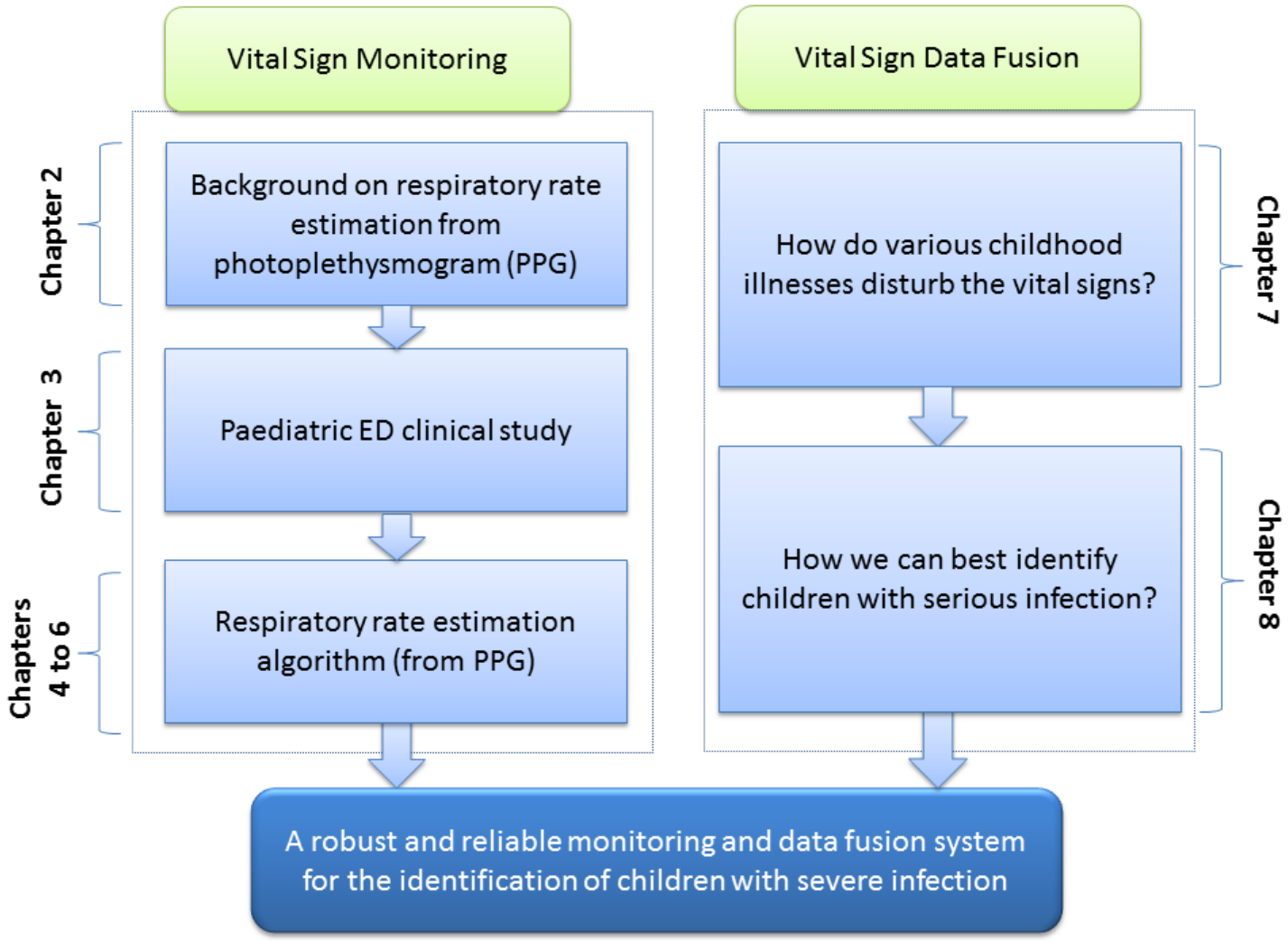


Figure 1.7: Overview of thesis

While heart rate, oxygen saturation and peripheral perfusion can be measured with a pulse oximeter, respiratory rate is currently estimated by counting chest wall movements. Signal processing algorithms will be developed to provide an estimate of respiratory rate using the photoplethysmogram (PPG) signal recorded by the pulse oximeter. Respiratory rate modulates both the amplitude (AM) [90] and the period of the PPG waveform (RSA) [150], and this can be exploited to estimate respiratory rate. Chapter 2 describes the physiological basis of respiratory rate estimation, reviewing the two phenomenon in detail. Previous studies for the estimation of respiratory rate from the PPG have mostly been undertaken in controlled settings and there is significant deterioration in performance when algorithms are evaluated on data acquired from subjects in clinical settings. The clinical trial set up to test methods to estimate respiratory rate from PPG recordings

acquired from a large number of children presenting to the Emergency Department is described in chapter 3.

Chapter 4 and 5 focus on the development and implementation of signal processing algorithms for the estimation of respiratory rate from PPG. The performance of these algorithms is likely to be affected by the quality of PPG signal, which is expected to deteriorate when motion artefact(s) occur. Manual selection of artefact-free PPG recordings is not applicable in a clinical setting where automated respiratory rate estimation is needed. Chapter 6 describes the development of an algorithm for automatically identifying high-quality PPG segments, so that the accuracy of respiratory rate estimation can be improved.

While the predictivity of individual vital signs for the identification of serious illness is low [140, 115], it is anticipated that combining different vital signs will improve it. The extent of vital sign disturbance is investigated in chapter 7, using both univariate and multivariate analyses. Finally, a score for quantifying the severity of infection using vital signs is developed and tested on independent datasets in Chapter 8.

Chapter 2

Measuring Respiratory Rate using the Finger Probe (PPG)

2.1 Introduction

Respiratory rate has traditionally been measured by counting chest wall movements using a watch for 30-second or 60-second periods. Although the method is simple and does not interfere with the breathing mechanism, it has been found to be inaccurate with errors up to 34% reported in the literature [87]. As explained in section 1.2 of chapter 1, there are a number of other methods including nasal thermistry, capnography, inductance plethysmography, and electrical impedance plethysmography that allow the measurement of respiratory rate. However, all of these methods require the use of extra hardware attached to the patient. This can be difficult and inconvenient in many clinical settings and impossible in the context of paediatric triage.

Photoplethysmography is a non-invasive procedure routinely used in hospital, including paediatric triage, to measure pulse rate[1] and oxygen saturation. The ability to extract respiratory rate from the photoplethysmogram (PPG) would be a major improvement in the triage process since it would release the clinician or triage nurse from holding a timer and observing chest wall movement, thereby allowing them to focus on the child's medical history and other symptoms. Furthermore,

[1] pulse rate and heart rate are considered to be the same in this thesis (they only differ in the measurement site and/or method)

the ability to measure respiratory rate automatically from the PPG waveform will eliminate human error and also allow respiratory rate to be combined with other vital signs for improved diagnosis of a child's health status.

Before presenting a review of the various signal processing algorithms that have been developed to estimate respiratory rate from the PPG, it is important to understand the physiological basis for the respiratory information in the PPG, as this will allow us to select the optimal signal processing method for respiratory rate estimation.

2.2 Physiological basis of respiratory information in the PPG

The origin of respiratory rate information in the PPG is not fully understood, although it is clearly related to changes in peripheral circulation , directly and/or indirectly, as a result of respiration.

Respiration causes changes in the baseline of the PPG (typically described as "respiratory induced intensity variation" [90]) as well as changes in the amplitude of the PPG. The distinction between the two is rarely made by the manufacturers of pulse oximeters and depends on the choice of cut-off frequency chosen to separate the PPG signal into a baseline component and a pulsatile component [53, 90]. In this thesis, the term "amplitude modulation" (AM) of the PPG will be used to refer to both phenomena.

Respiration also influences the period of the PPG due to changes in heart rate caused by respiration, a phenomenon known as Respiratory Sinus Arrhythmia (RSA). This may be described as"frequency modulation" (FM) of the PPG signal [34]. To avoid confusion later, this phenomenon will be referred to as RSA in the rest of the thesis.

2.2.1 Respiratory induced Amplitude modulation (AM) of the PPG

Fundamentally, the change in peripheral tissue perfusion in synchrony with respiration (successive cycles of inspiration and expiration) results in the amplitude modulation of the PPG. Three different reasons have been documented in the literature to explain the reduction in peripheral tissue blood volume with inspiration: an increase in the venous return of blood from the periphery to the heart

[73, 29, 55, 90], a reduction in stroke volume [73, 39, 90] and the neural activity of the autonomic nervous system resulting in the constriction of blood vessels [99].

Increased venous return

Two reasons have been proposed for an increased venous return synchronous with inspiration. Firstly, there is a reduction in intra-thoracic pressure during inspiration which leads to a reduction in the right atrial pressure. Consequently, there is an increased venous return from the periphery to the heart [73]. Secondly, the pressure on the abdominal vessels is increased due to the lowering of diaphragm during inspiration which also leads to an increased venous return [90]. The effect of the latter is likely to be minimal at normal respiratory rates [90].

Reduction in stroke volume

Three different reasons have been given for the phenomenon of reduced stroke volume during inspiration. Firstly, the compliant pulmonary vessels external to the lungs (extra-alveolar vessels) distend due to the reduction in intra-thoracic pressure during inspiration effectively pooling the blood [90, 52]. Consequently, the blood flow to the left side of the heart (left atrium) is reduced leading to a reduction in stroke volume. Secondly, there is an increased resistance to blood flow from the right side of the heart (right ventricle) due to compression of the alveolar capillaries as a result of the expanding alveoli during inspiration [90]. Lastly, the increased venous return during inspiration results in an increased blood flow to the right ventricle thereby expanding it. This causes a bulging of the intra-ventricular septum, reducing the volume of the left ventricle and decreasing its compliance, a phenomenon known as ventricular interdependence [14].

The relative contributions of reduced stroke volume and the increased venous return are not clear. A study by Nilsson et. al [97] investigating the relationship between peripheral venous pressure and the amplitude modulation of the PPG waveform (measured by a sensor positioned on the medial side on a forearm) found that although both are correlated, a causal relationship could not be demonstrated. A phase analysis comparing the respiratory-synchronous peak in the PPG signal (i.e. the amplitude modulation) and the peak of the peripheral venous pressure showed that

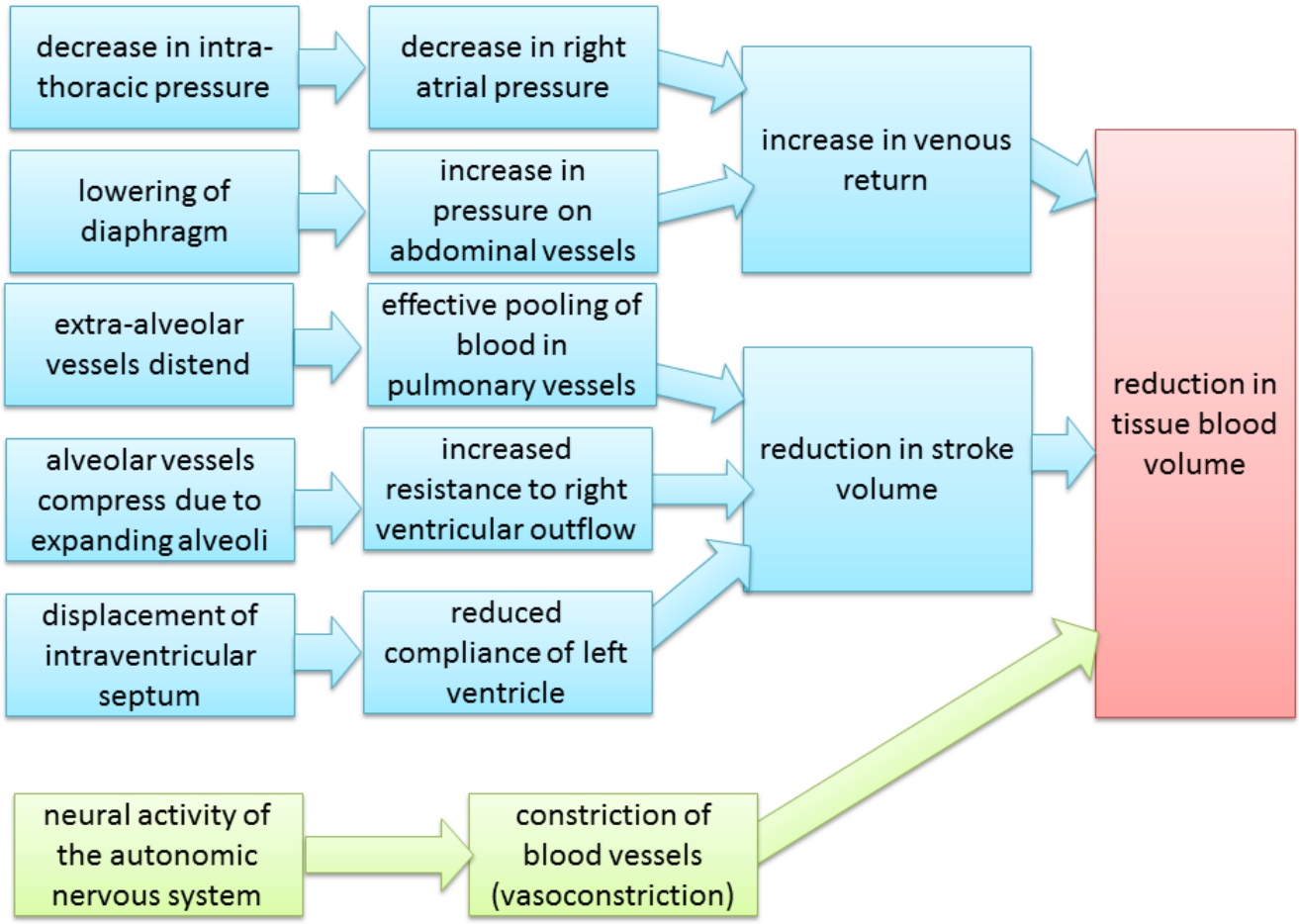


Figure 2.1: The various changes in physiology that lead to a reduction in tissue blood volume during inspiration. The major factors are likely to be mechanical (shown shaded in blue) while there is little evidence so far for the role of neural activity (shown shaded in green)

the former was usually leading the latter. However, due to large variations between the subjects, it was not conclusive. A later study by the same group [98] showed through phase analysis that a reduced arterial blood pressure (reflecting a reduction in stroke volume) preceded the respiratory-synchronous changes in the PPG signal (measured at the same site as the previous study) while the decreased central venous pressure was seen to occur after respiratory-synchronous changes in the PPG signal. However, it is still difficult to draw any conclusions from the study since the difference in the two cases did not reach significance.

Neural activity

It has also been suggested in the literature that respiratory activity is coupled with neural activity, independently of the mechanical effects of respiration which modulate the amplitude of the PPG [96]. Neural activity causes constriction of the peripheral blood vessels (vasoconstriction) leading to reduced peripheral blood perfusion [99]. The evidence for this comes mostly from a study by Nitzan et. al [99] who found respiratory-induced variations in the PPG signal (measured on the index finger) even though the corresponding artery was occluded by a pressure cuff to eliminate the possibility of mechanically-transmitted changes in peripheral tissue perfusion. This phenomenon is thought to occur due to the stretch receptors in the lungs as a result of the pressure variations during respiration [73, 97]. It is also possible that the stretch receptors in the aorta (arterial baroreceptors) react to reduced stroke volume contributing to reduced peripheral tissue perfusion and hence amplitude modulation in the PPG [6].

Figure 2.1 illustrates the various stages just described to show how tissue blood volume is reduced due to inspiration thereby causing “amplitude modulation” of the PPG due to respiration.

Based on the current evidence, the major contribution to amplitude modulation of the PPG seems to be of mechanical origin, arising from changes in intra-thoracic pressure, diaphragm movement, compression of alveolar vessels, distension of extra-alveolar vessels, and displacement of intraventricular septum (shown shaded in blue figure 2.1). There is little, if any, direct neural control that contributes to the amplitude modulation of the PPG (shown shaded in green in figure 2.1) and thus, it is unlikely to be affected by age or health status. The magnitude of AM is however, affected

by tidal volume (volume of air taken in during inspiration) and the type of breathing (thoracic or abdominal). A study by Johansan and Stromberg [55] reported that mainly thoracic breathing and large tidal volumes contribute most to the amplitude modulation of the PPG.

2.2.2 Respiratory Sinus Arrhythmia (RSA) manifested in the PPG

Beat-to-beat changes in heart rate (commonly termed heart rate variability) are controlled by the sympathetic (likened to a fight or flight response) and parasympathetic (likened to a rest and digest response) function of the autonomic nervous system [150]. Various cardiovascular control mechanisms provide frequent small adjustments in the heart rate [150]. One such control is provided by the vagus nerve, a paired nerve that starts from the brainstem and extends to different internal organs. The vagus nerve is stimulated during expiration, and this causes it to decrease the activity of the sinoatrial node (a group of cells in the heart that provides the stimulus which results in a heartbeat) resulting in a reduction in the heart rate [51]. During inspiration, the ability of the vagus nerve to stimulate the sinoatrial node is inhibited leading to a relative increase in the heart rate. Consequently, the heart rate fluctuates in synchrony with the respiratory rate [150]. This phenomenon is referred to as Respiratory Sinus Arrhythmia (RSA).

RSA is thought to originate from two major factors: direct modulation of the vagus nerve by the respiratory centre in the brain [155], and the inhibitory effect on the activity of the vagus nerve of stretch receptors during inspiration leading to an elevated heart rate [90]. Peripheral reflexes due to blood pressure and volume changes are also thought to contribute to RSA [150].

RSA is thought to provide a mechanism for optimising the gaseous exchange of oxygen and carbon dioxide in the lungs by matching the alveolar volume (high during inspiration due to reduced intra-thoracic pressure) and the blood flow in alveolar capillaries (high during inspiration due to elevated heart rate) [155]. This phenomenon was observed to be very pronounced in patients with Cheyne-Stokes syndrome [74], a condition characterized by periodic cessation of breathing (apnoea) and elevated breathing. In the study carried out by Lorenzi-Filho et al. [74], a clustering of heart beats was observed during the elevated breathing phase and a scattering was observed during the apnoeic phase, thereby optimising the gaseous exchange and saving cardiac energy [74]. This is

further corroborated by the study reported in [155], which was undertaken on seven anaesthetised dogs and found that the RSA mechanism improved oxygen intake by 4%.

RSA is thought to be the fastest internal physiological rhythm that influences the heart rate and it is therefore seen as a (relatively) high-frequency component (about 0.25 Hz) in the heart rate variability spectrum [138, 92]. Spectral analysis of heart rate variability has previously been suggested as one approach to obtain respiratory rate. However, there are slower physiological rhythms that influence heart rate and these appear as low-frequency components mediated by the sympathetic and parasympathetic functions [138, 92]. Slow respiration can also cause the frequency components due to RSA to overlap with these low-frequency components. Consequently, RSA will be affected by conditions that affect heart rate variability, known to be affected by the presence of neurological or cardiac disorders, smoking, alcohol, as well as certain drugs [110]. Furthermore, it has been reported in the literature that RSA decreases with age [130]. Tulppo et al. [146] reported that this reduction seems to be independent of any impairment of physical fitness. Lastly, Grossman et al. [46], in their study on investigating the relation between RSA and physical activity, reported a progressive attenuation of RSA with increasing levels of physical exertion.

Traditionally heart rate variability measures have been derived from the ECG, including RSA measurements for respiratory rate estimation [122, 27]. This is done by analysing the time series of beat-to-beat intervals in ECG using time domain or frequency domain techniques [110]. However, there are a few recent studies that have compared the time series of beat-to-beat intervals derived from the ECG and the PPG [13, 124, 77] (see section 5.1 for further details). These studies have reported a high correlation and concluded that the PPG can serve as an alternative to ECG for extracting heart rate variability information from the beat-to-beat interval time series. It is thus useful to develop signal processing algorithms for the estimation of respiratory rate from RSA extracted from the PPG recordings.

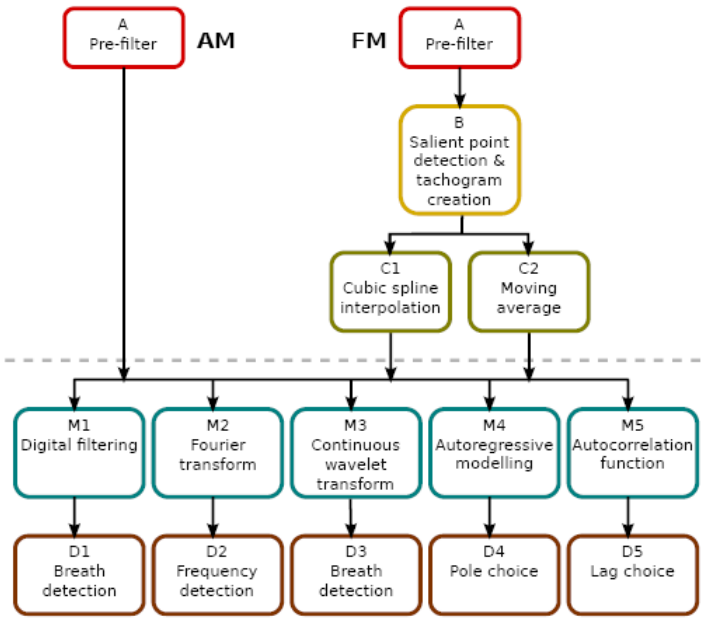


Figure 2.2: Various signal processing methods for obtaining respiratory rate from the PPG/ECG from AM and FM (referred to as RSA in the thesis). Blocks above the line correspond to pre-processing steps while blocks below the line correspond to respiratory waveform analysis (Reproduced from [34])

2.3 Signal processing algorithms for the estimation of respiratory rate from the PPG

There have been several studies that attempted to extract respiratory rate from the PPG. The various signal processing algorithms previously developed include digital filtering, autoregressive modelling based methods, time-frequency analysis, independent component analysis (ICA) , and neural network. There are relatively fewer papers that have described signal processing algorithms for extracting respiratory rate estimation from the PPG using RSA (in comparison to AM in the PPG). For illustration, figure 2.2 shows a summary of various methods developed for the estimation of respiratory rate from PPG and ECG from a literature review of different signal processing algorithms undertaken in a previous study [34]. Chapter 4 of this thesis will cover the development and implementation of different signal processing algorithms to extract the amplitude modulation of the PPG. Chapter 5 will then present the implementation results of extracting RSA for the PPG.

A number of performance evaluation metrics for respiratory estimation algorithms have been

Metric	References
error rate (sensitivity/true positive/false positive etc.)	[98, 70, 54]
correlation coefficient	[39, 125]
mean absolute error	[94, 71, 24, 37, 35, 34]
root mean square error (RMSE)	[66, 67]

Table 2.1: different performance evaluation metrics for respiratory rate estimation algorithms used in the literature

used in the literature, summarised in table 2.1. Metrics computing error rate can only be used if a reference respiratory waveform is available. The correlation coefficient was used in the literature to compute the degree of similarity between the reference respiratory waveform and the respiratory waveform derived from the PPG. Due to the non-availability of a reference respiratory waveform, it was not possible to use the error rate or the correlation coefficient for performance evaluation in this thesis. The "mean absolute error" metric was preferred over the RMSE for performance evaluation since the former has been used in four of the five studies undertaken in the clinical setting (one used true positive/false positive metric). In addition, it is intuitive and relevant in a clinical context to use "mean absolute error" rather than "root mean square error".

2.3.1 Digital Filtering based methods

Four papers were identified that used techniques based on digital filtering to extract respiratory rate from the PPG. Nilsson et al. [98] recorded the PPG from 12 patients (mean age of 66 years) and used a 3rd-order band-pass Butterworth filter with a passband of 0.1 to 0.3 Hz (6 breaths/minute to 18 breaths/minute). The breaths were identified manually in the filtered PPG signal. Nilson et al. [98] reported a sensitivity of 86% and a specificity of 96% for breath detection in comparison to the reference data obtained using end-tidal CO2.

The study by Nakajima et al. [94] included 11 healthy adult subjects (mean age of 26 years) whose PPG was measured at the earlobe. The analogue signal acquired was first preprocessed by using a low-pass Butterworth filter (with a cutoff frequency of 5 Hz) and a high-pass Butterworth filter (with a cutoff frequency of 0.1 Hz). Subsequently, a band-pass FIR filter with a pass-band extending from 1.5 Hz to 2.3 Hz was used to derive the heart rate while three low-pass FIR filters with

cutoff frequencies of 0.3 Hz, 0.4 Hz, and 0.55 Hz (corresponding to 18, 24 and 33 breaths/minute) were used to obtain the respiratory rate. The choice of low-pass filter cutoff was based on the value of the heart rate acquired, with a higher cutoff being chosen for higher heart rates. The heart rate was determined from the filtered signal by counting zero crossings while a peak detection method was employed to derive the estimate of respiratory rate. The results were compared with the reference respiratory rates obtained with transthoracic electrical impedance plethysmography and a maximum error of 7 beats/minute was reported.

The study by Johansson et al. [54] included 6 neonates who were continuously monitored for 8 hours. The PPG was band-pass filtered using a 16^{th} order Bessel filter. The choice of pass-band region was made adatptive by varying the sampling rate off-line. Breaths were also identified in the PPG signal using the zero-crossing method. For 29% of the total monitoring time for which the reference respiratory signal (obtained using electrical impedance plethysmography) was deemed to be of good quality, the respiratory signal obtained from the PPG was reported to miss 2.7% of the reference breaths and to detect 1.5% false positive breaths.

Foo and Wilson [39] analyzed PPG signals recorded from 8 children (mean age of 8.6 years) using zero-phase filtering. A 50th-order non-causal Wiener filter with a bandwidth of 0.1 to 5 Hz was used for this purpose. The peaks in the filtered PPG signal were identified by a peak detector with a moving threshold to minimize poor threshold choices due to baseline fluctuations. The reference respiratory rate was obtained from a calibrated air pressure transducer. A statistical analysis was carried out to compare the mean respiratory intervals derived from the PPG signal and the reference signal. A significant correlation ($r^2 > 0.75$) was reported after analyzing PPG segments not corrupted by motion artefact.

The method reported by Johnston and Mendelson [56] differs slightly from the three studies so far reviewed. The use of a low-pass filter with a cutoff frequency of 0.7 Hz was followed by spectral analysis using Fourier transform, instead of a zero-crossing or peak detection method. The study included four healthy adults with ages between 22 and 24 years. The accuracy of the method was not reported.

2.3.2 Time-Frequency Analysis based methods

Addison and Watson [1, 2] proposed the use of a technique based on the continuous wavelet transform (CWT) for the estimation of respiratory rate. A time-frequency representation of the PPG signal is first obtained by using complex Morlet wavelets. Ideally, this 3-D representation (called a scalogram) is expected to contain two distinct bands (called the ridges), the pulse band (corresponding to the heart rate) and the respiratory band (corresponding to the respiratory rate). The respiratory signal can be derived directly from the respiratory band by following the crest of the ridge in time and plotting the corresponding frequency. The authors argue that even in the presence of noise making it difficult to obtain the respiratory waveform directly from the respiratory band, the coupling between the two bands can be exploited to obtain a respiratory waveform from the pulse band. This is done by ridge-following the 3-D pulse band thereby extracting the amplitude-time signal (called the ridge amplitude perturbation (RAP) signal) and the frequency-time signal (ridge frequency perturbation (RFP) signal). Subsequently, the same wavelet transform technique is applied to these two signals and a respiratory waveform is obtained by ridge-following.

In the work by Leonard et al. [70], the CWT based technique was tested on PPG signals acquired from 22 healthy adults (aged 27 to 82 years) who were asked to breathe at a constant rate of 15 breaths per minute in synchronisation with an on-screen-bar-graph (metronome). A 90-second trace from each subject was manually analyzed using the three respiratory waveforms derived (the primary respiratory band, RAP and RFP) and the most accurate at any time was selected. 98.7% of the total breaths were correctly identified with no false breath identified.

A subsequent study by the same group [71] included 27 children attending an Accident and Emergency Department in a hospital with minor illnesses/injuries. 7 children (28%) were omitted due to distress. PPG recordings from the remaining 18 children (aged from 18 months to 12 years) were analyzed to obtain a respiratory rate using the CWT-based technique.The reference signal was obtained with the help of a trained observer who activated a push button in synchronization with a child's chest movements. These recordings were obtained for a period of 180 seconds after the subject had become comfortable with the probe in position. The recordings were manually observed first to identify sections least corrupted by movement artefact and these were subsequently

analyzed to obtain a respiratory rate estimate. It was not clear how the selection from the three respiratory waveforms (the primary respiratory band, RAP and RFP) was automated to obtain a single esimate. The range of reference respiratory rates was only 17 to 27 breaths per minute. A difference of less than one breath per minute was reported for all 18 children in the study. A very similar protocol was followed in the study by Clifton et al. [24], who recorded the PPG in 13 patients (aged 30 to 77 years) with various respiratory illnesses attending a chest clinic and in 3 healthy volunteers (aged 26 to 40 years). Again, only PPG recordings with minimal movement artefact were analyzed and an error of within 1 breath per minute was reported for all subjects.

In the study by Shelley at al. [125], the short-time Fourier transform (STFT) was used to estimate the respiratory rate from the PPG recorded from 30 patients undergoing general anaesthesia for abdnominal surgery. An 82-second 4096-point moving Hann window was used in the analysis followed by peak detection in a very limited frequency range (corresponding to 5-24 breaths/minute) deemed to be the frequency range of ventilation. A correlation of 0.89 was reported between the respiratory rate extracted from the PPG and the reference rate derived from analysis of the CO_2 waveform extracted from the ventilator.

A recent method for the estimation of time-frequency spectra, variable-frequency complex demodulation (VFCDM) [152], has been proposed for the estimation of respiratory rate from the PPG by Chon et al. [23]. In their study [23], PPG recordings were acquired from 15 healthy adults (mean age of 21 years) who were asked to breathe at a constant frequency for three minutes. Data was collected for the frequency range of 0.2 to 0.6 Hz (12 breaths to 36 breaths per minute) with an increment of 0.1 Hz in both supine and upright position. A reference respiratory rate was obtained using inductive plethysmogrpahy. The results were reported only for the respiratory rate estimation from the RSA component, as the estimate corresponding to AM was found to be unreliable. This method is similar to the CWT-based method and results from each method were compared using the median and the interquartile range of the percentage difference between the estimated and the reference respiratory rate. The VFCDM approach was found to be more accurate and more consistent than the CWT-based method. The mean absolute difference or the RMSE were, however, not reported.

2.3.3 Autoregressive Modelling based methods

AR modelling will be the basis of the signal analysis techniques developed in this thesis. The theoretical background for the techniques and their application to the processing of PPG data are reviewed in detail in Chapter 4. AR modelling was first applied to the problem of extracting respiratory rate information from the PPG by Fleming and Tarassenko [37]. In their original study, 14 five-minute records from seven patients were extracted from the MIMIC database in the Physiobank archive [43]. The PPG recordings, sampled at 125 Hz, were first detrended, filtered (anti-aliasing filter) and downsampled before applying an AR model to the processed signal. Subsequently, a criterion based on pole magnitude was used to find the pole corresponding to the respiratory rate. A downsampling frequency of 1 Hz and an 11^{th}-order AR model were experimentally found to be the optimal combination based on 6 of the 14 records. This meant that there were five poles in the upper half of z-plane corresponding to 0 to 30 breaths/minute, one of which would be the pole corresponding to the respiratory rate. This method was compared to two digital-filtering methods [94, 98] and the CWT-based method [71], all of which it outperformed. A mean absolute error of 0.04 breaths per minute was reported.

In a subsequent study by Fleming [34], data from 36 children (mean age of 9.9 years) were recorded using a nasal airflow probe and two inductance plethysmographic bands (thoracic and abdnominal) to compute the reference respiratory rate. The PPG was recorded using a pulse oximeter probe. The data were recorded before, during and after two to three minutes of vigorous exercise on an exercise bicycle with a typical recording length of 7 minutes. Due to poor signal quality, it was only possible to extract the reference respiratory rate from 19 children (53%) who had a good-quality reference signal throughout their recording. The results were generally poor with short periods of high accuracy and the mean error was consequently not reported. In a later study by Fleming et al. [35], a subset of the data (less than 20% of the total data) from 11 different children (31%) was used and a mean error of 3.4 breaths per minute was reported.

A study in a clinical setting by Fleming [34] included 44 children for whom PPG recordings and a respiratory rate estimated by observing chest wall movement were available. These children were potentially unwell and had attended an out-of-hours GP clinic. The respiratory rate was

separately estimated using the RSA and the AM components in the PPG utilising the AR-modelling approach. A post-processing stage with Kalman filter was also used in an attempt to improve the performance of the algorithms. Poor signal quality (due to motion artefact) was deemed to be the major cause of poor algorithm performance. As a result, an algorithm for measuring the signal quality was subsequently developed. This algorithm, based on probabilistic principal component analysis (PPCA), derived a novelty score for every cycle of the PPG waveform (using a salient point detection algorithm), with a high novelty score indicating a poor quality section. Incorporating signal quality information with Kalman filtering improved the mean absolute error from 11 to 9.2 breaths/minute using AM and from 13.3 to 10.2 breaths/minute using RSA. For the PPCA based algorithm with Kalman filtering, the number of patients for whom the mean absolute error was less than 5 breaths/minute was 13 (30%) using AM and 18 (41%) using RSA.

Lee and Chon [68] reported an improved AR-based respiratory rate estimation algorithm which used the optimal parameter search (OPS) technique for AR-model parameters estimation. Their study included 15 healthy adults (mean age of 21 years) breathing by following a metronome in the range of 0.2 to 0.6 Hz (12 to 36 breaths/minute) with an increment of 0.1 Hz and a further 8 healthy subjects (mean age of 28.4 years) with a natural respiratory rate in the range of 0.7 to 0.8 Hz (42 to 48 breaths/minute). Data were recorded in both upright and supine positions. This method was compared with the widely-used Burg algorithm and model order selection based on the Akaike information criterion (AIC). The OPS method was found to be better than the Burg method over the whole range of breathing rates studied. This conclusion was arrived at by comparing the median and the interquartile range of the percentage difference between the estimated and the reference respiratory rates for the two methods. However, a major problem with this study is that the model order selected for the Burg method (30) is too high. The PPG recordings were detrended, filtered and downsampled to 2 Hz which meant that there were 15 poles in the range of 0 to 60 breaths/minute. With such a high number of poles, the assumption that only the pole lying at a certain angle, θ, determines the shape of the frequency spectrum at $\omega = \theta$ is no longer valid since there will be other poles lying close (see Appendix A for details) leading to pole-splitting (see 4.6 for illustration). In such a case, selecting the pole with the highest magnitude as the pole

whose angle corresponds to the respiratory rate is not appropriate and thus provides a plausible explanation for the comparatively poor performance of the Burg algorithm. Furthermore, the mean absolute error and the RMSE were not reported in the study. A subsequent study by Lee and Chon [67] reported the mean of the RMSE for these 23 subjects using the OPS technique in different breathing rate ranges. The lowest value was 3.8 breaths per minute in the upright position and the highest was 13.0 breaths/minute in the supine position, both in the frequency range corresponding to 42 to 48 breaths/minute.

In the later study, Lee and Chon [67] proposed the use of particle filtering to deal with situations in which an incorrect pole corresponds to the highest magnitude, most likely when the PPG is corrupted by movement artefact. The particle filtering algorithm allows the use of a Bayesian framework for respiratory rate estimation where the particle distribution (representing possible respiratory rates) is sequentially improved when a new set of poles and the associated magnitude becomes available. This study used the data from the 23 subjects in [68] and from a further 10 subjects (mean age of 26.7 years) who were instructed to breathe at 60, 72 and 90 breaths per minute. The proposed algorithm starts with an initial set of particles (one of which represents a potential respiratory rate) with uniform weights. Based on the pole magnitudes and their corresponding angles as determined by the OPS technique, the weights of the particles are reassigned using a likelihood function. Five different likelihood functions were investigated. The new weights are used to resample the particles to generate new particles so that particles with small weights can be eliminated. The choice of initial set of particles and the variance parameters (which determines the spread of the regenerated particles) for the likelihood function affects the accuracy of the method.

The use of particle filtering with the OPS technique was reported to improve the performance of the respiratory rate estimation algorithm. The lowest mean RMSE, 1.3 breaths/minute, was achieved in the frequency range of 42 to 48 breaths/minute in the upright position. For the same likelihood function, however, the mean RMSE was 7.7 breaths/minute in the same frequency range in the supine position. A further study by Lee and Chon [66] incorporated a time-varying AR model with particle filtering to account for the cases where the subject's respiratory rate is not constant and varies either slowly or abruptly. This study was undertaken using PPG recordings from 16

healthy adult subjects (mean age of 21 years) who were instructed to breathe at four rates: 24, 18, 12, and 6 breaths/minute with abrupt changes from one rate to the next.

2.3.4 Other methods

Apart from digital filtering, AR-modelling and time-frequency analysis based approaches, there have been attempts to use neural networks [53], and independent component analysis (ICA) [157, 80] to estimate respiratory rate from the PPG.

The study by Johansson [53] included 15 healthy subjects with 10 minutes of spontaneous breathing without instructions and a further 10 minutes of metronome-controlled breathing in the range of 6-18 breaths/minute. The measurements were taken with the subjects in the supine position after allowing them a resting period of 5-10 minutes. The analysis was subsequently carried out by manually selecting 12 minutes of data for each patient (60% of total duration) that were deemed to be free from movement artefact. Five features were extracted from the PPG waveform: peak value, trough value, instantaneous heart rate, baseline wander and pulse amplitude. A neural network of 25:5:2 topology was employed, the five features from five consectuive heartbeats resulting in 25 inputs, a 5 unit hidden layer and the two outputs encoded to indicate whether the inputs followed inspiration, expiration or neither (00, 01 and 10). The reference respiratory rate was obtained using an airflow humidity sensor positioned adjacent to the nasal/oral airflow. Analysis was also carried out to extract the respiratory rate from the five features separately. The results were reported in terms of false positive and false negative breath detection. The error rates (sum of the two errors) ranged from 9.7% to 14.5% for the separate analysis of the five features, and the corresponding values for the combination of the five features with the neural network ranged from 9.5% to 9.6% thus suggesting a marginal improvement by fusing the different features.

The study by Zhou et al. [157] proposed the use of ICA to extract the respiratory rate information from the PPG signal. Their study was based on PPG recordings from 10 healthy adults (aged from 27-40 years) in the supine position. The subjects went through both spontaneous breathing and apnoea simulation during the study. Sound recorded from a microphone placed beneath the mouth was processed to derive the reference respiratory rate. While the results suggested a corre-

lation between the reference respiratory rate and the PPG, high-frequency noise was found to be present in the respiratory waveform extracted from the PPG using ICA. The use of a low-pass filter was suggested to remove the high frequency noise. Zhou et al. [157] did not report any errors in their paper.

An ICA-based technique was also described by Madhav et al. [80]. They analysed PPG recordings from 10 healthy volunteers (aged fom 22-35 years) from which the motion artefact component was extracted using ICA. 8 clean records from the MIMIC database [43] were also processed and the motion artifacts from the recorded signals were linearly mixed with these 8 records. Statistical measures were subsequently reported to suggest that a strong correlation existed between the extracted respiratory rate from the synthesized PPG recordings and the reference respiratory rate derived, most likely, from impedance plethysmography.

2.3.5 Summary of methods

Consderable overlap exists between the different approaches described in the previous section. In addition, the techniques for extracting respiratory rate estimation using RSA are similar to those that use AM, although the former requires additional pre-processing. Figure 2.3 provides a pictorial summary of the signal processing algorithms reviewed in the previous section.

2.4 Limitations of current methods

The majority of the studies discussed above have been undertaken in a controlled setting using PPG recordings from healthy subjects, in many cases using metronome controlled-breathing rates. Even under these conditions, motion artefact is the main cause of poor performance. In many studies, manual selection of clean sections of PPG recordings was undertaken before assessing the accuracy of the signal processing algorithms. Such selection procedures are not applicable in a clinical setting where automated respiratory rate estimation using PPG signal is needed.

The performance of the algorithms reviewed in this chapter deteriorates significantly when the studies were undertaken in a clinical setting. There were five such studies. The study by

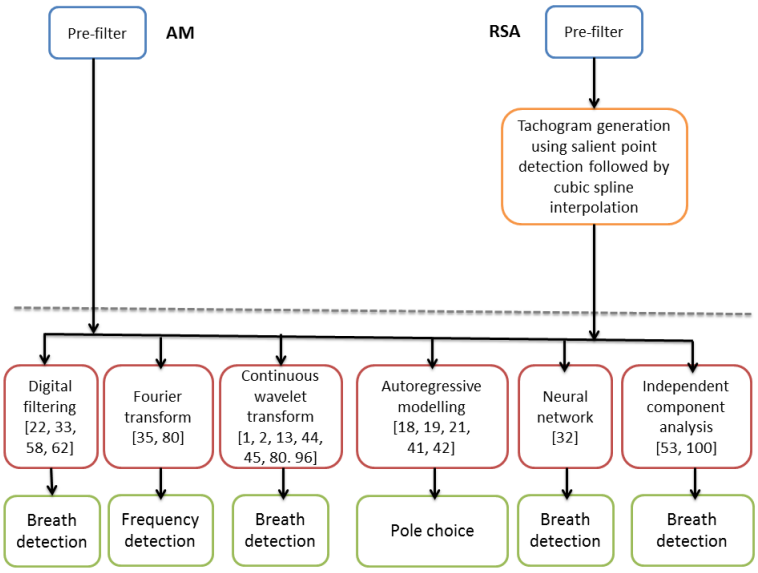


Figure 2.3: Summary of signal processing algorithms for the estimation of respiratory rate from PPG using AM and RSA. The dashed line separates the pre-processing steps and the respiratory waveform analysis using various methods.

Johansson et al. [54] was able to identify only 29% of good quality data (based on reference signals acquired with electrical impedance plethysmography). The CWT technique developed by Addison and Watson [2, 1] was evaluated on PPG recordings from only 18 children (72% of total consented) [71] recruited in an Accident and Emergency Department. The PPG recordings were manually reviewed to select for analysis those sections with least movement artefact. The same limitations exist in another study [24] utilising the CWT-based technique. Finally, the range of reference respiratory rates was very small (17 to 27 breaths/minute) in [71], given that the data were recorded from children attending an Emergency Department.

The AR-modelling based technique described in [37] has reported a mean absolute error of only 0.04 breaths/minute when used to analyse the MIMIC data from patients monitored in ICU. However, only a subset of the database was manually selected for the analysis. The same AR-modelling based technique (with appropriate modifications to account for a different population) applied to PPG recordings from children in an out-of-hours GP clinic [34] resulted in a mean absolute error of 11 breaths/minute. Instead of manual selection of clean sections for analysis, an automated signal quality algorithm, developed in [34], reduced the mean absolute error to 9.2 breaths/minute. This error is however, still comparatively much higher than the errors reported in

studies using PPG recordings in a controlled setting and/or with manual selection of clean sections of PPG.

2.5 Conclusion

This chapter has shed light on the physiological basis for estimating respiratory rate information from the PPG. A literature review of the various signal processing algorithms developed for the estimation of respiratory rate was presented. The majority of the studies were undertaken using PPG recordings from healthy subjects in a controlled setting. Motion artefact was identified as the major reason for the poor performance of the processing algorithms. There was a significant deterioration in performance when algorithms were evaluated on data acquired from subjects in clinical settings without manual selection of high-quality PPG data.

In the context of paediatric triage, it will be important to acquire PPG recordings from a large number of children, including unwell children, from all paediatric age groups. The time available during triage is limited and manual selection of good-quality signals is not an option. The next chapter describes the clinical study that was undertaken to provide the PPG data appropriate for developing signal processing algorithms capable of reliably estimating respiratory rate in unwell children presenting in the Emergency Department.

Chapter 3

The Paediatric Triage Study

3.1 Introduction

The photoplethysmogram (PPG) is routinely recorded using pulse oximetry (see section 1.2) in various clinical settings to measure pulse rate and peripheral arterial oxygen saturation. As discussed in section 2.3, previous studies suggest that it may be possible to obtain respiratory rate from the PPG from children. However, only the studies by [71] and [34] have investigated the measurement of respiratory rate in children in a clinical setting. The study by Leonard et al. [71] was also undertaken in an Accident and Emergency Department, but it was limited to only 18 children, and the reference respiratory rates varied over a very narrow range, from 17 breaths per minute to 27 breaths per minute. The study undertaken in the Paediatric Emergency Department at the John Radcliffe Hospital, Oxford, UK deliberately included children from different age groups (see next section) who were likely to have a wide range of respiratory rates.

The overall aim of the study was to develop and test the effectiveness of algorithms to process the PPG signals recorded from the pulse oximetry finger probe and derive an estimate of the respiratory rate in these children. The finger probe also measures the pulse rate and the peripheral arterial oxygen saturation of these children.

The study was carried out between January and August 2011. The protocol for the study was approved by the South Birmingham NHS Research Ethics Committee. PPG recordings as well as temperature were recorded from children who presented to the Paediatric Emergency Department

using a monitoring system that was specifically developed for this study (see section 3.3 below).

3.2 The Study Protocol

The original aim was to collect data from about 100 patients from each of four age groups, with each age group chosen to be the same as in the APLS[1] guidelines[2] [79]: 0-1 years, 1-2 years, 2-5 years, and 5-12 years. Children with physical trauma and those who were clinically very unwell were excluded as it was felt that their involvement in the research study might cause an unnecessary delay in their receiving appropriate clinical care.

Upon arrival in the Emergency Department (ED), the vital signs were measured by the paediatric nurses who had previously been trained to use the monitoring equipment. This involved attaching the Nonin pulse oximeter (see section 3.3), taking the child's temperature and measuring the respiratory rate by counting the number of chest wall movements for either 15 seconds or 30 seconds. This information was also recorded in the child's medical record for the clinicians providing care to the child. The compliance of the child during monitoring was also recorded. This was done according to a qualitative assessment by the nurse, with four levels of compliance: compliant, intermittent movement, frequent movement and non-compliant.

Consent from the parents was usually obtained by the research nurse after the child had been seen by a clinician. A participant information sheet was provided which explained the purpose and procedure of the study for the parents, which they read before making a decision. The study ID assigned by the monitoring system was recorded on the consent form so that only data collected from children whose parents consented was used for research purposes.

[1] Advanced Paediatric Life Support is a program created by the American Academy of Paediatrics and the American College of Emergency Physicians.

[2] These guidelines are used in emergency and paediatric departments by healthcare providers including nurses and physicians.

(a) Nonin Probe

(b) Tablet PC

Figure 3.1: Hardware used in the Paediatric Triage Study

3.3 Hardware and Software

Figure 3.1 shows the hardware used in the clinical study. The Nonin 4100 pulse oximeter (Nonin Medical, Plymouth, MN, USA) measured pulse rate, peripheral arterial oxygen saturation and peripheral perfusion and transmited the data using a wireless Bluetooth radio link. A tablet PC was used to record and store this data. The software running on the tablet PC also allowed the nurse to enter the date of birth of the patient, the temperature recorded using a standard axillary temperature probe and the respiratory rate estimated by counting the number of chest wall movements for either 15 or 30 seconds.

To allow for use during busy times, two sets of equipment were used. One set of equipment (the Nonin 4100 pulse oximeter and a tablet PC, supplied by OBS Medical Ltd, Abingdon, Oxon) was placed in the standard triage room. The second set of equipment (the Nonin 4100 pulse oximeter and a tablet PC, Toughbook CF-X, Panasonic) was portable and therefore allowed monitoring of children outside the triage room including the waiting room.

The software automatically assigned a sequential identification number for each child. The same identification number was recorded on the consent forms that were subsequently used to seek consent from parents for the use of the child's data for research purposes. This method ensured that anonymity was maintained while verifying that consent had been obtained before using any data

for analysis.

The software for this study was developed in Visual Studio using an existing code base provided by OBS Medical Ltd., Abingdon, Oxon, UK. The structure of the software was based on a Client-Server model as illustrated in figure 3.2. All the communication with the Nonin probe was handled by the client that received the PPG signal, as well as the values of oxygen saturation, pulse rate and peripheral perfusion. This information was then sent by the "client" to the "server". Subsequently, this information was sent to the graphical user interface (GUI) that displayed it on the tablet touch screen. The "client" and the "server" were both designed in a Visual Studio environment with the code for the "client" written in the C# language while the code for the "server" was written using a combination of the C# and C languages.

The Nonin probe was configured to send data at 75 frames every second, with each frame consisting of 5 bytes (Nonin, 2010). Two bytes in every frame were used to store the PPG signal. Thus the PPG signal had 16-bit accuracy and its sampling frequency was 75 Hz. The Nonin probe uses two channels, the red channel at 660 nanometers and the infrared channel at 910 nanometers. The PPG signal is acquired from the infrared channel. The DC offset is removed from the signal. Subsequently various prefilters are used to remove undesirable signal components including an averaging filter to eliminate noise artefacts in low perfusion and center the waveform such that the peak values are not clipped. However, there is no detailed information available from the manufacturer regarding the various processing stages that the output from the two photodetectors go through before being output from the oximeter.

The noise in a signal depends on many factors like perfusion, excessive ambient light, high intensity and high ambient light sources, motion, and compromised connections. An attempt was made to esimate the signal to noise ratio of the PPG signal acquired from the Nonin pulse oximeter. One approach to estimate the signal to noise ratio is to equate it to the ratio of mean to the standard deviation for parts of the signal where the PPG is expected to be constant (e.g. when the finger is removed from the probe). However, it was not possible to identify such segments in the PPG signal acquired because the Nonin probe is configured to set an "Out of Track" flag as soon as a finger is removed from the probe and thus no PPG signal is subsequently sent via the bluetooth connection.

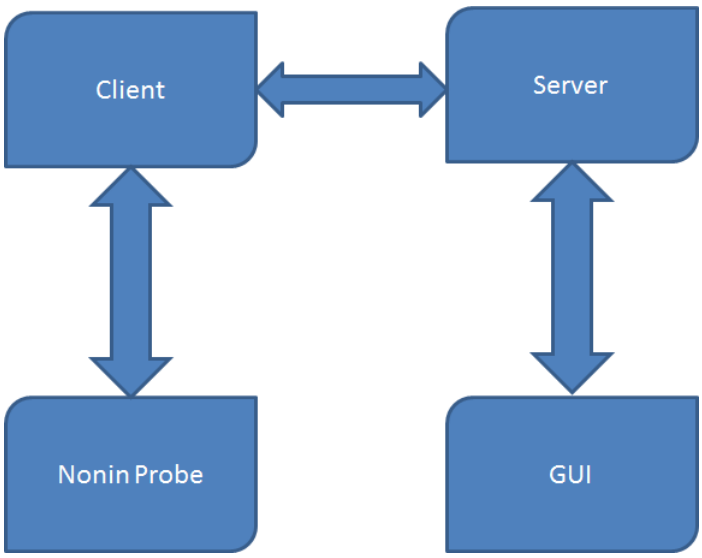


Figure 3.2: Structure of the software used for the Paediatric Triage Study

Upon contacting Nonin, the noise was found to be about 25 counts (up to 5 bits) in low perfusion case (the noise would be lower in high perfusion cases).

The range occupied by the PPG signal was found by finding the minimum and maximum PPG magnitude for all patients and estimating the bits required to represent that range. It was found to be 14.6 (mean number of bits required) implying that up to 15 bits are used to represent the PPG signal output. This means that typically we get a PPG signal where about 10 to 11 bits represent the PPG signal while about 4 to 5 bits are noise.

The main screen of the user interface is shown in figure 3.3. Recording data for a new patient was triggered by pressing the "New Patient" button which then prompted the nurse to enter the child's date of birth. As soon as valid data was received from the pulse oximeter, a countdown timer started to indicate to the nurse how much of the two-minute recording period had elapsed. Two minutes were deemed to be a sufficiently long time to obtain a PPG recording from which respiratory rate could be derived without introducing excessive delay into the triage practice.

The GUI on the tablet PC screen displayed the sequential ID that was assigned automatically after clicking on "New Patient". This allowed the nurse to record the same ID on the consent form while the child was being monitored. The PPG signal was also displayed on the screen so that any occurrence of the oximeter probe not being properly attached to the child's finger could easily be identified. There was provision to enter respiratory rate estimates and temperature recordings

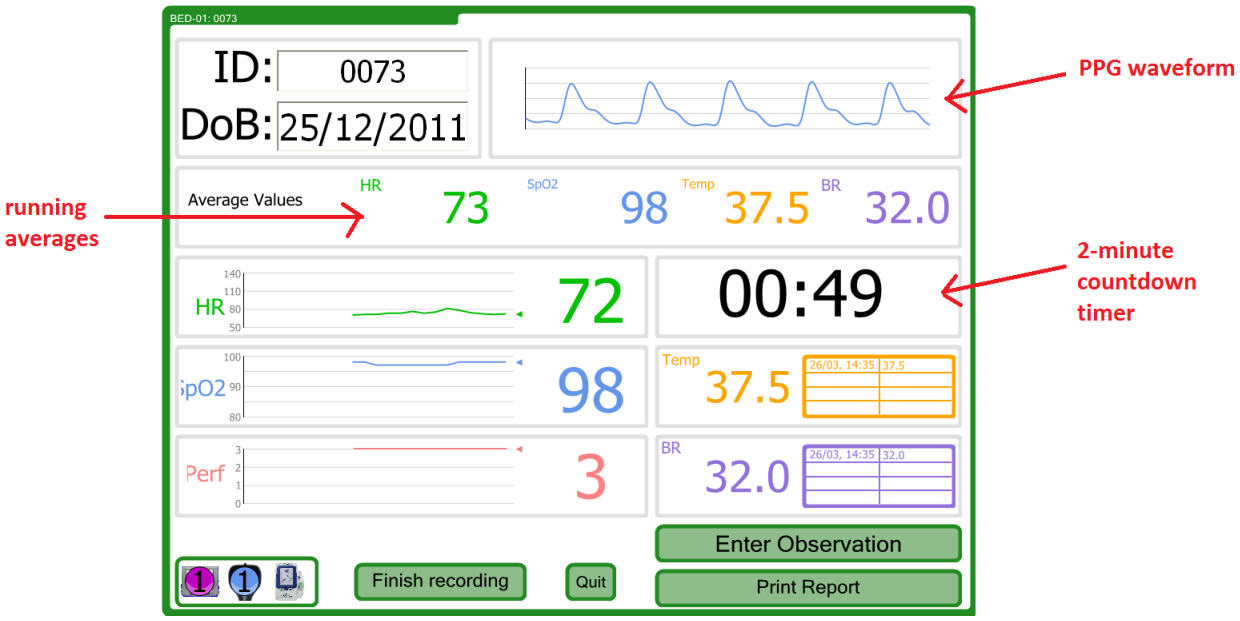


Figure 3.3: Screenshot of the GUI used in the Paediatric Triage Study. "Perf" gives a measure of peripheral perfusion i.e. the extent of supply of blood to the peripheral tissues such as the skin and limbs. This measure is automatically provided by the Nonin probe but it is unlikely to be useful due to the short monitoring period. [34]

as many times as the nurse wished to do so during the recording. The pulse rate, and oxygen saturation averaged over the most recent 4 beats (standard setting on Nonin probe oximeters) were displayed on the screen. The peripheral perfusion index updated every beat was also displayed. The screen also displayed the running mean of the pulse rate and the oxygen saturation using the 4-beat averages of heart rate and oxygen saturation provided by the Nonin probe oximeter. values of that were automatically updated every beat. At the end of the monitoring, the temperature recordings and estimates of respiratory rate entered manually by the nurse, were averaged and displayed on the screen. The average values of these vital signs displayed on the screen allowed the nurse to enter these into the child's medical record.

If the Nonin oximeter was not attached to a child's finger for ten minutes, it turned off automatically to conserve battery power. The GUI indicated to the nurse if no valid data was being received (by flashing an "Out of Track" icon on the screen) and showed a 9-minute countdown timer after one minute of not receiving any valid data.

All the vital signs, along with the patient ID, date of birth, and the 16 bit-PPG waveform were stored in text files on the hard drive of the tablet PCs. The filenames included the patient's study

Table 3.1: Summary of Paediatric Data collected classified according to the duration of PPG recordings

Duration of PPG	Number of Patients
duration > 8 minutes	3 (1%)
4 minutes <= duration < 8 minutes	19 (7%)
3 minutes<= duration < 4 minutes	31 (11%)
1 minute <= duration < 3 minutes	197 (72%)
30 seconds <= duration < 1 minute	12 (4%)
duration < 30 seconds	10 (4%)
Total	272

ID, thereby ensuring that only the data of the consented patients was accessed.

3.4 Paediatric Triage Study Database

The total number of children who were recruited in the Paediatric Triage study from January 2011 till August 2011 was 272. Classification according to the duration of PPG recorded is reported in Table 3.1.

The data from children who had PPG recordings of at least one minute were selected to develop the respiratory rate algorithm. This reduced the total number of patients available for analysis to 250. Out of these, one child's reference respiratory rate was not recorded in the database and therefore the number of children selected for respiratory rate algorithm development was 249 out of a total of 272 patients (92%).

The distribution according to age and reference respiratory rate for these 249 children is summarised in table 3.2 and table 3.3 respectively.

Table 3.2: Distribution of the 249 patients selected according to age

Age Range (years)	Number of Patients	Males	Females
0 to 1 years	59	31	28
1 to 2 years	40	21	19
2 to 5 years	69	37	32
5 to 12 years	81	53	28
Total	249	142	107

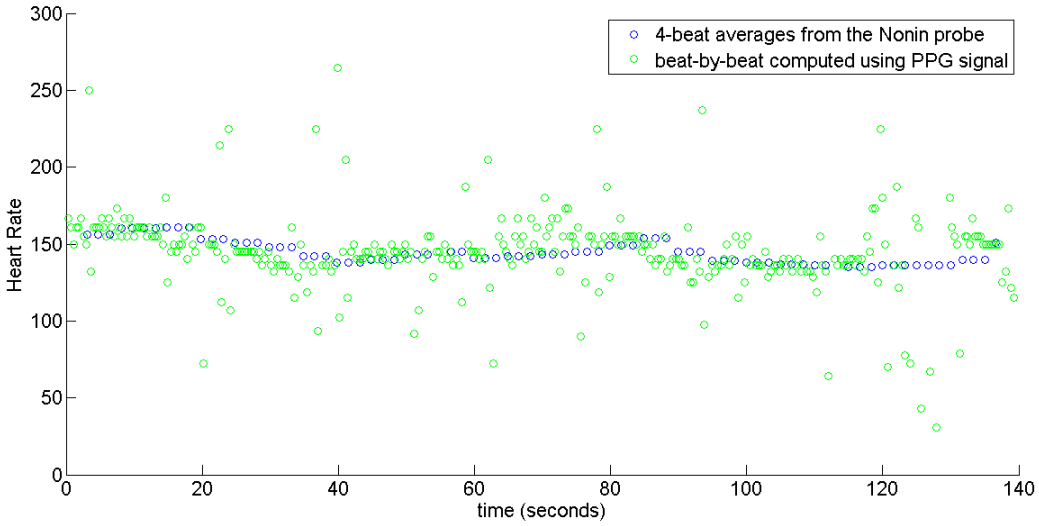
Table 3.3: Distribution of the 249 patients selected according to reference respiratory rate

Reference Respiratory Rate (RRR, breaths per minute)	Number of Patients	Males	Females
RRR >= 20	31	17	14
20 < RRR <=30	90	54	36
30 < RRR <= 40	73	40	33
40 < RRR <= 60	48	30	18
RRR > 60	7	1	6
Total	249	142	107

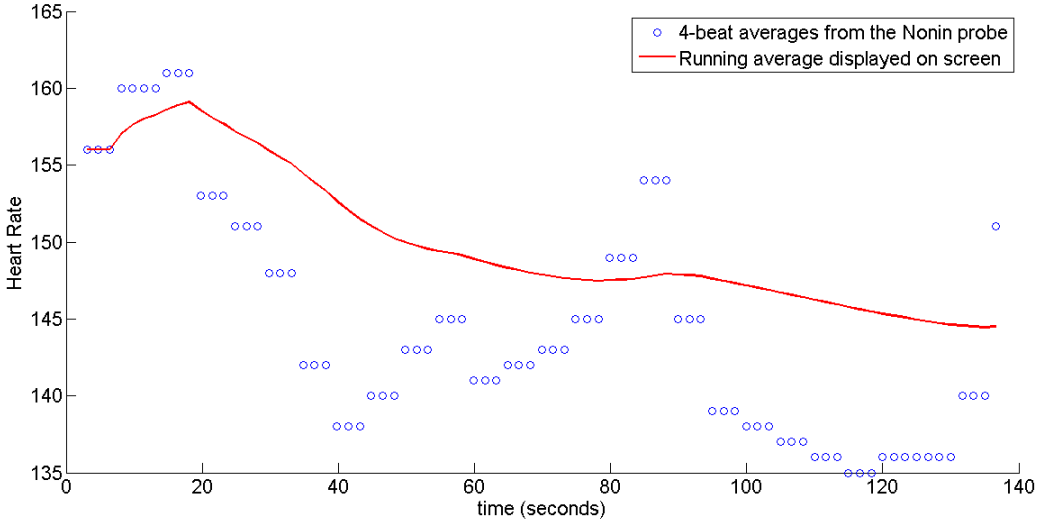
3.5 Heart rate recorded by the triage nurses

Figure 3.4(a) shows the 4-beat average heart rates in blue calculated by the Nonin pulse oximeter and updated every beat, for a patient monitored for about 140 seconds. The beat-by-beat heart rates determined using the interbeat interval in the PPG signal using a beat onset detection algorithm (see section 6.3 for details) are shown in green. The Nonin pulse oximeters use internal signal processing algorithms to capture the 4-beat averages. The running mean using values of the 4-beat heart rate averages was also captured, as shown in figure 3.4 (b). It can be seen that there is a natural physiological variability during the monitoring period which can lead to an error depending on when the nurse chooses to record the heart rate in the child's medical record from observation of the display on the monitoring equipment.

To quantify this error, the running mean heart rate closest and furthest to the heart rate recorded by the triage nurse in the child's medical record were found for every patient in the database. These comparisons are shown in the scatter plots in figure 3.5. The root mean absolute difference in the two cases is 2.5 beats per minute and 9.2 beats per minute respectively. This error reflects the inaccuracy of the method employed during the triage process when the nurses record a single heart rate.

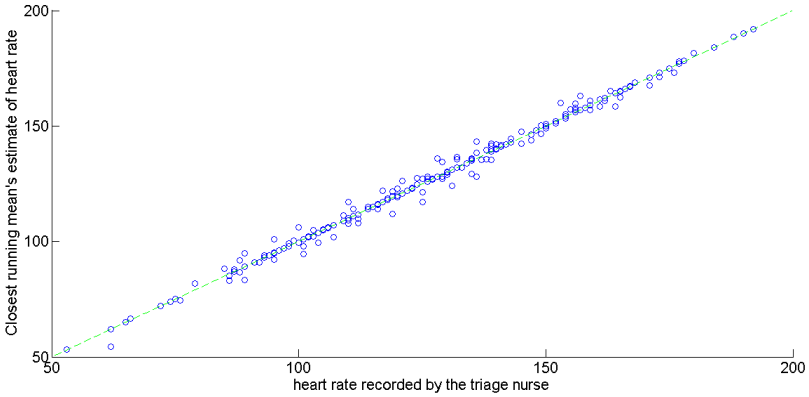


(a) beat-by-beat heart rates and the 4-beat averages (blue) of heart rates for a single patient

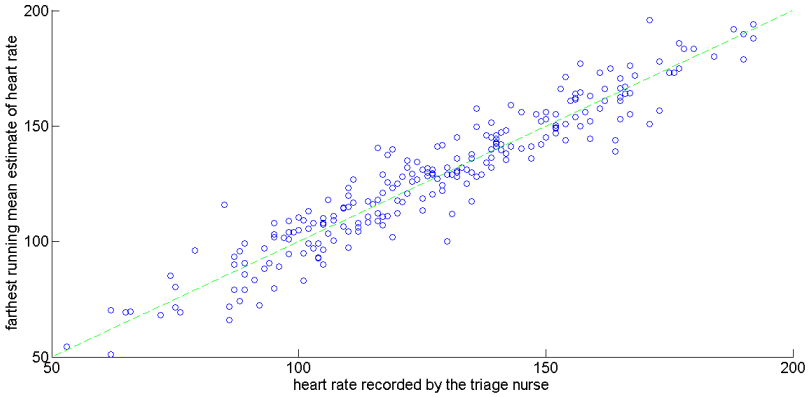


(b) 4-beat averages (blue) and the running mean (red) of heart rate for a single patient

Figure 3.4: Heart rate variability for a single patient over a 130-second duration (Patient 0321)



(a) the running mean of the 4-beat heart rate averages closest to the value recorded by the nurse in the child's medical record



(b) the running mean of the 4-beat heart rate averages furthest from the value recorded by the nurse in the child's medical record

Figure 3.5: Scatter plot of heart rate recorded by the triage nurse versus the heart rate determined from the running mean of 4-beat heart rate averages over the monitoring period during triage

3.6 Respiratory rates estimated by the triage nurses

As with heart rate, it is likely that the natural variations in respiratory rate in a 2-minute period will also contribute to errors when a single respiratory rate is recorded by the triage nurses. Section 6.11 will quantify this once respiratory rate algorithms have been developed. In addition, there will also be an error in the respiratory rate estimate recorded by the triage nurses due to the way that these are estimated.

Respiratory rate is estimated by the triage nurse by counting chest wall movements for either 15 seconds or 30 seconds. In the former case, an error of 1 breath in 15 seconds will lead to an error of 4 breaths in one minute and in the latter case, an error of 1 breath in 30 seconds will lead to an error of 2 breaths in one minute.

A study of 97 children (under 5 years of age) by Simoes et al. [128] compared respiratory rate estimation by comparing the quantification of chest wall movement with the automated method of electrical impedance pneumography. Not surprisingly, they found that the estimate becomes more accurate as the counting period is increased from 30 seconds to 60 seconds. They also found that that the standard deviation of the difference between the respiratory rate estimate from counting chest wall movement with respect to the electrical impedance pneumograph was 6.85 breaths per minute if a counting period of 60 seconds was used[128]. The error increased to 8.60 breaths per minute with a counting period of 30 seconds [128]. The counting period of 15 seconds was not investigated but it is logical to assume that the error will increase further with a counting period of 15 seconds. In comparison to the paediatric clinical study, the study by Simoes et al. [128] was more controlled as the timings of the two methods of respiratory rate estimation was matched, artefactual signal segments (recorded with impedance pneumography) were manually removed and the exclusion criteria included presence of any chronic illness, current illness greater than two weeks and children who were inconsolably agitated.

Figure 3.6 shows the scatter plot of respiration rates and pulse rates estimated by the triage nurses. It can be seen from the figure that there are gaps of 2 breaths per minute or 4 breaths per minute between consecutive vertical lines reflecting the quantization error in the determination of the respiratory rate in breaths per minute by the triage nurse. Moreover, the vertical lines that are

located at multiples of 4 breaths per minute can be seen to be more dense than the vertical lines located at the remaining even numbered respiratory rates (e.g. compare the vertical lines at 24, 28 or 32 breaths per minute with those at 26, 30 or 34 breaths per minute). This suggests that, on a significant number of occasions, the nurses counted chest wall movement for 15 seconds and then multiplied the result by 4 to determine the respiratory rate. This is further corroborated by the fact that there were 260 patients (96%) in the database with an even value for the reference respiratory rate. Out of these 260 patients, only 88 patients (34%) had a reference value that was not divisible by 4. If the nurses had been counting the chest wall movement for 30 seconds, about 50% patients with a reference respiratory rate that was not divisible by 4 would have been expected.

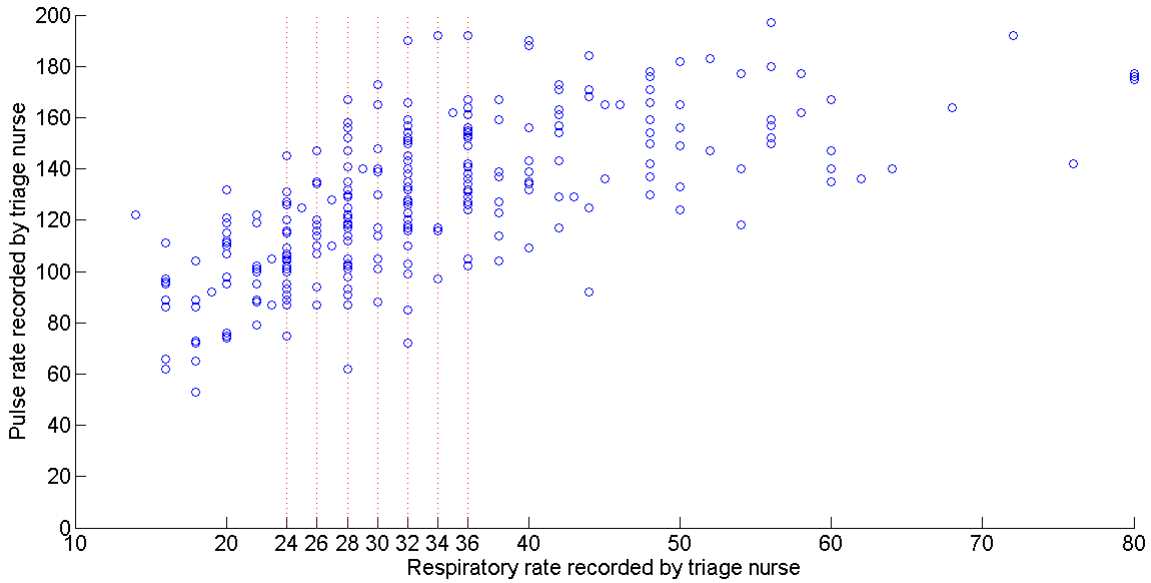


Figure 3.6: Scatter plot of the respiratory rate recorded by the triage nurse with the heart rate recorded by the triage nurse

3.7 Summary

This chapter started by explaining the rationale for the paediatric triage study and the protocol that was followed during the 8-month period of the study. The hardware acquired and the structure of the software designed for the study were described. A summary of the data collected categorised according to the monitoring period, age group, gender and the reference respiratory rate (i.e. the value estimated by the triage nurse) was provided. The effect of recording a single value of heart

rate in the medical record when the patient's heart rate varies during the recording period was investigated. It was also argued that the single respiratory rate estimate from the triage nurse will not only contain errors due to the variation in respiratory rate during monitoring but also because chest wall movements are counted for a period of, at best, 30 seconds only and in a significant number of instances, of 15 seconds only.

Chapter 4

Respiratory Rate Estimation with a Finger Probe during Paediatric Triage (Amplitude Modulation)

4.1 Introduction

It is not desirable to use existing electronic methods (see section 1.2) for measuring respiratory rate in the context of paediatric triage as it is likely to over-burden the limited human resources available, and it is also be difficult, if not unethical, to apply equipment that may cause distress to the unwell child. Accurate determination of respiratory rate from the PPG signal would be a major improvement over current paediatric triage systems, which invariably rely on the counting of chest wall movement for 15 or 30 seconds. It would also allow nurses to focus on other aspects of the triage process, thereby freeing up finite medical resources.

This chapter presents the signal processing algorithms that were developed for estimating the respiratory rate from the PPG data that were collected from children during triage in the emergency department at the John Radcliffe hospital (see Chapter 3). Section 4.2 describes the testing procedure that was used for this investigation. This chapter will present the development and implementation of the algorithms based on amplitude modulation (AM) of the PPG (see section 2.2).

The next chapter will then describe the performance of these algorithms using the Respiratory Sinus Arrhythmia (RSA) which is present in the PPG (see section 2.2).

4.2 Testing procedure

The testing procedure was designed to evaluate the performance of algorithms for respiratory rate estimation. The child's respiratory rate may vary during the monitoring period (see section 6.11) but this variation is not captured with the manual method of determining the respiratory rate. Metrics were designed to take account of this possible variation in respiratory rate (see section 4.2.2 below).

4.2.1 Windowing of data

A typical PPG recording during the study lasted for about 120 seconds. Since the aim eventually is to develop a real-time respiratory rate algorithm in the context of paediatric triage, a 60-second window was chosen for initial investigations. Longer windows will generally produce more accurate results since the effects of artefact and poor signal quality will be mitigated as a longer-duration signal will contain more respiratory cycles. However, the increased accuracy comes at the expense of an additional delay.

In this work, the performance of the algorithms when using 60-second windows or 30-second windows was evaluated. The separation between consecutive windows was chosen to be 10 seconds to provide a large amount of overlap between consecutive windows. A separation of 10 seconds between two consecutive 60-second windows will lead to an overlap of 50 seconds and provide six estimates of respiratory rate in a 60-second period.

Choosing a large overlap between consecutive windows also mitigate large edge effects (rapid transitions present at the edges). Another solution to mitigate such effects is by using a suitable window function in order to attenuate the edges (see section 6.7.1.2 for more details).

4.2.2 Error metrics

The reference respiratory rate was the respiratory rate estimated by the triage nurse by counting the chest wall movement for 15 seconds (if the respiratory rate was high) or 30 seconds. Although the reference rate is prone to error, this was the only available standard (currently used in practice) which was available and feasible to acquire in the clinical study without introducing any new hardware or affecting the clinical workflow. The mean absolute difference between the respiratory rate estimated by processing the PPG waveform and the reference respiratory rate was used to quantify the performance of the algorithms (see section 2.3). Due to intra- and inter-observer variation, the reference respiratory rate is unlikely to be accurate; however, it is the best information available on the child's respiratory rate.

At the moment, there is no evidence to suggest the level of accuracy of respiratory rate that is needed in effective triage. Based on expert opinion, an error above 5 breaths/minute was considered to be clinically significant. Various error metrics were therefore developed to compare and evaluate the various respiratory rate estimation algorithms developed as part of the research described in this thesis. This includes the number of patients for whom the estimated respiratory rate is within 5 or 10 breaths per minute of the reference, and the number of patients for whom the estimated respiratory rate for at least one 60-second window is within 5 or 10 breaths per minute of the reference value. The latter metric is important in the cases for which the quality of the PPG signal varies during the PPG recording and also those for which the child's respiratory rate varies significantly during the monitoring period.

4.2.3 Overview of methods

Figure 4.1 shows the components of the respiratory rate estimation algorithms developed: a PPG signal from the pulse oximeter is first pre-processed to extract the "respiratory signal" (shown within the blue rectangle) followed by the extraction of respiratory rate from the "respiratory signal" (shown within the green rectangle).

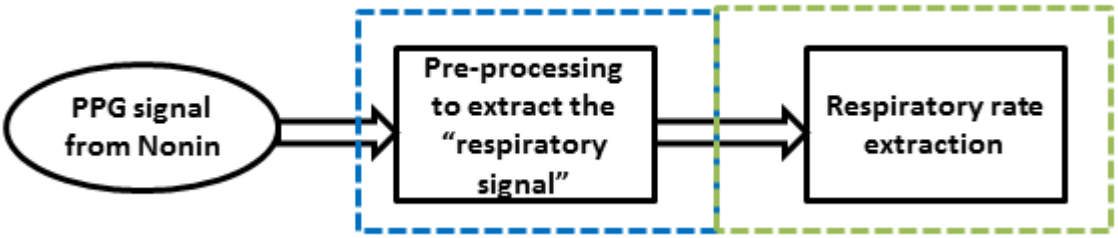


Figure 4.1: Stages of respiratory rate estimation algorithm

Since little work has previously been done on developing and validating a respiratory rate estimation algorithm for children, it was decided to start with a simple time-domain method (a peak detection algorithm) applied to the PPG waveform before eventually investigating more complex algorithms. Firstly, the PPG signal is preprocessed using a filter and a basic peak detection algorithm (see Section 4.3) to extract the "respiratory signal". However, due to erroneous results in the presence of dicrotic notches, a Complex peak detection algorithm is developed to extract the "respiratory signal" (see Section 4.4). Subsequently, the respiratory rates can be estimated using the basic peak detection algorithm to find out the instantaneous respiratory rates. The afore-mentioned stages are shown in figure 4.2.

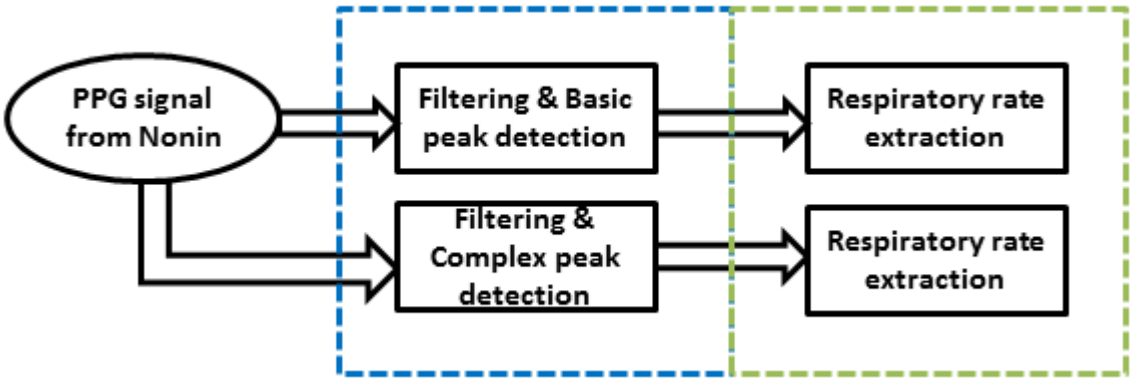


Figure 4.2: Basic peak and Complex Peak detection algorithms for respiratory rate estimation

Instead of peak detection on the "respiratory signal" however, a frequency domain based method can also be used to find the spectrum of the "respiratory signal" and esimate the respiratory rate. Thus, various methods based on Autoregressive modelling (AR), which has previously been applied to a number of physiological signals, are developed to esimate the respiratory rate from the

respiratory signal. To fulfill the requirements to have an evenly sampled time-series and to avoid aliasing for the application of an AR modelling based approach, the PPG signal is preprocessed using a filter and downsampling to extract the "respiratory signal" instead of peak detection. This is illustrated in figure 4.3 (see Sections 4.5-4.6 for details).

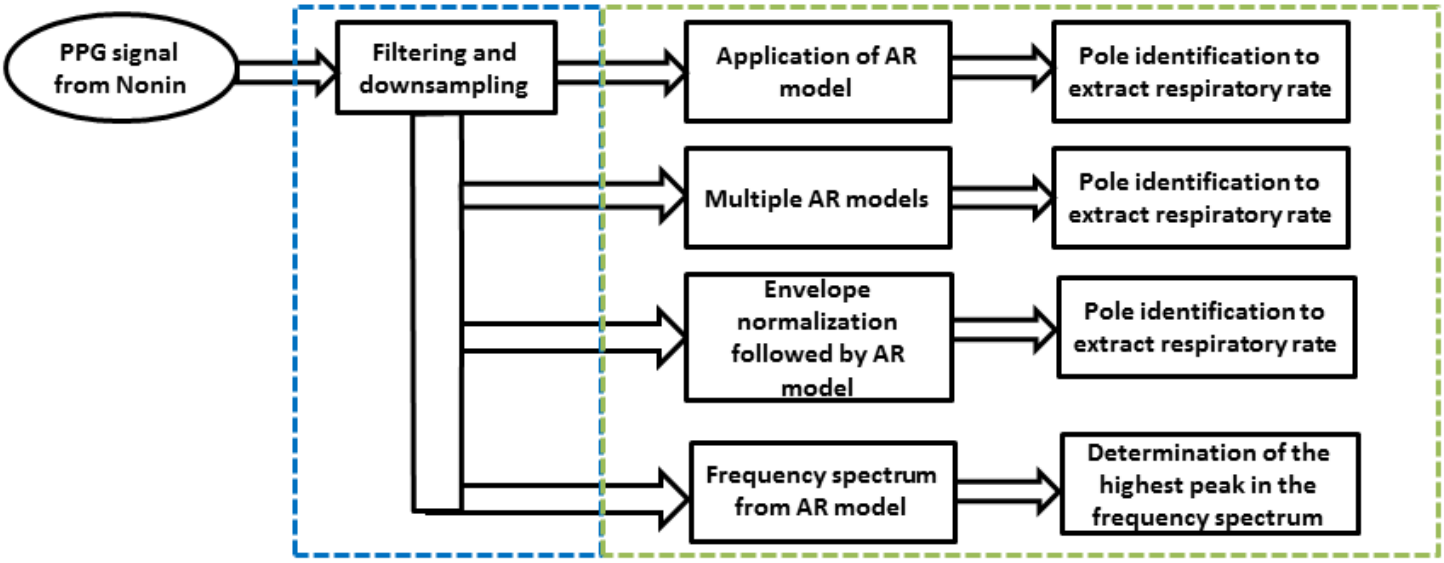


Figure 4.3: AR model based algorithms for respiratory rate estimation

4.3 Basic Peak Detection Algorithm

Figure 4.4 shows a10-second PPG segment to illustrate how the amplitude of the PPG is modulated by respiratory rate. The PPG signal was first detrended and low-pass filtered using the 11-point parabolic filter whose characteristics are shown in figure 4.5. Subsequently, a 3-point peak detection algorithm was applied to the PPG signal in a 60-second or 30-second window. The peak detection algorithm is:

$$if\ (x(n) > x(n-1))\ \&\ (x(n) > x(n+1))$$

$$\text{or } if\ (x(n) > x(n-1))\ \&\ (x(n) = x(n+1))\&\ (x(n) > x(n+2))$$

$$\text{assign peak at } n \text{ where the magnitude of the PPG signal is } x(n)$$

The peaks identified in the original PPG signal are used to obtain the envelope of the PPG signal by linear interpolation, as illustrated in figure 4.4. The same basic 3-point peak detection algorithm is then applied to the envelope to find the locations of its peaks. The time interval, T,

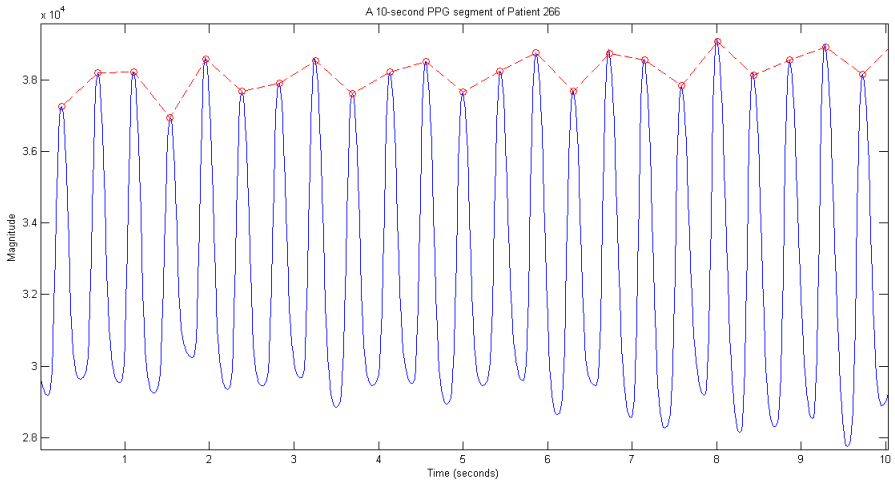


Figure 4.4: Illustration of the basic peak detection algorithm with a 10-second PPG segment with peaks identified and linearly interpolated to form an envelope

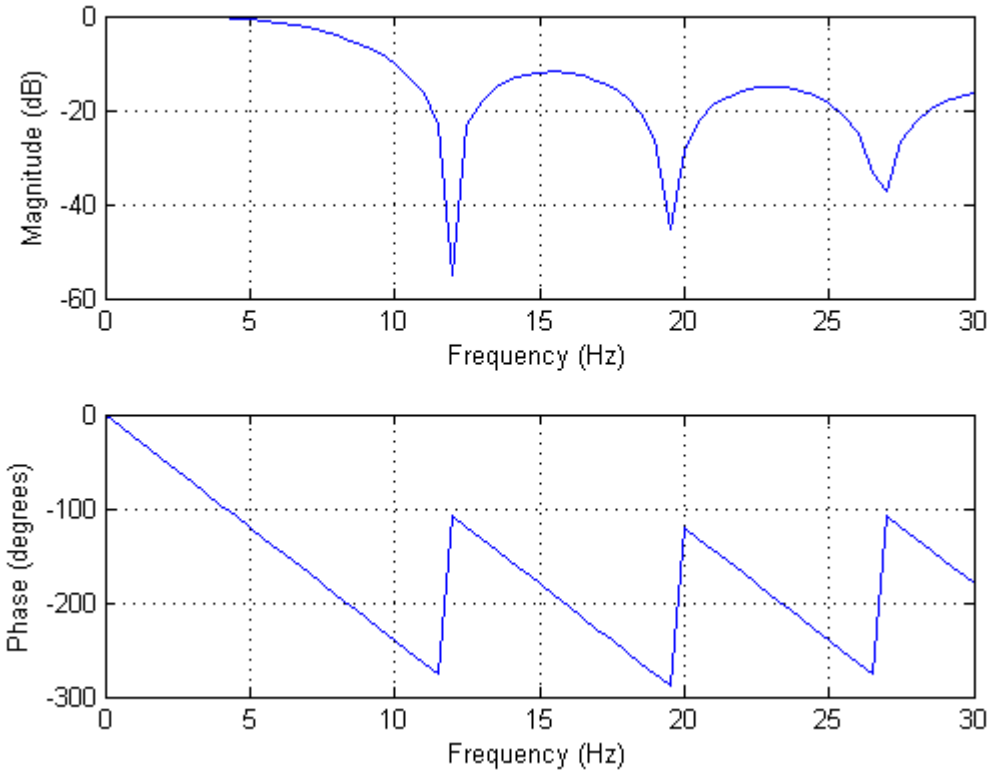


Figure 4.5: Magnitude and Phase Response of 11-point parabolic filter (sampling frequency = $75Hz$)

Table 4.1: Results on the 249-patient database using Basic Peak Detection Algorithm

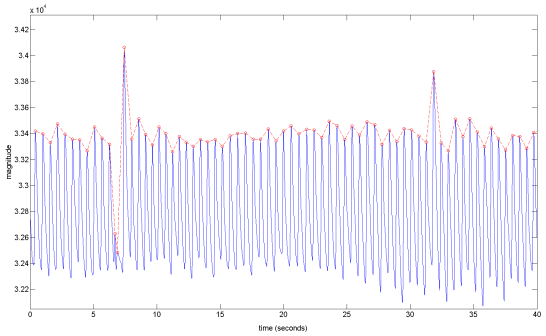
Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	55 (22%)	94 (38%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	25 (10%)	55 (22%)
Mean Absolute Error (breaths per minute)	24.5	

between consecutive peaks in the envelope is the interval between two breaths. The instantaneous respiratory rate, defined to be at the second of two consecutive peaks in the envelope, is then $60/T$ breaths/minute. For every 60-second sliding window, a mean respiratory rate is computed by taking the average of all the instantaneous respiratory rates within that window.

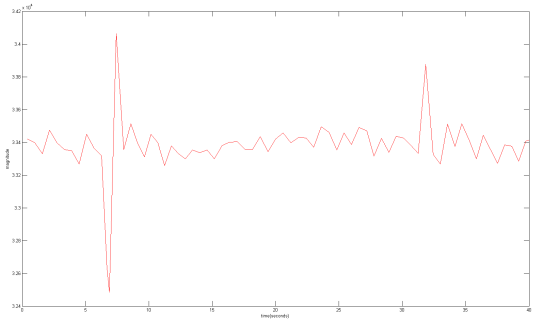
For illustration, figure 4.6 (a) shows the raw PPG waveform for a single patient with the peaks identified and part (b) of the figure shows the corresponding envelope ("the respiratory signal") extracted. Subsequently, the peak detection is applied to the envelope to identify the breaths in order to compute the instantaneous respiratory rates as illustrated in figure 4.6 (c). Finally, the instantaneous respiratory rate is calculated using the timing of the breaths identified in the evelope. In figure 4.6 (d), the instantenous respiratory rates are shown in blue while the mean respiratory rate is shown in red. The figure also shows the reference respiratory rate in green, as assessed by the triage nurse by counting the number of chest wall movements during a period of 15 or 30 seconds. These steps were followed using a 60-second sliding window and for every patient, the average of the mean estimates from all the 60-second sliding windows was also computed and taken to be the estimated respiratory rate for that patient. Finally, the absolute difference between the estimated respiratory rate and the measurement provided by the triage nurse was computed for every patient, the mean of which is taken as the value of one of the error metrics (see section 4.2.2) for the algorithm, the mean absolute error in breaths per minute.

Table 4.1 shows the results obtained when applying the basic peak detection algorithm to the PPG waveforms in the Paediatric Triage database.

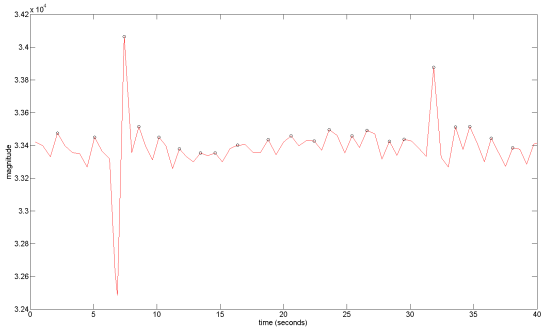
The large number of dicrotic notches detected (see section 1.2) as PPG peaks gives rise to extra



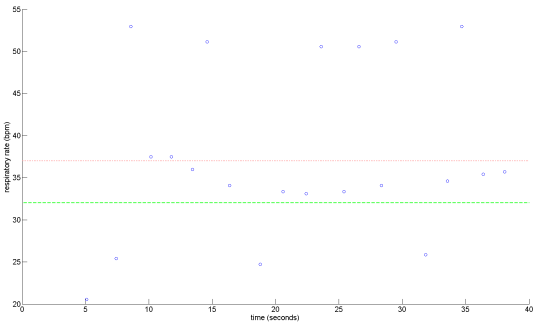
(a) Raw PPG waveform with peaks identified



(b) signal envelope (“respiratory signal”) after the basic peak detection



(c) breath detection using peak detection on the envelope



(d) instantaneous respiratory rate (blue), mean respiratory rate (red) and reference respiratory rate (green)

Figure 4.6: Illustration of respiratory rate estimates on a single patient with the Basic Peak Detection Algorithm (Patient 0005)

peaks and this can make the estimation of respiratory rate inaccurate. This is illustrated in figure 4.7 with an 8-second PPG segment for a child for whom these dicrotic notches were prominent.

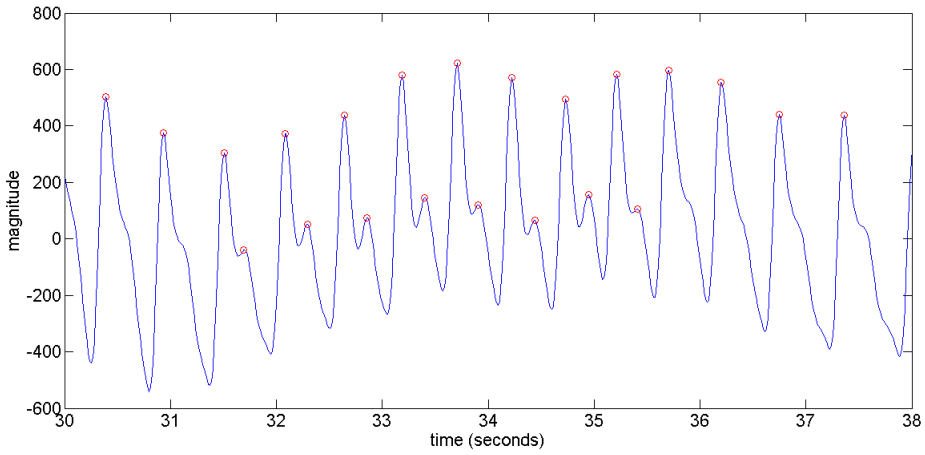
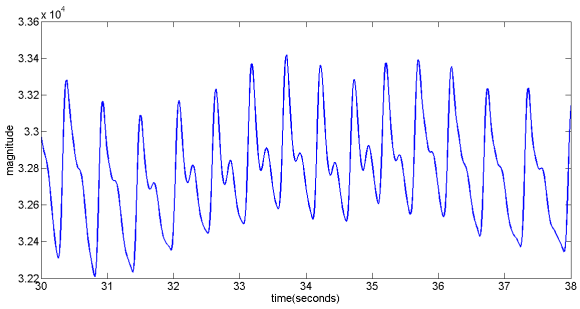


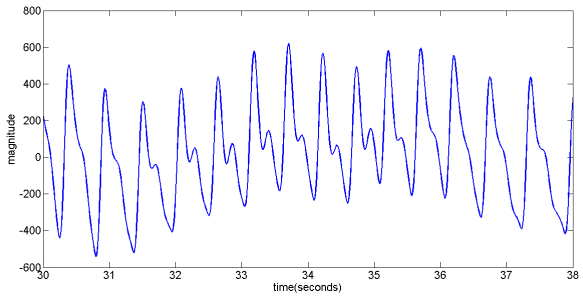
Figure 4.7: An 8-second PPG segment with peaks identified to illustrate how some of the dicrotic notches are also detected as peaks by the Basic Peak Detection algorithm (Patient 0049)

4.4 Complex Peak Detection Algorithm

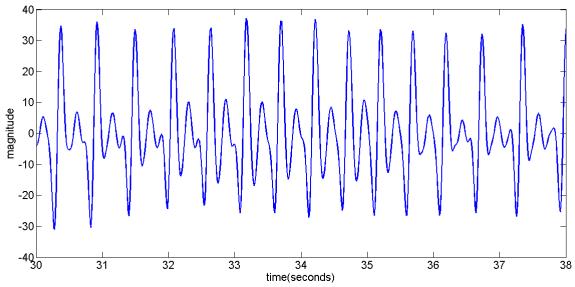
Additional pre-processing steps were introduced to address the problem caused by the detection of the extra peaks due to the dicrotic notch. The PPG waveform is again detrended and filtered using an 11-point parabolic filter. This signal is then differentiated using a 5-point differentiator to find its gradient, followed by cubing to emphasize these gradients, and half-wave rectification so that only upslopes are eventually identified. Peaks in the processed waveform correspond to the points of maximum upslope (salient points) in the original PPG waveform. These are identified by applying the 3-point peak detector to the processed waveform. Subsequently, the 3-point peak detector is applied to the original PPG waveform, but forwards from the salient point only. Figure 4.8 illustrates the outputs from the various processing steps of the complex peak detection algorithm applied to a 5-second PPG segment.



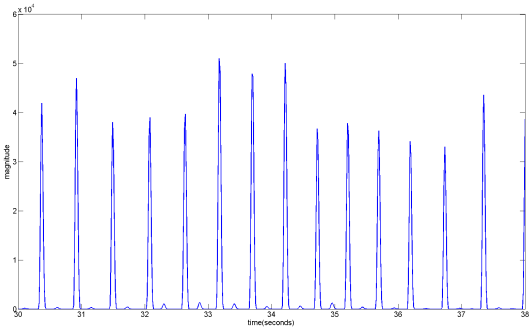
(a) A 5-second raw PPG segment



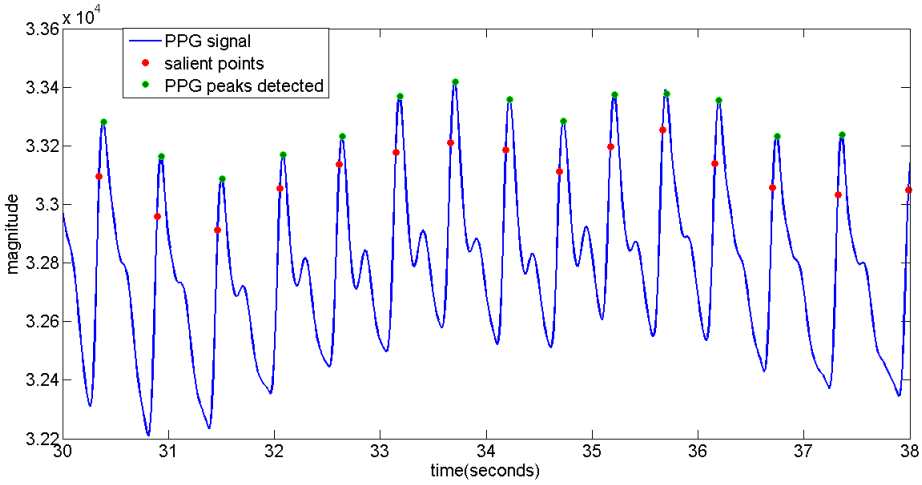
(b) PPG waveform after detrending and filtering



(c) Signal after differentiation



(d) Signal after cubing and half-wave rectification



(e) Raw PPG Signal with Peaks identified shown in green and the salient points identified shown in red

Figure 4.8: Illustration of the outputs of the various steps of the Complex Peak Detection Algorithm (Patient 0049)

Comparing figure 4.8 (e) with figure 4.7, it can see that the dicrotic notches in the PPG signal that were incorrectly identified using the Basic Peak Detection algorithm are no longer identified as major peaks in the PPG waveform.

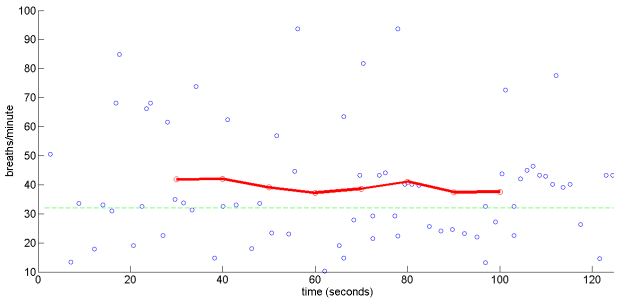
Table 4.2 shows the results on the 249-patient database after applying the Complex Peak Detection algorithm.

Table 4.2: Results on the 249-patient database using Complex Peak Detection Algorithm

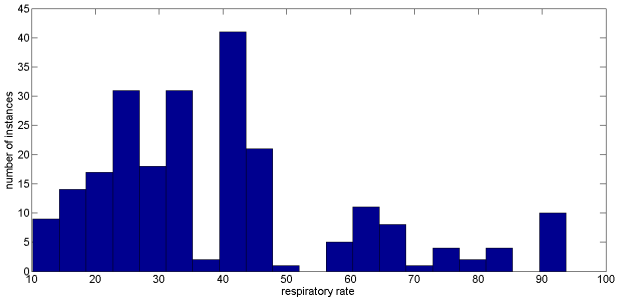
Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	114 (46%)	176 (71%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	56 (22%)	114 (46%)
Mean Absolute Error (breaths per minute)	12.3	

Outlier Rejection

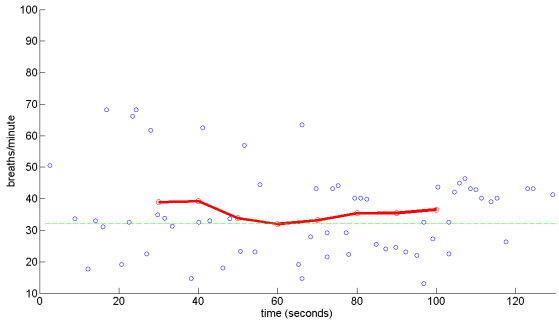
Incorrectly identified peaks result in an erroneously high instantaneous respiratory rate and missed peaks result in an erroneously low instantaneous respiratory rate. An outlier rejection procedure was introduced to mitigate this problem. The outlier rejection procedure is illustrated in figure 4.9. The upper panel shows the mean and the instantaneous respiratory rates prior to the outlier rejection procedure. For each 60-second window, the mean and standard deviation of the respiratory rate estimates were determined and any instantaneous respiratory rate estimate outside z standard deviations was rejected. The lower panel shows the instantaneous respiratory rates, the mean and the distribution of the respiratory rates after removal of outliers with $z = 1$. Figure 4.10 shows the variation in the mean absolute error (in blue) with various choices of standard deviation selected (from $z = 0.5$ to $z = 5.0$ in steps of 0.2) for the data collected for the 249 paediatric patients (see Table 3.2). The mean error without any outlier rejection is shown by the red line. The results indicate that outlier rejection improves the accuracy of the respiratory rate estimates by about 2.4 breaths per minute (in terms of mean absolute error), the greatest improvement being seen when z=1.2. Table 4.3 shows the results on the 249-patient database after applying the Complex Peak Detection algorithm to the PPG waveform with the outlier rejection procedure ($z = 1.2$).



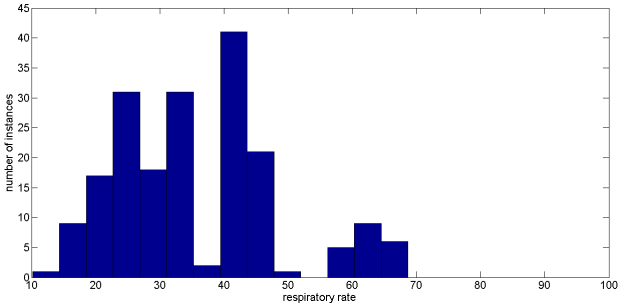
(a) instantaneous respiratory rate (blue), mean respiratory rate (red) and reference respiratory rate (green)



(b) histogram of distribution of instantaneous respiratory rates



(c) same as (a) with outlier rejection ($z = 1.2$)



(d) same as (b) with outlier rejection ($z = 1.2$)

Figure 4.9: Illustration of effect of Outlier Rejection Procedure (patient 0079)

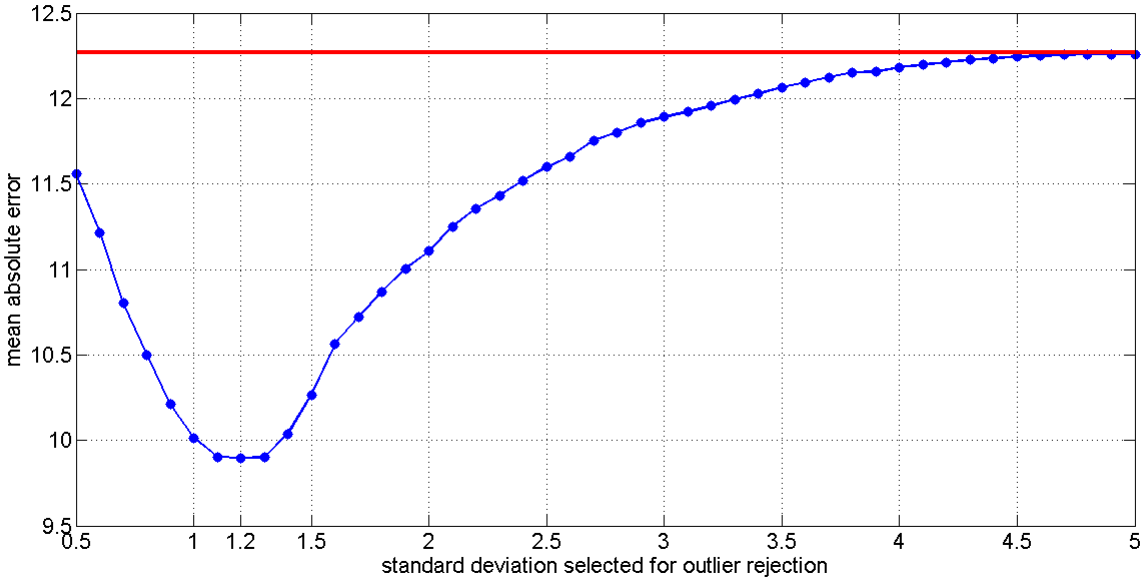


Figure 4.10: Variation of Mean Absolute Error with different choices of standard deviation for outlier rejection shown in blue and the mean absolute error without any outlier rejection shown in red

Table 4.3: Results on the 249-patient database using Complex Peak Detection Algorithm with an Outlier Rejection Procedure ($z = 1.2$)

Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	151 (61%)	203 (82%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	84 (34%)	147 (59%)
Mean Absolute Error (breaths per minute)	9.9	

4.5 Autoregressive Modelling (AR) based algorithm

Autoregressive modelling (AR) is an alternative method for estimating the power spectrum of a time series. Unlike the traditional periodogram, the frequency resolution does not depend on the number of data points. To increase the frequency resolution of the traditional periodogram, longer sequences are required. However, this is not feasible for physiological signals since stationarity[1] cannot be assumed over longer periods. AR-based spectral estimation has been previously applied to a number of physiological signals, including the EEG [104], the intrapartum cardiotogram [19] and the PPG [37].

Autoregressive modelling assumes that the current value of a time series can be predicted as a linearly weighted sum of the preceding p terms of the same series. The parameter p is the model order which is usually much smaller than the length of the sequence N.

$$y(n) = \sum_{k=1}^{p} a_k y(n-k) + e(n) \tag{4.1}$$

where $y(n)$ is the current output of the time series ($t = n$), $y(n-k)$ is the time series output at $t = n-k$, a_k are the parameters of the system quantifying the linear relationship between the current output (i.e. at $t = n$) and the previous outputs at $t = (n-k)$, and $e(n)$ is the error term which is assumed to be normally distributed with zero mean and a variance of σ^2. The above equation can be visualized with the help of figure 4.11 as a recursive filter (i.e. infinite impulse response filter)

[1] statistical properties do not change over time

with the predicted values $y'(n)$ representing the true underlying signal and the actual values $y(n)$ considered to be the output corrupted by additive white noise giving rise to the prediction errors $e(n)$.

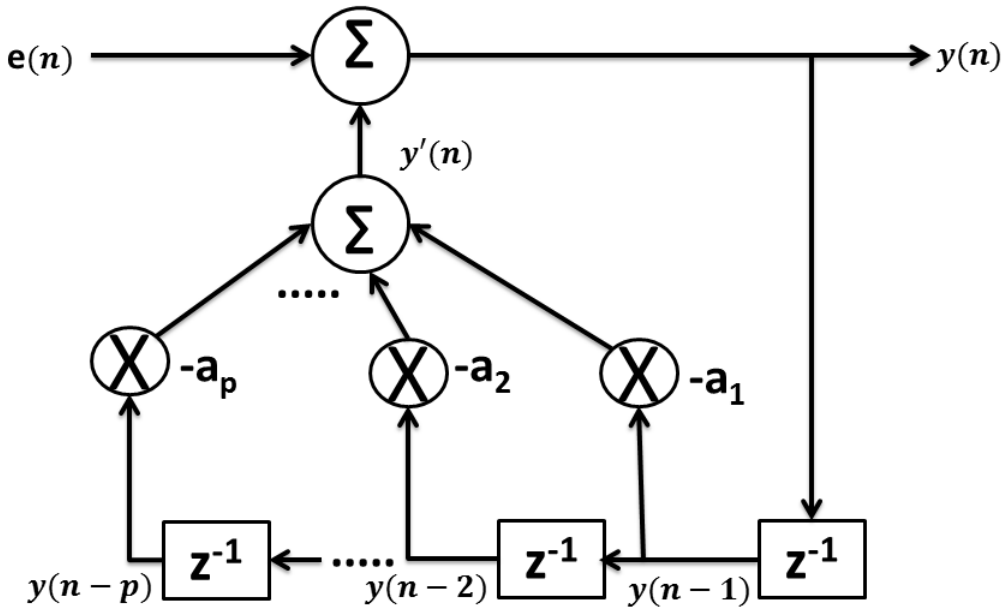


Figure 4.11: Autoregressive model seen as an all-pole filter (modified from [104])

Applying the z-transform to equation 4.1,

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

$$Y(z)(1 - \sum_{k=1}^{p} a_k z^{-k}) = E(z)$$

$$\frac{Y(z)}{E(z)} = H(z) = \frac{1}{1 - \sum_{k=1}^{p} a_k z^{-k}} \tag{4.2}$$

Factorising the k^{th} order polynomial in the denominator in equation 4.2 gives

$$H(z) = \frac{1}{\prod_{k=1}^{p}(1 - p_k z^{-1})} \tag{4.3}$$

Multiplying both the numerator and denominator of equation 4.3 by z^p gives

$$H(z) = \frac{z^p}{\prod_{k=1}^{p}(z - p_k)} \tag{4.4}$$

The AR model can thus be viewed as a system with an input $E(z)$ and an output $Y(z)$ with a transfer function $H(z)$ given by equation 4.4. This equation is the transfer function of an all-pole model since there are no zeros of $H(z)$ away from the origin. Hence the behaviour of the system is described in the frequency domain using only the poles. The poles occur in complex conjugate pairs and they define the spectral peaks in the power spectrum. The magnitude and the location of a pole in the complex plane determine the magnitude of the spectral peak and the frequency at which it occurs [102]. The nearer the pole is to the unit circle on the pole-zero plot in the z-domain, the higher and the sharper is the corresponding peak in the frequency domain [102]. The frequency f associated with a given pole is related to the phase angle θ of that pole according to equation 4.5, where ∇t is the sampling interval of the original sequence.

$$\theta = 2\pi f \nabla t \tag{4.5}$$

For details on how the pole location is related to the frequency spectrum, see Appendix A.

Application of the AR model to PPG analysis

The PPG signal was first detrended and filtered to reduce the magnitude of spectral components at the cardiac frequency as these would otherwise tend to dominate the spectrum and decrease the accuracy of placement of the pole associated with the respiratory frequency (the respiratory pole). The filter chosen for this purpose was a low-pass FIR filter using a Kaiser window such that it had a 5% ripple in the passband and a 40dB attenuation in the stopband. The cutoff frequencies were decided based on the age of a child. For children up to 2 years of age, the transition band was chosen to be 1.25 Hz (75 breaths per minute) to 2 Hz (120 breaths per minute) while for children

over 2 years of age, a transition band of 0.7 Hz (42 breaths per minute) to 2 Hz (120 breaths per minute) was chosen. Application of the FIR low pass filter with the chosen transition bands limits the bandwidth of the PPG signal to 2 Hz thereby imposing a lower limit of 4 Hz for downsampling to avoid aliasing [109].

Subsequently, the PPG signal was downsampled to 5 Hz, which ensured that all the poles model frequencies up to 2.5 Hz (corresponding to 150 breaths/minute), an upper limit imposed by the Nyquist criterion. The original PPG waveforms were sampled at 75 Hz and choosing a downsampling frequency that is an exact divisor of 75 Hz (i.e. 5 Hz) ensured that no interpolation was required for the downsampling step. The downsampling process enhances the ability to resolve spectral components (by allowing a given number of degrees in the pole-zero plot to represent a smaller frequency range) in the expected range of respiratory frequencies.

The model order p determines the number of poles that are available to model the frequency spectrum of the PPG signal. For models with $p > 3$, there can be multiple poles with phase angles that lie in the expected range of respiratory frequencies. The criterion used to identify the pole corresponding to the respiratory frequency is based on the pole's magnitude, as it is expected that the dominant frequency in the PPG signal in the expected frequency range will be due to breathing. Therefore, the pole corresponding to this dominant frequency would have the highest magnitude in the AR model in the range of respiratory frequencies. The expected range of frequencies is shown in figure 4.12.

Tests were carried out to investigate the performance of the AR model-based algorithm with respect to the model order p. A two-step pole searching criterion was adopted for this investigation: firstly an attempt is made to find a pole in the range 12 to 50 breaths per minute and if no pole is found in this range, only then an attempt is made to find a pole in the range 6 to 80 breaths per minute.

Both the Yule-Walker approach [126] and the Burg method [126] were implemented and the mean absolute error for the 249-patient database was computed. Based on the mean absolute error with respect to the reference respiratory rate, the AR model with $p = 7$ was found to be optimal using the Burg algorithm for parameter estimation. A 7^{th} order AR model constructed using the

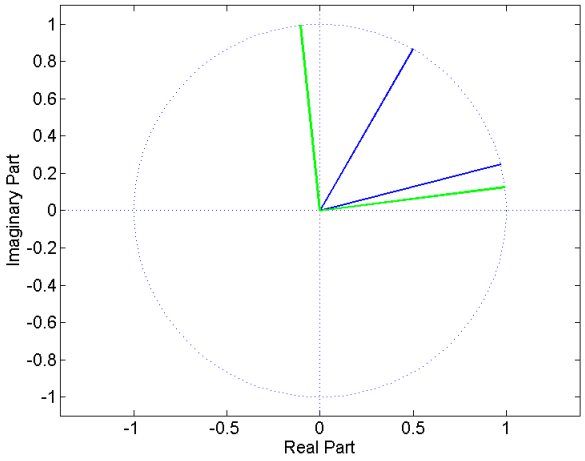
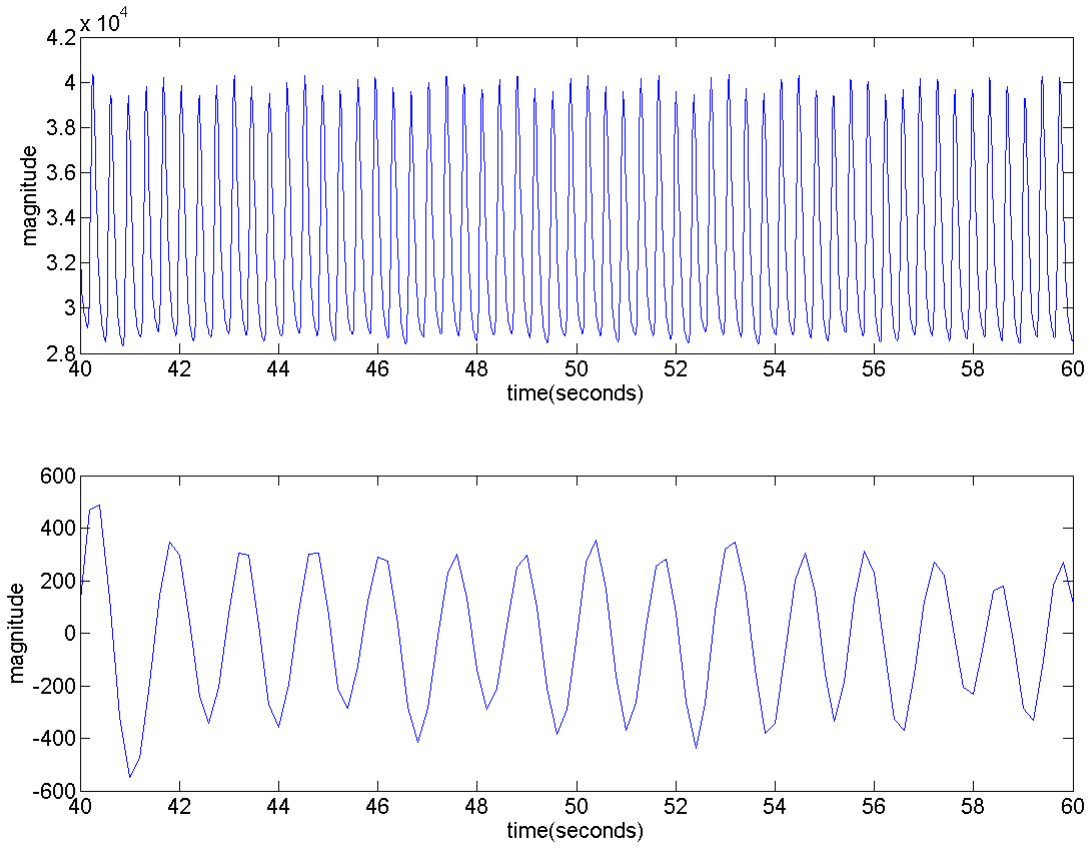


Figure 4.12: Region of interest for pole searching to find the pole that corresponds to the respiratory rate. The region corresponding to 6-80 breaths per minute is shown by the minor sector enclosed by the radii in green and the region corresponding to 12-50 breaths per minute is shown by the minor sector enclosed by the radii in blue.

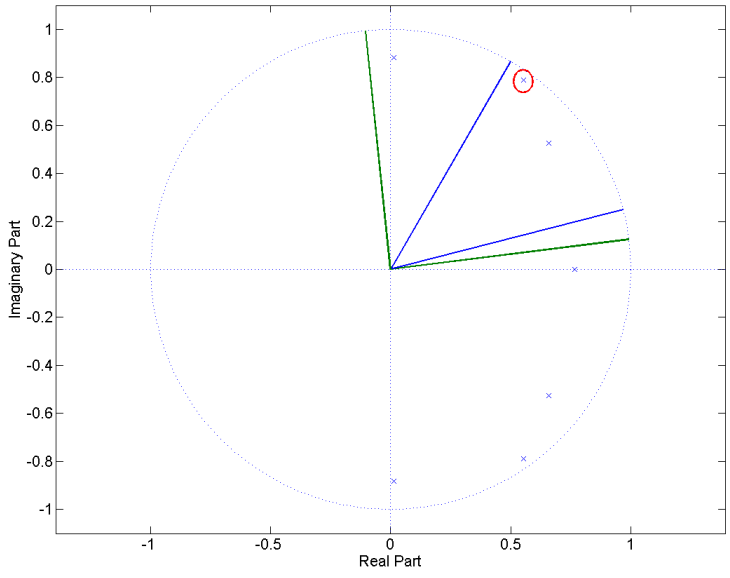
Burg method is therefore applied to consecutive 60-second sections of the downsampled PPG signal with an overlap of 50 seconds. This gives an updated respiratory rate estimate every 10 seconds, based on the previous 60 seconds.

Figure 4.13 illustrates the outputs of the various stages of the AR model based algorithm for respiratory rate estimation implemented on a 60-second PPG segment.

Table 4.4 shows the results obtained on the 249-patient database after applying the AR modelling based algorithm with a single pole searching criterion. This criterion was to select the pole with the highest magnitude in the range 6 to 60 breaths/minute to be the respiratory pole. Table 4.5 shows the results when applying the AR modelling based algorithm with a two-step pole searching criterion. As before, an attempt was first made to find a pole in the range 12 breaths/minute to 50 breaths/minute and if no pole was found in this range, a subsequent attempt was made to find a pole in the range 6 to 80 breaths/minute.



(a) Illustration of raw PPG segment and the output after detrending, low-pass filtering and downsampling to 5 Hz



(b) Poles of the AR model, with the respiratory pole circled in red and the sector of interest identified by the radii in green and blue

Figure 4.13: Example of AR modelling on a 20-second PPG segment for patient 0402 to find the "respiratory pole", using a model order of 7. The respiratory pole estimated in this example corresponds to a respiratory rate of 45.7 breaths/minute. The reference respiratory rate for this patient was 44 breaths/minute.

Table 4.4: Results on 249-patient database using a 7^{th} order AR model based algorithm using a single step pole search criterion. The model coefficients were obtained using the Burg Method.

Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	149 (60%)	196 (79%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	71 (29%)	127 (51%)
Mean Absolute Error (breaths per minute)	11.9	

Table 4.5: Results on the 249-patient database using an 7^{th} order AR model based algorithm using a two-step pole search criterion.

Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	153 (61%)	195 (78%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	93 (37%)	165 (66%)
Mean Absolute Error (breaths per minute)	9.2	

The results in the tables show that the two-step search criterion reduces the mean absolute error by about 2.7 breaths per minute and the number of patients for whom the estimated respiratory rate is within 5 breaths per minute of the reference respiratory rate, increases from 71 (29%) to 93 (37%).

4.6 Improvements to the AR-based Algorithm

The model order of 7 which was found to be the optimal choice of model order across the 249-patient database may not be optimal for individual patients. Figure 4.14 shows the estimated respiratory rate for a patient in the database (Patient 0402) with an 11^{th}-order and 14^{th}-order AR model based respiratory rate estimation algorithm using the two-step pole search criterion. It can be seen

that at $t = 50$ seconds, there is a large error in the estimation of the respiratory rate with a 14^{th} order AR model while the estimate with the 11^{th} order AR model is more accurate. However, at $t = 100$ seconds, the opposite occurs: there is a large error in the estimation of the respiratory rate with the 11^{th} order AR model while the 14^{th} order AR model is more accurate. Figure 4.15 shows the location of the poles on a pole-zero plot after applying the 11^{th} order AR model to a 60-second PPG segment for patient 0402 centered on $t = 50$ in figure 4.14. Figure 4.16 shows the same for the 14^{th}-order AR model. The pole angles that are within the expected respiratory range for the 11^{th}-order and the14^{th}-order AR models are tabulated in Table 4.6 with the corresponding magnitude arranged in descending order. It can be seen that for the $14^{th}-$order AR model, the pole corresponding to 15.8 breaths per minute has the highest magnitude and will therefore be erroneously identified as the respiratory pole.

The frequency spectra, however, for both the 11^{th}-order and the 14^{th}-order AR model are similar and have a dominant peak at about 44 breaths per minute as shown in figure 4.17 (see Appendix A for an explanation of how the frequency spectra were obtained). The reason for the presence of dominant pole at about 44 breaths per minute (even though the pole corresponding to it does not have the highest magnitude) after applying the 14^{th}-order AR model is due to the presence of poles corresponding to 57.2 and 37.9 breaths per minute (see Figure 4.16 and Table 4.6) which result in an overall dominant peak at 44 breaths per minute in the frequency spectrum (see section 4.6.2 for further explanation). This phenomenon, called pole-splitting, occurs when a model order too high for the data to be modelled is chosen [19][58].

Pole selection on magnitude alone can lead to the wrong choice of respiratory pole and the following three different approaches were adopted to address this:

1. Use multiple models for each patient, and then select the best estimate using either the mean, median or a probability density based approach.

2. apply the AR model and determine the frequency spectrum from the poles and then identify the highest peak which is expected to correspond to the respiratory rate.

3. investigate the reason for the secondary peak present at the lower frequency and eliminate it using envelope normalization before applying an AR model.

Approaches 1-3 are further explained in Sections 4.6.1, 4.6.2 and 4.6.3 in that order.

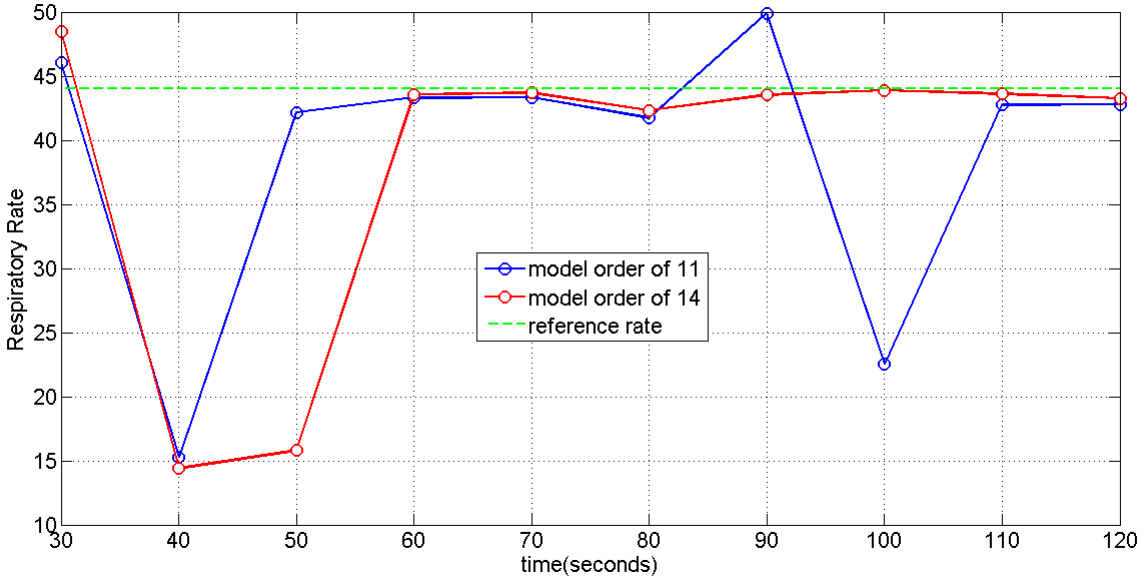


Figure 4.14: Estimated Respiratory Rate for Patient 0402 with a 14^{th}order AR model (red) and an 11^{th}order AR model (blue). The reference respiratory rate is shown in green (44 beats/minute).

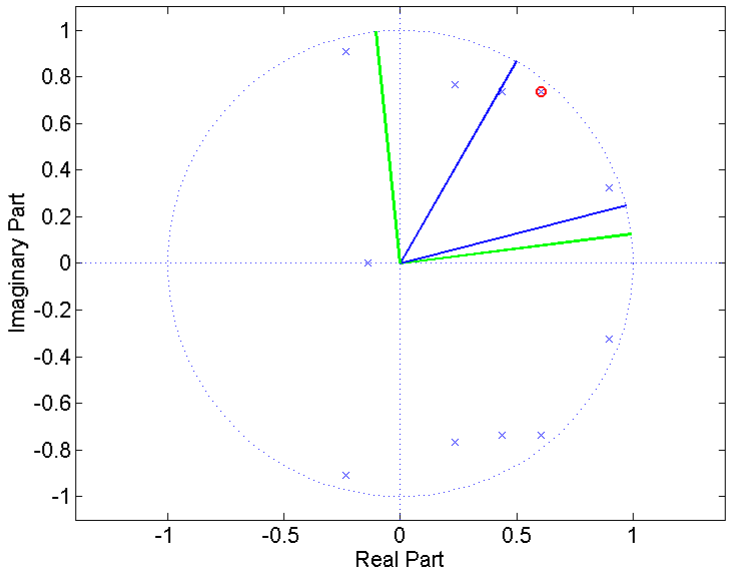


Figure 4.15: Poles of the 11^{th}order AR model on a 60-second PPG segment of patient 0402, with the respiratory pole identified circled in red and the regions of interest lying between the green or blue radii.

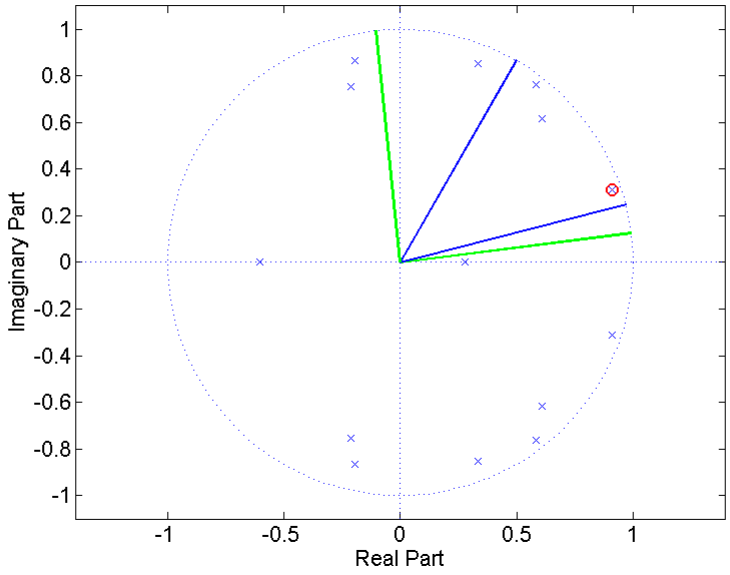


Figure 4.16: Poles of the 14^{th} order AR model on a 60-second PPG segment of patient 0402, with the respiratory pole identified circled in red and the regions of interest lying between the green or blue radii.

pole angle (radians)	corresponding breaths per minute	Magnitude
0.33	15.8	0.962
0.92	43.9	0.957
1.20	57.2	0.915
0.79	37.9	0.864

(a) model order of 14

pole angle (radians)	corresponding breaths per minute	Magnitude
0.88	42.2	0.954
0.35	16.6	0.951
1.03	49.4	0.855
1.27	60.8	0.801

(b) model order of 11

Table 4.6: Table of the poles (those that are within the expected respiratory range) with the corresponding magnitude in descending order.

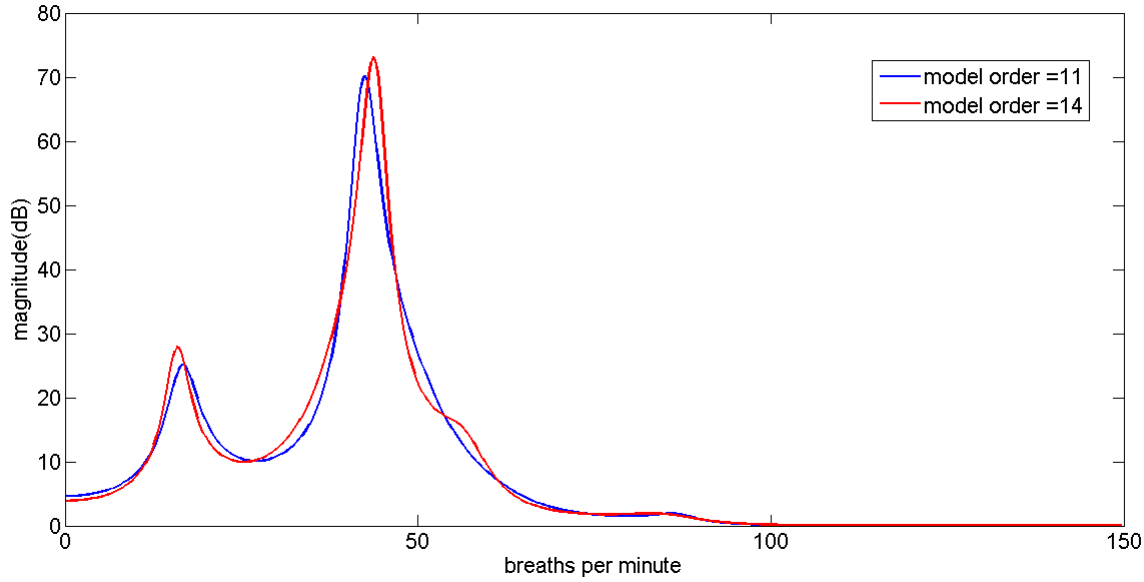


Figure 4.17: Frequency Spectrum corresponding to the 11^{th} order (blue) and the 14^{th}order (red) AR models for patient 402 at $t = 50$

4.6.1 Multiple AR models for Respiratory Rate Estimation

Once the PPG signal was preprocessed and downsampled to 5 Hz as described in section 4.5, the coefficients of AR models with $p = 4$ to $p = 20$ were estimated using the Burg method. For every 60-second window, the respiratory rate was estimated using the two-step pole search criterion. This meant that for 120 seconds of PPG signal, there would be seven 60-second windows with an overlap of 50 seconds, and 17 respiratory rate estimates for each of these windows. Therefore, the total number of respiratory rate estimates for a 120-second period would be 7x17 = 119. Different strategies were implemented to select the optimal estimate from these:

1. take the median of all the estimates
2. take the mean of all the estimates
3. take the median of all the estimates that are within w standard deviations of the mean varying w from 0.8 to 5
4. take the mean of all the estimates that are within w standard deviations of the mean varying w from 0.8 to 5.

5. determine a probability density distribution using a normal kernel function centered at every estimate with the width of the kernel smoothing function determined by the number of estimates. The estimate of the respiratory rate is then the one corresponding to the highest peak in the probability density estimate.

Criteria (1) and (2) are based on determining an estimate from the histogram of respiratory rate estimates. For (3) and (4), the error was seen to reduce with increasing w before flattening out at $w = 3.5$ where it effectively becomes equal to that estimated with criteria (1) and (2).

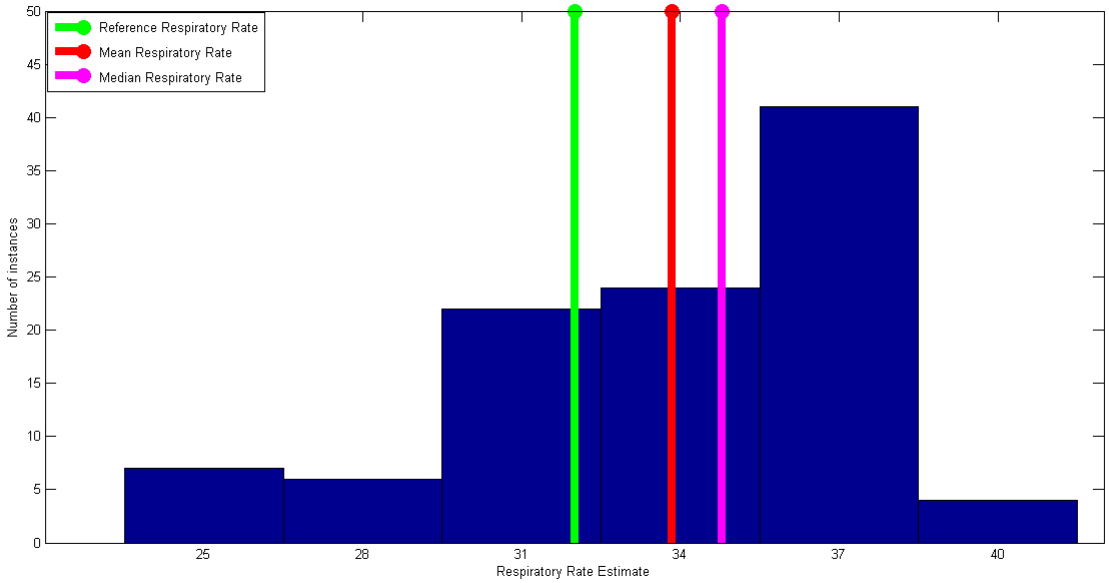
Figure 4.18 illustrates this process for a single patient. Firstly the mean and median of all the estimates are computed, then the mean and median of all the estimates that are within $w = 1$ standard deviation are computed.

Figure 4.19 illustrates option (5) for a single patient by computing a probability density estimate using a normal kernel function. The bandwidth of the kernel smoothing window is chosen to be optimal for estimating normal densities. This is based on the approximation provided by [127] given by equation 4.6.

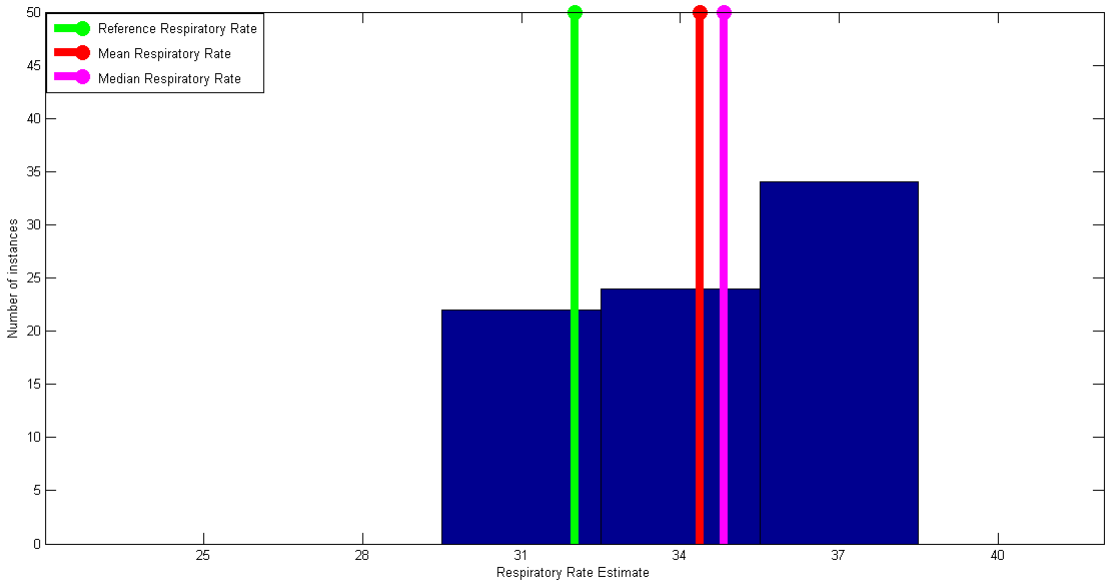
$$h = (\frac{4\sigma'^5}{3n})^{1/5} \approx 1.06\sigma' n^{-1/5} \tag{4.6}$$

where σ' is the standard deviation of the n samples for which the probability density is estimated.

Table 4.7 shows the results obtained on the 249-patient database after applying the multiple AR models based algorithms using the five criteria described above. Both the median and the mean based approach are found to be better than the probability density based approach and both are seen to perform better than the 7^{th} order AR model based algorithm followed by the two-step pole searching criterion. The median-based multiple AR model based approach can be seen to increase the number of patients whose respiratory rate is within 5 breaths per minute from increased from 93 (37%) to 99 (40%).



(a) Distribution of estimates of respiratory rates for AR models from $p = 8$ to $p = 20$. The reference respiratory rate is shown in green, the mean of all these estimates is shown in red and the median of all these estimates is shown in pink.



(b) Distribution of the same estimates within 1 standard deviation of the mean respiratory rate. The reference respiratory rate is shown in green, the mean of all these estimates is shown in red and the median of all these estimates is shown in pink.

Figure 4.18: Illustration of respiratory rate estimation based on multiple AR models (histogram-based selection) using patient 0005

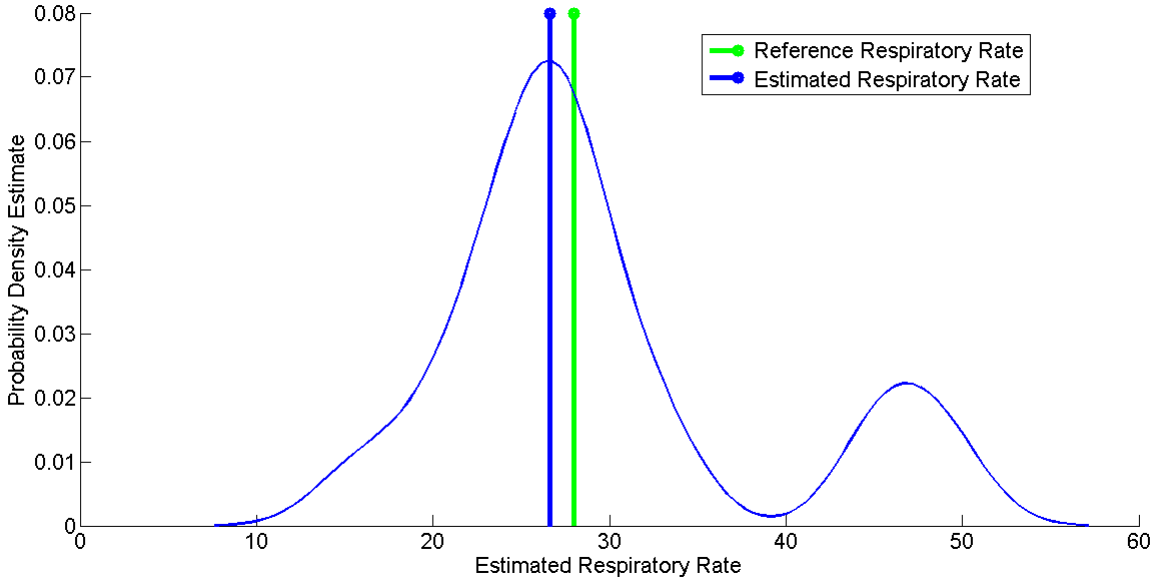


Figure 4.19: Probability density distribution computed using the respiratory rate estimates from AR models for $p = 8$ to $p = 20$ for patient 0088. The reference respiratory rate is shown in green and the estimated respiratory rate corresponding to the highest peak in the probability distribution is shown in blue.

Table 4.7: Results on 249-patient database using Multiple AR models for respiratory rate estimation

Criterion	Mean Absolute Error	Number of patients whose respiratory rate estimate is within 5 breaths per minute of the reference	Number of patients whose respiratory rate estimate is within 10 breaths per minute of the reference
take the median of all the estimates to be the respiratory rate	9.1	99 (40%)	170 (68%)
take the mean of all the estimates to be the respiratory rate	9.1	96 (39%)	170 (68%)
take the respiratory rate to be the estimate corresponding to the highest peak in the probability density estimate	10.0	89 (36%)	154 (62%)

4.6.2 Frequency Spectrum based Respiratory Rate Estimation

It is possible that the highest peak in the frequency spectrum of a processed PPG segment does not correspond to the pole with the highest magnitude. Figure 4.16 showed the pole-zero plot after applying a 14th order AR model to a processed 60-second PPG segment for patient 0402 whose reference respiratory rate was 40 breaths per minute. As figure 4.17 showed, the highest peak in the frequency spectrum corresponds to 44 breaths per minute while the pole with the highest magnitude corresponds to 15.8 breaths per minute. Estimating the respiratory rate based on the pole's magnitude in this case will therefore result in an erroneous estimate (see Figure 4.16 and Table 4.6). The reason for this observation is that the dominant peak in the frequency spectrum does not only depend on the magnitude of a single pole but also the presence of other closely-located poles.

In an attempt to investigate this problem, an AR model is applied to the processed PPG waveform and the frequency spectrum is determined from the poles found. A peak detection algorithm is then used on the frequency spectrum to determine the highest peak which is taken to correspond to the estimated respiratory rate. Table 4.8 show the results obtained after using this approach with a 7^{th} order AR model. Comparing with the results from table 4.5, it can be seen that the frequency spectrum based respiratory rate estimation improves the performance of the algorithm and the number of patients whose respiratory rate is within 5 breaths per minute is increased from 93 (38%) to 100 (40%).

Table 4.8: Results on the 249-patient database using an 7^{th} order AR model based algorithm using frequency spectrum based respiratory rate estimation

Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	157 (63%)	196 (79%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	100 (40%)	173 (69%)
Mean Absolute Error (breaths per minute)	9.1	

4.6.3 Envelope Normalization based Respiratory Rate Estimation

Figure 4.20 (a) shows the 60-second processed PPG segment before applying an AR model for the data from patient 0402 that corresponds to $t = 50$ in figure 4.14. The frequency spectrum after applying a 14^{th}order AR model is shown in figure 4.20 (b) which shows a dominant peak at 44 breaths per minute but a secondary peak at 15.8 breaths per minute that is erroneously selected as the respiratory rate if the pole with the highest magnitude is selected as the one corresponding to the respiratory rate (see the discussion at the beginning of section 4.6).

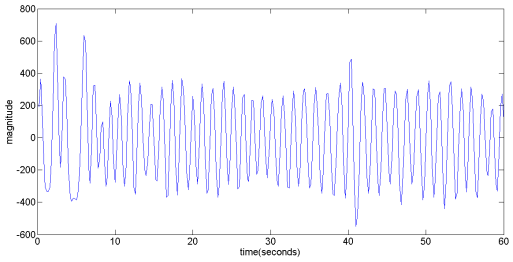
In this method, the peaks and the troughs of the PPG signal are rescaled such that the signal lies between -1 and $+1$. The envelope normalization algorithm consists of two stages. The first stage is standard extremum detection (similar to the algorithm explained in section 4.3) to find the peaks and troughs. Subsequently, the algorithm traverses every point of the signal starting from the left keeping track of the closest trough and closest peak. Each point traversed is normalized according to equation 4.7

$$y = \frac{(y - midy)}{(0.5 * pk2tr)} \tag{4.7}$$

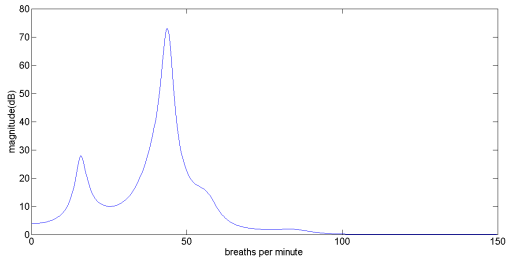
where y is the y-axis value of the current point being traversed, $midy$ is the average of the y-axis value for the closest trough and the y-axis value for the closest peak and $pk2tr$ is the distance between the closest peak and the closest trough.

Figure 4.20 (c) shows the same 60-second PPG segment after normalization and figure 4.20 (d) shows its corresponding frequency spectrum. It can be seen from the figure that the secondary peak in the frequency spectrum is eliminated and envelope normalization thus respresents a potential solution to the problem that was discussed at the beginning of section 4.6.

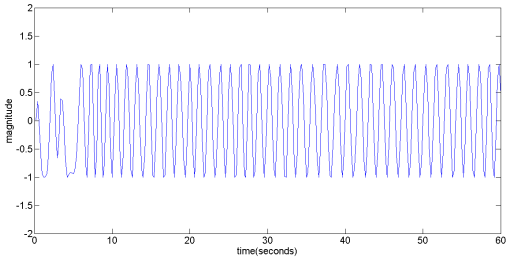
The AR model based respiratory rate algorithm was modified to include the envelope normalization step followed by a two-step pole search criterion and applied to the 249-patient database. Table 4.9 shows the results obtained. Comparing with the results from Table 4.5, it can be seen that the envelope normalization step improves the performance of the algorithm and the number of patients whose respiratory rate is within 5 breaths per minute is increased from 93 (38%) to 96 (39%).



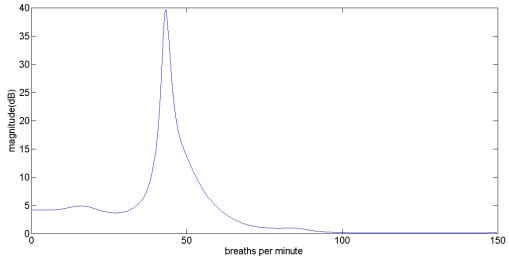
(a) 60-second downsampled PPG segment ('the respiratory signal') of patient 0402 corresponding to $t = 50$ in figure 4.14



(b) frequency spectrum evaluated using poles estimated from a 14th-order AR model applied to the processed PPG waveform



(c) 60-second downsampled and envelope normalized PPG segment ('the respiratory signal') of patient 0402 corresponding to $t = 50$



(d) corresponding frequency spectrum evaluated using poles estimated from a 14th-order AR model

Figure 4.20: downsampled PPG waveforms ('respiratory signal') with and without envelope normalization and their corresponding frequency spectrum to illustrate the reason for the erroneous secondary peak (see Figure 4.16 and Table 4.6)

Table 4.9: Results on the 249-patient database using a 7^{th} order AR model based algorithm using envelope normalization

Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	177 (71%)	221 (89%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	96 (39%)	163 (65%)
Mean Absolute Error (breaths per minute)	9.0	

Since the AR-based respiratory rate estimation algorithm, envelope normalization based AR algorithm and the frequency spectrum based AR algorithm depend on the choice of model order, figure 4.21 illustrates the performance of these with respect to model order choice. For comparison, the mean absolute error with the multiple AR-models based approach is also shown. The figure shows that envelope normalization performs better (by about 0.4 breaths per minute) at higher model orders. This is because the envelope normalization based algorithm is least susceptible to motion artefacts which introduce undesired low-frequency components, thereby reflecting an improvement in terms of mean absolute error across the 249-patient database (for details, see section 6.6).

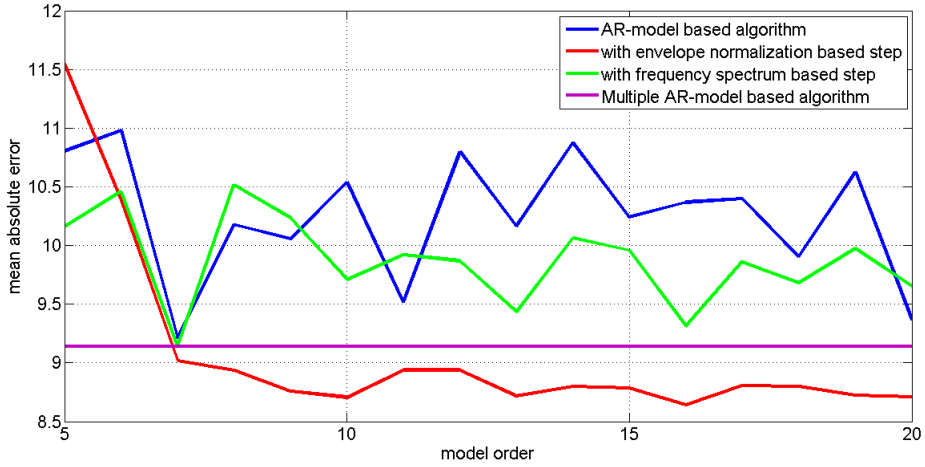
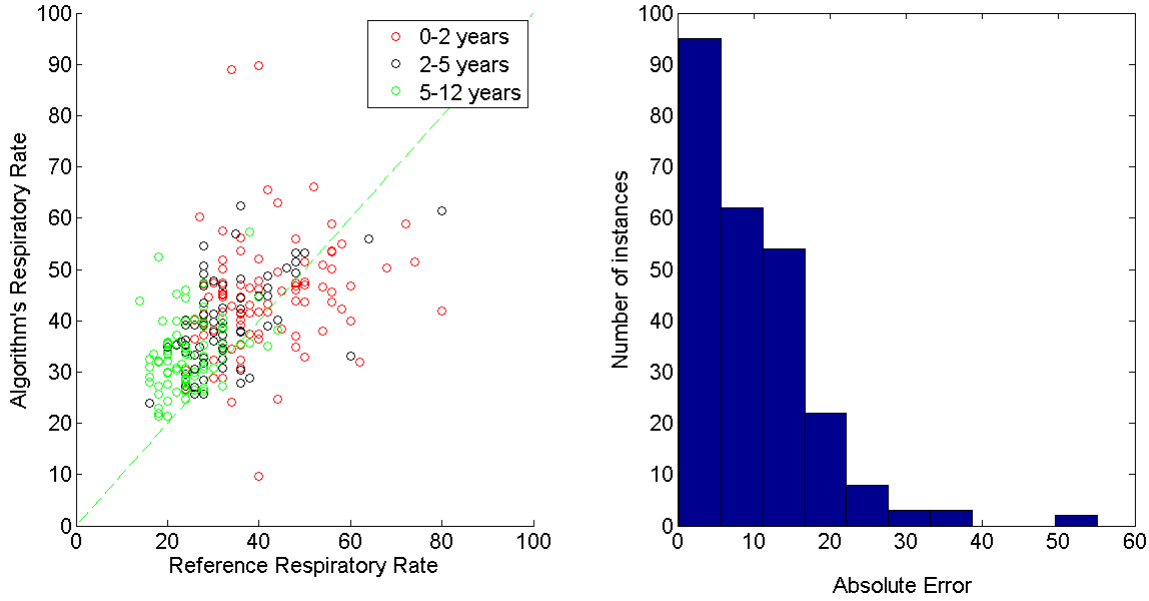


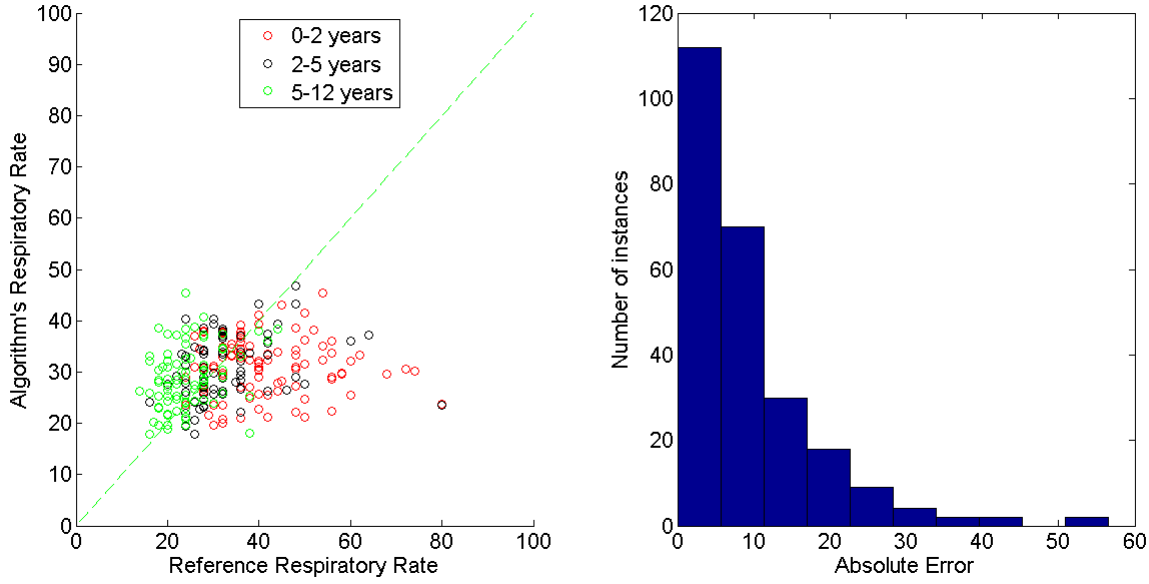
Figure 4.21: Comparison of the performance of the various types of AR-based algorithms with the choice of model order

4.7 Age-based respiratory rate estimation algorithm

Table 4.10 reports the mean absolute error in the four age groups for the various methods that had the overall lowest mean absolute error for estimating the respiratory rate from the amplitude modulation of the PPG signal. The results show that in the age group of 0-1 years, the mean absolute error is amongst the lowest for the Complex Peak Detection algorithm and the Envelope Normalization based AR algorithm while in the age group of 5-12 years, the mean absolute error is the lowest for the two Multiple AR models based algorithm. For illustration, figure 4.22 shows the scatter plot of reference respiratory rates with respect to the respiratory rate estimated by the Envelope Normalization based algorithm and the Multiple AR models based algorithm (median based). The points on the scatter plot are coloured according to three age groups. These age groups are 0-2 years, 2-5 years and 5-12 years. The figure also shows the distribution of mean absolute errors.



(a) Complex Peak Detection Algorithm with Outlier Rejection step



(b) Multiple AR model

Figure 4.22: Scatter plot of reference respiratory rate versus estimated respiratory rate by the algorithm (left panel), histogram of mean absolute errors for the estimated respiratory rate (right panel)

The AR-based algorithms (except the Envelope Normalization based method) can be seen to perform, in comparison to the Complex Peak Detection algorithm, comparatively poorly on younger children who tend to have higher respiratory rates. This is because the AR-based algorithms are most likely to find a pole in the first step (from 12 to 50 breaths per minute) during

the two-step pole search criterion due to the presence of low-frequency components due to motion artefacts. Envelope normalization performs relatively better than the other AR-based algorithms since it is more robust to low frequency noise (for further details, see Section 6.6). It is anticipated that a hybrid algorithm using the Complex Peak Detection algorithm for children belonging to the 0-1 year age group and the multiple AR model based algorithm (median-based) for children of the remaining age groups would improve the overall accuracy. A machine learning based approach for incorporating age was not considered since that would require splitting the data into training and testing sets. This was deemed to be inappropriate for the current data as the reference respiratory rates are only estimated by triage nurses with varying degrees of accuracy.

Method	0 - 1 years ($n = 59$)	1 - 2 years ($n = 40$)	2 - 5 years ($n = 69$)	5 - 12 years ($n = 81$)
Complex Peak Detection based algorithm	11.9	9.3	9.1	9.6
Single Model AR based algorithm (two-step criterion)	13.9	8.4	8.3	8.0
Multiple AR model based algorithm (median-based)	14.8	8.1	7.6	6.7
Multiple AR model based algorithm (mean-based)	14.2	8.2	7.2	7.3
Envelope normalization based algorithm	12.2	7.5	6.6	8.5
Frequency Spectrum based algorithm	13.1	8.0	7.6	8.1

Table 4.10: Mean Absolute Error for both the AR based algorithm and the Complex Peak Detection algorithm in different age groups

Figure 4.23 shows the scatter plot and the mean absolute error distribution after applying a hybrid algorithm such that the Complex Peak Detection algorithm is used, with an outlier rejection step, if the child's age is from 0 to 1 year and the multiple AR model based algorithm is used, with the two-step pole search criterion, if the child's age is greater than 1 year. The absolute mean error and other error metrics for the hybrid algorithm are provided in Table 4.11. Comparing the results with those in Table 4.7 and Table 4.3, it can be seen that the mean absolute error is reduced to 8.4 breaths per minute and that the number of patients whose respiratory rate is estimated to be within 5 breaths per minute of the reference is increased to 105 (42%) after using the hybrid algorithm.

Figure 4.24 makes a comparison of the respiratory rate estimated using the hybrid algorithm with the reference respiratory rate on the 249-patient database on a Bland-Altman plot. The mean line is also shown in the figure which was computed to be at -0.1. The data points in the figure are

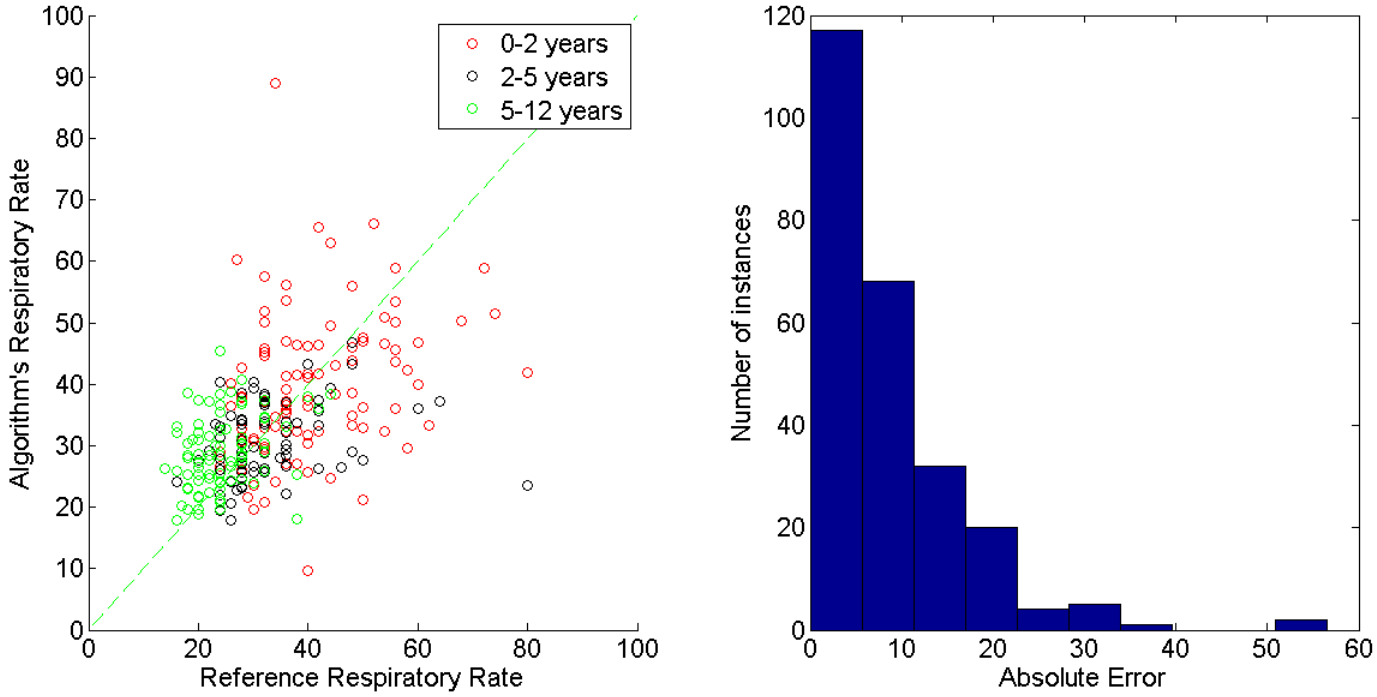


Figure 4.23: Scatter plot of reference respiratory rate versus the respiratory rate estimated by the algorithm (left panel); histogram of mean absolute errors for the estimated respiratory rate (right panel) for the hybrid algorithm.

found both above and below the zero line which suggests that there is no consistent bias. Furthermore, the data points tend to spread out more with magnitude on the $x-axis$ which suggests that there are larger errors for higher respiratory rates (respiratory rate $\geqq 40$ breaths per minute)

Quality Metric	$y = 5$	$y = 10$
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	103 (41%)	173 (69%)
Mean Absolute Error (breaths per minute)	8.4	

Table 4.11: Results on the 249-patient database using age as the criterion to decide between "Complex Peak Detection" with outlier rejection" algorithm and "Multiple AR models based algorithm (median based)" for respiratory rate estimation

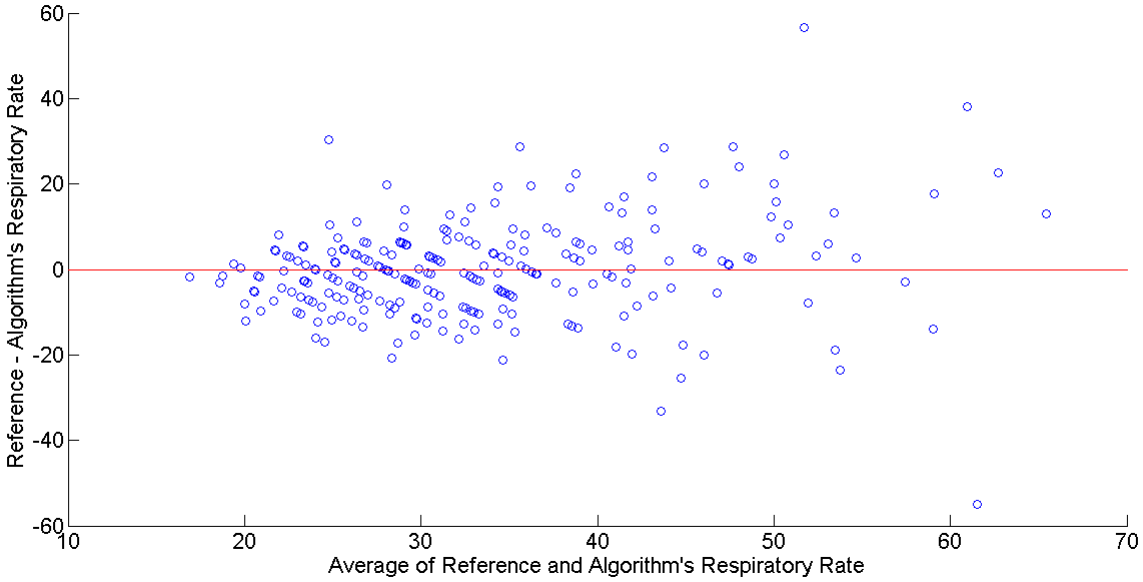


Figure 4.24: Bland-Altman plot for comparing respiratory rate recorded by the triage nurse and the estimate using the hybrid algorithm

4.8 Conclusion

In this chapter, respiratory rate algorithms have been developed and validated, based on the analysis of PPG waveforms acquired from children using a pulse oximeter finger probe. These children were monitored during routine triage in the Emergency Department (ED) at the John Radcliffe Hospital, Oxford, for an average of 2 minutes.

The investigation described in this chapter focused on respiration-induced amplitude modulation (AM) of the PPG waveform. Taking the estimate of the respiratory rate by the triage nurse as the reference, the optimal algorithm was found to have a mean absolute error of 9.1 breaths per minute with 40% of patients having an error within 5 breaths per minute of the reference. Incorporating age information led to an improved algorithm with a mean absolute error of 8.4 breaths per minute with 41% of patients having an error within 5 breaths per minute of the reference.

Chapter 5

Respiratory Rate Estimation with a Finger Probe during Paediatric Triage (Respiratory Sinus Arrhythmia)

The previous chapter presented various signal processing techniques that were developed to exploit the amplitude modulation of the PPG due to respiration. This chapter presents the results of applying those techniques to the phenomenon of Respiratory Sinus Arrhythmia (RSA) in the PPG (see 2.2). This requires the generation of a tachogram to determine heart rate/pulse rate variability, as explained in Section 5.1. Section 5.2 then shows the implementation results using time-domain methods (peak detection algorithms) and Section 5.3 gives the results obtained with frequency-domain methods (the previously developed AR-based algorithms). The performance of the various signal processing algorithms in the four age groups (0-1, 1-2, 2-5 and 5-12 years) is presented in secton 5.4. Lastly, a comparison of the results obtained in the previous chapter (AM) and in this chapter (RSA) is given in Section 5.5.

5.1 Tachogram generation

As mentioned in Section 2.2, respiratory rate estimation using RSA requires the extraction of a time series reflecting beat-to-beat changes in heart rate or pulse rate. Tradtionally, when the ECG

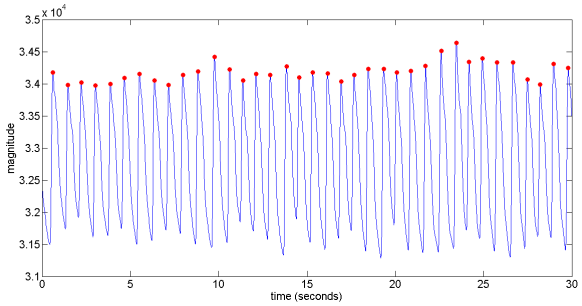
is used to extract a tachogram, an appropriate QRS detector is used to locate R peaks [48] and the time difference between consecutive R peaks (the RR intervals) is then computed. It is possible to extract an equivalent time-series from the PPG waveform by selecting a salient point which can be automatically detected every cycle. The choice of the salient point in previous studies has included the dicrotic notch [13], the point of steepest descent [77], and the systolic peak [124].

In this work, the systolic peak was chosen as the salient point. The systolic peaks were detected as part of the "Complex Peak Detection" algorithm developed earlier and such a choice will therefore provide a consistent framework for comparison of respiratory rate estimation algorithms using either AM or RSA. The dicrotic notch was not a feasible option since it was absent in many PPG recordings. Finally the use of steepest descent as in the work described in [77] was not considered since it required complex preprocessing with empirical mode decomposition which did not seem to provide any added benefit.

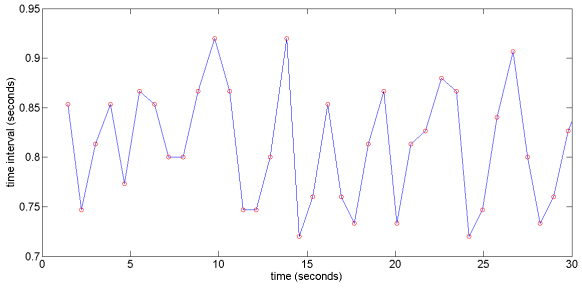
Figure 5.1 illustrates the procedure adopted for the generation of a tachogram using a 30-second PPG segment. The "Complex Peak Detection" algorithm developed previously (see Section 4.4 for details) was utilised to detect the systolic peaks as shown in figure 5.1 (a). Part (b) of the figure shows the time series of intervals (the tachogram) obtained from the salient points identified in part (a), with each interval taken to occur at the second of the two salient points in the interval. To ensure that any artefactually long intervals that appear in the raw tachogram due to missed systolic peaks (most likely during noisy PPG segments) are eliminated before subsequent processing, a 30-point sliding window with a 20-point overlap was implemented and any interval that was 200% times greater than the median interval in that window was discarded. The choice of 200% was made empirically after assessing different thresholds for the reliable removal of artefactual peaks while allowing for variations in pulse rate during the 30-point window. This is similar to the approach described in [34].

Since the tachogram obtained is not an evenly-sampled time series, it is not possible to apply the signal processing algorithms developed in the previous chapter for respiratory rate estimation as these require an evenly-sampled time series. Therefore, cubic spline interpolation was used to resample the tachogram to obtain an evenly-sampled time series. To provide a consistent framework

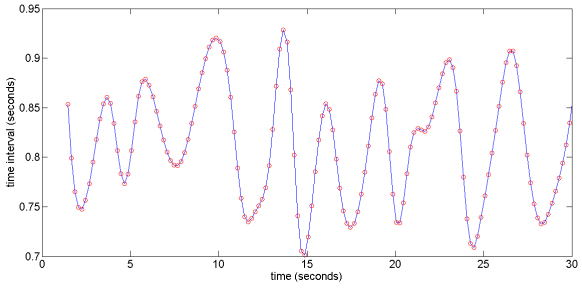
for comparison of AM and RSA-based methods, a resampling frequency of 5 Hz was chosen (the downsampling frequency for the AM signals, see Section 4.5). Part (c) of figure 5.1 shows an evenly-sampled tachogram using cubic spline interpolation obtained from the original 30-second PPG segment.



(a) Raw PPG waveform with systolic peaks identified as shown by the red markers



(b) Raw tachogram with each interval sample taken to occur at the second salient point in the corresponding interval



(c) evenly sampled time-series at 5 Hz using cubic spline interpolation

Figure 5.1: Illustration of tachogram generation using 30-second PPG segment (Patient 0065)

Figure 5.2 gives an overview of the various methods that were used to estimate the respiratory rate using the RSA.

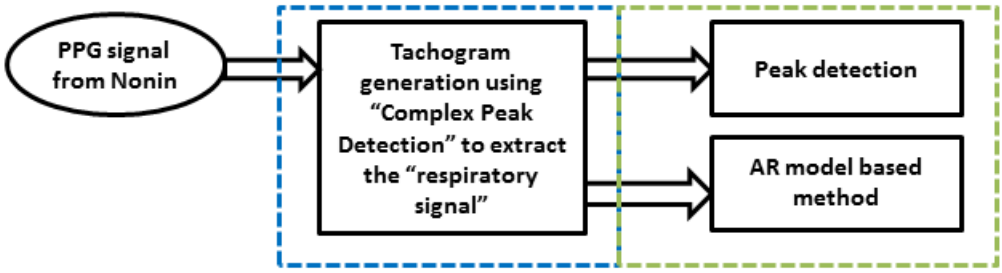


Figure 5.2: Overview of respiratory rate estimation algorithms using RSA. The AR model based methods refer to the 4 methods illustrated in figure 4.3

5.2 Time-domain based approaches

It was not possible to extract the RSA component reliably using the "Basic Peak Detection" algorithm (i.e. detrending, low-pass filtering and 3-point peak detection) because of the erroneous detection of dicrotic notches as systolic peaks. Instead of using the 3-point peak detector on the low-pass filtered PPG signal, the waveform was first differentiated to find its gradient, followed by cubing to emphasize these gradients, and half-wave rectification so that only upslopes are identified. The results of this modified algorithm (previously called the "Complex Peak Detection" algorithm, see Section 4.4 for details) are reported in Table 5.1. To mitigate the problem of incorrectly identified peaks and missed peaks, the results of the "Complex Peak Detection" algorithm were further processed with an outlier rejection algorithm.

Once the "respiratory signal" was acquired with the "Complex Peak Detection" algorithm, the simple 3-point peak detection algorithm ("Basic Peak Detection") was used to determine the respiratory rate. Since no dictrotic notches are present in the "respiratory signal" unlike the raw PPG, there was no need to use the "Complex Peak Detection" algorithm on the "respiratory signal" to estimate the respiratory rate. The results obtained on the 249-patient database are given in Table 5.2. The outlier rejection step during tachogram generation improves the accuracy of the respiratory rate estimation by about 1.5 breaths per minute (in terms of mean absolute error), the greatest improvement being seen with $z = 0.9$ in this case. For further details on the algorithms and how the results were computed, see Section 4.4 of the previous chapter.

Table 5.1: Results on the 249-patient database using Complex Peak Detection Algorithm with no Outlier Rejection Procedure

Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	145 (58%)	193 (78%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	76 (31%)	138 (55%)
Mean Absolute Error (breaths per minute)	10.8	

Table 5.2: Results on the 249-patient database using Complex Peak Detection Algorithm with an Outlier Rejection Procedure ($z = 0.9$)

Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	166 (67%)	212 (85%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	95 (38%)	155 (62%)
Mean Absolute Error (breaths per minute)	9.3	

5.3 Frequency-domain based approaches

The results on the 249-patient database are given in Table 5.3 after applying the AR algorithm with a single step pole search criterion on the RSA signal ("the respiratory signal"), as previously developed in the context of AM. The results of using the two-pole search criterion instead on the 249-patient database are given in Table 5.4. Comparing the two tables, it can be seen that the two-step pole search criterion reduces the mean absolute error by about 1.9 breaths per minute and that the number of patients for whom the estimated respiratory rate is within 5 breaths per minute of the reference respiratory rate increases from 77 (31%) to 101 (41%).

Table 5.3: Results on the 249-patient database using a 7^{th} order AR model based algorithm and a single step pole search criterion

Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	181 (73%)	221 (89%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	77 (31%)	126(51%)
Mean Absolute Error (breaths per minute)	10.8	

Table 5.4: Results on the 249-patient database using a 7^{th} order AR model based algorithm and a two-step pole search criterion

Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	184 (74%)	216 (87%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	101 (41%)	159 (64%)
Mean Absolute Error (breaths per minute)	8.9	

It is possible that the model order chosen for the entire database is not optimal for an individual patient. A poor choice of AR model can lead to pole-splitting. Furthermore, the presence of undesirable low-frequency components can lead to erroneous respiratory rate estimation. Three approaches previously developed to address these issues were to (i) use a peak detection algorithm on the frequency spectrum found from the AR model (instead of tracking poles); (ii) use envelope normalization to mitigate the effect of the undesirable low-frequency components; or (iii) use multiple AR models (see Section 4.6 for details). Tables 5.5, 5.6 and 5.7 give the results of these approaches using the RSA signal.

Table 5.5: Results on the 249-patient database using a 7^{th} order AR model based algorithm using frequency spectrum based respiratory rate estimation

Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	182 (73%)	213 (86%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	100 (40%)	168 (67%)
Mean Absolute Error (breaths per minute)	8.9	

Table 5.6: Results on the 249-patient database using a 7^{th} order AR model based algorithm using envelope normalization

Quality Metric	$y = 5$	$y = 10$
Number of patients with at least one 60-second window's respiratory rate estimate within y breaths per minute of the reference respiratory rate	196 (79%)	227 (91%)
Number of patients whose respiratory rate estimate is within y breaths per minute of the reference respiratory rate	97 (39%)	168 (67%)
Mean Absolute Error (breaths per minute)	8.5	

Table 5.7: Results on 249-patient database using Multiple AR models for respiratory rate estimation

Criterion	Mean Absolute Error	Number of patients whose respiratory rate estimate is within 5 breaths per minute of the reference	Number of patients whose respiratory rate estimate is within 10 breaths per minute of the reference
take the median of all the estimates to be the respiratory rate	8.7	101 (41%)	175 (70%)
take the mean of all the estimates to be the respiratory rate	8.4	100 (40%)	176 (71%)
take the respiratory rate to be the estimate corresponding to the highest peak in the probability density estimate	9.8	91 (37%)	155 (62%)

5.4 Performance of the RSA-based algorithms with age

Table 5.8 reports the mean absolute error in the four age groups for the various methods that had the overall lowest mean absolute error for estimating respiratory rate from the RSA extracted from the PPG. As with the results obtained for the analysis of the amplitude modulation (AM) of the PPG, the mean absolute error is the lowest for the Complex Peak Detection algorithm and the Envelope Normalization based algorithm for younger children (0 to 1 year of age) and the mean absolute error is the lowest for the two Multiple AR models based approaches for older children (5 to 12 years of age).

Method	0 - 1 years ($n = 59$)	1 - 2 years ($n = 40$)	2 - 5 years ($n = 69$)	5 - 12 years ($n = 81$)
Complex Peak Detection based algorithm	12.3	10.0	8.4	7.6
Single Model AR based algorithm (two-step criterion)	12.8	6.6	6.9	8.9
Multiple AR models based algorithm (median-based)	13.5	7.4	7.2	7.0
Multiple AR models based algorithm (mean-based)	12.9	7.0	6.7	7.5
Envelope normalization based algorithm	11.8	8.9	7.1	7.9
Frequency Spectrum based algorithm	13.4	8.2	6.6	8.0

Table 5.8: Mean Absolute Error for both the AR based algorithm and the Complex Peak Detection algorithm in different age groups

5.5 Comparison of RSA-based methods with AM-based methods

Figure 5.3 compares the mean absolute error obtained on the 249-patient database with the various signal processing algorithms developed so far and tabulated in Table 5.9, using the AM and RSA components in the PPG. The algorithms are also compared in terms of the number of patients for whom the estimated respiratory rate is within 5 breaths/minute of the reference in figure 5.4. The results show that the Complex Peak Detection algorithm peforms the poorest in both cases, fol-

lowed by the single-step AR-based algorithm. For both the AM and RSA methods, the addition of the outlier rejection step to the Complex Peak Detection algorithm leads to improved performance. However, the performance of the modified AR-based algorithms with the RSA-based methods is consistenly marginally better than the corresponding AM-based methods.

Method	Label
Complex Peak Detection with no outlier rejection	A
AR model with single step pole search criterion	B
Complex Peak Detection with an outlier rejection step	C
AR model with two-step pole search criterion	D
Envelope normalization with AR algorithm	E
Peak Detection on Frequency Spectrum with AR algorithm	F
Multiple AR models based algorithm(mean based)	G
Multiple AR models based algorithm (median based)	H

Table 5.9: Signal processing algorithms from chapters 4 and 5, with labels for comparing their performance using the AM and RSA components in the PPG

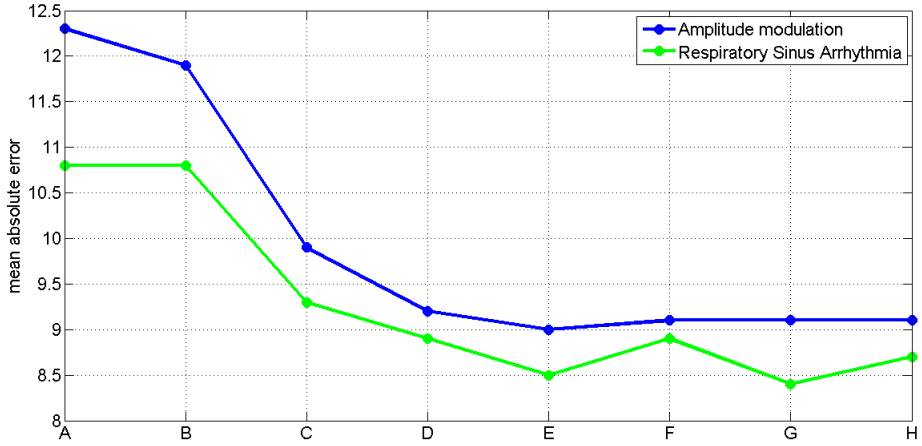


Figure 5.3: Variation of mean absolute error with various signal processing algorithms using the AM and RSA components in the PPG

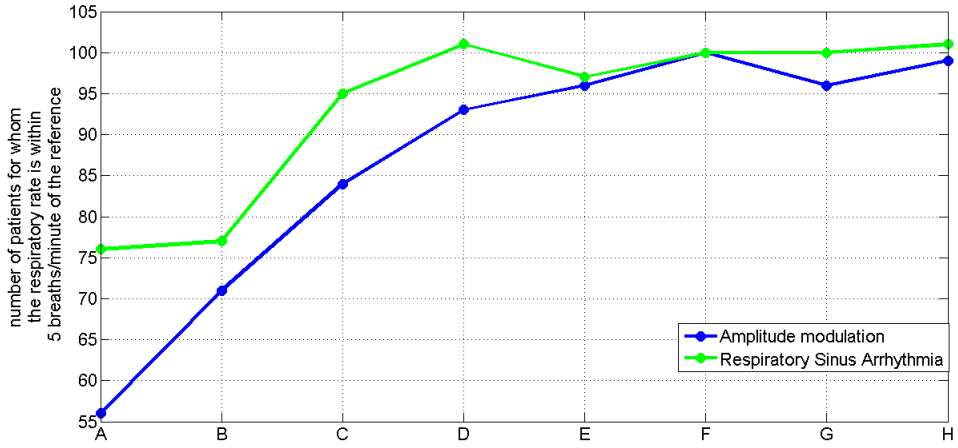


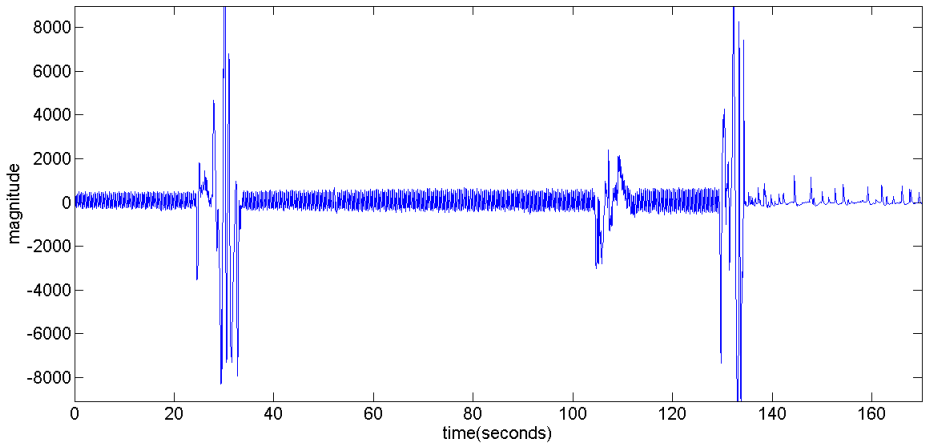
Figure 5.4: The number of patients for whom the estimated respiratory rate is within 5 breaths/minute of the reference with various signal processing algorithms using the AM and RSA components in the PPG

5.6 Conclusion

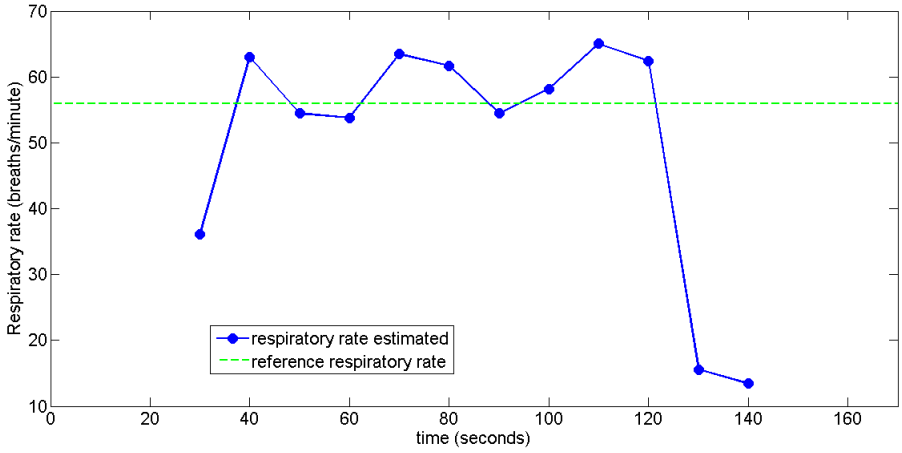
In this chapter, techniques were developed and implemented using the frequency modulation (FM) of the PPG caused by respiration. The optimal algorithm was found to have a mean absolute error of 8.4 breaths per minute, with 40% of patients having an error within 5 breaths per minute of the reference. The distribution of the mean absolute error for both frequency-domain and time-domain approaches in different age groups revealed that the performance of the former deteriorated for younger children (because of the high respiratory rates of younger children).

The general trend in improvement in performance with various signal processing algorithms was found to be similar for both the AM and RSA-based algorithms, with the latter giving consistently better results.

The performance of the various algorithms developed in this chapter is likely to be affected by the quality of the PPG signal which is expected to deteriorate when motion artefact occur. By way of illustration, figure 5.5 (a) shows a PPG recording from one of the patients in the Paediatric Triage study and part (b) of the figure shows the respiratory rate estimated using the Complex Peak Detection algorithm. The respiratory rates are estimated using a 60-second sliding window with the corresponding estimate plotted at the centre of the window. In the figure, the estimate at $t = 30$



(a) PPG recording



(b) Respiratory rate estimated by the "Complex Peak Detection" algorithm showing periods where the estimated respiratory rate is erroneous corresponding to poor PPG quality signal

Figure 5.5: Illustration of how the performance of the respiratory rate estimation algorithm is affected by poor signal quality (patient 0453)

is much lower than the reference respiratory rate (shown in green) as a result of the corrupted PPG segment between $t = 20$ and $t = 40$. As the sliding window moves forward and the PPG segment is no longer affected by motion artefact, the difference between the estimated respiratory rate and the reference decreases. The estimates at $t = 130$ and $t = 140$ are again erroneous, as the PPG segment is again corrupted by motion artefact.The next chapter investigates algorithms for identifying the sections of the PPG signal affected by artefact using a signal quality index.

Chapter 6

Quality Index based Respiratory Rate Estimation

6.1 Introduction

The paediatric triage study included a qualitative assessment of compliance by the triage nurse based on the movement of the child's finger once the pulse oximeter probe was attached to it. The four levels of compliance in order of increasing movement were "compliant", "intermittent movement", "frequent movement" and "non-compliant" (see 3.2.1). The number of patients in each of these four levels is summarised in table 6.1.

Metric	Compliant	Intermittent movement	Frequent movement	non-compliant
Males	83	36	22	1
Females	59	30	17	1
Total	142 (57%)	66 (27%)	39 (16%)	2 (1%)

Table 6.1: Distribution of children in the 249-patient database based on the qualitative compliance metric

The mean absolute error values using the optimal respiratory rate estimation algorithm based on amplitude modulation (see Chapter 4) and respiratory sinus arrhythmia (see Chapter 5) of the PPG signal in the different compliance groups are summarised in figure 6.1. Children in the "non-compliant" category are combined with children in the "frequent movement" category since there

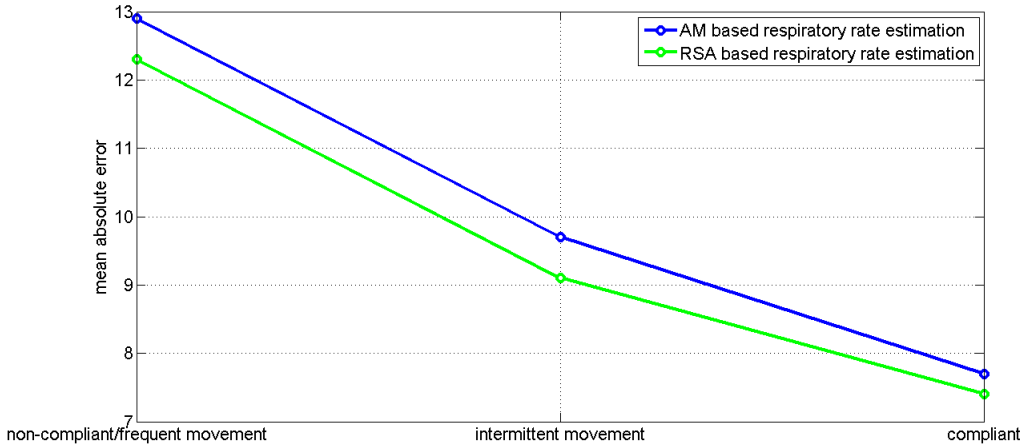


Figure 6.1: Variation of mean absolute error with movement for the best AM based and RSA based respiratory rate estimation algorithms

are only 2 children in the former group. The figure shows that the mean absolute error increases with increasing movement for both the AM-based method and the RSA-based method thereby suggesting that motion artefact causes deterioration in the performance of the respiratory rate estimation algorithms.

6.2 Previous work dealing with motion artefact

Corruption of the PPG signal as a result of motion artefact is a phenomenon which has been recognised in several previous studies [69, 95, 117]. As shown above, motion artefact is a barrier to the efficient and reliable use of the PPG for extracting vital sign information. Previous techniques dealing with motion artefact in the PPG can be divided into two main types: (i) techniques that suppress the motion artefact , (ii) techniques that identify good-quality segments of PPG and subsequently process only those segments.

A number of approaches depend on finding a reference signal to provide input to an adaptive filtering scheme in order to suppress the artefact introduced in the PPG due to motion (typically referred to as the noise signal). Such an approach include the use of accelerometers [113, 5, 41] or an additional optoelectronic sensor [21] to quantify the noise signal. There have also been attempts to find the reference signal without attaching extra hardware. This includes identifying the clean

parts of a PPG signal and then using it to estimate a synthetic reference signal [25]. Finding a reference signal without any extra hardware is also the methodology implemented by the Masimo Corporation which assumes that the venuous blood is a major contributor to the noise signal during motion and then attempts to extract a venuous noise reference which is then used to suppress artefacts in PPG using adaptive noise cancellation [44]. Another technique is to suppress motion artefact is through the use of of a filter bank and a matched filter [69]. The techqniues adopted in [44, 69] use both the red and infrared channels to estimate the reference signal. Other techniques include modelling the motion artefact and taking the inverse of the model to eliminate noise [49], use of wavelet transforms [65], singular value decomposition [112], fourier series analysis [111] and independent component analysis [59].

An alternative approach for dealing with motion artefact is to detect segments of signal that are contaminated by artefact and consider only good-quality segments for further analysis. Signal quality analysis includes the use of higher-order statistics [63], morphological analysis [133, 28] and template matching [153].

An approach based on signal quality assessment is more suitable here than approaches based on suppression of motion artefact as extracting respiratory rate information from only good-quality signals will not only improve the accuracy of the estimate but also increase the confidence on it.

Krishnan et al. [63] used skew and kurtosis measures and the detection of quadratic phase coupling using bispectral analysis to differentiate between clean and corrupted PPG segments. However, their study included data from only 2 subjects and a window of fixed length was used in the analysis. The duration of the time window was not mentioned in the paper. The work by Deshmane et al. [28] and Gil et al. [42] used Hjorth parameters to determine thresholds for making decisions on signal quality. The Hjorth parameters characterise the morphology of a signal by its activity, mobility and complexity [50]. Gil et al. [42] recorded PPG data from 9 paediatric patients and used a fixed window length of 5 seconds to determine Hjorth parameters and associated thresholds. The problem with using a fixed window length is that patients with different heart rates will have a different number of beats lying within a window. This will lead to a large spread of these morphological parameters making it challenging to find an optimal threshold. Deshmane et

al. [28] partly addressed this problem by finding thresholds on adult patients under four different conditions, namely patients with asytole, extreme bradycardia, extreme tachycardia and ventricular tachycardia using 2-second non-overlapping windows. Investigation into the use of Hjorth parameters on the 249-patient database using 1-second and 2-second windows revealed that fixed window lengths could not be used because the heart rate variation is very high (from 53 beats per minute to 197 beats per minute) unlike the expected variation in adult patients. Such fixed window length based approaches will therefore lead to poor signal quality classification for the PPG data of paediatric patients.

One solution to circumvent the problem caused by the variation in heart rate is to assess the PPG signal beat by beat. Sukor et al. [133] used beat-by-beat morphology to analyse the PPG recordings from 13 healthy subjects. In their work, the PPG signal was first pre-processed to detect the beat onset for every beat and then various morphological parameters in the time domain were determined and a threshold-based criterion was employed. Finally, the pool of good-quality beats identified was further processed by computing the average Euclidean distance of each beat from a mean beat shape. Beats were classified as good, poor or bad quality using heuristic thresholds. The work by Weng et al. [153] also computed a mean PPG pulse waveform and then a cross-correlation coefficient was calculated to identify segments of PPG corrupted by motion artefact. However, the beats were grouped in several categories based on heart rate, without any attempt to analyse the PPG waveform on an individual beat basis.

6.3 Algorithm for deriving a Signal Quality Index (SQI)

This work is a modification of the work by Qiao et al. [72] that combines beat-by-beat analysis with template matching to derive a signal quality index. Figure 6.2 shows the flow diagram of the algorithm that was implemented in this thesis.

Beat onset detection

The beat onset detection is based on the work by Zong et al. [158], who describe an algorithm for detecting the onset of arterial blood pressure pulses. A second-order low-pass recursive filter is applied to the input signal, $x(n)$, to suppress high-frequency noise that might otherwise affect the beat onset detection. The transfer function, frequency response, and the difference equation of the low-pass filter are given by equations 6.1 - 6.3.

$$H(z) = \frac{(1-z^{-5})^2}{(1-z^{-1})^2} \tag{6.1}$$

$$|H(\omega T)| = \frac{sin^2(5\omega T/2)}{sin^2(\omega T/2)} \tag{6.2}$$

$$y(n) = 2y(n-1) - y(n-2) + x(n) - 2x(n-5) + x(n-10) \tag{6.3}$$

In this study, the sampling frequency, f_s, is 75Hz and therefore $T = 1/75$ seconds. Since $\omega = 2\pi f$, equation 6.2 becomes:

$$|H(f)| = \frac{sin^2(\pi f/15)}{sin^2(\pi f/75)} \tag{6.4}$$

The gain of the filter at 0 Hz is 25, the 3dB cut-off is at about 5 Hz, and the phase shift is 4 samples (about 53 ms at 75 Hz). Figure 6.3 shows the magnitude and phase response of this filter.

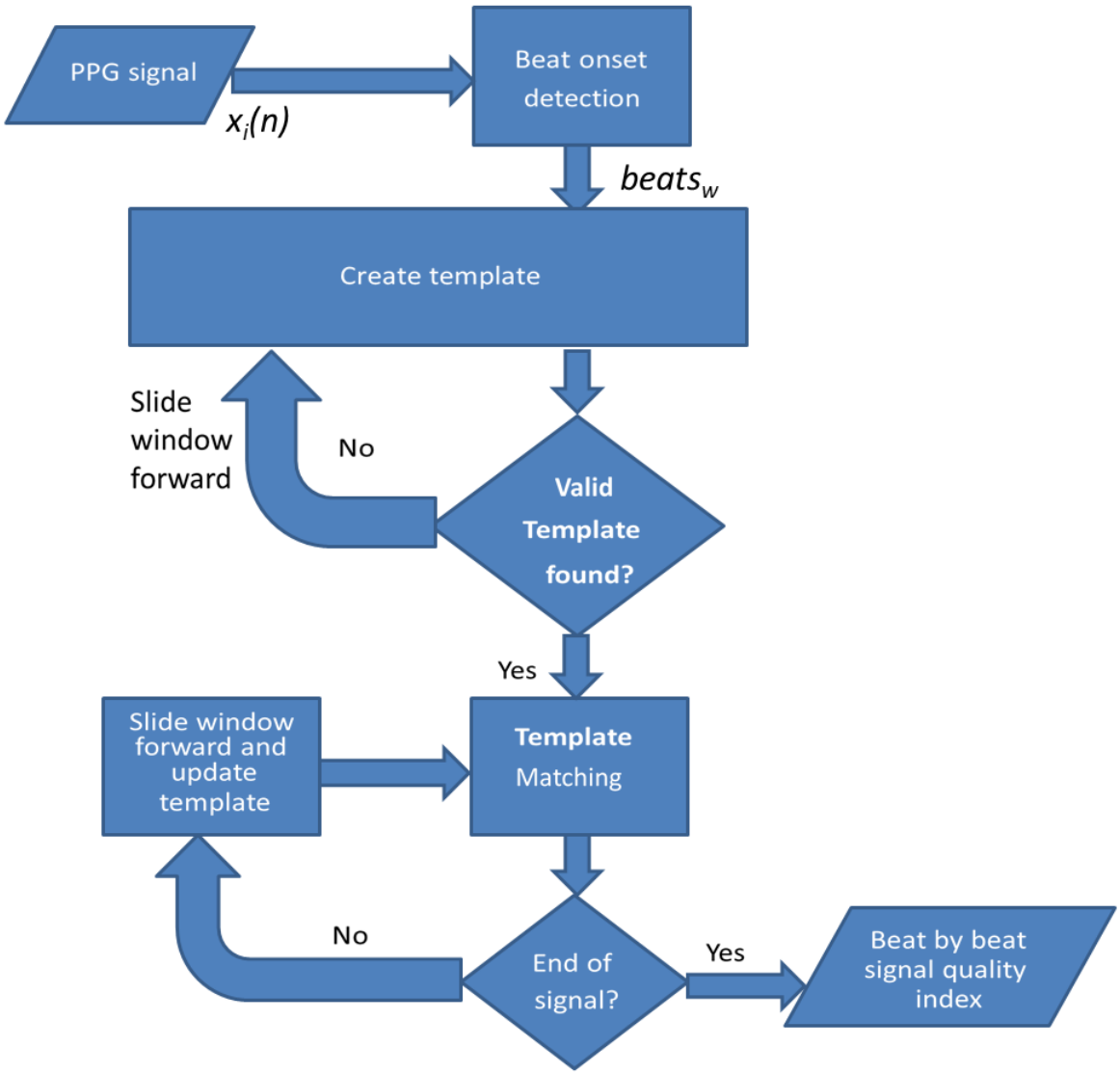


Figure 6.2: Flowchart of the algorithm for determining a Signal Quality Index (SQI)

Subsequently, the slope sum function (SSF) of the filtered signal, y_n, is computed to enhance the upslope of the PPG signal and suppress the remainder of the waveform. The SSF, z_i, at time i is computed using the preceding w samples of the filtered signal y_n using the following equations:

$$z_i = \sum_{k=i-w}^{i} \triangle u_k$$

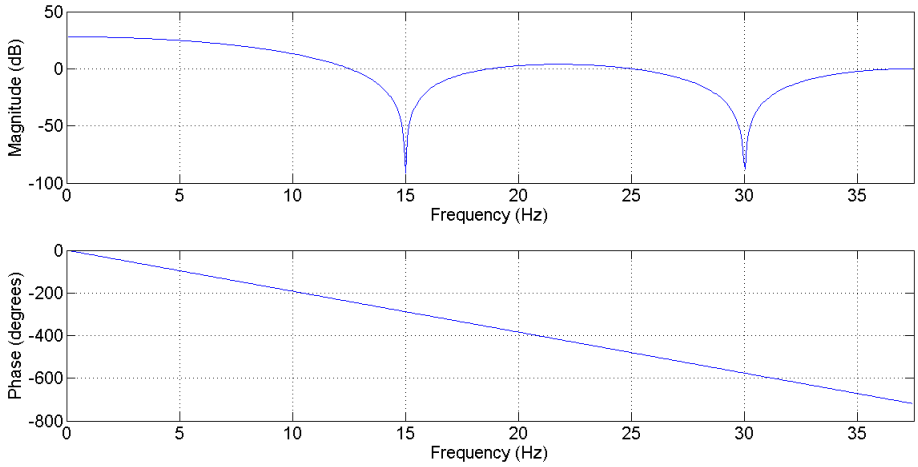


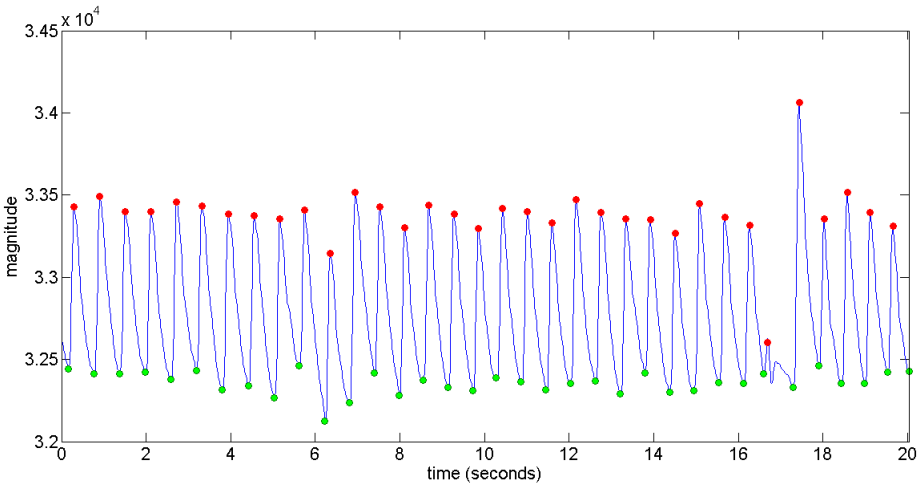
Figure 6.3: Magnitude and phase response of the low pass filter used for beat onset detection

$$\triangle u_k = \begin{cases} \triangle y_k & : \triangle y_k > 0 \\ 0 & : \triangle y_k < 0 \end{cases}$$

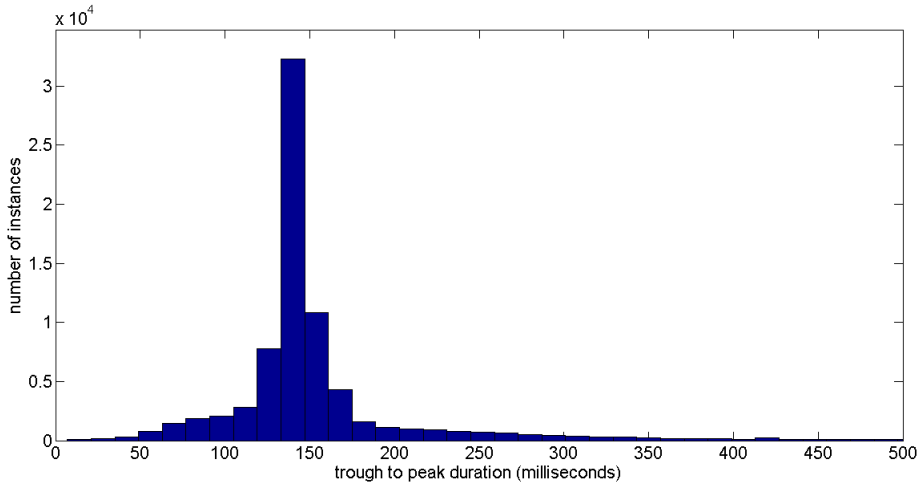
$$\triangle y_k = y_k - y_{k-1}$$

The value of w chosen in this thesis is 11 samples which corresponds to147 ms for the sampling frequency of 75 Hz. This value was the median of the trough-to-peak intervals for all the PPG records in the 249-patient database and it was chosen with a view to maximizing the SSF. For finding the trough-to-peak interval, the complex 3-point peak detection algorithm developed earlier (see Chapter 3) was used. For every peak found in the PPG signal, a corresponding trough was searched for by traversing backwards from the peak detected and taking the minimum point before the previous peak. Figure 6.4 (a) shows a typical 30-second PPG segment with the peaks and troughs detected shown in red and green respectively. Figure 6.4 (b) shows the histogram of distributions of the trough-to-peak intervals for the PPG records in the entire database (over 75,000 trough-to-peak intervals)

Finally, the SSF is passed through a decision rule to allow the detection of beat onset. The decision rule consists of adaptive thresholding to detect SSF pulses of appropriate amplitude followed by a local search strategy within a small window preceding and following the threshold-crossing window. The adaptive threshold is initialised based on the average SSF signal in the first 10 sec-



(a) 20-second PPG segment recorded from Patient 0005, showing the peaks and troughs detected using the 3-point Complex Peak Detection algorithm



(b) Distribution of trough-to-peak intervals over the entire database of 249 patients. The median value was found to be 148 ms.

Figure 6.4: Determination of median trough-to-peak intervals for all 249 patients using Complex Peak Detection, to provide input to the beat onset algorithm developed by Zong et al. [158]

onds and is subsequently updated using the maximum SSF value for each SSF peak detected. A refractory period of 300 ms (corresponding to a maximum heart rate of 200 beats per minute) is applied once a beat onset is identified, to avoid double detection. This algorithm is explained in detail in [158]. Figure 6.8 (a) shows a 30-second PPG segment with the beat onsets detected using this algorithm marked in red.

Template creation

A sliding 30-second window with no overlap is then used to create a template. As can be seen in figure 6.2, the beat onset timings $beats_w$ as well as a 30-second PPG signal, $x_i(n)$ are used in the template creation stage. The algorithm consists of the following steps:

Step 1

find the median inter-beat time interval, μ_{ii} using the beat onset times, $beats_w$

Step 2

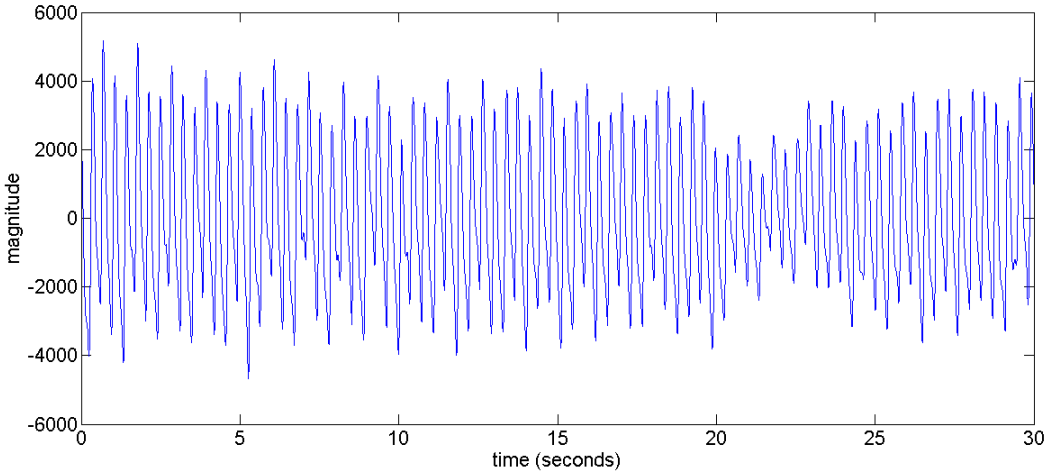
compute the autocorrelation function of the input signal using equation 6.5 [103].

$$R_{xx}(\tau) = \begin{cases} \frac{1}{N}\sum_{n=0}^{N-\tau-1} x(n)x(n+\tau) & : 0 \leq \tau \leq N-1 \\ R_{xx}(-\tau) & : -(N-1) \leq \tau \leq -1 \end{cases} \tag{6.5}$$

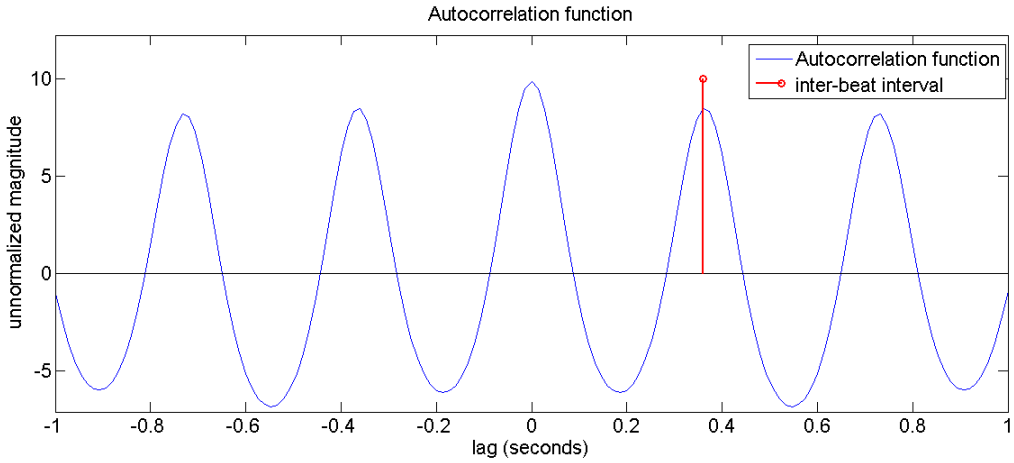
Step 3

Use a 3-point peak detection algorithm to determine the location of peaks in $R_{xx}(\tau)$. The corresponding value of τ, τ_c, that is closest to μ_{ii} is determined. If $(T_{lowerA} * \mu_{ii}) \leq \tau_c \leq (T_{upperA} * \mu_{ii})$, then the period of the template, *cycle*, is assigned τ_c otherwise the time period between the first two peaks in $R_{xx}(\tau)$, τ_{pp}, is taken as the period of the template provided that $(T_{lowerB} * \mu_{ii}) \leq \tau_{pp} \leq (T_{upperB} * \mu_{ii})$, If this condition is not satisfied, then *cycle* is assigned μ_{ii}. The thresholds, T_{lowerA}, T_{upperA}, T_{lowerB}, T_{upperB}, were heuristically determined with the constraint that $T_{lowerA}, T_{lowerB} < 1$ and $T_{upperA,} T_{upperB} > 1$.

The rationale for further checks once τ_c is determined will become clear with the help of figures 6.5 to 6.7. Figure 6.5 shows a 30-second PPG segment and the corresponding autocorrelation function. In this case, the first peak in $R_{xx}(\tau)$ is close to μ_{ii} and is therefore correctly taken to be the period of the template. Sometimes, however, the first peak in the autocorrelation function is not clear and the closest peak to μ_{ii} in $R_{xx}(\tau)$ that the algorithm finds is at a lag of two cycles and will approximately correspond to twice the heart rate. Such a case is illustrated in figure 6.6 where (a) shows a 30-second PPG segment and (b) shows the corresponding $R_{xx}(\tau)$. The value corresponding to the median inter-beat interval is shown in red and the closest peak to μ_{ii} in $R_{xx}(\tau)$ detected by the algorithm is marked by the green circle. If no further checks were applied and the value of τ, τ_c, (marked green in the figure) corresponding to the closest peak was chosen as the template period, then the subsequent analysis would be erroneous. The mean template in this case is shown in figure 6.7 before and after applying further checks. The correct template is generated in the latter case only.

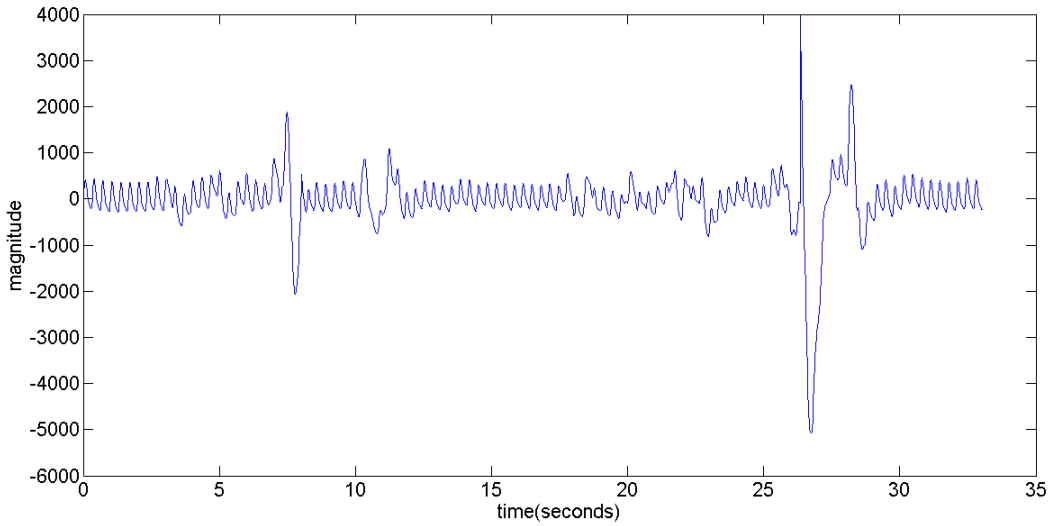


(a) Raw PPG-segment

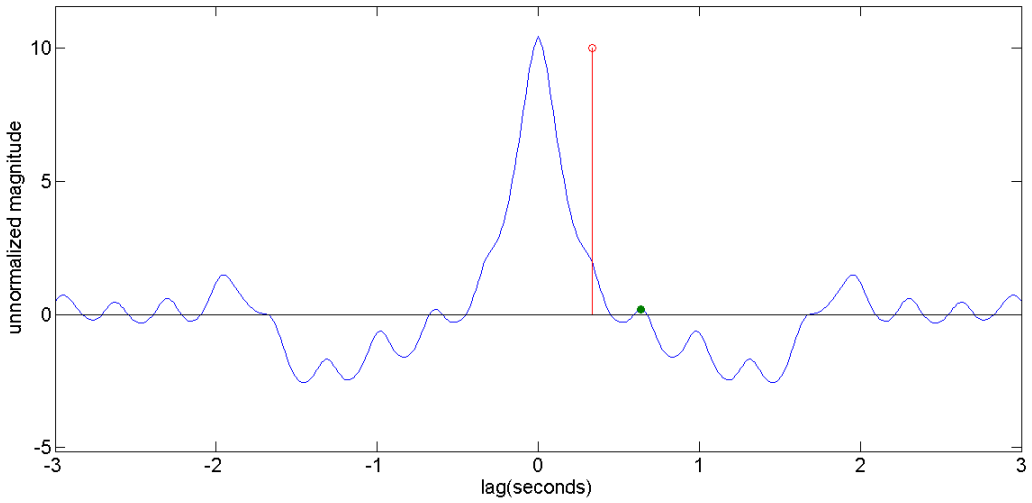


(b) $R_{xx}(\tau)$ of the PPG signal with the median heart rate, μ_{hr} marked in red

Figure 6.5: Illustration of the template creation stage using a 30-second PPG data segment recorded for patient 0037. In this case, the first peak in the autocorrelation is very close to the median inter-beat interval and it is therefore taken as the period of the template.



(a) Raw PPG segment

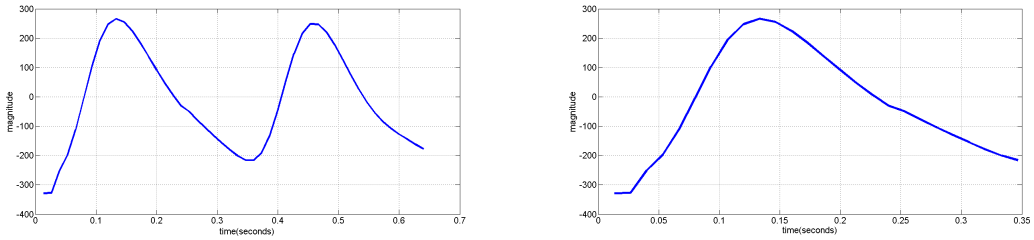


(b) $R_{xx}(\tau)$ of the PPG signal with the median heart rate, μ_{hr} marked in red, and the closest peak to μ_{hr} marked by the green marker

Figure 6.6: Illustration of the template creation stage using a 30-second PPG data segment for patient 0057. In this case, the first peak in the autocorrelation shown by the green marker is almost twice the median heart rate.

Step 4

Check that the period of the template found in (iii) is within acceptable physiological limits. In this case, the limits chosen were from 20 beats per minute to 300 beats per minute.

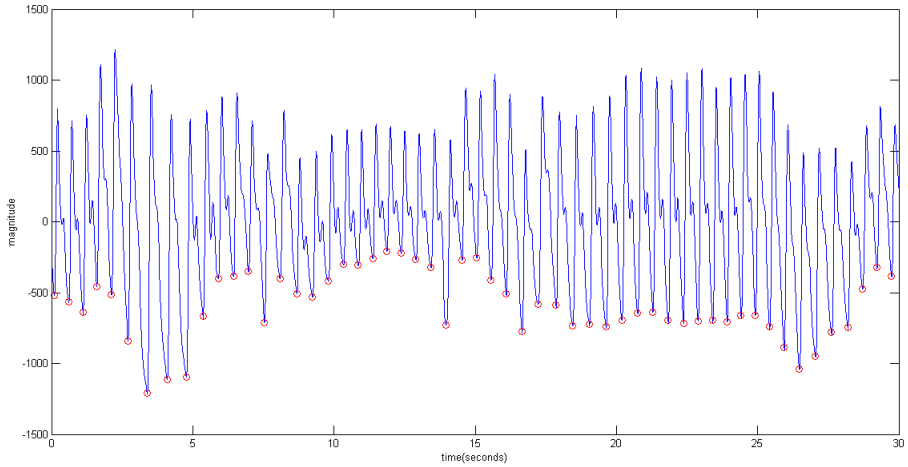


(a) without additional checks (b) after additional checks

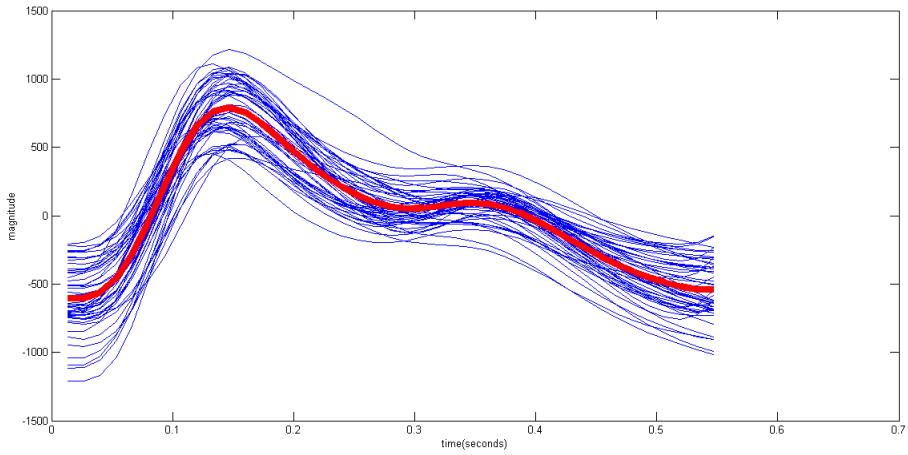
Figure 6.7: Illustration of the importance of additional checks before selecting the period of a template in a 30-second window, showing the two possible mean templates created for a 30-second window of PPG signal

Step 5

Finally, a mean template, $t(n)$, is created by collecting all the beats in the 30-second window, with the start of each defined by $beats_w$ and the end taken to occur at $beats_w + cycle$. This is shown for a representative patient in figure 6.8. Figure 6.8 (a) shows a 30-second PPG segment with the beat onsets, $beats_w$, identified by the beat onset stage while figure 6.8 (b) shows all the beats that are collected from the 30-second window in blue. The mean of these, which is the template for these 30 seconds, is shown in red.



(a) A 30-second PPG segment with the beat onsets, $beats_w$ marked in red



(b) individual PPG cycles in blue and the mean template taken to be the mean of all those shown in red

Figure 6.8: creation of mean template from individual cycles after the template period is determined

Template matching

Once a valid *template* is created, every beat in a 30-second window is compared with the template using the cross-correlation function defined by equation 6.6 below:

$$CC(t,s) = \frac{\sum_{i=1}^{N}(t_i-\mu_t)(s_i-\mu_s)}{\sqrt{(t_i-\mu_t)^2}\sqrt{(s_i-\mu_s)^2}} \tag{6.6}$$

where t_i is the ithsample of the template, s_i is the i^{th} sample of the beat that is being analyzed, μ_t is the mean of the template, μ_s is the mean of the beat and N is the total number of samples in the

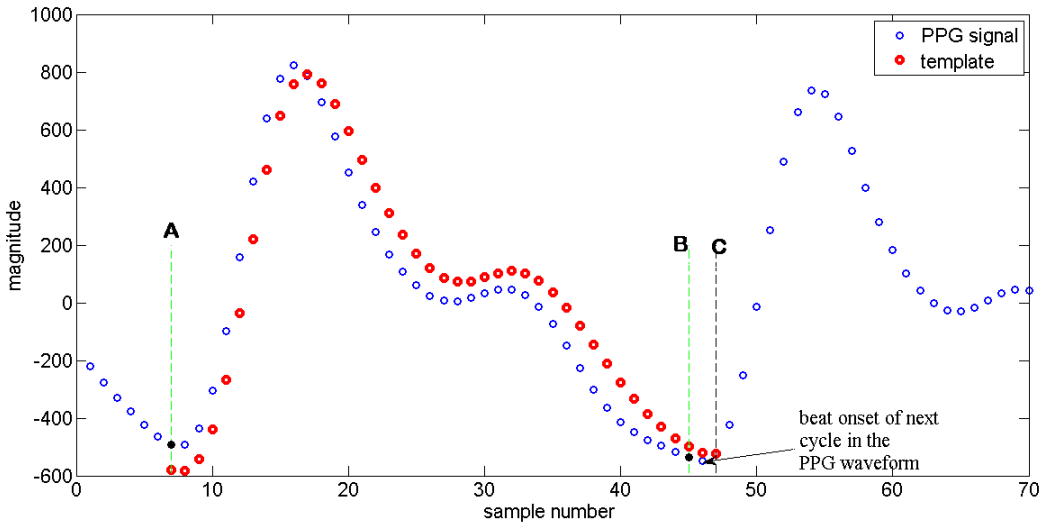


Figure 6.9: Template (red) and a typical PPG segment to illustrate how the ends of the segment in the PPG signal are chosen for template matching. Beat onset points are marked by black markers. The sampling interval is $1/75$ seconds.

template. Figure 6.9 shows a typical PPG segment in blue and the corresponding template in red.

The beat segment, s, is taken to start from the beat onset (marked A in the figure). However, for the end of segment, two different strategies were implemented. In method 1, the end point of s is taken to be the point where the template, t, ends (labelled C in the figure) and the correlation coefficient computed. In method 2, the end point of s is defined to be the sample before the next beat onset (labelled B in the figure). With the latter method, linear interpolation is performed so that the number of samples in s is made equal to N, the number of samples in the template, before the correlation coefficient is computed.

A beat is considered to be of good quality if either of the following conditions are met, otherwise it is labelled as a bad quality signal:

(i) the correlation coefficient for both methods of determining the end of the segment is at least 0.7.

(ii) the mean of the coefficients computed with the two methods is at least 0.8 and the correlation coefficient for the first method (without linear interpolation) is also at least 0.8.

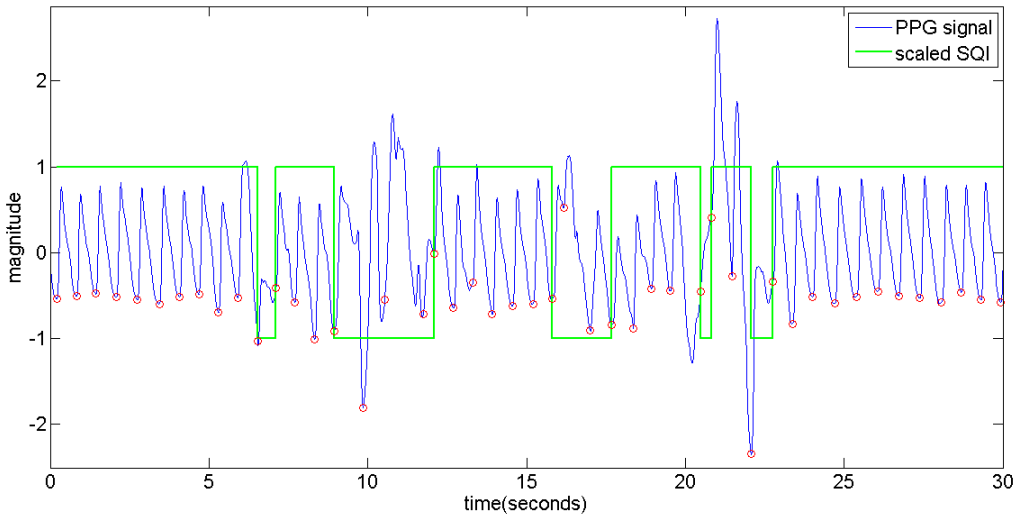


Figure 6.10: Signal quality index for patient 0081 shown in green,PPG signal shown in blue and the beat onset instants shown in red

6.4 Signal quality of patients in the paediatric triage study

Figure 6.10 shows the performance of the signal quality algorithm on a single patient for a period of 30 seconds. The raw PPG signal is shown in blue and the beat-by-beat quality index is used to generate a quality index with a value of +1 for a good quality beat and a value of -1 for a poor quality beat. The beat onsets are also shown in the figure marked in red. It can be seen in the figure that the signal quality index (SQI) drops to -1 in the presence of artefact and stays at +1 whenever the waveform is clearly a PPG waveform.

The algorithm for deriving signal quality was applied to the PPG data record from the 249 patients in the database and the percentage of good-quality waveforms as determined by this algorithm computed. The distribution of signal quality is shown in the histogram in figure 6.11. Over 35% of the patients have PPG data with good signal quality for at least 90% of the time, over half of the patients have PPG data with good signal quality for at least 80% of the time while 85% patients have PPG data with good signal quality for at least 50% of the time.

Table 6.2 provides the mean percentage of good-quality signal in the four classes categorised according to the level of compliance (see table 6.1). The results show that the mean percentage of good quality signals decreases significantly with decreasing compliance as determined by the subjective judgement of the paediatric nurses. These results indicate that the SQI algorithm can be

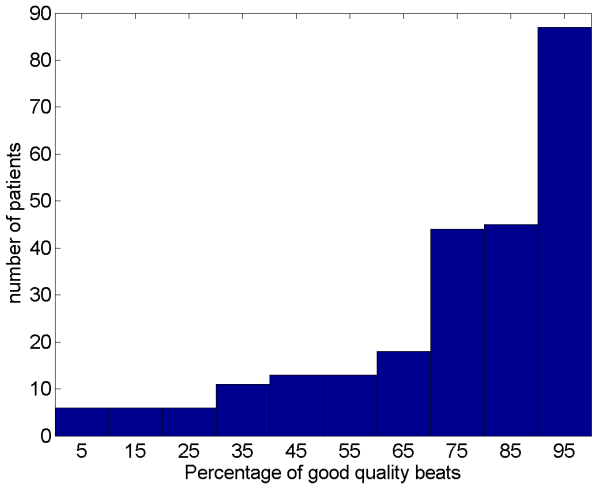


Figure 6.11: Distribution of signal quality as percentage of good quality beats across the 249-patient database

Category	Compliant (n=142)	Intermittent movement (n=66)	Frequent movement (n=39)	Non-Compliant (n=2)
Mean percentage of good quality pulses based on SQI algorithm	84.3%	72.6%	46.8%	26.6%

Table 6.2: Mean percentage of good quality signal in the different classes of patients categorised according to their compliance

used to give a measure of a child's compliance with pulse oximetry recording in a clinical setting.

Similar to the approach described in the previous chapter, sections 6.5 to 6.7 will present algorithms to incorporate computation of the SQI when extracting the amplitude modulation (AM) component of the PPG data. Section 6.8 will then show the implementation results with extracting the frequency modulation (FM) component (due to RSA) in the PPG data.

percentage cutoff	mean absolute error after using data from patients with percentage of good quality signal $\geq T\%$ (number of patients)	mean absolute error after using data from patients with percentage of good quality signal $< T\%$ (number of patients)
$T = 90$	8.2 (87)	9.8 (161)
$T = 70$	9.1 (176)	10.4 (72)
$T = 50$	9.2 (207)	11.9 (41)

Table 6.3: Distribution of mean absolute error using Complex Peak detection algorithm in various groups of patients categorised according to the percentage of good-quality signal

6.5 The use of SQI to improve the accuracy time-domain respiratory rate estimation

There was just one patient in our 249-patient database for whom the percentage of good quality beats was 0%.The distribution of mean absolute errors using the best respiratory rate estimation algorithm based on time-domain analysis (previously found to be the complex peak detection algorithm with outlier rejection, see Chapter 3) for the remaining 248 patients is shown in table 6.3. The table shows that the mean error progressively increases with decreasing percentage of good signal quality. This suggests that embedding signal quality information into the complex peak detection algorithm will improve its performance.

Two strategies were implemented to investigate the improvement in performance after signal quality information is introduced. In the first case, the percentage of good-quality beats in every sliding window was computed. These percentages were then normalized and used as weights to compute a weighted mean of respiratory rate for the entire recording, with a single respiratory rate computed for every sliding window. For example, a window for which the percentage of good-quality beats is over 90% will have a much higher weight than a window for which the percentage of good-quality beats is only 20%. Previously, a simple mean was used to estimate the respiratory rate for the entire recording, since no signal quality information was available. In the second case, instead of computing a weighted mean, the estimate from the window with the highest percentage of good-quality beats was taken as the estimate of respiratory rate for that recording.

Figure 6.12 shows the PPG data (in blue) from a patient who was monitored for more than 3

minutes and the signal quality index (SQI) in green. The reference respiratory rate for this patient measured by the triage nurse was 66 breaths per minute. Without using the SQI, the estimated respiratory rate is simply be the mean of estimates from all the sliding windowsm, 50.2 breaths per minute in this case. After using the SQI to compute the estimate of respiratory rate, the estimated respiratory rate is 65.7 breaths per minute using the weighted mean approach and 68.8 breaths per minute using the estimate from the window with the highest percentage of good-quality signal.

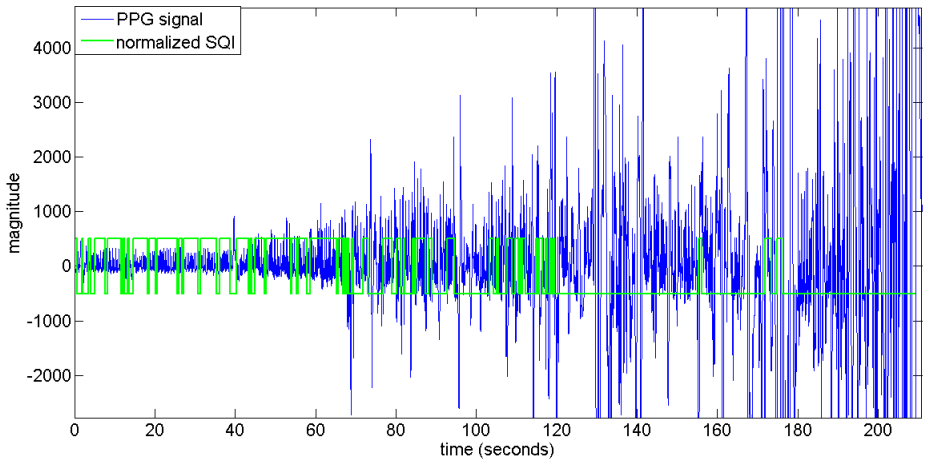
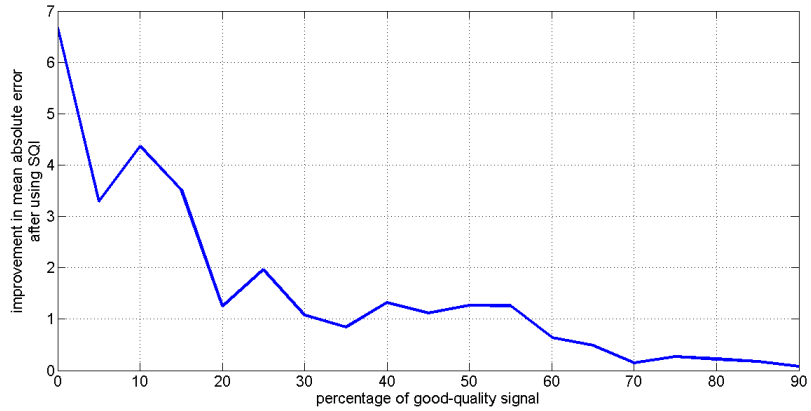


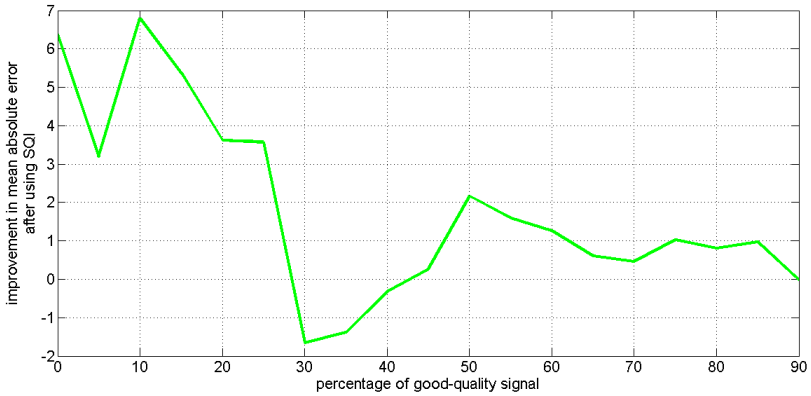
Figure 6.12: PPG signal (blue) with the SQI determined (green) for patient 3214. The respiratory rate estimated by the nurse for this patient was 66 breaths per minute. The respiratory rate estimated by the Complex Peak detection algorithm was 50.2 breaths per minute.

Both approaches were tested on the 248 patients and the results are summarized in figure 6.13. This shows the improvement in mean absolute error as a result of introducing the SQI in various groups of patients classified according to the percentage of good-quality signal. A data point at $x1\%$ on the x-axis is for patients with a good signal percentage which varies from $(x1-10)\%$ to $(x1+10)\%$. For example, a data point at 60% on the x-axis corresponds to data from patients with a good signal percentage between 50% and 70%. Figure 6.13 (a) shows the improvement in mean absolute error after using the weighted mean approach and figure 6.13 (b) shows the improvement in the mean absolute error after using the estimate from the window with the highest percentage of good-quality signal only.

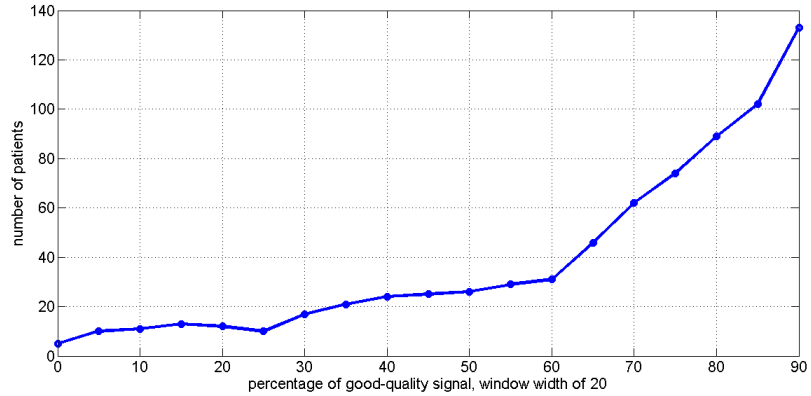
The number of patients whose PPG data is used to compute each of the points in figures 6.13 (a) and (b) is shown in figure 6.13 (c). The results in (a) show that the incorporation of the SQI



(a) Improvement in mean absolute error as a result of computing a weighted mean with the weights computed using the SQI. Patients are grouped in various categories according to the percentage of good-quality signal.



(b) Improvement in mean absolute error as a result of using the estimate from the window with the highest percentage of good-quality signal based on the SQI. Patients are grouped in various categories according to the percentage of good-quality signal.



(c) Number of patients in each window of width 20% that were used to compute the data points in (a) and (b) in the figure

Figure 6.13: Improvement in mean absolute error after embedding SQI information into the best time-domain based algorithm for respiratory rate estimation (i.e. complex peak detection algorithm with outlier rejection)

into the time-domain respiratory rate estimation algorithm leads to an improvement in the accuracy of the estimates, with the extent of improvement progressively increasing as the percentage of good-quality signal decreases. The results in (b) also show improvement for patients with good-quality PPG data for at least 50% of the time. The large variance in the left half of figure 6.13 (b) (percentage of good-quality signal below 50%) is likely to be due to the small number of patients (see 6.13 (c)). The reason for the higher variance in 6.13 (b) with respect to 6.13 (a) is because the former uses the estimate from a single window per patient while the latter uses all windows and hence is able to capture the general trend with fewer number of patients.

6.6 Performance of the frequency-domain respiratory rate estimation algorithms

This section investigates the performance of the three best frequency-domain based respiratory rate estimation algorithms. The three methods were: AR-modelling with a 2-step pole search criterion, peak detection in the frequency spectrum estimated from the AR-model to address the pole-splitting phenomenon, and envelope normalizaton with AR model (see Chapter 3 for more details).

Table 6.4 (a-c) shows the distribution of mean absolute error for the three AR-based respiratory rate estimation algoithms (for model order of 7). It can be seen from the three tables that there is a reduction in the mean absolute error for all three methods as the signal quality decreases. Taking the AR-method with a 2-pole search criterion as the reference, it can be seen that the frequency spectrum based AR-method improves the accuracy of respiratory rate estimation for patients with very high quality data ($T = 90\%$) and does not make any significant difference for patients with a lower percentage of good-quality data. In contrast, the envelope normalization method performs slightly worse for high-quality data but it performs significantly better on low-quality data ($T = 50\%$) in comparison to the AR method with the 2-pole search criterion. The reasons for this are best understood with the help of table 6.5. Figure 6.14 illustrates how envelope normalization behaves in the presence of undesired low-frequency components and undesired high-frequency components

.

percentage cutoff	mean absolute error after using data from patients with percentage of good-quality signal $\geq T\%$ (number of patients)	mean absolute error after using data from patients with percentage of good-quality signal $< T\%$ (number of patients)
$T = 90$	7.1 (87)	10.3 (161)
$T = 70$	8.5 (176)	10.8 (72)
$T = 50$	8.8 (207)	11.6 (41)

(a) Results after using a 7^{th} order AR model based algorithm with a two-step pole search criterion

percentage cutoff	mean absolute error after using data from patients with percentage of good-quality signal $\geq T\%$ (number of patients)	mean absolute error after using data from patients with percentage of good-quality signal $< T\%$ (number of patients)
$T = 90$	6.4 (87)	10.6 (161)
$T = 70$	8.5 (176)	10.9 (72)
$T = 50$	8.7 (207)	11.5 (41)

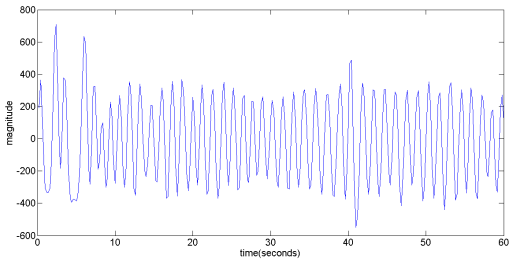
(b) Results after using a 7^{th} order AR model based algorithm with peak detection in the frequency spectrum

percentage cutoff	mean absolute error after using data from patients with percentage of good-quality signal $\geq T\%$ (number of patients)	mean absolute error after using data from patients with percentage of good-quality signal $< T\%$ (number of patients)
$T = 90$	7.7 (87)	9.7 (161)
$T = 70$	8.9 (176)	9.3 (72)
$T = 50$	9.0 (207)	9.3 (41)

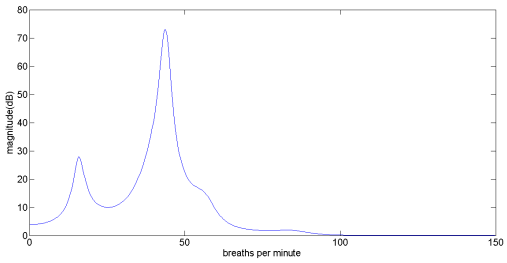
(c) Results after using a 7^{th} order AR model based algorithm with envelope normalization

Table 6.4: Distribution of mean absolute error using various AR based respiratory rate estimation algorithms. Patients are grouped in various categories according to the percentage of good-quality signal.

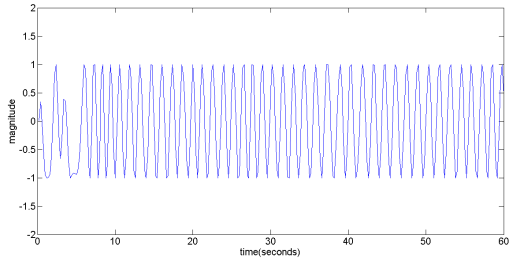
The distribution of the low and high-frequency components was investigated for all patients in this study to help explain why the envelope normalization method performs significantly better with low signal quality and slightly worse with high-quality data. To find the relative frequency components in the low- and high-frequency ranges, the power spectral density of each patient in every sliding window was calculated using the same AR model that was used for determining the respiratory rate (model order of 7) and then divided the frequency band into three regions based on



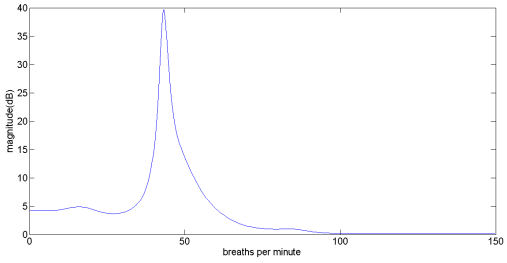
(a) 60-second downsampled PPG segment ('respiratory signal') of patient 402



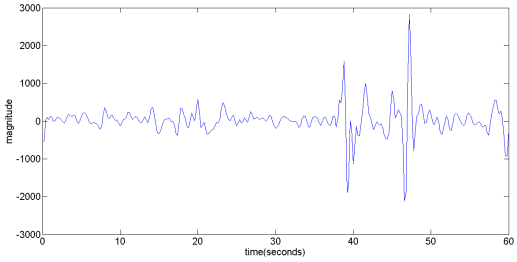
(b) frequency spectrum using poles from an AR model



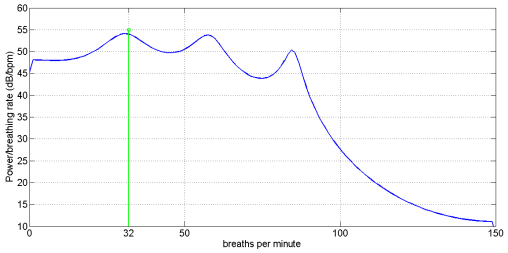
(c) 60-second downsampled PPG segment ('respiratory signal') after envelope normalization



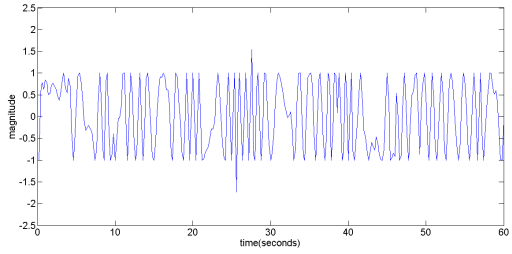
(d) corresponding frequency spectrum estimated from an AR model



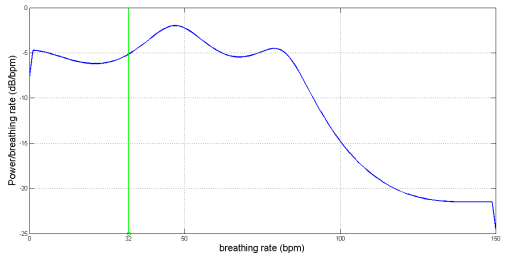
(e) 60-second PPG segment ('respiratory signal') of patient 3135



(f) corresponding frequency spectrum using a 7th order AR model



(g) 60-second envelope normalized PPG segment ('respiratory signal') of patient 3135



(h) corresponding frequency spectrum using a 7th order AR model

Figure 6.14: Illustration of how envelope normalization is able to address the problem of an undesired low- frequency component (a-d) and how its performance is degraded in the presence of undesired high- frequency content (e-h)

Method	Pole-Splitting	Undesired low-frequency components	Undesired high-frequency components
peak detection in frequency-spectrum from AR-modelling	eliminates the problem (see chapter 3 for illustration)	makes no difference	makes no difference
Envelope normalization	makes no difference	addresses the problem by eliminating occasional large deviations in the signal that give rise to undesired low-frequency components (see figure 6.14(a-d) for illustration)	presence of undesired high-frequency components can lead to deterioration in the performance of the algorithm (see figure 6.14 (e-h) for illustration)

Table 6.5: How the AR-based methods deal with the three common issues that arise with the use of AR-based respiratory rate estimation algorithms

the reference respiratory rate. Figure 6.15 illustrates this process for a single window of 60 seconds for patient 0082. The three regions colour-coded in the figure were computed as:

$$A_{low} = \sum_{k=a}^{b} |H_k(e^{j\omega})|$$

$$A_{resp} = \sum_{k=b}^{c} |H_k(e^{j\omega})|$$

$$A_{high} = \sum_{k=c}^{d} |H_k(e^{j\omega})|$$

where $|H_k(e^{j\omega})|$ is the estimate of the power spectral density, A_{low}, A_{resp} and A_{high} are the regions which define the low-, respiratory and high-frequency bands and the limits were determined to be $a = 6$ breaths per minute since a breathing rate below 6 breaths per minute is not expected, $b = Ref - 2$, $c = Ref + 2$ and $d = 150$ breaths per minute. Ref is the respiratory rate determined by the triage nurse. b and c were chosen such that the expected bandwidth of respiratory rate is 5 breaths per minute while d is 150 breaths per minute due to the limit imposed by the Nyquist criterion for a sampling frequency of 5 Hz. These areas were computed using the trapezium rule for numerical integration [120] and were subsequently normalized by the total area under the curve. The means of the estimates of the areas from each sliding window of 60 seconds for every patient

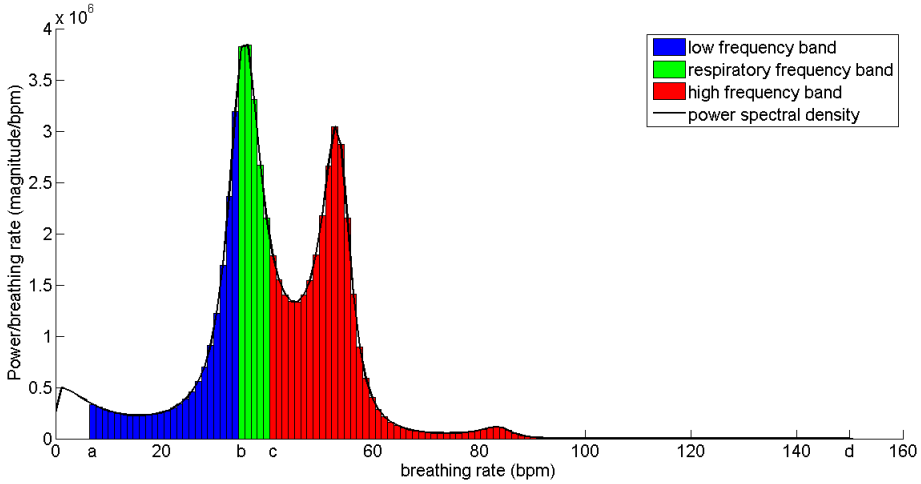


Figure 6.15: Distribution of power in the three frequency bands: low-frequency band, respiratory frequency band and high frequency band

Band	Compliant	Intermittent movement	Frequent movement	Non-Compliant
low-frequency	0.26	0.35	0.41	0.52
high-frequency	0.55	0.47	0.41	0.37

Table 6.6: Distribution of normalized power in the low and high frequency bands in four groups of patients categorised according to their compliance as measured by the triage nurse

were computed.

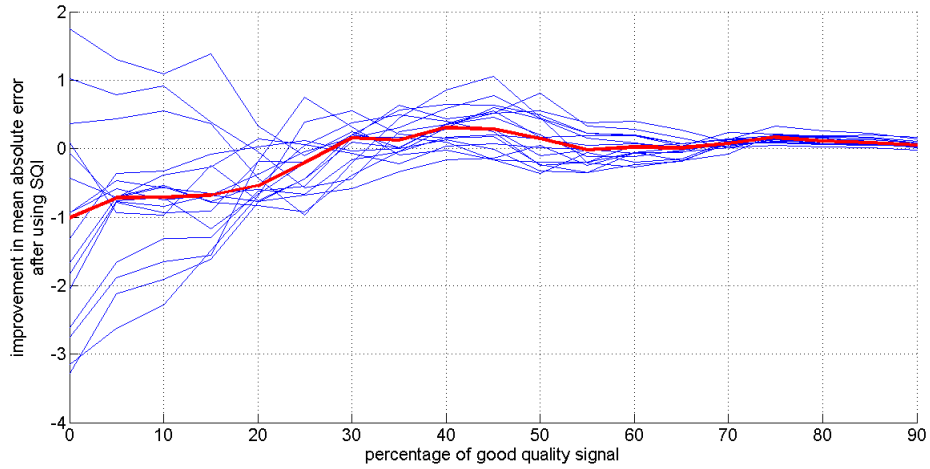
Table 6.6 summarises the mean values for the power in the low frequency and high frequency components for PPG data recorded from the four groups of patients: compliant, intermittent movement, frequent movement and non-compliant. The results show that motion artefact introduces undesired frequency components in the low-frequency band and thus, the algorithm with envelope normalization performs significantly better than without envelope normalization for the PPG data recorded from patients with low compliance. Furthermore, the spectral power in the high-frequency band is higher in the PPG data recorded from patients who are more compliant and thus envelope normalization performs slightly worse in such cases.

6.7 Accuracy of frequency-domain respiratory rate estimation with SQI

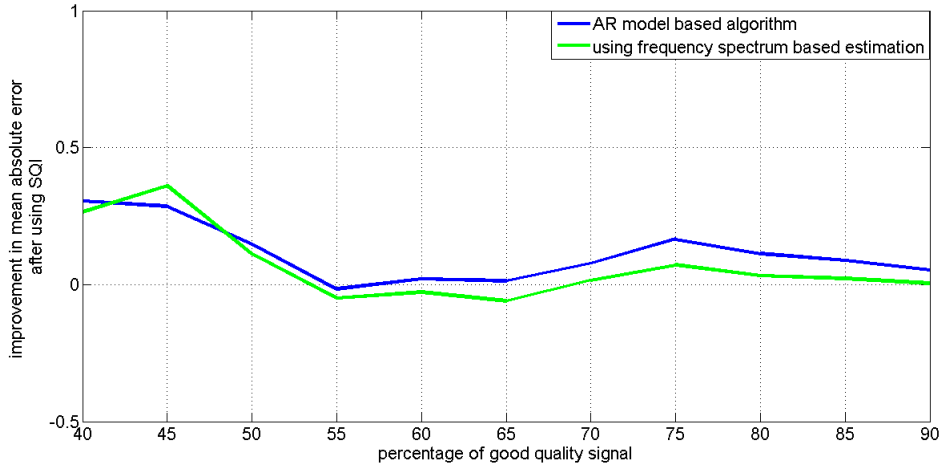
The previous section showed that envelope normalization is useful in the presence of unwanted low-frequency components which arise due to motion artefact thereby reducing the signal quality. Envelope normalization is thus a useful tool in situations where no signal quality information is available and the PPG signal is expected to be corrupted by motion artefacts. However, if signal quality information is available, it is possible to identify signals corrupted by motion artefact thereby eliminating the need for envelope normalization. This section aims to develop algorithms for incorporating the signal quality information into the two remaining AR based respiratory rate estimation algorithms to improve their performance.

Firstly, the two strategies employed with the time-domain based algorithms were investigated to find out if they led to any improvement in the performance of the two frequency-domain based respiratory rate estimation algorithms. Based on the mean absolute error metric, no overall improvement was found if only the estimate from the window with the highest percentage of good-quality signal was used. For the weighted mean approach, figure 6.4 (a) shows the improvement in the mean absolute error for the AR based respiratory rate algorithm using a 2-step pole search criterion for model orders $p = 4$ to $p = 20$ in blue. The median of these is also shown in red. The number of patients for computing each of the points in the figure is shown in 6.13 (c). The figure shows that performance based on the mean absolute error is more sensitive to the choice of model order for sections of PPG data with a lower percentage of good-quality signal leading to a larger variation between the results for various choices of model order in the left-hand most section of the figure. The variability in the results that arises from the use of different model orders is much less for PPG data with signal quality of over 40% as shown in the right half of the figure. However, that part of the result shows that the improvement using SQI is marginal, if any. Figure 6.4 (b) shows the improvement in the mean absolute error when including the SQI for the three AR based respiratory rate estimation algorithms for sections of PPG data recorded from patients with a percentage of good-signal quality of over 40% . The results show that the improvement in the mean absolute

error as a result of the use of the SQI is marginal, if any, for the three AR based methods.



(a) AR model based respiratory rate estimation algorithm using a 2-pole search criterion, model order 4 to 20 (blue) and the median of those (red)



(b) AR model based algorithm with 2-pole search criterion, frequency spectrum based respiratory rate estimation, and AR based model with envelope normalization (median of results from using pole order 4 to 20)

Figure 6.16: Improvement in the mean absolute error after embedding SQI information for the three AR based respiratory rate estimation algorithms

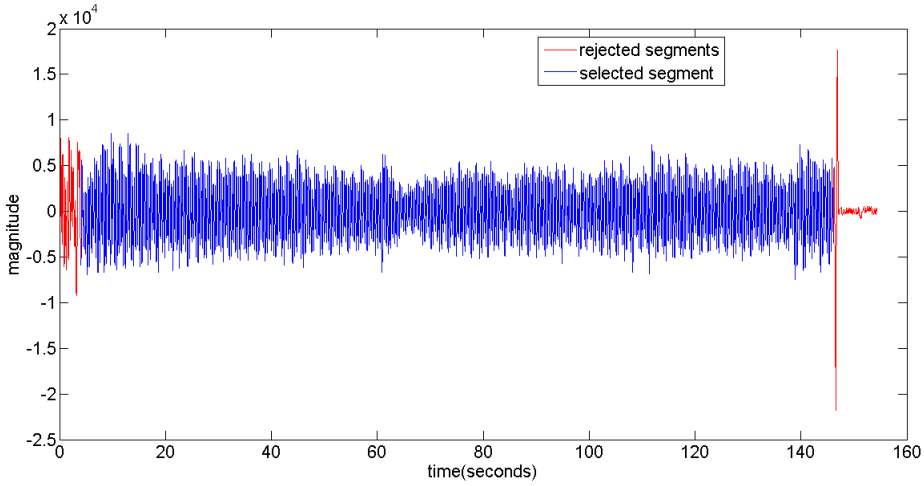
The reason for the difference in the two cases is the difference in sensitivity to artefactual PPG segments. The AR-based methods are more sensitive to the presence of artefactual segments than the time-domain methods. The complex peak detection with outlier rejection method computes a respiratory rate estimate for every beat in a window and then rejects the outliers before finding an estimate for each window. Thus, any estimate from an artefactual segment within a sliding window is more likely to be an outlier and will be rejected in the outlier step. However, the AR-

based method finds the frequency spectrum for every sliding window and then computes a single estimate. As the shape of the spectrum will be affected by the occurrence of artefact, the respiratory rate estimate for that window will also be affected. Thus, a better approach for incorporating signal quality information is to ensure that the input to the respiratory estimation algorithm consists of only those segments of PPG signal that were considered to be of good quality by the signal quality estimation algorithm.

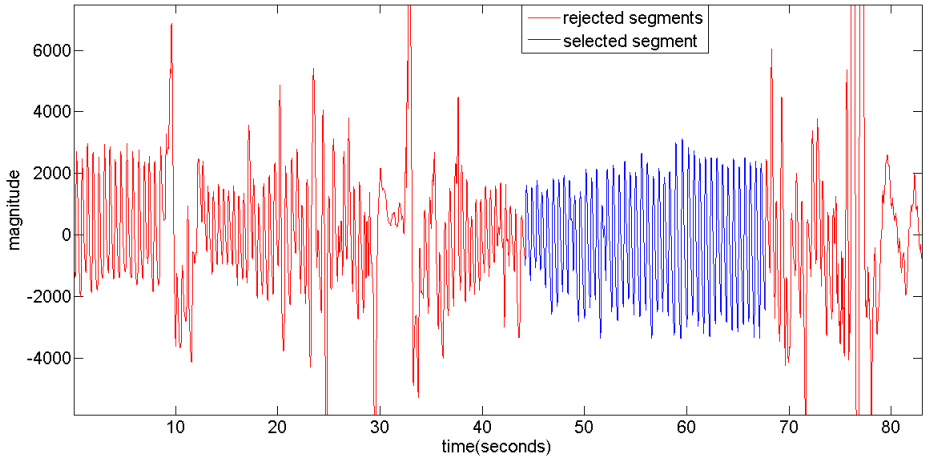
6.7.1 Algorithm based on longest high-quality PPG segment

The signal quality information was used to find the PPG segments of high quality and then the segment with the highest number of high-quality beats (i.e. the highest number of consecutive beats for which SQI=1.0) was selected as the input to the AR-based respiratory rate estimation algorithms. Figure 6.17 (a) illustrates this selection process for the data record from a patient with a relatively high overall PPG signal quality as the selected PPG segment (shown in blue) has a duration of 142 seconds. Figure 6.17 (b) illustrates this process for the data recorded from another patient with a relatively low overall signal quality in which case the selected PPG segment only has a duration of 23 seconds.

There are 178 patients from the 249-patient database (71%) who have a high-quality PPG segment which lasts at least 15 seconds. The PPG data from these patients was subsequently analysed using the two AR-based respiratory rate estimation algorithms. The cutoff of 15 seconds was chosen since this is the minimum duration that a nurse is known to use to estimate respiratory rate during routine triage of patients. Table (6.7) shows a comparison of the two AR-based respiratory estimation algorithms with and without the use of the signal quality index (SQI) using the same two metrics as before: mean absolute error and the number of patients for whom the estimated respiratory rate is within 5 breaths per minute. On the basis of the number of patients whose estimated respiratory rate is within 5 breaths per minute, there is an improvement of between 8% and 9% when the SQI is used with the two AR-based algorithms. On the basis of the mean absolute error however, the use of the SQI does not give improved results. The discrepancy between the two metrics is due to the following:



(a) for patient 0055



(b) for patient 0171

Figure 6.17: Illustration of selection of the longest PPG segment of good quality

Method	number of patients with estimate within 5 breaths per minute of reference (without SQI)	number of patients with estimate within 5 breaths per minute of reference (with SQI)	mean absolute error (without SQI)	mean absolute error (with SQI)
AR-based method with two-step pole search criterion	70 (38%)	83 (47%)	8.4	8.5
AR-based algorithm with frequency spectrum based estimation	71 (40%)	86 (48%)	8.4	8.5

Table 6.7: Comparison of the two AR-based respiratory estimation algorithms with and without the use of the signal quality index (178 patients)

(a) The SQI-based algorithm uses a single window and it is therefore more susceptible to measurement errors. In a large number of cases, where the measurement errors are small, only one of the two metrics, the "mean absolute error", will show degradation in performance. One common problem that can degrade the performance of the SQI-based algorithm is illustrated with an example in section (6.7.1.2).

(b) There is an inherent physiological variability in respiratory rate during the 120 seconds of monitoring which will be discussed in more detail in section 6.11. The "mean absolute error" metric is sensitive to the inherent physiological variability while the "number of patients within 5 breaths per minute" is not under certain physiologically plausible assumptions. Section (6.7.1.1) below discusses this in detail.

6.7.1.1 Effect of inherent physiological variability on quality metrics

Let $f(t)$ be the true respiratory rate of a patient at time t (see figure (6.18)). The function, f, is an unknown function of time which will depend on the individual patient being monitored. Let $t_1$and t_2 be the start and end times when the patient of the monitoring and let t_{n1} and t_{n2} be the times when the nurse measures the respiratory rate. Assuming that the nurse makes no error in measurement, then the respiratory estimate from the nurse would be $f_n(t)$ where $f_n = \frac{1}{t_{n2}-t_{n1}} \int_{t_{n1}}^{t_{n2}} f(t)dt$. The typical value of $(t_{n2} - t_{n1})$ is either 15 or 30 seconds (although there is evidence to suggest that

most of the time it is 15 seconds). The respiratory rate estimated from the AR-based algorithms will be the mean of the estimates over the whole monitoring period, i.e. $f_{al} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} f(t)dt$. The mean absolute error will only be zero when $f_{al} = f_n$. In contrast, the condition, $|f_{al} - f_n| \leq 5$ is less stringent and it is unlikely to be affected by the inherent physiological variability over two minutes.

As an illustration, let us assume that the patient's respiratory rate monotonically increases over the period of monitoring at a constant rate as shown in figure (6.19). Then, $f_{al} = \frac{f(t_2)+f(t_1)}{2}$ and the requirement is that:

$$|\frac{f(t_2) - f(t_1)}{2} - f_n| \leq 5 \tag{6.7}$$

Assuming that the nurse measurement, f_n, and the estimation by the algorithm, f_{al} are both error-free, the difference between the two will depend on when the nurse chooses to measure the respiratory rate. In the extreme case, if the nurse esimates the respiratory rate at the start of monitoring period, t_1, then equation (6.7) reduces to $|f(t_2) - f(t_1)| \leq 10$ and $|f(t_1) - f(t_2)| \leq 10$. A similar result is obtained for the case when the nurse chooses to measure the respiratory rate at the end of the monitoring period, t_2. Thus, it can be concluded that the metric of "number of patients within 5 breaths per minute" is independent of physiological variability provided that the total change in the true respiratory rate is less than 10 breaths per minute, which is likely to be the case for a monitoring period that will typically last around 2 minutes. On the contrary, the mean absolute error is sensitive to the physiological variability and thus, it is possible that it shows a decrease in performance even if there is no actual degradation in the accuracy of the respiratory rate estimation algorithm.

6.7.1.2 Potential problem with the single segment based approach

Figure (6.20) (a) shows a 55-second PPG segment with the section of PPG selected using the SQI shown in blue. Figure (6.20) (b) shows the output of the low-pass filter (pre-processing FIR filter before application of AR-model, see 4.5 for details) with the segment selected shown in blue (A to C) while figure (6.20) (c) shows the corresponding frequency spectrum of this segment in blue. The reference respiratory rate is shown in red in (c). It can be seen that no visible peak exists close to the reference respiratory rate and thus the algorithm will provide an erroneous estimate. The reason

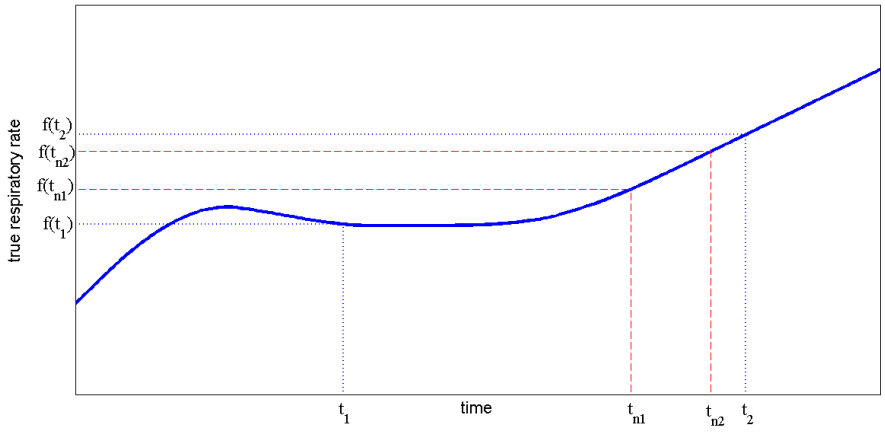


Figure 6.18: Illustration of how errors are introduced in the respiratory rate estimate due to physiological variability during the monitoring period

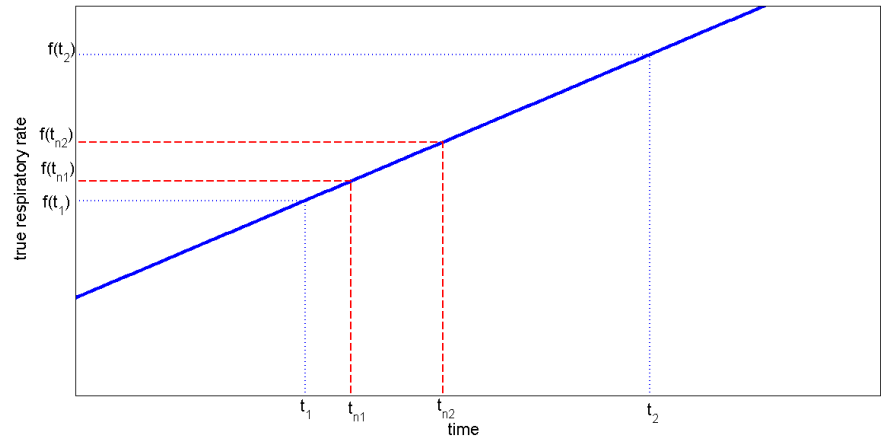
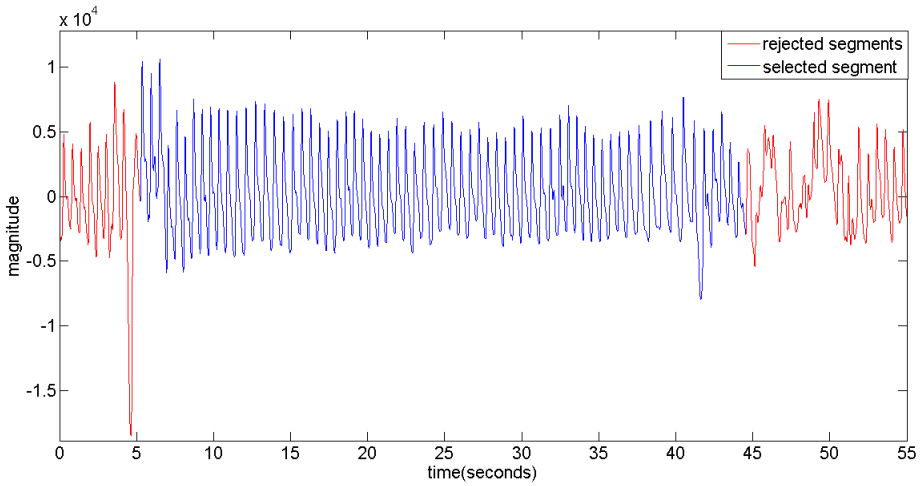


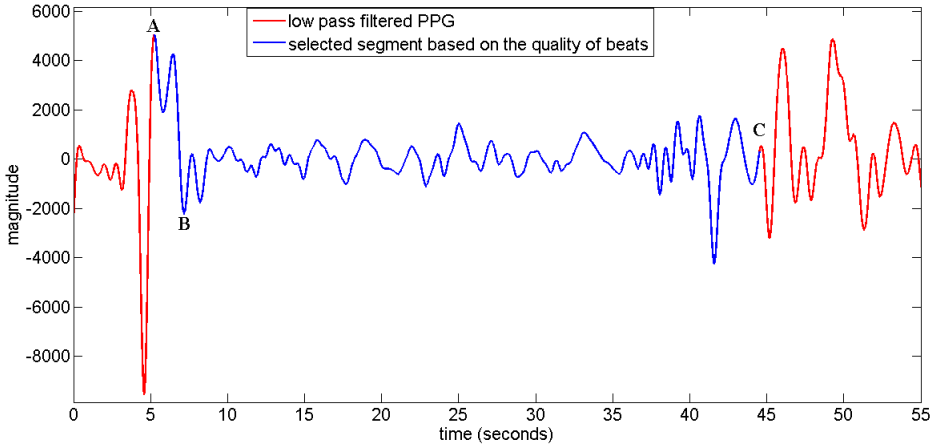
Figure 6.19: Illustration to find the effect of variability of respiratory rate on error metrics used in this thesis

for this error is the section of the PPG waveform from A to B in (a). If this section is removed, then the frequency spectrum of the remaining segment (B to C) is shown in green in (c) where it can be seen that the peak in the spectrum is now close to the reference respiratory rate. In the algorithm without the use of SQI, the estimates were computed over consecutive sliding windows with overlap for the entire length of the monitoring period and rapid transitions at the start or end of the window will only affect a proportion of the windows.

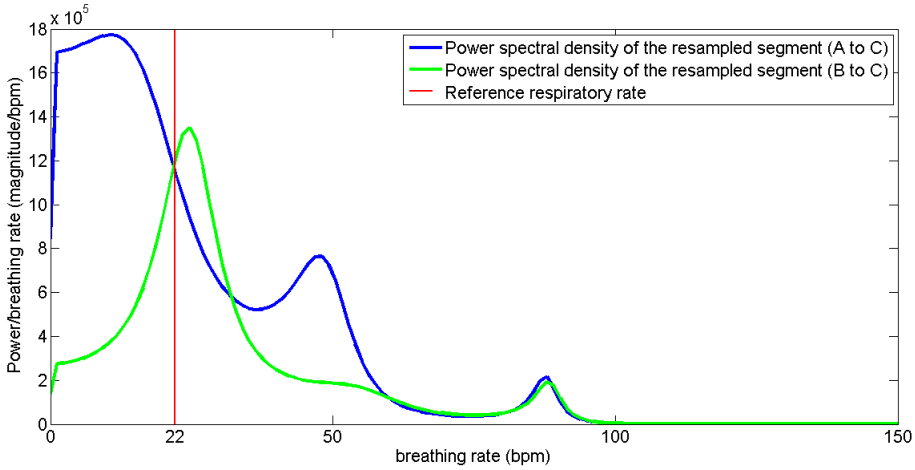
The problem described above can be resolved by using a window function equal to the length of the PPG signal that preserves most of the signal but attenuates the edges. Various window functions and their properties in the context of filtering and spectrum analysis have been extensively studied [101, 57]. For the study presented in this thesis, Kaiser window is an appropriate choice than other commonly available window functions because it allows us the flexibility to choose the amount of attentuation at the edges by changing the parameter β [101] and can thus conveniently be incorporated into a recursuve algorithm where the extent of tapering is progressively increased in order to determine an optimal value of β. In this thesis, an attempt was made to find the optimal value of β in the range of 0 (corresponding to a rectangular window) to 8. Figure (6.21) shows the shape of different Kaiser windows depending on the choice of β. It can be seen that with increasing β, there is an increasing attenuation at the edges of the signal to which the window will be multiplied. For illustration, figure (6.22) shows the improvement in respiratory rate estimation algorithm of the PPG recording from patient 0031 (shown earlier in figure (6.20)) with the use of Kaiser window.



(a) A 55-second of PPG segment after detrending



(b) low-pass filtered ('respiratory signal')



(c) Frequency spectrum of the PPG segment (A-C) in blue, PPG segment (B-C) in green and the reference respiratory rate in red

Figure 6.20: Illustration of how a single window measurement with the SQI can lead to erroneous results (Patient 0031)

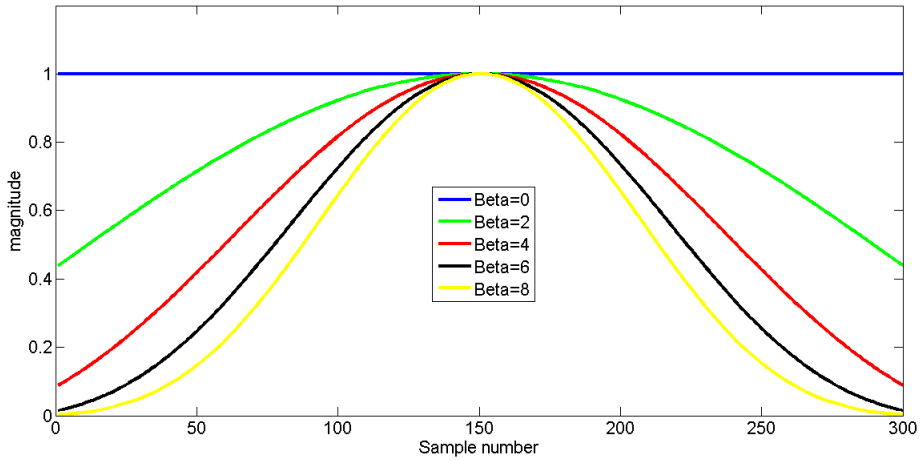


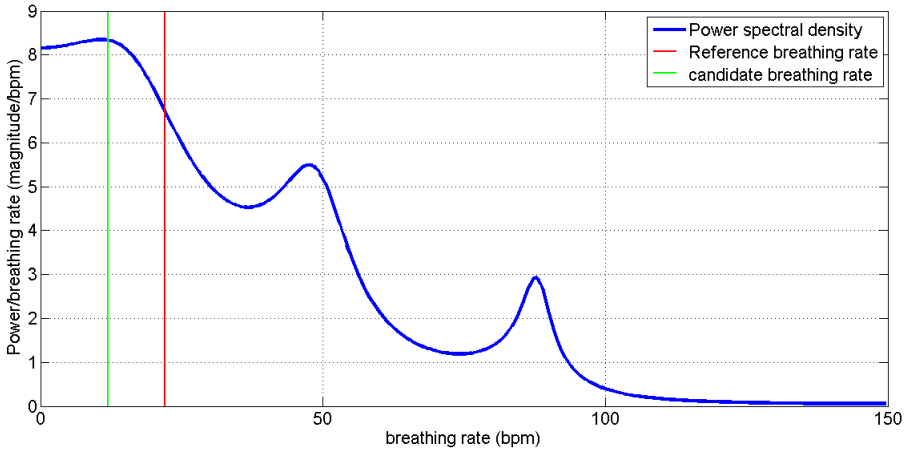
Figure 6.21: Kaiser windows with various choices of β

Table (6.8) shows the results obtained on the 178 patients with multiplying the PPG waveforms with a Kaiser window (β=2.4) before applying an AR model. For comparison, the results without using SQI, and using SQI without a windowing function are also given in the table.

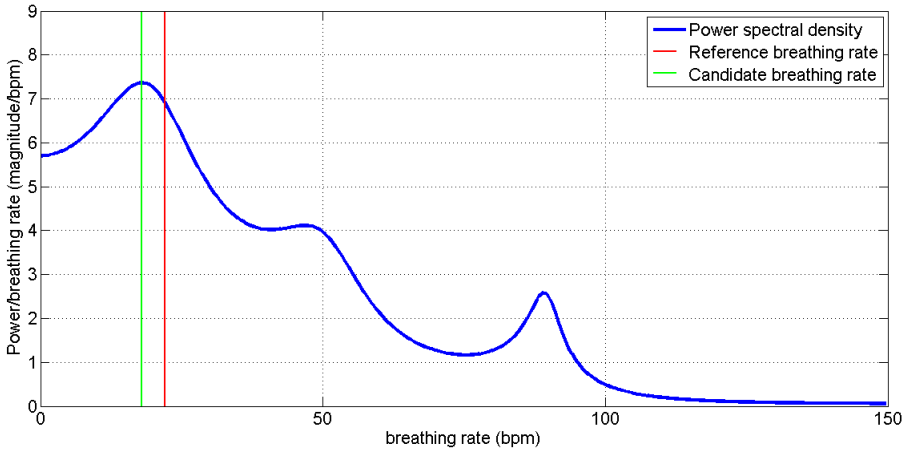
6.7.2 Algorithm based on reconstruction of several high-quality PPG segments

The previous section showed that if the longest PPG segment with high quality signal is selected for respiratory rate estimation, then this leads to an improved respiratory rate estimate. There are, however, only 178 patients (71%) in the 249-patient database with a high-quality signal of 15 seconds or more. It was therefore decided to modify the algorithm so that it identified all high-quality PPG segments of at least 5 beats duration. A high-quality PPG segment is then reconstructed from these segments for subsequent processing. Figure (6.23) illustrates the reconstruction process for the PPG data from one of the patients.

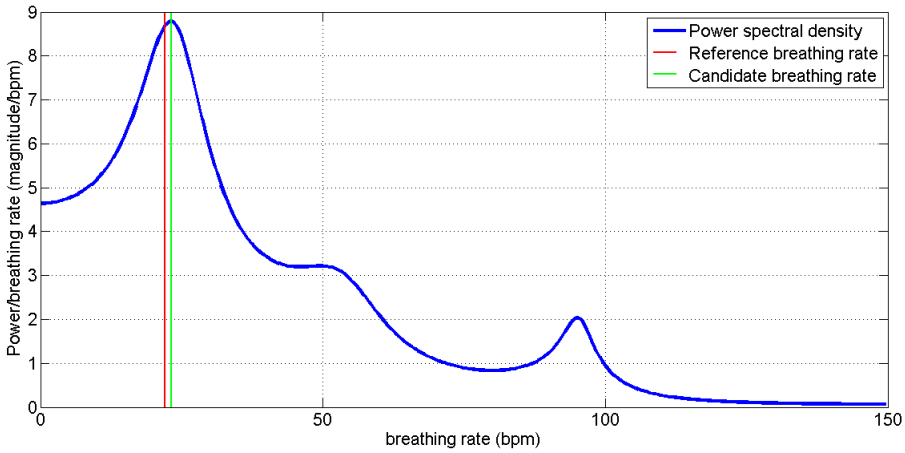
With the introduction of this process, the number of patients with a high-quality PPG signal for least 15 seconds rises to 220, an increase of 17%. Furthermore, since now there is a longer PPG signal for each patient, the sliding window can be used with overlap. As the nurses are known to estimate respiratory rate over a minimum time of 15 seconds, window size of 15 seconds with an overlap of 1 second was chosen to estimate the respiratory rate. The results for the 220 patients



(a) β=0, same as the power spectrum shown in figure (6.20) (c) in blue

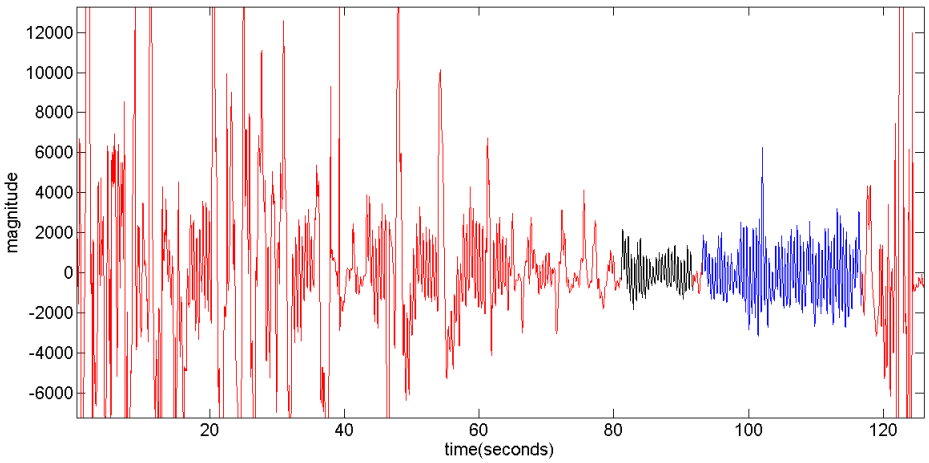


(b) β=2

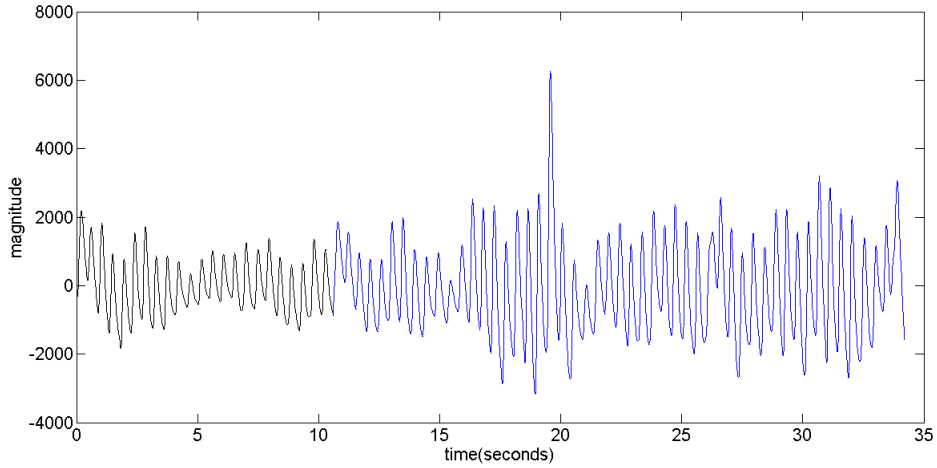


(c) β=4

Figure 6.22: Illustration of how the use of Kaiser window to attenuate edges improve the respiratory rate estimation algorithm (for patient 0031, compare with figure (6.20))



(a) Raw PPG signal with good quality segments shown in blue and black while the segments corrupted by motion are shown in red



(b) reconstructed PPG segment

Figure 6.23: Illustration of how high-quality PPG segment can be reconstructed from a PPG signal corrupted by motion artefacts (Patient 0160)

Method	Mean absolute error	number of patients within 5 breaths per minute of reference
AR-based method with two-step pole search criterion	8.4	70 (39%)
AR-based algorithm with frequency spectrum based estimation	8.4	71 (40%)
Single-window SQI based method with frequency spectrum based estimation	8.5	86 (48%)
Single-window SQI based method with Kaiser window	8.1	86 (48%)

Table 6.8: Comparison of the two AR-based respiratory estimation algorithms with and without the use of signal quality information on 178 patients

who have at least 15 seconds of high-quality PPG signal are summarised in Table (6.9) below. As Table 6.9 shows, using a sliding window with a large overlap eliminates the need to introduce a Kaiser window to mitigate the effects of any rapid transitions at the start or the end of the windows unlike the results in the previous section where a single window was used (see Table 6.8).

6.8 Incorporating SQI to improve RSA-based Respiratory Rate Estimation

Similar improvements were observed when the SQI was incorporated into the RSA-based respiratory rate estimation algorithms. For illustration, figure (6.24) shows the improvement in mean absolute error after incorporating the SQI in the Complex Peak Detection algorithm (with outlier rejection) using RSA (compare with figure 6.13). The error bars represent one standard error. This improvement was achieved after assigning weights to each sliding window and then using a weighted mean of all the sliding windows to compute the respiratory rate (see Section 6.5 for details).

Method	Mean absolute error	number of patients within 5 breaths per minute
AR-based method with two-step pole search criterion	8.6	88 (40%)
AR-based method with frequency spectrum based estimation	8.6	95 (43%)
SQI based method using reconstructed signal	8.3	103 (47%)
SQI based method with Kaiser window	8.3	103 (47%)

Table 6.9: Comparison of the AR-based methods with and without SQI on 220 patients (89%) for whom a high-quality PPG signal is available for at least 15 seconds. A window size of 15 seconds with 1-second overlap is used in this case.

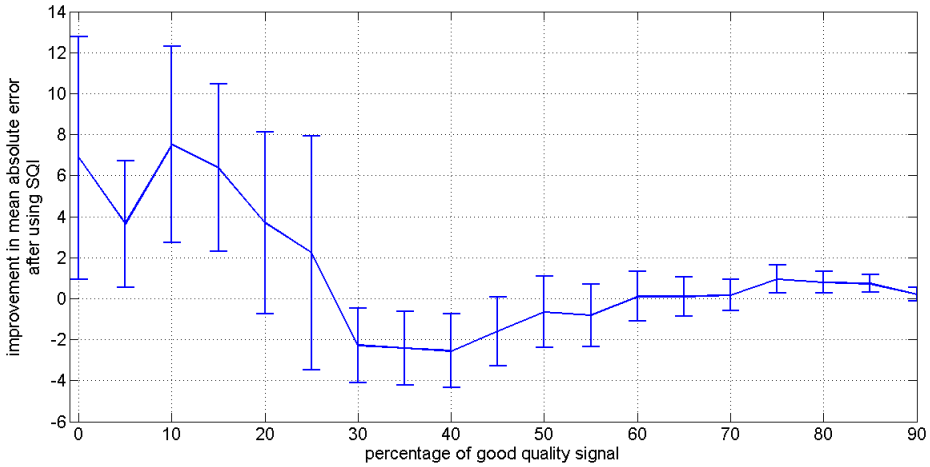


Figure 6.24: Improvement in mean absolute error with SQI after using RSA-based Complex Peak Detection algorithm with outlier rejection

Method	Mean absolute error	number of patients within 5 breaths per minute
AR-based method with two-step pole search criterion	8.4	93 (42%)
AR-based method with frequency spectrum based estimation	8.3	94 (43%)
SQI based method using reconstructed signal	7.7	101 (46%)

Table 6.10: Comparison of the RSA-based AR algorithms with and without SQI on 220 patients (89%) for whom a high-quality PPG signal is available for least 15 seconds. A window size of 15 seconds with 1-second overlap is in this case.

As with Table 6.9 in the previous secction, Table 6.10 provides a comparison to show the improvement in the RSA-based AR algorithms when using a PPG segment reconstructed from high-quality segments. These results are comparable to those presented in the previous section.

6.9 Combining time-domain and frequency-domain approaches

Figure 6.25 shows the performance of the AM and the RSA based algorithms (with SQI) with respect to the reference respiratory rate. This figure shows a significant difference between the frequency-domain and the time-domain approaches when the reference respiratory rate is greater than 50 breaths per minute, with the mean absolute error for the frequency-domain AR algorithms increasing rapidly beyond 50 breaths per minute. This is due to the use of the two-step pole search criterion that first attempts to find a pole in the region corresponding to 12 to 50 breaths per minute. Most of the time, the frequency-domain methods do identify to find a pole below 50 breaths per minute leading to large errors in those cases for which the reference respiratory rate is greater than 50 breaths per minute. The use of a single step pole search criterion gives much worse performance overall, however, and so is not a suitable option to address this limitation.

An alternative approach is to use a simple decison rule whenever the respiratory rate estimated by the time-domain method is greater than 50 breaths per minute and the frequency-domain method estimate is below 50 breaths per minute. The improvements obtained with this data fusion approach

are given in Table 6.11 for the AM-based methods and in Table 6.12 for the RSA-based methods. These combined algorithms will be referred to as "Hybrid AM" and "Hybrid RSA" in the remainder of this thesis.

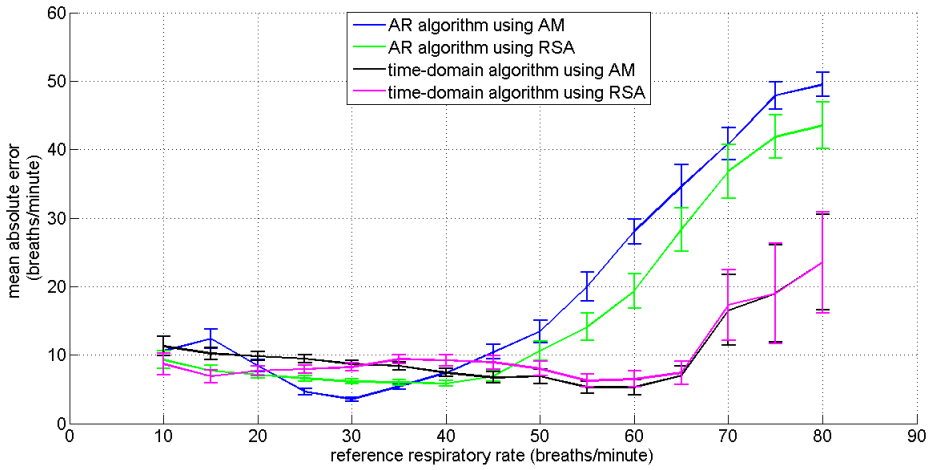


Figure 6.25: Variation of mean absolute error with the reference respiratory rate for the SQI-based time-domain and frequency domain respiratory rate estimation algorithms

Method	Mean absolute error	Number of patients for whom the estimated respiratory rate is within 5 breaths per minute of the reference
time-domain method with SQI	8.9	81 (37%)
frequency-domain method with SQI	8.3	103 (47%)
Hybrid method (AM)	7.6	107 (49%)

Table 6.11: Improvement in the AM-based method after fusing the time-domain and the frequency-domain respiratory rate estimation algorithms

Method	Mean absolute error	Number of patients for whom the estimated respiratory rate is within 5 breaths per minute of the reference
time-domain method with SQI	8.6	91 (41%)
frequency-domain method with SQI	7.7	101 (46%)
Hybrid method (RSA)	7.4	103 (47%)

Table 6.12: Improvement in the RSA-based method after fusing the time-domain and the frequency-domain respiratory rate estimation algorithms

6.10 Combining AM and RSA-based methods based on Pulse Rate

Investigation of the performance of the "Hybrid AM" and the "Hybrid RSA" methods revealed that there is a significant difference in the performance of the two algorithms at low values of pulse rate, as shown in figure 6.27. The reason for this lies in the inherent limitation of the signal processing algorithms rather than physiology. The AM-based algorithms attempt to filter out any spectral components in the PPG signal above the low-pass cut-off, with the expectation that the spectral components due to heart rate (or pulse rate) will be eliminated. However, at low heart/pulse rates (below 90 beats/minute), the cardiac spectral component will not be fully eliminated (see section 4.5 in Chaper 4 for a discussion on the choice of filtering cut-offs for analysing the PPG signals in the database) thereby affecting the performance of the AM-based respiratory rate estimation algorithms. This limitation is not present in the RSA-based algorithms since the tachogram is generated from a single salient point from every beat, there is an upper limit of half the cardiac frequency (imposed by the Nyquist criterion). This also ensures that cardiac spectral components are not present in the RSA signal even if the heart rate is low.

A simple decision rule whereby the estimate from the RSA-based algorithm is selected whenever the pulse rate is below 90 beats/minute and the AM-based algorithm otherwise leads to a significant improvement. It reduces the mean absolute error to 6.6 breaths per minute and increases the number of patients for whom the estimated respiratory rate is within 5 breaths per minute of the reference respiratory rate to 119 (54%). Figure 6.26 shows the scatter plot and the distribution of the mean absolute error for the algorithm.

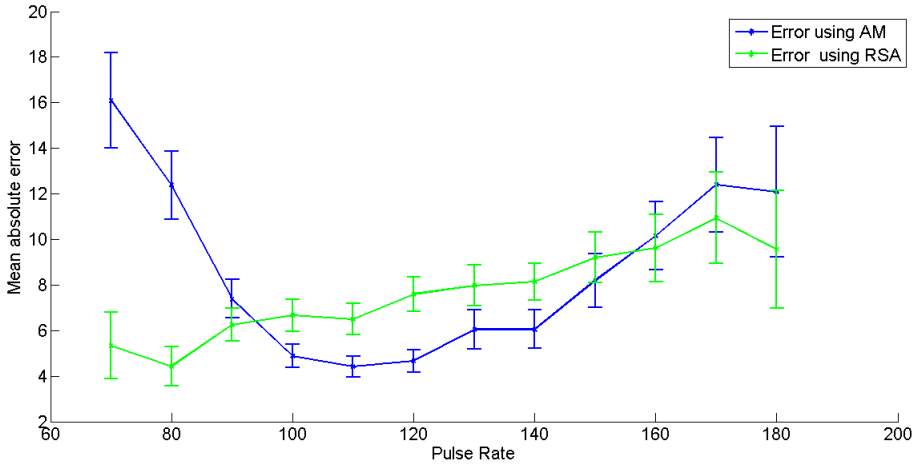
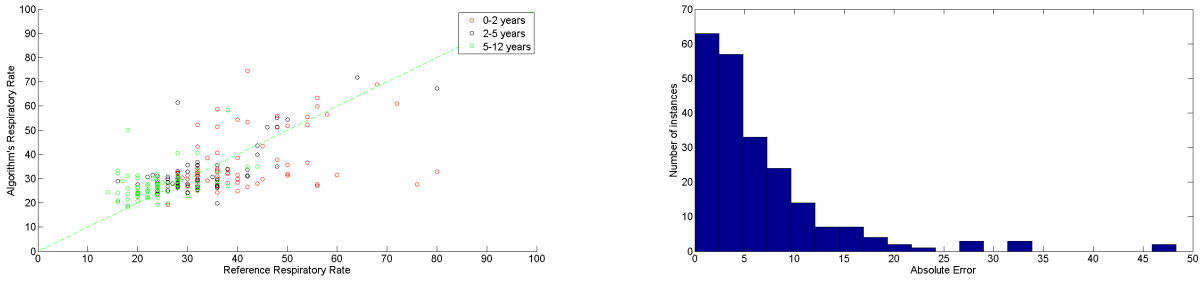


Figure 6.27: Variation of mean absolute error with the pulse rate for the "Hybrid AM" and the "Hybrid RSA" respiratory rate estimation algorithms



(a) (b)

Figure 6.26: Scatter plot of reference respiratory rate versus the respiratory rate estimated by the algorithm (left panel); histogram of mean absolute errors for the estimated respiratory rate for the algorithm developed (right panel)

The variation in performance with pulse rate is independent of the HR/BR ratio (Heart Rate/Breathing Rate) as confirmed by figure 6.28 which shows that the performance of the "Hybrid AM" and the "Hybrid RSA" algorithms are both equally affected by the Nyquist criterion. Due to the limit imposed by the Nyquist criterion, it is not possible to accurately estimate the respiratory rate in those cases where the reference respiratory rate is more than half of the heart/pulse rate. This is confirmed in figure 6.28 which shows the mean absolute error increasing rapidly for both the "Hybrid AM" and "Hybrid RSA" methods as the HR/BR ratio drops below 2.

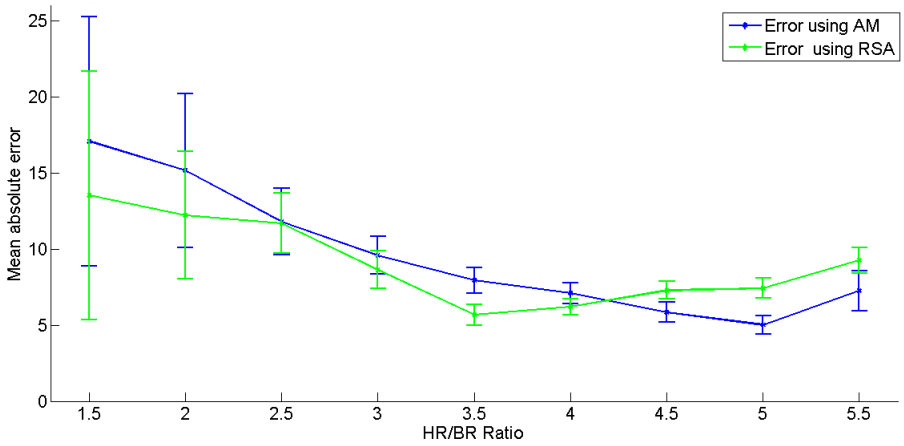


Figure 6.28: Variation of mean absolute error with the HR/BR ratio (i.e. heart rate to respect to reference breathing rate) for the "Hybrid AM" and the "Hybrid RSA" algorithms

Method	Mean absolute error	Number of patients for whom the estimated respiratory rate is within 5 breaths per minute of the reference
Hybrid AM and RSA combined using the Heart Rate	6.6	119 (54%)

Table 6.13: Improvement in the respiratory rate estimation algorithm after combining the RSA and the AM based methods based on pulse rate

6.11 Physiological variability of respiratory rate during monitoring

As with heart rate (see section 3.5 in Chapter 3), it is anticipated that there will be some physiological variability in the respiratory rate during triage. To quantify this variability, PPG recordings from 86 patients who had a mean percentage of high-quality signal of over 90% were analysed in more detail. Figure 6.29 shows the variation in respiratory rate for one of the patients during triage (in blue) and the reference respiratory rate as estimated by the triage nurse. This plot was obtained by using 30-second windows, sliding by 1 second, followed by an 11-point moving average filter (effectively using 40 seconds of data for computing every estimate). This figure shows that inherent physiological variability contributes to the mean absolute error, which depends to some extent on

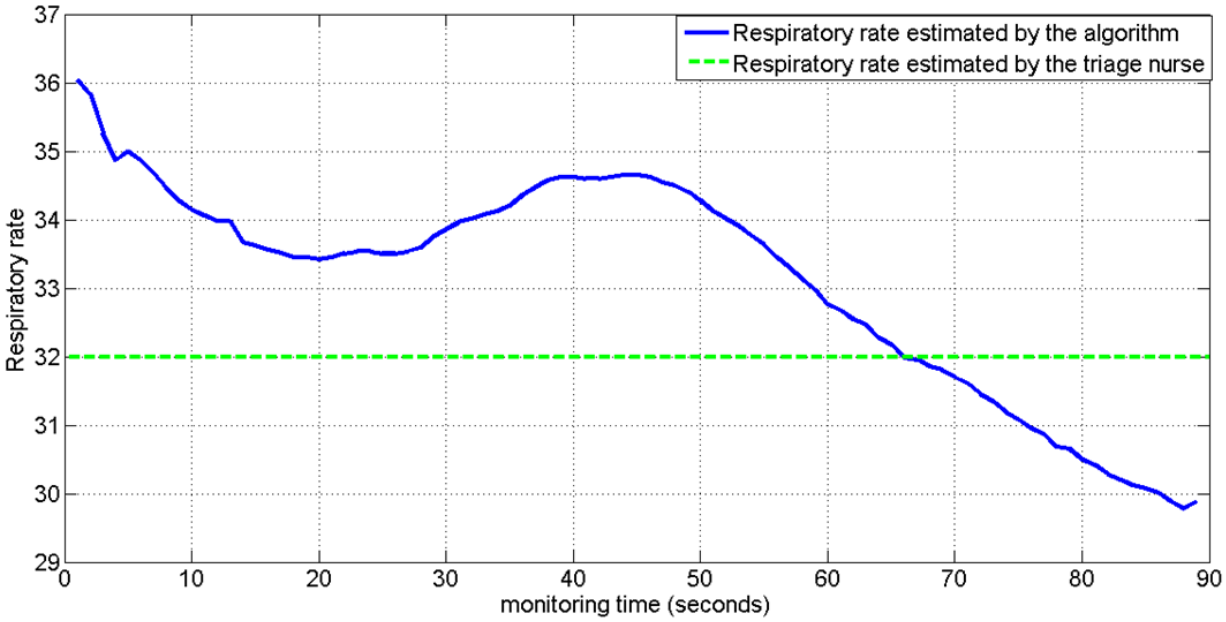


Figure 6.29: Illustration of physiological variability of a single patient during triage (patient 0005)

median filter order,N	mean of $(max - min)$	mean absolute error if respiratory rate closest to the reference is chosen	mean absolute error if respiratory rate farthest from the reference is chosen
11	8.6 breaths/min	2.1 breaths/min	9.4 breaths/min
21	7.2 breaths/min	2.5 breaths/min	8.7 breaths/min
31	6.6 breaths/min	2.7 breaths/min	8.3 breaths/min

Table 6.14: Mean range and the limits of the mean absolute error in respiratory rate estimation due to the inherent physiological variability during triage

when the nurse chooses to measure the respiratory rate.

The maximum and minimum respiratory rates during triage was determined for the 86 patients, after post-processing with a median filter to eliminate outliers. Both the respiratory rate closest and farthest from the reference respiratory rate were identified to get an estimate of the range of variation in the mean absolute error depending on the timing of nurse's measurement. The results, which depend on the choice of median filter used for post-processing, are given in Table 6.14. These results confirm that the inherent physiological variability in respiratory rate during triage contributes significantly to the overall mean absolute error, as the timing of the respiratory rate estimation by the triage nurse is unknown.

6.12 Accuracy of manual counting for respiratory rate estimation

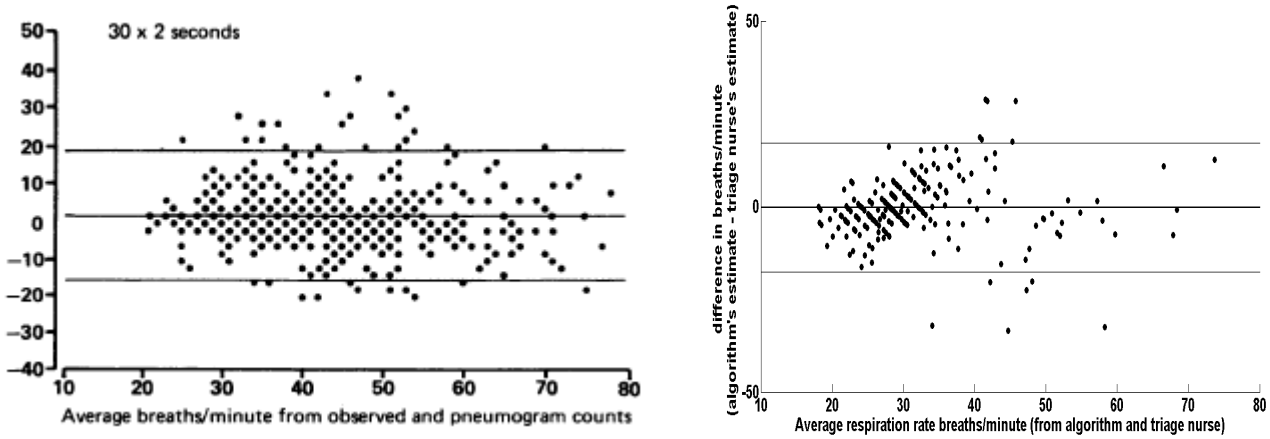
In the study by Simoes et al. [128], the respiratory rate was simultaneously estimated from 97 children using a pneumogram (based on electrical impedance measurement using chest electrodes) and counting chest wall movement. The study excluded children with any chronic illness, or presence of any illness for longer than two weeks, as well as children who were inconsolably agitated. The study found that the accuracy of manual counting depends on the counting period, with higher accuracy attained with a 60-second counting period as opposed to a period of only 30 seconds.

Table 6.15 provides the standard deviation and the mean error found in their study [128] and compares it with the results obtained using the quality index based respiratory estimation algorithm (see section 6.10) developed in this thesis.

Method	Mean difference (breaths per minute)	Standard deviation of the difference
Simoes et al. [128] results using 60-second counting period	1.79	6.85
Simoes et al. [128] results using 60-second counting period	1.72	8.60
automated quality index based respiratory rate estimation results (counting period of 15 or 30 seconds)	0.15	8.80

Table 6.15: Comparison of the quality index based respiratory rate estimation algorithm with the results given in [128] showing that the accuracy of the respiratory rate estimation algorithm is most likely limited by the manual method of counting chest wall movement to obtain a reference rate

The inherent physiological variability discussed in section 6.11 is unlikely to have affected the accuracy of the results given in [128] since the two methods for respiratory rate estimation were synchronised in that study. Moreover, the impedance pneumogram recordings were displayed on a strip chart (and respiratory rate estimated from measuring the interval between peaks); any signal corruption due to movement artefact was manually removed. It is thus likely that the inaccuracy of respiratory rate estimation in the study was primarily caused by human error, for example whenever the child was crying, coughing, yawning, agitated or being fed. This is further corroborated by the investigation of the counting failures in the study [128] which found that respiratory disease did



(a) Reproduced from [128] with permission from BMJ Publishing Group Ltd

(b)

Figure 6.30: Comparison of the results with those obtained in [128]

not increase the number of counting failures. On the other hand, 19% of all counting failures were due to agitation and 9% of all counting failures were due to the observer getting distracted and/or losing track of time.

Figure 6.30 compares the Bland-Altman plots for the results obtained in [128] and for the results reported in this thesis (comparing respiratory rate estimation by an observer counting chest wall movement with the one obtained automatically from a PPG finger probe).

6.13 Conclusion

In this chapter, an algorithm has been developed for deriving a signal quality index (SQI) based on adaptive template matching. This identifies the highest-quality PPG segments for estimating respiratory rate, thereby eliminating the need for subjective analysis (e.g. identifying times when the child is not agitated/coughing/feeding etc) .

A total of 220 patients (88%) had a PPG signal of high quality for at least 15 seconds. The respiratory rate estimation algorithm was improved by fusing both the time-domain and the frequency-domain methods after identifying the limitation of the latter at high respiratory rates. Further improvement was possible after fusing the AM-based algorithm and the RSA-based algorithm according to pulse rate.

The inherent physiological variability in respiratory rate was quantified using PPG recordings from 86 patients with high-quality data. It was shown that the variability in respiratory rate also contributes to the overall mean absolute error as the timing of the respiratory rate estimation by the triage nurse was unknown in this study. Nevertheless, the errors from the best respiratory rate estimation algorithm from the PPG finger probe were found to be as low as had previously been reported using ECG electrodes and impedance pneumography [128], in a study by Simoes et al. in which any potential error due to inherent physiological variability was eliminated by matching the timing of the respiratory rate estimation from pneumography with that of the manual counting of chest wall movement.

The final algorithm developed had a mean absolute error of 6.6 breaths/minute, with 119 patients (54%) having a respiratory rate estimate within 5 breaths per minute of the reference.

Since a commercially available pulse oximeter was used in the ED study, the scope for dealing with movement artefact was limited with the PPG signal acquired from the infra-red channel. While sensor design was beyond the scope of this thesis, it is anticipated that the effect of movement artefact(s) can be mitigated with improvements in sensor design thereby simplifying the task of respiratory rate estimation. For example, using motion-tolerant wearable probes employing MEMS accelerometers proposed in [5] might improve the performance of respiratory rate estimation algorithms during motion in the context of paediatric triage. In addition, currently, commercial oximeter designs only aim to ensure accurate oxygen saturation readings in the presence of motion. It is proposed however, that future investigations should aim to not only improve the measurement of oxygen saturation but also ensure that the respiratory rate information is not lost due to the various signal processing stages introduced to improve the estimation of oxygen saturation readings while dealing with motion artefact and low perfusion.

Chapter 7

Assessment of Vital Sign Disturbance in Childhood Illnesses using Novelty Detection

7.1 Introduction

Chapter 1 identified the two stages of the triage process illustrated in figure 7.1. Chapters 2 to 6 focused on the first stage of this process. The remainder of this thesis will focus on the second stage of the triage process, i.e. the use of vital signs for diagnosis and to help identify children with serious infection.

Vital signs are known to be predictive of serious infection [84, 108, 20, 140, 149]. The National Institute for Health and Clinical Excellence (NICE) recommends that the measurement of temperature, heart rate, respiratory rate and capillary refill time be made part of the routine assessment of children with fever [115]. However, the measurement of vital signs in primary care is still rare [141, 140]. This could partly be due to time constraints, poorly defined normal ranges of vital signs (especially since the normal range of values of some vital signs vary with age [36]), difficulty in achieving accurate readings, and the difficulties faced by clinicians in interpreting combinations of vital signs [141, 140]. In Emergency Departments, on the other hand, vital signs are routinely measured as part of the triage process in combination with complaint-specific flow charts [140] (e.g. the Manchester Triage System, see section 1.1 for details). However, the predictive value of such complicated triage systems is reported to be low to moderate [116, 140, 151, 45]. This could,

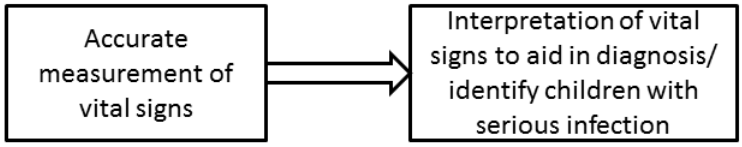


Figure 7.1: the two stages of a paediatric triage system

again, be due to poorly defined normal ranges of vital signs, and difficulty in achieving accurate readings.

In order to address the constraints faced by clinicians during initial assessment of a child presenting to primary care or ED, multivariate analysis techniques were applied to paediatric data to address the following questions:

(i) how do various childhood illnesses disturb the vital signs (investigated in this chapter)?

(ii) how to best identify children with serious infection (investigated in chapter 8)?

Ideally the same paediatric dataset collected from the ED (Oxford) would have been used to answer both questions, but it was not practical to do so because of the small number of patients in that database and because of the extremely small number of children who were admitted with an illness . Instead, another set of databases, the "Nottingham dataset" (described later in this chapter) and the "Walsgrave dataset" and the "Pinderfields dataset" (described in chapter 8), were obtained to develop and test algorithms for identification of illnesses. These datasets are summarised in Table 7.1. As there were large numbers of patients with clearly labelled final diagnosis in this database, the Nottingham dataset was selected to investigate the extent of separation of vital signs between the various diagnosis groups.

Dataset	Population characteristic	Number of children
Oxford	generally unwell children presenting to an ED	272 with various illnesses
Nottingham	generally unwell children presenting to an ED	1403 with various final diagnoses
Walsgrave	children attending a PAU suspected of having an acute infection	568 in four classes based on severity of infection
Pinderfields	ill children admitted in hospital	621 in four classes with four vital signs and a further 363 with only three vital signs

Table 7.1: Summary of datasets used for Data Fusion, PAU=Paediatric Assessment Unit

7.2 The Nottingham Database

This database consists of data collected from 2776 children as part of a study of paediatric presentations to the Queen's Medical Centre Emergency Department in Nottingham, UK between September 2000 and March 2002. Neonates and children requiring immediate resuscitation were excluded from this study. The work in this chapter uses a subset of 1,386 children for whom measurements of heart rate, temperature, oxygen saturation and respiratory rate were available, along with the age and discharge diagnosis.

Diagnosis	Total number	Temperature(C^o)	SpO_2(%)	Heart rate (beats/min)	Respiratory rate (breaths/min)
Pneumonia	95	38.00 (1.04)	94.02 (4.07)	147.41 (27.00)	43.01 (14.20)
Sepsis	1	35.70 (0)	100.00 (0)	138.00 (0)	40.00 (0)
Croup	146	37.09 (1.00)	96.62 (3.98)	137.88 (25.71)	31.26 (8.01)
Bronchiolitis	198	36.99 (1.07)	95.17 (3.79)	155.04 (21.46)	53.28 (13.97)
Viral Wheeze	173	37.04 (1.05)	94.12 (7.54)	149.12 (22.60)	47.63 (12.63)
URTI[1]	229	37.51 (1.10)	96.88 (2.34)	146.23 (24.42)	37.99 (11.69)
UTI[2]	10	37.52 (1.16)	97.10 (1.66)	142.30 (33.58)	39.60 (19.16)
Viral Illness	135	37.41 (1.06)	97.41 (2.12)	142.18 (27.68)	36.48 (12.03)
Gastroenteritis	59	36.83 (0.66)	97.59 (1.97)	131.97 (26.34)	33.85 (9.33)
Febrile fit	18	38.27 (0.98)	95.83 (3.84)	150.94 (26.04)	31.61 (12.99)
Pertussis	3	36.63 (0.06)	98.33 (1.53)	136.33 (4.93)	38.00 (8.72)
Tonsillitis	29	38.11 (1.08)	96.34 (4.25)	137.17 (25.41)	32.55 (11.46)
Asthma	24	36.64 (0.67)	96.13 (2.61)	116.46 (29.15)	31.71 (10.09)
1st presentation of asthma	15	36.78 (0.63)	93.87 (4.07)	133.27 (21.88)	35.47 (13.57)
Acute exacerbation of asthma	118	36.92 (0.67)	93.09 (4.45)	134.76 (24.63)	37.35 (12.97)
Ingestion	2	35.90 (2.12)	94.50 (2.12)	104.00 (56.57)	37.50 (26.16)
Diabetes	3	36.97 (0.25)	97.67 (1.53)	117.67 (26.73)	34.67 (11.72)
Afebrile fit	18	36.29 (0.62)	97.56 (1.62)	114.00 (38.02)	26.89 (8.00)
Cyanotic episode	11	36.95 (1.04)	97.91 (1.64)	155.64 (29.33)	40.18 (11.12)
Choking episode	8	36.60 (0.34)	98.63 (1.69)	143.25 (22.78)	36.50 (7.84)
Other	88	36.70 (0.68)	97.44 (1.82)	124.92 (31.83)	32.07 (12.46)
Well children	20	36.60 (0.42)	97.85 (1.73)	136.60 (30.42)	36.85 (9.95)
Total	1403				

Table 7.2: Mean (standard deviation) of each variable in the Nottingham database for the different diagnosis groups

[1]URTI=Upper Respiratory Tract Infection
[2]UTI=Urinary Tract Infection

7.2.1 Data statistics

Table 7.2 shows the means and the standard deviations for the Nottingham database in the various diagnosis groups for whom the four vital signs were available.

7.2.2 Age correction of heart rate and respiratory rate

Normal values of heart rate and respiratory rate vary with age, with a normal range of 110-160 beats per minute for the heart rate at birth reducing to 60-100 beats per minute by late adolescence and a normal range of 30-40 breaths per minute for the respiratory rate reducing to 15-20 breaths per minute by late adolescence [78]. Consequently, the use of raw heart rate or respiratory rate, without age correction, is not appropriate for assessing the severity of a child's illness.

Fleming et al. [36] used kernel regression, a non-parametric form of curve fitting, to derive centile charts as part of a meta-analysis using heart rate data from 143,346 children and respiratory rate data from 3,881 children. The advantage of kernel regression is that it does not assume any specific analytical form (e.g. linear or exponential) of the curve to be fitted to the data and thus avoids putting excessive constraint on the curve. These centile charts can be used to correct heart rate and respiratory rate for age so that the age dependency is eliminated from these variables. For example, to perform age correction for heart rate for a given child with age x and heart rate y, the kernel regression estimators for heart rate will be used to obtain the mean μ_x and standard deviation σ_x of heart rate at age x. The corrected heart rate y_c can then be calculated using equation 7.1.

$$y_c = \frac{y - \mu_x}{\sigma_x} \tag{7.1}$$

For example, a y_c of 1 would correspond to a child with a heart rate that is one standard deviation above the mean heart rate for a child of that age. The same process is used for the age correction of respiratory rate.

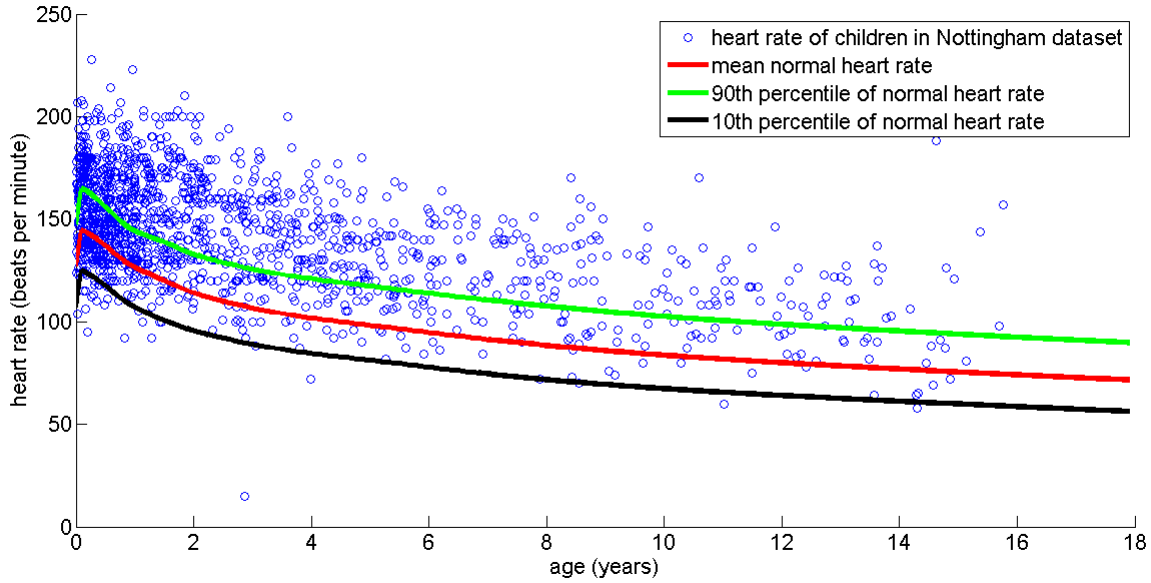
Figure 7.2 (a) shows the distribution of heart rates in the Nottingham database with age. The mean and the 10th and the 90th percentiles of heart rates for normal children [36] are also shown in the figure. It can be seen that the heart rates follow the general trend with age, reported in [36],

for healthy children. It can also be seen that the mean heart rate of children in the Nottingham database is higher than the mean value for healthy children which is expected since the children in the former category are those who attended an emergency department in hospital and therefore many of them are unlikely to be healthy. After applying equation 7.1 to each of the data points in the figure, an age corrected heart rate is obtained for every child in the Nottingham database. The corrected heart rate (i.e. the number of standard deviations the heart rate is away from the normal value for that age) is plotted with age for the 1403 children in the Nottingham database and shown in figure 7.2 (b). A similar trend is found for age-corrected respiratory rates shown in figure 7.3.

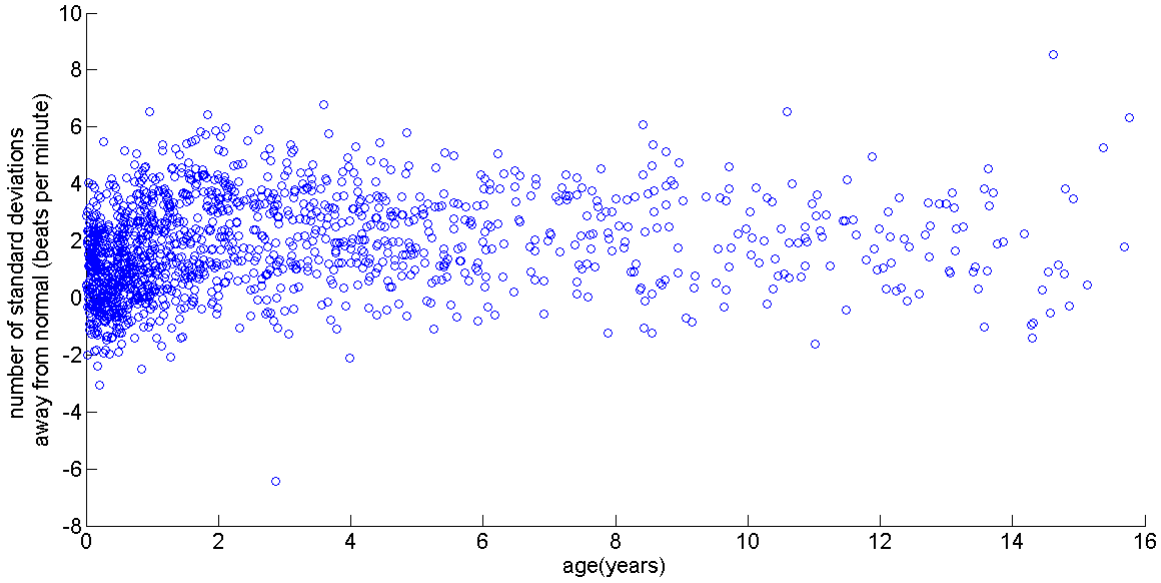
Table 7.3 provide the mean and standard deviation of corrected heart rate and respiratory rates along with the mean age and the age range for the diagnosis groups which had at least 10 patients for subsequent analysis.

Label	Final diagnosis	Number	Mean age (months)	Age range (months)	Corrected heart rate	Corrected respiratory rate
A	Pneumonia	95	45.3	0.2-190.9	2.55 (1.62)	3.41 (3.35)
B	Croup	146	40.4	3.8-161.2	1.92 (1.51)	0.93 (2.14)
C	Bronchiolitis	198	6.0	0.4-71.9	1.15 (1.43)	1.60 (1.74)
D	Viral Wheeze	173	25.3	1.3-191.9	2.10 (1.69)	3.14 (3.08)
E	URTI	229	27.0	0.4-159.8	1.81 (1.59)	1.17 (2.07)
F	UTI	10	40.4	0.8-129.2	1.69 (2.04)	1.39 (2.68)
G	Viral Illness	135	29.8	0.2-148.0	1.49 (1.66)	0.90 (2.34)
H	Gastroenteritis	59	28.1	0.6-115.7	0.82 (1.47)	0.34 (1.60)
I	Febrile Fit	18	30.1	12.2-69.1	2.58 (1.65)	0.60 (2.64)
J	Tonsillitis	29	64.7	3.3-166.0	2.41 (1.40)	1.89 (2.61)
K	Asthma	24	102.4	14.0-184.1	1.84 (1.47)	3.78 (3.67)
L	1st Presentation of Asthma	15	66.7	12.3-154.8	2.35 (1.39)	3.04 (3.24)
M	Acute Exacerbation of Asthma	118	81.8	13.4-187.1	2.88 (1.38)	5.12 (4.01)
N	Afebrile Fit	18	70.1	5.1-153.1	0.78 (2.31)	0.53 (1.87)
O	Cyanotic Episode	11	11.9	0.3-56.0	1.41 (1.76)	0.08 (0.80)
P	Other	88	58.4	0.1-180.8	0.95 (1.42)	0.77 (2.01)
Q	Well Children	20	21.2	0.2-164.0	0.40 (1.32)	-0.03 (1.1)
	All Children	1,386	36.5	0.1-191.9		

Table 7.3: List of diagnoses with corrected heart rate and respiratory rate and mean and corresponding mean and range of age in months.

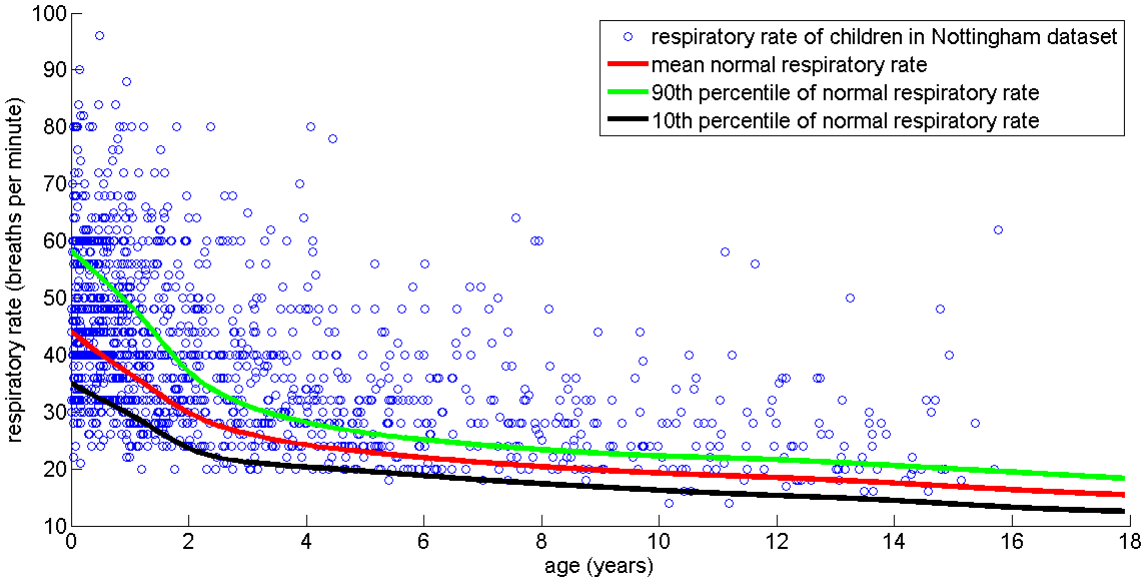


(a) distribution of heart rates for children with age in the Nottingham database, before age correction

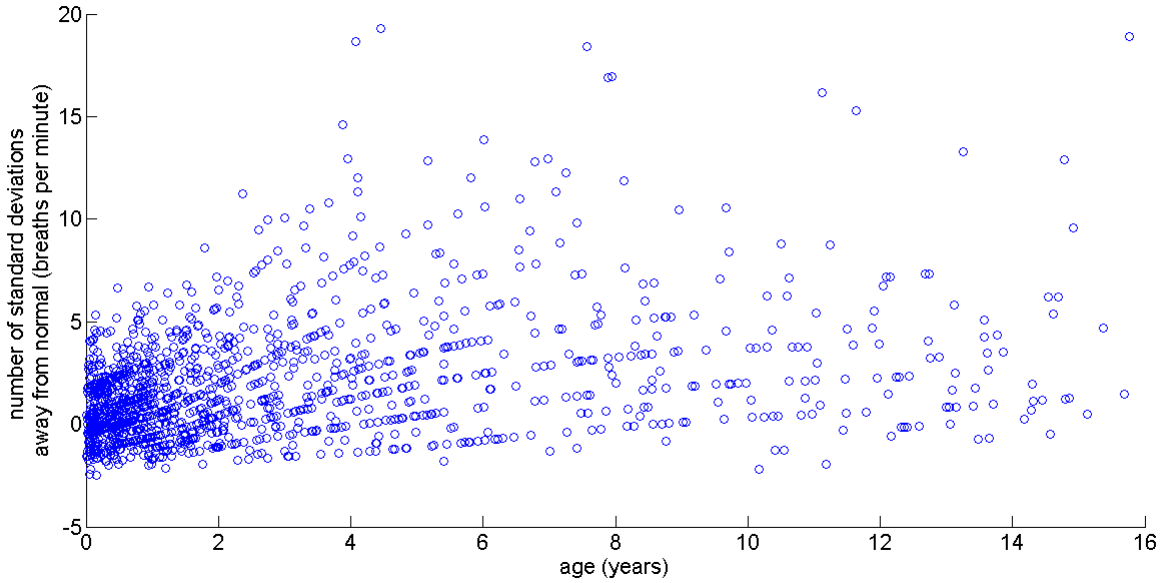


(b) distribution of the number of standard deviations a child's heart rate is away from the normal (for that age) with age

Figure 7.2: Distribution of heart rates in the Nottingham dataset to illustrate the effect of age correction



(a) distribution of respiratory rates for children with age in the Nottingham database, before age correction



(b) distribution of the number of standard deviations a child's respiratory rate is away from the normal (for that age) with age

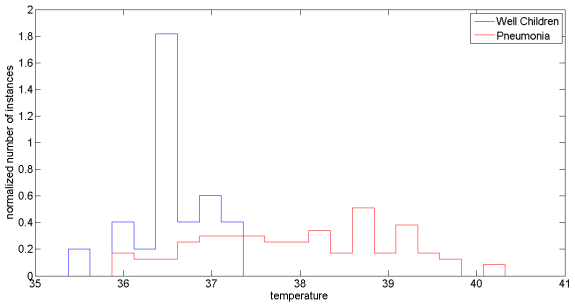
Figure 7.3: Distribution of respiratory rates in the Nottingham dataset to illustrate the effect of age correction

7.2.3 Data visualisation and separation based on individual vital signs

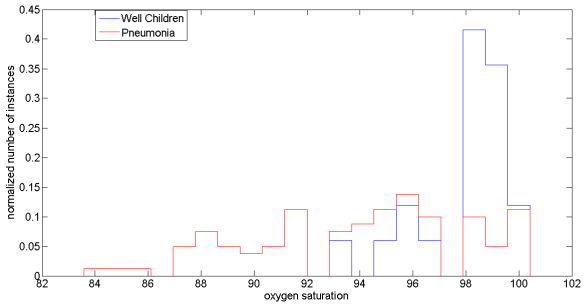
This section compares the distribution of vital signs (heart rate, respiratory rate, oxygen saturation and temperature) for children with a given diagnosis against well children. For illustration, two diagnoses, "pneumonia" and "acute exacerbation of asthma", are chosen for comparison with well children. Pneumonia is chosen since it is one of the most frequent serious infection that has a high chance of life-threatening complications and which can be life-threatening if left untreated [148, 140]. In contrast, "acute exacerbation of asthma" also chosen to illustrate the fact that vital signs can still be significantly disturbed in children even in the absence of any infection.

Figure 7.4 shows the distribution of the four vital signs (heart rate, respiratory rate, oxygen saturation and temperature) using histograms for children with "pneumonia" shown in red against well children in blue. The histograms were normalized so that the area under both the red and blue plots is equal to 1. A 2-D visualisation plot comparing distribution of values of the four vital signs for children with "pneumonia" with those of "well children" is also shown in part (e) of the figure. This plot is generated using the Neuroscale algorithm [76], a method based on Sammon's mapping [119]. The projection of 4-dimensional data onto 2-dimensional space is carried out so as to ensure that the Euclidean distance between two image patterns in the 2-D visualisation space is as close as possible to the Euclidean distance in the 4-D space between the two corresponding patterns. A similar comparison is shown in figure 7.5 for children with "acute exacerbation of asthma" where the values for the vital signs of children with "acute exacerbation of asthma" are shown in red and the values of the vital signs of "well children" are shown in blue.

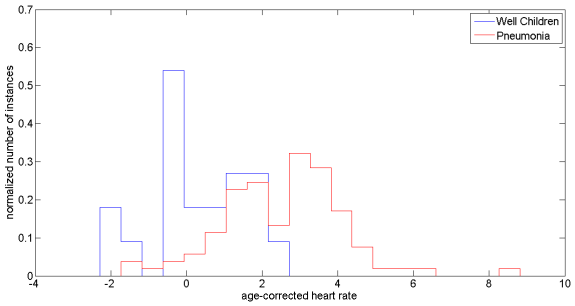
Comparing the Neuroscale plots in part (e) of figure 7.4 and figure 7.5, it can be seen that there is a significant separation of vital signs for both the children with acute exacerbation of asthma and pneumonia in comparison to children who are deemed to be normal. However, comparing the plots in part (a) of figure 7.4 and figure 7.5, it can be seen that only in the case of children with pneumonia is there a significant increase in temperature. It is therefore important to recognise that childhood diseases have a different effect on the vital signs.



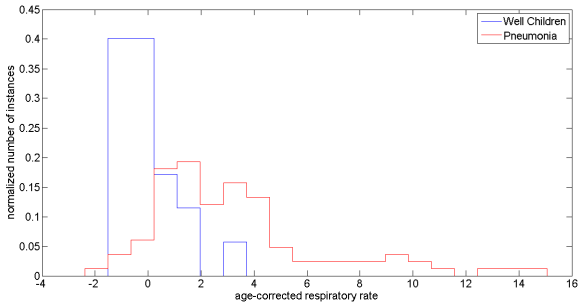
(a) Normalized histogram for temperature



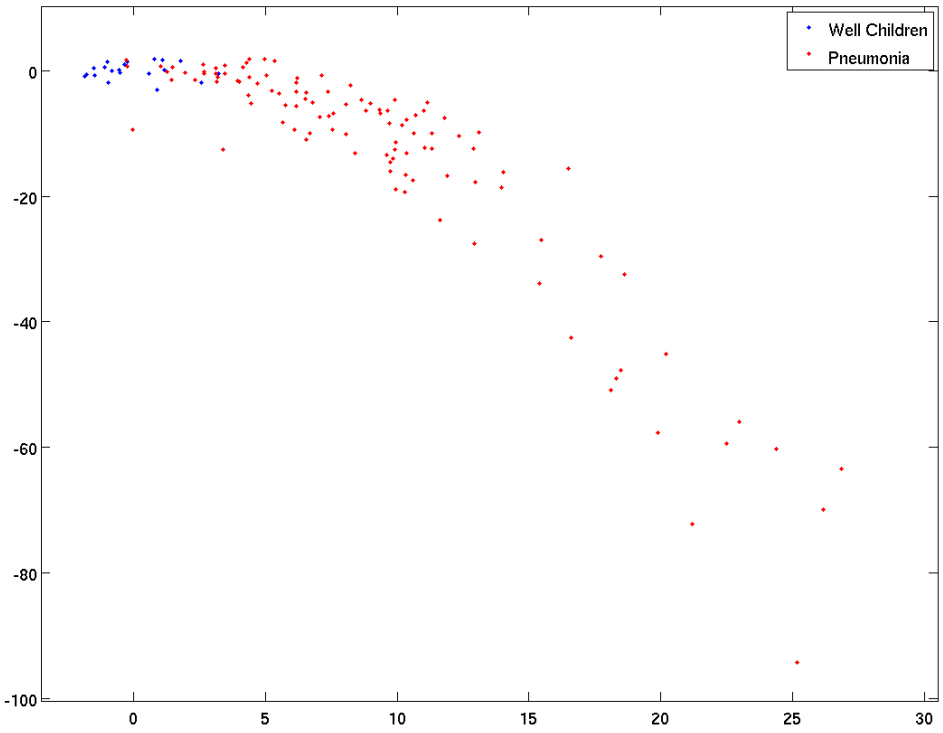
(b) Normalized histogram for Oxygen Saturation



(c) Normalized histogram for Age Corrected Heart Rate

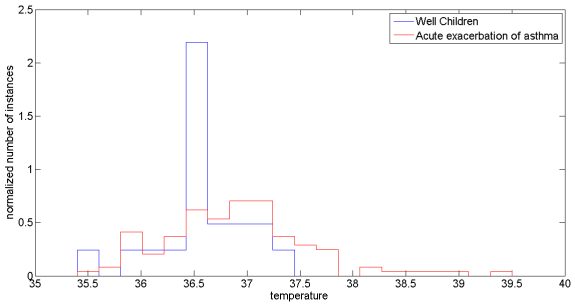


(d) Normalized histogram for Age Corrected Respiratory Rate

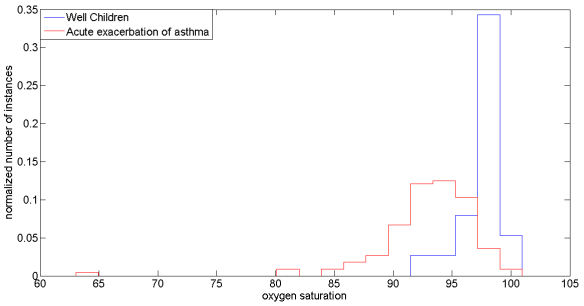


(e) Neuroscale visualization plot

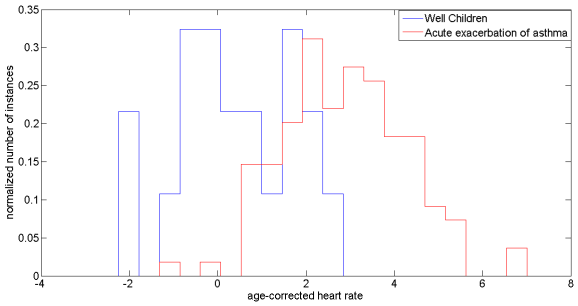
Figure 7.4: Distribution and Neuroscale plot of vital signs for "Well Children" in blue in comparison to the vital signs for children with "Pneumonia" in red



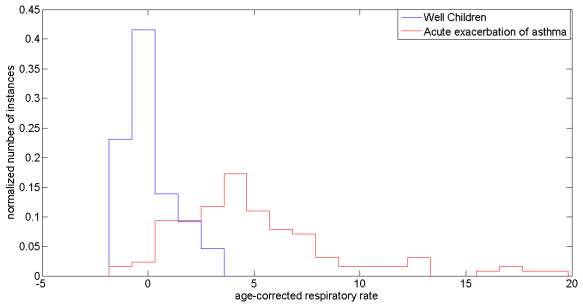
(a) Normalized histogram for temperature



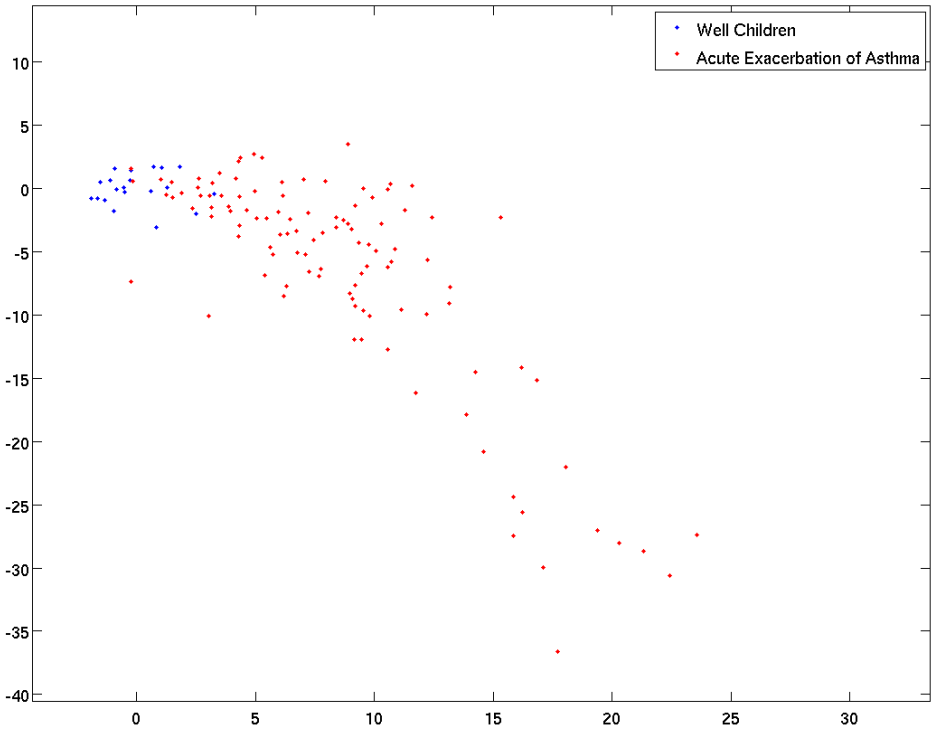
(b) Normalized histogram for Oxygen Saturation



(c) Normalized histogram for Age Corrected Heart Rate



(d) Normalized histogram for Age Corrected Respiratory Rate



(e) Neuroscale visualization plot

Figure 7.5: Distribution and Neuroscale plot of vital signs for "Well Children" in blue in comparison to the vital signs for children with "Acute exacerbation of asthma" in red

For each diagnosis group, the extent of separation with respect to healthy children can be quantified based on each individual vital sign in turn. A number of metrics have been described in the literature to quantify the separation between two distributions, $p(x)$ and $q(x)$, such as the Kolmogorov-Smirnov (KS) metric [40], the Kullback-Leibler (KL) divergence [64], and the Battacharyya coefficient (BC) [9].

Table 7.4 shows the ranking of the diagnosis groups based on the amount of overlap with the "well children" group for each of the four vital sign, using the Battacharyya coefficient (BC) given by:

$$BC(p,q) = \sum_{x \in X} \sqrt{p(x)q(x)}$$

where $p(x)$ and $q(x)$ are the pdf of the two distributions compared. These results confirm that childhood diseases have a different effect on vital signs. For example, febrile fit is ranked in the top two in Table 7.4 (a) and (b) but in the lower half in 7.4 (c) and (d) indicating that children with febrile fit are likely to have fever increasing the heart rate. However, the effect on the respiratory rate and oxygen saturation is much less.

While the results given in Table 7.4 uses the Battacharyya coefficient, the KL divergence and the KS metric will give a similar set of results.

Diagnosis	% overlap
UTI	20.0
Febrile Fit	21.1
Tonsillitis	27.2
Pneumonia	29.2
Cyanotic Episode	38.2
Viral Illness	40.0
URTI	44.7
Croup	47.9
Bronchiolitis	56.1
Acute Exacerbation of Asthma	57.2
Viral Wheeze	58.9
Gastroenteritis	60.7
Asthma	61.7
Afebrile Fit	63.3
1st Presentation of Asthma	65.0

(a) Ranking of diagnosis groups based on temperature only

Diagnosis	% overlap
Febrile Fit	31.1
1st Presentation of Asthma	31.7
Acute Exacerbation of Asthma	38.0
Asthma	38.3
UTI	40.0
Tonsillitis	47.2
Pneumonia	47.9
Viral Wheeze	56.0
Afebrile Fit	56.1
Croup	60.8
Cyanotic Episode	61.4
Viral Illness	63.7
URTI	64.1
Bronchiolitis	70.8
Gastroenteritis	73.7

(b) Ranking of diagnosis groups based on age corrected heart rate only

Diagnosis	% overlap
1st Presentation of Asthma	26.7
Acute Exacerbation of Asthma	31.2
Viral Wheeze	37.8
Pneumonia	47.1
Asthma	50.8
Bronchiolitis	54.3
Croup	55.5
URTI	55.8
Viral Illness	60.3
Febrile Fit	62.8
Tonsillitis	62.9
Cyanotic Episode	64.1
Gastroenteritis	65.6
UTI	70.0
Afebrile Fit	74.4

(c) Ranking of diagnosis groups based on SpO_2 only

Diagnosis	% overlap
Acute Exacerbation of Asthma	17.3
1st Presentation of Asthma	20.0
Asthma	23.3
Pneumonia	35.8
Viral Wheeze	36.3
Bronchiolitis	51.3
Cyanotic Episode	52.3
Tonsillitis	54.8
UTI	55.0
Febrile Fit	57.8
URTI	62.1
Viral Illness	62.4
Afebrile Fit	63.3
Croup	67.1
Gastroenteritis	72.2

(d) Ranking of diagnosis groups based on age corrected respiratory rate only

Table 7.4: Ranking of Diagnosis groups for each vital sign based on the Battacharyya coefficient, measure of the amount of overlap between two distributions expressed in percentage

7.3 Novelty detection to assess vital sign disturbance

Novelty detection is the identification of abnormal data based on deviation from a model of normality constructed from normal data. Novelty detection has been extensively applied in a large number of applications, including fault detection [137], detection of masses in mammograms [136], condition monitoring of patients in ICU [135] and several others [85].

A number of techniques exist for novelty detection which can broadly be classified into statistical approaches [85] and neural network based approaches [86]. There is no single best method and the performance of a technique depends on the application and the statistical properties of the data to which it is applied.

A common statistical approach used in novelty detection is to model the probability density function of the normal class and then compute the probability of a test sample belonging to the normal class. This approach is ideally suited when attempting to find the extent of vital sign disturbance in various childhood diseases in comparison to a model of normality based on the vital signs of well children. Firstly, a normal group needs to be identified so that a model of normality can be constructed. Subsequently, the probability computed for every patient needs to be converted into a novelty score such that the lower the probability, the higher will be the novelty score suggesting that the vital signs of the patient are more abnormal, as they have a lower probability of belonging to the normal group. Finally, the mean novelty score for each diagnosis group can then be computed and ranked to find out which illness affects the vital signs and to what extent, if any.

7.3.1 Selection of children to form the normal class ("Nottingham Normal")

The "Nottingham database" contains only 20 children who are labelled as being "well". It was decided that in order to select a statistically meaningful model of normality across the entire age range, the diagnosis group labelled "other" shall be investigated to increase the number of children used in constructing the model of normality. The following procedure was adopted for the selection of children from the "other" group for inclusion in the normal class:

- i) Find the standard deviation for each of the four vital signs for the 20 "well" children i.e. σ_{hr}

for heart rate, σ_{temp} for temperature, σ_{resp} for respiratory rate, and σ_{sat} for oxygen saturation.

- ii) Short-list the diagnoses of children from the "other" group whose heart rate, temperature, respiratory rate and oxygen saturation all lie within 3 standard deviations of σ_{hr}, σ_{temp}, σ_{resp} and σ_{sat} respectively.

- iii) Select the diagnoses from the short-listed diagnoses found in step (ii) for which two experienced primary care physicians independently agree that they represent neither serious infections nor do they have any serious impact on vital signs.

The diagnoses identified in (iii) were: benign floppy larynx, colic, faint, cord, impetigo, maternal cracked nipples, migraine, muscular pain, night terror, period pain, petechiae, poor feeding, pyloric stenosis, rash, reflux, scarlatina, sinusitis, thread worms, vaginal pain, and erythema margintum.

Table 7.5 lists the age distribution of children in the "well" group and those in the "Nottingham normal" group, who were selected for inclusion in the model of normality. The latter group consists of 28 children selected according to steps (i) to (v) above from the "other" group plus the 20 children from the "well" group.

Age range (years)	Number in the "Nottingham well" group	Number in the "Nottingham normal" group
0-1	16 (80%)	29 (60%)
1-5	1 (5%)	5 (10%)
5-10	2 (10%)	7 (15%)
10-15	1 (1%)	7 (15%)
Total	20	48

Table 7.5: Comparison of age distributions of "Nottingham well" and "Nottingham normal" groups

7.3.2 Selection of children to form the normal class ("Expanded Normal")

Out of the 272 patients monitored during the paediatric triage study at the John Radcliffe Hospital, Oxford (see Chapter 3 for details), 252 had the four vital signs (heart rate, oxygen saturation, temperature and respiratory rate) recorded. Out of these 252 patients, the record of the Manchester Triage category [151] assigned was available for 236 patients. These categories, in the order of decreasing urgency, were "emergency" (4 patients) , "very urgent" (70 patients), "urgent" (70

Database	number of patients	temperature (celsius)	oxygen saturation (%)	age-corrected heart rate (beats/min)	age-corrected respiratory rate (breaths/min)
"Nottingham normal"	48	36.6 (0.50)	97.6 (1.6)	0.62 (1.2)	0.25 (1.1)
"Oxford normal"	26	36.6 (0.52)	98.2 (1.0)	0.36 (0.9)	0.40 (1.5)
"Expanded normal"	74	36.6 (0.50)	97.8 (1.5)	0.53 (1.1))	0.29 (1.2)

Table 7.6: Mean (standard deviation) of the four vital signs in the "Nottingham normal", "Oxford normal" and the "Expanded normal" groups

patients) and "standard" (92 patients). The urgency category is determined based on 49 flowchart diagrams specific to a patient's presenting problem and each category has a maximum waiting time with 0 minutes for "emergency", 10 minutes for "very urgent", 60 minutes for "urgent" and 120 minutes for the "standard" triage category [151].

Using a similar approach to what was described in section 7.3.1, a total of 26 patients were selected to form the "Oxford normal" group. The final diagnosis of these 26 selected patients were: irritable hip (12 patients), growing pain (1 patient), painful right foot with no injury (1 patient), impetigo (1 patient), constipation (1 patient), muscular pain (1 patient), ENT problem (2 patients), Limp with unclear cause (1 patient), paroxysmal palpitations (1 patient), intolerance to cow's milk (1 patient), dermatological problem (1 patient), rheumatological problem (1 patient), "well child" (1 patient) and infected wound on knee (1 patient). Out of these 26 patients, 13 patients (50%) were assigned to the "standard" triage category while the remaining 13 patients (50%) were assigned to the "urgent" triage category.

The "Oxford normal" group and the "Nottingham normal" group were combined to form a larger group of normal children, labelled the "Expanded normal" group. Table 7.6 provides the mean and standard deviation of the four vital signs in these three groups while table 7.7 provides their age distribution.

Age range (years)	Number in the "Nottingham normal" group	Number in the "Oxford normal" group	Number in the "Expanded normal" group
0-1	29 (60%)	2 (8%)	31 (42%)
1-5	5 (10%)	7 (27%)	12 (16%)
5-10	7 (15%)	12 (46%)	19 (26%)
10-15	7 (15%)	5 (19%)	12 (16%)
Total	48	26	74

Table 7.7: Comparison of age distribution of "Nottingham normal", "Oxford normal" and the "Expanded normal" groups

7.3.3 Building a model of normality

A vital sign model of normality for the N children in a normal group where $N = 48$ if the "Nottingham normal" group is used, and $N = 74$ if the "Expanded normal" group is used, can now be constructed using a probability density estimation technique. This model of normality will then be used to compute the novelty scores for the different diagnosis groups of Table 7.3. An $Nx4$ dimension matrix of data is constructed using the four vital signs of the normal group chosen. The four variables (temperature, age-corrected heart rate, age-corrected respiratory rate, and oxygen saturation) are each normalized to have zero mean and unit variance. This is done by using:

$$x_n = \frac{x - \mu}{\sigma}$$

where x_n refers to the normalized vital sign value, x is the unnormalized vital sign value, μ is the mean of each variable and σ is the standard deviation of each variable in the normal group. Normalization ensures that each variable has equal importance as input to the model of normality.

Probability density estimation techniques can be classified into parametric and non-parametric techniques [139]. The former assumes that the probability to be estimated comes from a class of known distribution and a small number of parameters are computed to fit data, using an optimization technique such as least squares fitting, to this distribution. Such an approach cannot be used in the application since it is not practical to make an assumption about the distribution of the data to be modelled. The alternative approach, non-parametric probability density estimation technique, is thus adopted to build a model of normality.

In this work, the model of normality is constructed by using the Parzen windows technique, a non-parametric probability density estimation technique which does not impose any limitation on the shape of distribution to be estimated [114]. A window function (also called the kernel function) is centered at each point in the training set in the 4-dimensional space (each dimension corresponding to one of the vital signs).

In this work, equi-covariant Gaussians are used as the window functions which are centered on the corresponding point, z_i, in the training set so that the mean of the i^{th} Gaussian is z_i. There are two advantages of using equi-covariant Gaussian functions as the kernel functions. Firstly, the Gaussian function is smooth and the resulting probability density estimated will also be smooth [85]. Secondly, a Gaussian kernel function can be completely characterised by a single variance factor [85].

The equation of the resulting probability density function at point z is given by equation 7.2 where d is the dimensionality of the model (in this case, 4). The variance σ^2 defines the smoothness of the resulting window function. A larger value of σ^2 will result in a smoother window function. If σ^2 is too small, it will lead to over fitting of the data. It is therefore, important to choose an appropriate value for σ^2. Local variance is estimated by calculating, for each point in the training set (the data points which are used to construct the model of normality), the mean squared distance to the m nearest neighbours. Then σ^2 is set to be the mean of these local variances, with $m = 10$ as suggested by Bishop [10].

$$p(z) = \frac{1}{n}\sum_{n=0}^{j}\frac{1}{2\pi^{d/2}\sigma^d}exp\left\{-\frac{|x-x_j|^2}{2\sigma^2}\right\} \tag{7.2}$$

7.3.4 Ranking of diagnosis groups based on novelty score

For each child in the database, $p(z)$ is calculated using equation 7.2. A higher value of $p(z)$ means that the data lies close to the distribution of data points from the normal group (i.e. the training set). An index of severity , I, is then calculated according to equation 7.3 which ensures that the lower the value of $p(z)$, the higher will be the severity score (referred to as the novelty score) indicating that it lies further away from the region of space most populated by the "normal" group [134].

$$I = ln\frac{1}{p(z)} = -ln(p(z)) \tag{7.3}$$

The mean novelty score is then calculated for every diagnosis group, which is equal to the mean of the severity score of each child in the corresponding diagnosis group. This process is repeated twice, first after building a model of normality of 48 children using the "Nottingham normal" group and then after building a model of normality of 74 children using the "Expanded normal" group.

7.3.4.1 Ranking based on the "Nottingham normal" group

To ensure that the contribution of any oxygen saturation above the mean oxygen saturation of the "Nottingham normal" group (97.6%, see table 7.6) to the novelty score is zero, any oxygen saturation above 97.6% is assigned 97.6% before computing $p(z)$ for that patient.

Table 7.8 shows the ranking of the various diagnosis groups based on the mean novelty score which is illustrated in figure 7.6 showing the error bars (representing one standard error on either side) associated with each point. Figure 7.7 shows the difference in the mean novelty score of each diagnosis group in comparison to the mean novelty score of the "Nottingham normal" group. Not surprisingly, the mean novelty score for the "Oxford normal" group can be seen to be equivalent to the mean novelty score of the "Nottingham normal" group.

Label	Number of patients	Diagnosis group	Mean novelty score (standard deviation)
M	118	Acute exacerbation of asthma	15.4 (13.4)
A	95	Pneumonia	14.2 (10.0)
J	29	Tonsillitis	11.6 (8.2)
L	15	1st presentation of asthma	11.2 (9.3)
I	18	Febrile fit	11.1 (51.)
D	173	Viral wheeze	10.8 (9.5)
K	24	Asthma	10.5 (8.8)
F	10	UTI	8.7 (4.2)
E	229	URTI	8.4 (3.8)
C	198	Bronchiolitis	8.2 (3.9)
G	135	Viral illness	8.2 (4.5)
B	146	Croup	7.7 (4.0)
O	11	Cyanotic episode	6.6 (2.6)
H	59	Gastroenteritis	6.4 (1.5)
N	18	Afebrile fit	6.2 (0.9)
S	26	"Oxford Normal"	5.8 (0.7)
R	48	"Nottingham Normal"	5.8 (0.4)

Table 7.8: Ranking of diagnosis groups based on the mean novelty score and standard deviation using the model of normality based on the "Nottingham normal" group

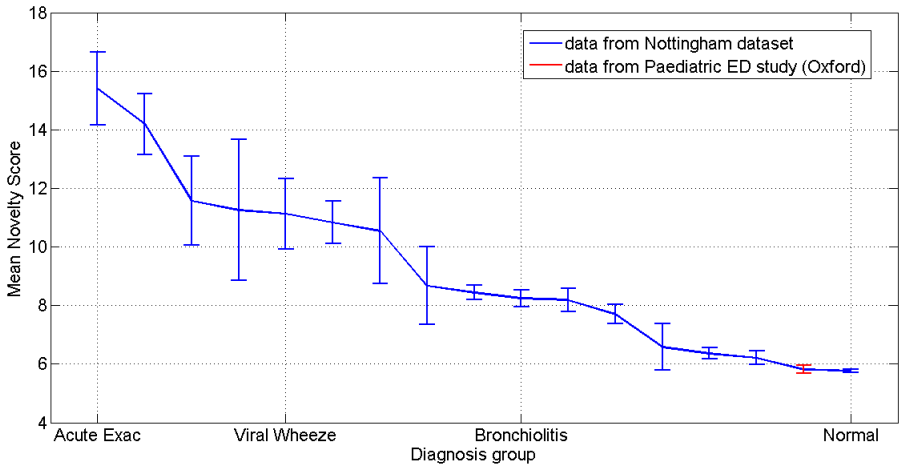


Figure 7.6: mean novelty score of various diagnosis groups ranked using a model of normality built from the "Nottingham normal" group. The novelty score of the "Oxford normal" group is shown in the figure in red

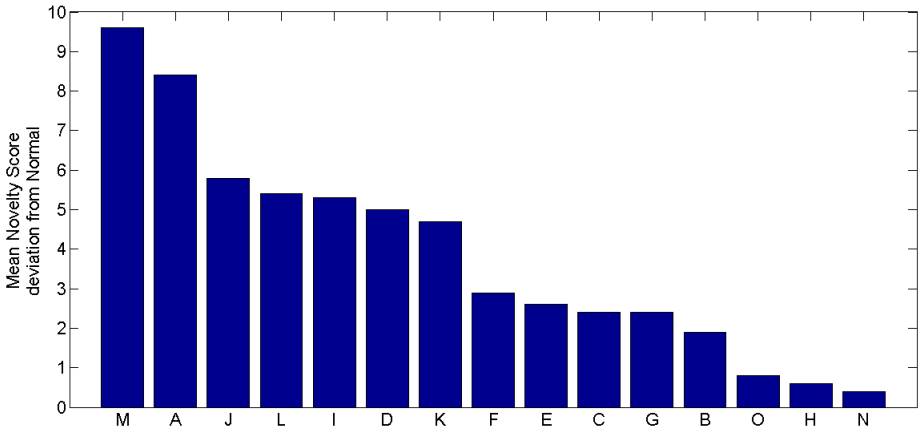


Figure 7.7: Mean novelty score deviation from "Nottingham normal" group (same labelling as in Table 7.3)

The dataset in the Paediatric triage study database was analyzed but unlike the "Nottingham database", there were not enough patients with distinctly clear final diagnosis to form further classes. However, the Manchester triage category was available for 236 patients and excluding the "emergency" triage category (since only 4 patients were present in that category), the mean novelty score for the remaining three groups is shown in table 7.9.

Triage category	Number of patients	Mean novelty score (standard deviation)
standard	92	6.3 (1.5)
urgent	70	7.6 (4.4)
very urgent	70	10.7 (9.3)

Table 7.9: Mean novelty score (standard deviation) of the Paediatric ED study database categorised into different triage categories

7.3.4.2 Ranking based on the "Expanded normal" group

As before, to ensure that the contribution of any oxygen saturation above the mean oxygen saturation of the "Expanded normal" group (97.8%, see table 7.6) to the novelty score is zero, an oxygen saturation of children above 97.8% is assigned a value of 97.8% before computing $p(z)$ for that patient.

Table 7.10 shows the ranking of the various diagnosis groups based on the mean novelty score which is illustrated in figure 7.8 showing the error bars (representing one standard error on either

side) associated with each point. Figure 7.9 shows the difference in the mean novelty score of each diagnosis group in comparison to the mean novelty score of the "Expanded normal" group.

Label	Number of patients	Diagnosis group	Mean novelty score (standard deviation)
M	118	Acute exacerbation of asthma	16.2 (13.8)
A	95	Pneumonia	14.6 (10.3)
J	15	1st presentation of asthma	12.1 (10.7)
L	29	Tonsillitis	11.9 (10.1)
I	18	Febrile fit	11.5 (5.6)
D	173	Viral wheeze	11.3 (10.1)
K	24	Asthma	10.5 (8.5)
F	10	UTI	8.6 (4.6)
E	198	Bronchiolitis	8.6 (4.8)
C	229	URTI	8.4 (3.7)
G	135	Viral illness	8.1 (4.5)
B	146	Croup	7.8 (4.3)
O	11	Cyanotic episode	6.6 (3.1)
H	59	Gastroenteritis	6.3 (1.7)
N	18	Afebrile fit	6.1 (1.1)
R	48	"Nottingham Normal"	5.6 (0.6)
T	74	"Expanded Normal"	5.6 (0.4)
S	26	"Oxford Normal"	5.5 (0.6)

Table 7.10: Ranking of diagnosis groups based on the mean novelty score and standard deviation using the model of normality based on the"Expanded normal" group

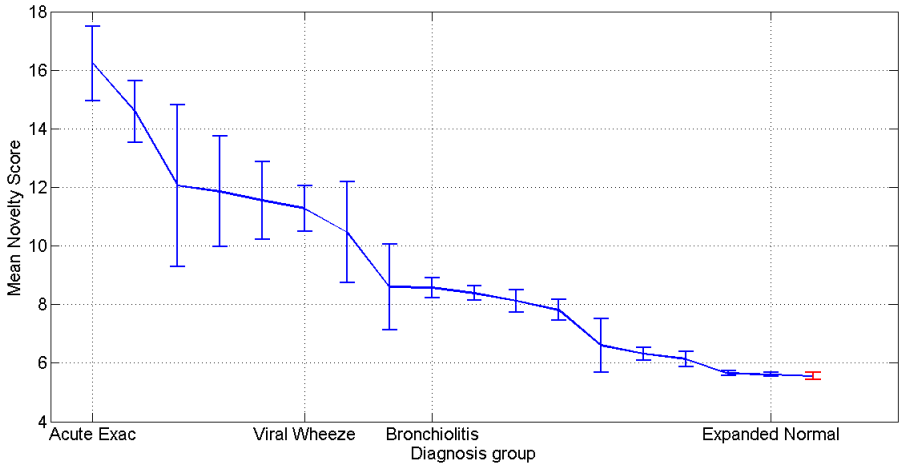


Figure 7.8: mean novelty score of various diagnosis groups ranked using a model of normality built from the "Expanded normal" group. The novelty score of the "Oxford normal" group is shown in the figure in red

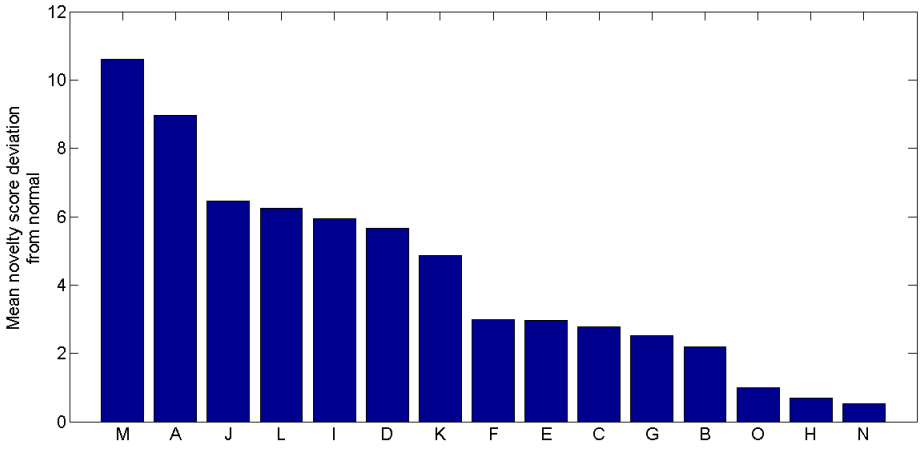


Figure 7.9: Mean novelty score deviation from the "Expanded normal" group (same labelling as in Table 7.3)

Comparing these results with the results in subsection 7.3.4.1, it can be seen that the results are very similar except for two small differences. The diagnosis group labelled "1st presentation of asthma" is third in table 7.10 while fourth in Table 7.8, and consequently "tonsillitis" is fourth in 7.10 and third in Table 7.8. And the diagnosis group labelled "bronchiolitis" is ninth in Table 7.10 and tenth in Table 7.8, and consequently "URTI" is tenth in Table 7.10 and ninth in Table 7.8. However, the mean novelty scores of these groups are very close to each other in both sets of results.

7.4 Summary

Previous studies have attempted to identify patients with serious illness using labels (e.g. serious infection/ no infection etc.) provided by the clinicians based on their final diagnosis. However, significant overlap exists in the distribution of vital signs of children in such pre-labelled data and it is therefore important to take a step back and try to understand the extent of separation between different diagnosis groups based on their vital signs. The initial investigation was carried out using a subset of a database consisting of 1,386 children attending an Emergency Department in a hospital in Nottingham, the "Nottingham database".

Firstly, the extent of separation between different diagnosis groups based on individual vital

signs (temperature, respiratory rate, heart rate and oxygen saturation) was determined and ranked. Then a model of normality was constructed using vital sign data from a normal group of children (the "Nottingham normal") and the mean novelty scores for various diagnosis groups were then determined and ranked.

There were only 48 normal children in the Nottingham database whose data was used for constructing a model of normality. The paediatric ED study database collected at John Radcliffe Hospital, Oxford was therefore investigated to see if it was possible to expand the normal group. Based on the final diagnosis, 26 patients were selected from this database (the "Oxford normal"). The mean novelty score of these 26 patients was computed using the model of normality built using the "Nottingham normal" data and it was found to have the same novelty score as that of the "Nottingham normal" group.

Subsequently, the two normal groups were combined to form a larger group of 74 normal children (the "Expanded normal" group) and a model of normality constructed from this data. The mean novelty scores for various diagnosis groups were then determined and ranked. The results in both cases were very similar.

It was not possible to form further diagnosis groups from the paediatric ED study database to determine their novelty score. However, the triage category assigned by the nurses during triage was available for 232 patients, labelled "standard", "urgent" and "very urgent". The model of normality built using the "Nottingham normal" data was used to determine the novelty score of these three groups and a correlation was found between the novelty score and the level of urgency based on the Manchester Triage Score.

The work presented in this chapter has provided insights into which vital signs are affected by a given diagnosis group and to what extent this occurs. The use of two databases has enabled us to make use of an independent dataset for novelty score ranking.

Chapter 8

Data Fusion for Paediatric Triage

8.1 Introduction

This chapter extends the work in the previous chapter by developing data fusion models based on the values of their vital signs to identify children with serious infection. Serious infection is understood by the medical community to refer to those illnesses that are are likely to be life-threatening if untreated, or with a high chance of life-threatening complications [140]. The most frequent illnesses in this category include pneumonia, sepsis, meningitis, pyelonephritis, and bacterial gastroenteritis [148].

Most of the children presenting to GP surgeries and/or the Emergency Department have self-limiting infections. However, a small percentage of these children will have a serious infection and distinguishing the children with a serious infection from the children with self-limiting infection is a major challenge for paediatricians. Typically, the occurrence of serious infection varies according to the clinical settting being lowest in primary care and self-referers to the ED, higher in children in the ED following a GP referral and highest in children admitted to hospital [148, 131, 16, 147].

Figure 8.1 illustrates the typical stages involved in studies that attempt to derive a model for classifying patients according to the severity of their illness, using the values of their vital signs. Children presenting in a clinical setting are diagnosed by clinicians who use symptoms and physiological data (including vital signs), and various clinical tests if necessary. Subsequently, the datasets are labelled (e.g. whether the infection is serious or not) based on disease knowledge.

Apart from the extent of vital sign disturbance caused by the illness, classifier performance built also depends on the errors introduced by (i) the measurement errors when acquiring vital signs; and (ii) the inaccuracy of the clinical labelling (diagnoses).

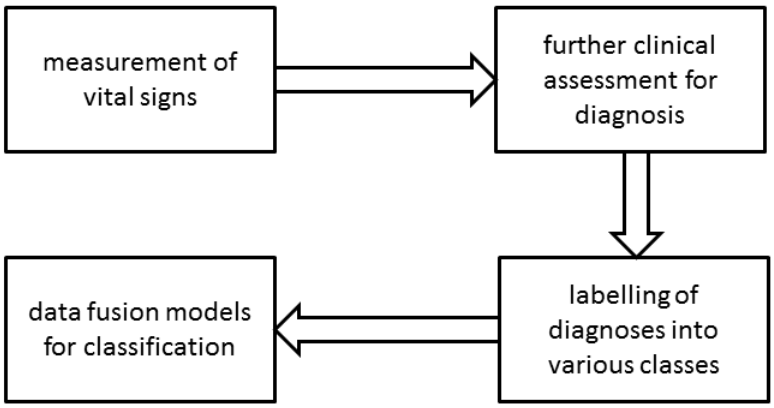


Figure 8.1: Typical stages for studies which attempt to derive a model for classifying patients according to severity of illness, using vital sign data

A previous study, attempting to build data fusion models for classification of the Walsgrave dataset (see Chapter 7), used pre-labelled data and in-sample validation using a jack-knifing technique [34]. It is well known, however, that prediction models tend to perform better when using in-sample validation data and significantly worse when using external validation data [12, 11]. It was therefore decided to validate the work described in the previous chapter using novelty detection firstly by using in-sample validation and secondly by using independent datasets.

		True Classes	
		Normal	Abnormal
Model Classification	Normal	True Negative	False Negative
	Abnormal	False Positive	True Positive

Figure 8.2: Possible outcomes from a binary classification

8.2 Performance evaluation

The model of normality described in the previous chapter is used to generate a novelty score for every validation or test patien. To classify a patient into a serious infection class (positive class) or no /non-serious infection class (negative class) requires the choice of a threshold. The performance of the classifier can be assessed using sensitivity and specificity as defined in figure 8.2 and equations 8.1-8.2.

$$Sensitivity = \frac{TP}{(TP+FN)} \times 100\% \tag{8.1}$$

$$Specificity = \frac{TN}{(TN+FP)} \times 100\% \tag{8.2}$$

Sensitivity quantifies the ability of a classifier to identify the positive class i.e. the percentage of patients that are correctly classified as positive (true positive) from all the patients who truly belong to the positive class (true positive + false negative). Specificity quantifies the ability of a classifier to identify the negative class i.e. the percentage of patients that are correctly classified as negative (true negative) from all the patients who truly belong to the negative class (true negative + false positive).

There is typically a trade-off between sensitivity and specificity depending on the choice of threshold. A single set of sensitivity and specificity values is associated with a given threshold. The variation of sensitivity and specificity threshold is best visualised using a Receiver Operating Characteristic (ROC) curve. The area under the ROC curve (AUC) gives a measure of the accuracy of a classifier. An AUC can be likened to the probability that a given randomly chosen positive sample is correctly classified. An AUC of 0.5, therefore, corresponds to a classifier that performs

no better than chance, whereas a perfect classifier will have an AUC of 1 [60, 15].

Choosing an appropriate operating point on an ROC curve depends on the requirements of the application. For example, in primary care settings it is more likely that a higher sensitivity is chosen at the expense of reduced specificity so that children with serious infections are more certain to be identified. However, in a tertiary setting (hospitals), it is more likely that a higher specificity will be chosen at the expense of reduced sensitivity so that children who do not have a serious infection can be identified. In this chapter, the sensitivity and specificity corresponding to the point closest to (0,1) in the ROC curve are reported since (0,1) corresponds to the perfect classifier (i.e. one with 100% sensitivity and specificity).

Since the datasets used for testing are not balanced, the following iterative algorithm (similar to jack-knifing) was developed so that the results from using a testing set with equal numbers of normal and abnormal examples can be reported.

(i) Select N random samples from both the positive and the negative class

(ii) Determine the corresponding ROC curve and the area under the curve (AUC)

(iii) Repeat (i) and (ii) k times and find the 10^{th}, 50^{th} and 90^{th} percentile ROC and the corresponding AUC.

In the subsequent analysis, results based on using (i) the complete testing dataset and (ii) a subset of the testing dataset using the iterative algorithm are reported.

8.3 Independent datasets for testing

Two different datasets were available for testing the novelty detection-based classification algorithm: the Walsgrave dataset and the Pinderfields dataset.

8.3.1 Walsgrave dataset

The Walsgrave dataset [140] included a total of 700 children with an age range of 3 months to 16 years who attended the Paediatric Assessment Unit (PAU) at the University Hospital Coventry and Warwickshire NHS Trust, UK. Children were recruited for the study only if they were suspected

of having an acute infection. Exclusion criteria included children with diseases that were expected to cause repeated serious bacterial infection (including haematological malignancies and iatrogenic immunosuppression) and infections due to trauma. The four vital signs (heart rate, respiratory rate, temperature, oxygen saturation) as well as the child's age, were available for 568 children in the Walsgrave dataset analysed in this thesis.

Patients were categorised into four classes based on the final diagnosis made by the senior paediatricians: "serious infection" (associated with life-threatening conditions); "intermediate infection" (associated with conditions that are not life-threatening but are likely to last for 10 days or cause non-life threatening complications); "minor infection" (associated with conditions from which a child is likely to recover without any complications); and "no infection". The total number of patients in each class along with the final diagnoses is summarised in Table 8.1.

Class Label	Diagnoses	Number
Serious infection	Pneumonia, Appendicitis, Meningococcal disease, Mastoiditis, Other bacteraemia, Meningitis	93
Intermediate infection	Non-specific viral illness/Viral respiratory illness/Viral gastroenteritis requiring hospital intervention, Febrile seizure, Lower respiratory tract infection, Cellulitis	155
Minor infection	URTI, Viral gastroenteritis, Non-specific viral illness, Uncomplicated UTI/Cellulitis	279
No infection	no infection	41

Table 8.1: Final diagnoses of children in the Walsgrave database

As part of their study [140], Thompson et al. developed a clinical decision rule based on taking into account all of the vital signs in turn. They found that this provided comparable diagnostic performance to more complicated triage methods (like the Manchester Triage Score, reviewed in section 1.1 of chapter 1). They assigned cut-off values to temperature ($\geq 39^oC$), oxygen saturation ($\leq 94\%$) as well as age-dependent cut-offs for heart rate and respiratory rate. The age ranges were the same as those defined in the APLS guidelines [78] and any child with at least one abnormal vital sign was classified as abnormal. The cut-off values were determined using in-sample validation. With these values, Thompson et al. [140] reported a sensitivity of 86% and a specificity of 34% when attempting to discriminate between serious infection and intermediate/minor/no infection.

Section 8.7 will provide the results of applying this clinical decision rule to an independent dataset.

8.3.2 Pinderfields dataset

The Pinderfields dataset is a case-control study involving acutely ill children, hospitalised at Pinderfields Hospital (Wakefield, UK) between 2000 and 2005. Children were classified into four groups based on the final diagnosis (determined through structured clinical assessment, and various clinical tests including full blood count, C-reactive protein, blood and urine culture and chest X-ray in the presence of respiratory distress) by a consultant paediatrician. The four groups in the database were: a case-controlled group consisting of patients who were admitted but subsequently turned out not to have any serious infection (labelled Control); a group of patients consisting of those diagnosed with pneumonia (labelled Pneumonia); a group of patients diagnosed with meningitis/sepsis/bacteraemia/osteomyelitis (labelled Sepsis); and a group of children diagnosed with urinary tract infection (labelled UTI). The total number of patients in each class, along with the final diagnoses, is summarised in Table 8.2.

Class Label	Diagnoses	Number
Pneumonia	Pneumonia	78
Sepsis	Meningitis, Abscess, Bacteraemia, Osteomyelitis	94
UTI	UTI	31
Control	Tonsilitis, Viral infection, URTI (self-limiting or mild infection)	781

Table 8.2: Final diagnoses of children in the Walsgrave database

8.3.3 Summary of datasets

Table 8.3 provides a summary of the various datasets used for data fusion. The model of normality is constructed from a subset of patients from the first two datasets (Nottingham and Oxford study) as discussed in the previous chapter and the remaining two datasets, Walsgrave and Pinderfields, form a completely independent test set.

Dataset	Population characteristic	Number of children
Nottingham	generally unwell children presenting to an ED	1403 with various final diagnoses
Oxford Study	subset of children presenting to an ED	26 "normal" used
Walsgrave	children attending a PAU suspected of an acute infection	568 in four classes based on severity of infection
Pinderfields	ill children admitted in hospital	621 in four classes with four vital signs and a further 363 with only three vital signs

Table 8.3: Summary of datasets used for Data Fusion

8.4 Data fusion using in-sample validation

Out of all the diagnoses in the Nottingham database, only pneumonia is considered to be a serious infection. Novelty detection was employed to see if patients with pneumonia can be identified from the rest of the database in the Nottingham dataset. Table 8.4 gives the results using novelty detection to identify patients with serious infection using the model of normality trained with the "Expanded Normal" dataset. The sensitivity and specificity corresponding to the point closest to the top left corner (0,1) are reported in the table. The classification results are provided both with and after removing patients with asthma.

Serious infection (94) vs.	Sensitivity	Specificity	AUC
rest of the database (1185)	74%	57%	0.703
rest of the database except patients with asthma (1030)	74%	61%	0.732

Table 8.4: Identification of serious infection in the Nottingham database using novelty detection with model of normality built using data from the "Nottingham" database and the Oxford Paediatric Triage study. Sensitivity and Specificity of the point closest to (0,1) on the ROC curve are reported. "Serious infection" in this case is defined as patients with pneumonia.

To investigate the results of having a balanced dataset (same number of positive and negative examples), the iterative algorithm described earlier was implemented with $N = 50$ (number of samples) and $k = 1000$ (number of iterations) and the results are reported in Table 8.5.

Serious infection (94) vs.	Sensitivity %	Specificity %	AUC
rest of the database (1185)	68 (64-70)	62 (54-71)	0.706 (0.636-0.768)
rest of the database except patients with asthma (1030)	70 (66-76)	65 (57-70)	0.735 (0.666-0.795)

Table 8.5: Identification of serious infection in the Nottingham database using novelty detection with model of normality built using data from the "Nottingham" database and the Oxford Paediatric Triage study. Sensitivity and Specificity of the point closest to (0,1) on the ROC curve are reported. "Serious infection" in this case is defined as patients with pneumonia.

Variation in performance with different choices for variance for Parzen windows model of normality

The choice of $\sigma^2$for the Parzen windows model of normality (give by equation 7.2) is based on first estimating the local variance by calculating, for each point in the training set, the mean squared distance to its m nearest neighbours. Then σ^2 is set to be the mean of these local variances. Previously, $m = 10$ was chosen as suggested by Bishop [10]. The variation of AUC with various choices of m (6 to 73) was investigated using a subset of the Nottingham dataset (pneumonia vs. URTI). These results were also compared with taking σ^2 to be equal to the square root of the maximum distance between the training points. The results are summarised in figure 8.3. The optimal choice of m depends on the distribution of the training sets. While the variation in error is very small, it is still clear from the figure that $m = 10$ is not an optimal choice. From these results, it was decided to use $m = 50$ for the subsequent analysis.

8.5 Data fusion using independent test set (Walsgrave)

Although the Walsgrave dataset was collected in a different clinical setting than the Nottingham dataset, it was still deemed useful to investigate classification performance with novelty detection on this dataset, using the previously built model of normality on the "Expanded Normal" group from the Nottingham and the Oxford Paediatric Study databases (see Chapter 7 for details).

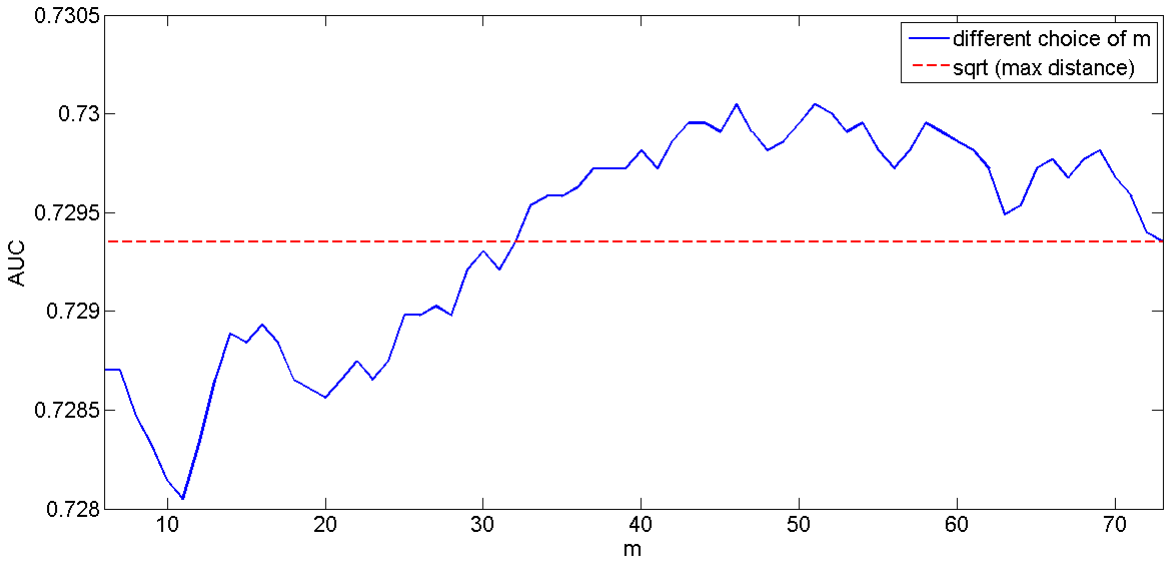


Figure 8.3: choice of variance for Parzen window model

8.5.1 Novelty score distribution of children from the Walsgrave database

The novelty scores (using the previously built model of normality) for the children in each of the four groups in the Walsgrave database are shown in figure 8.4 using normalized histograms along with the mean and the standard deviation. As expected, the mean score is highest for the "serious infection" group, followed by the "intermediate infection" group, the "minor infection" group and the "no infection" group. Not surprisingly, five of the six outliers in the "no infection" group were diagnosed with asthma, confirming the results of the previous chapter that the vital signs of patients with asthma are significantly disturbed (see figure 8.5). Table 8.6 reports the mean and the standard deviation of the novelty scores in each of the four classes in the Walsgrave database. The results of removing the five patients with asthma in the "no infection" group are also reported in the table.

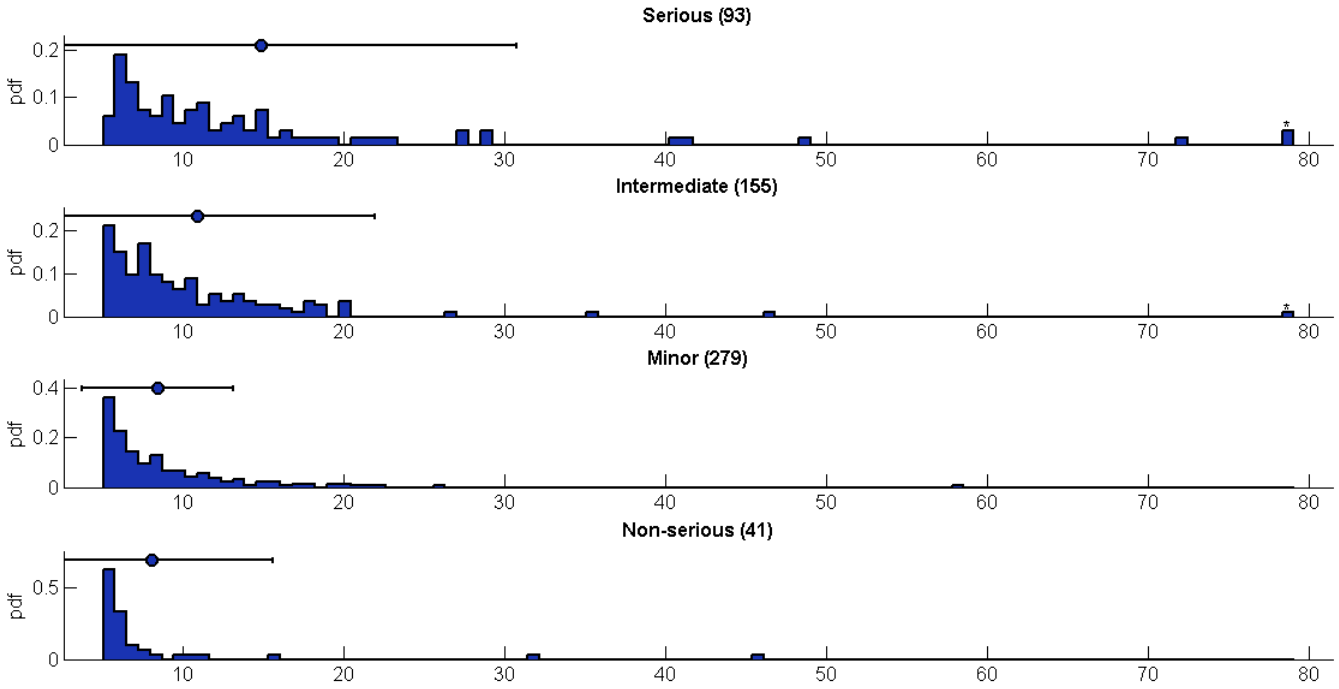


Figure 8.4: distribution of novelty scores for children in the Walsgrave database using the model of normality built from data from children in the Nottingham and the Oxford Paediatric triage study. The mean novelty score and the standard deviation are also shown.

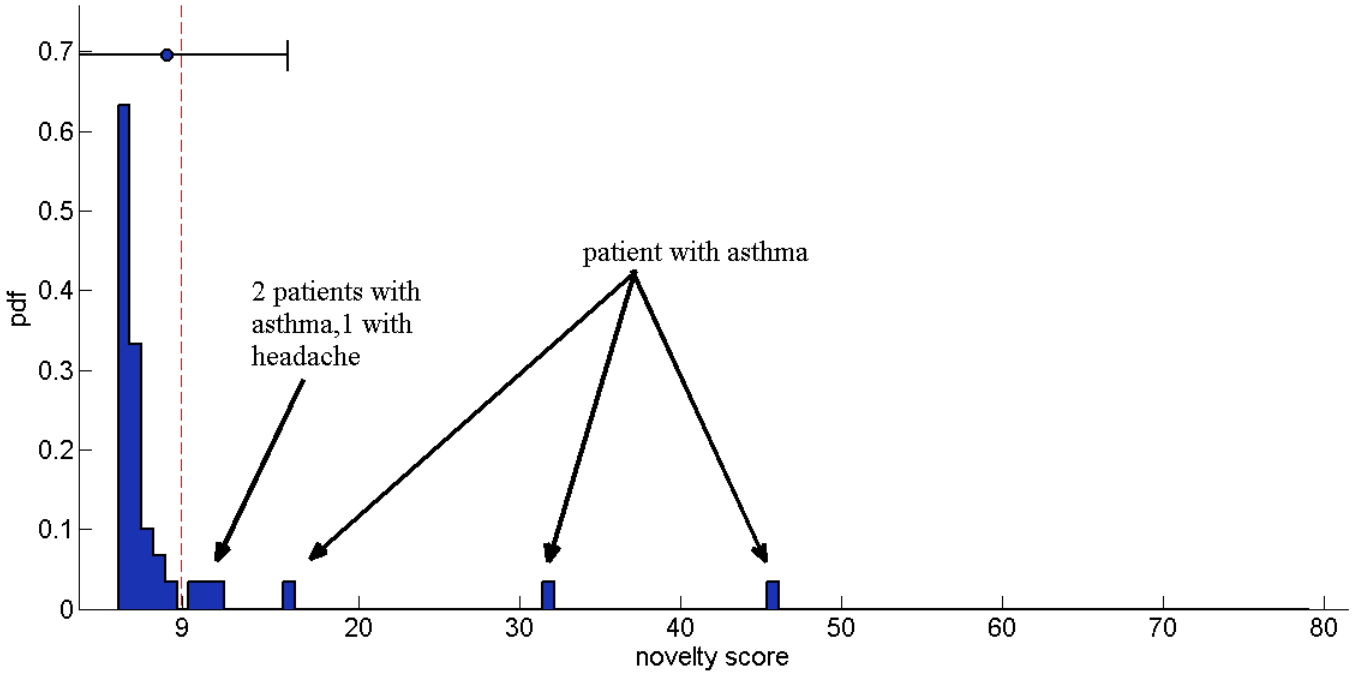


Figure 8.5: distribution of novelty score for the "non-infection" group, illustrating that asthma also contributes to high novelty scores (above 9) and accounts for the outliers in the group

Class (number)	Mean (Standard deviation)
Serious infection (93)	14.85 (15.87)
Intermediate infection (155)	10.87 (11.07)
Minor infection (279)	8.43 (4.71)
No infection (41)	8.07 (7.48)
No infection after removing patients with asthma (36)	6.01 (1.11)

Table 8.6: Mean (standard deviation) of novelty score in the four classes of the Walsgrave database using the model of normality built with data from the Nottingham and Oxford Paediatric Study database

While significant overlap exists between the four classes, there is a gradual shift from low novelty score to high novelty score as the infection becomes more serious.

Novelty-detection based classification is a viable and practical approach to using vital signs for paediatric triage. It is not possible to use multiclasss classification approaches since the overlap between the various classes is very large, with "normal" children the only class which is well-defined. As an illustration, figure 8.6 shows the overlap of the vital signs for children with asthma (labelled M), pneumonia (labelled A), viral illness (labelled G), bronchiolitis (labelled C) and the

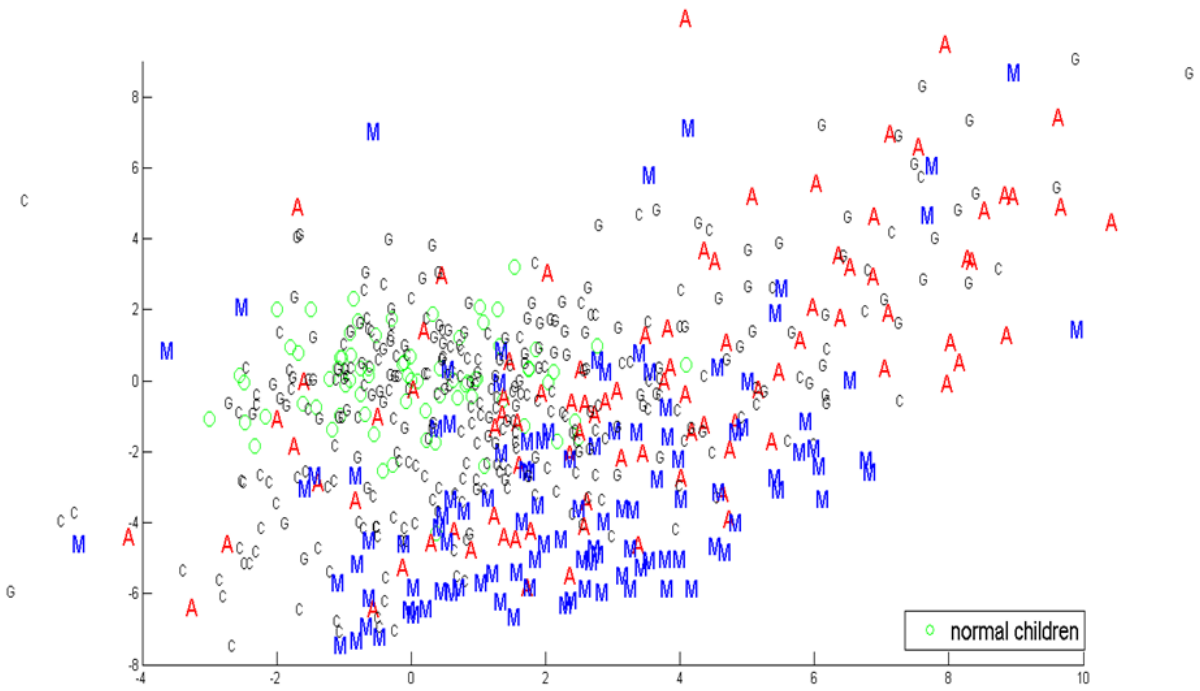


Figure 8.6: illustration to show the large extent of overlap among the different classes of children in the Nottingham database. This figure was obtained using a Neuroscale visualisation algorithm using the data from the normal children's class as the training set.

set of normal children (shown in green circles). This represents less than half of the total number of children in the Nottingham dataset but, even then, there is a significant overlap among the various classes thus making it impossible to use a multiclass classification approach to the identification of childhood illnesses using vital signs only.

It was thus concluded that the four vital signs can best be used to find the two extremes i.e. patients with serious infection (rule-in) and patients who are healthy (rule-out) (see figure 8.7).

8.5.2 Data fusion for the identification of children with serious infection (Rule-in)

As shown in the previous chapter, the values of the vital signs of patients with asthma are significantly affected (especially those with acute exacerbation of asthma). Out of the 41 patients labelled as having "no infection", there were 5 patients with asthma. The classification results are provided both with and without the 5 patients with asthma in the "no infection" group. Table 8.7 gives the

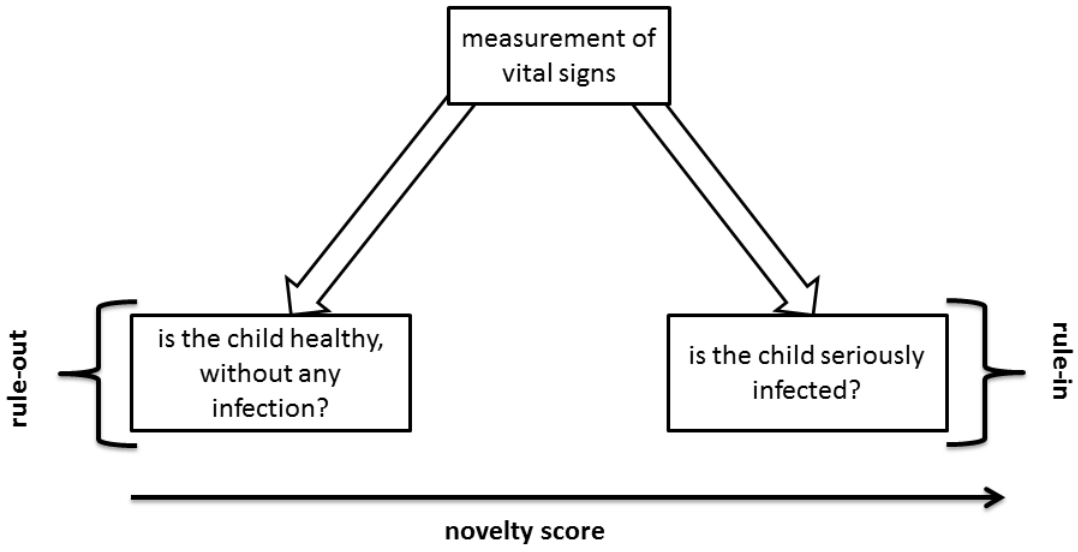


Figure 8.7: A vital signs-based classification model developed to aid clinicians during initial assesment of children

results using novelty detection to identify patients with serious infection using the model of normality trained with the "Expanded Normal" group. The sensitivity and specificity corresponding to the point closest to the top left corner (0,1) are reported in the table. For illustration, figure 8.8 shows the ROC curve with and without including patients with asthma.

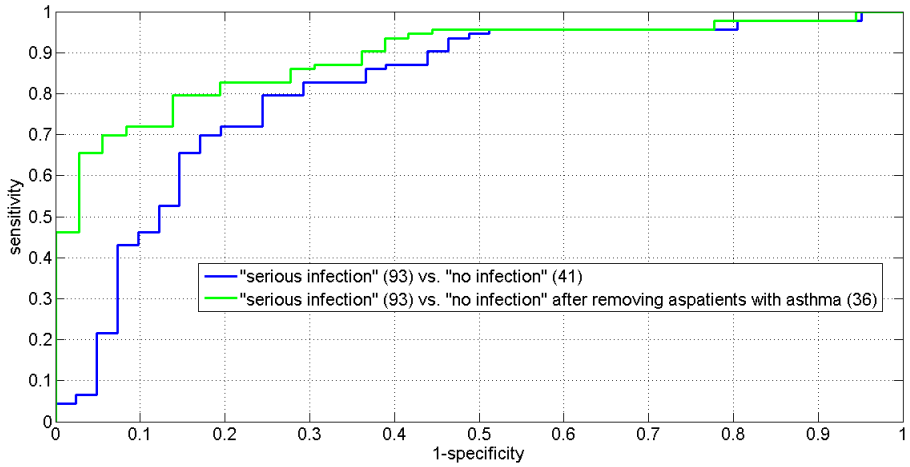


Figure 8.8: ROC curve to illustrate the performance of novelty detection in identifying patients with serious infection in a subset of the Walsgrave dataset consisting of patients with no infection (with and without those diagnosed with asthma)

Serious infection (93) vs.	Sensitivity %	Specificity %	AUC
No infection (41)	80	76	0.811
No infection after removal of patients with asthma (36)	80	86	0.888
No/Minor/Intermediate infecion (475)	63	65	0.682
No/Minor/Intermediate infecion after removal of patients with asthma (470)	63	66	0.686

Table 8.7: Identification of serious infection in the Walsgrave database using novelty detection with model of normality built using data from the "Nottingham" database and the Oxford Paediatric Triage study. Sensitivity and specificity of the point closest to (0,1) on the ROC curve are reported.

As in the previous section, an iterative algorithm was implemented to report results on a test set containing equal numbers of normal and abnormal patients with $N = 50$ (number of samples) and $k = 1000$ (number of iterations) and the results are reported in Table 8.8.

Serious infection (93) vs.	Sensitivity (%)	Specificity (%)	AUC
No/Minor/Intermediate infecion (475)	62 (54-69)	66 (60-70)	0.683 (0.611-0.745)
No/Minor/Intermediate infecion after removal of patients with asthma (470)	62 (56-66)	66 (60-73)	0.687 (0.618-0.748)

Table 8.8: Identification of serious infection in the Walsgrave database using novelty detection with model of normality built using data from the "Nottingham" database and the Oxford Paediatric Triage study. Median (10-90%) sensitivity and specificity of the point closest to (0,1) on the ROC curve are reported.

8.5.3 Data fusion for the identification of healthy children (Rule-out)

It is expected that the vital signs of children with no infection will be least disturbed, if at all and thus a threshold on novelty score could be set such that patients with novelty score below this threshold are classified as healthy. Table 8.7 gives the results using novelty detection to identify healthy patients (i.e. patients in the "no infection" group) using the model of normality trained using the "Expanded Normal" group. The sensitivity and specificity corresponding to the point closest to the top left corner (0,1) are reported in the table. As before, an iterative algorithm with $k = 1000$ and $N = 50$ was implemented and the 10^{th}, 50^{th} and 90^{th} percentile ROC and the corresponding AUC are reported in Table 8.10. For illustration, figure 8.9 shows the 10^{th}, 50^{th} and 90^{th} percentile ROC for separating healthy children (without asthma) from children with infection.

Comparison	Sensitivity (%)	Specificity (%)	AUC
No infection excluding asthma patients (36) vs rest of the database	65	86	0.782
No infection including asthma patients (41) vs rest of the database	65	76	0.706

Table 8.9: Identification of children with no infection in the Walsgrave database using novelty detection with model of normality built using data from the "Nottingham" database and the Oxford Paediatric Triage study. The figures for sensitivity and specificity for the point closest to (0,1) on the ROC curve are reported.

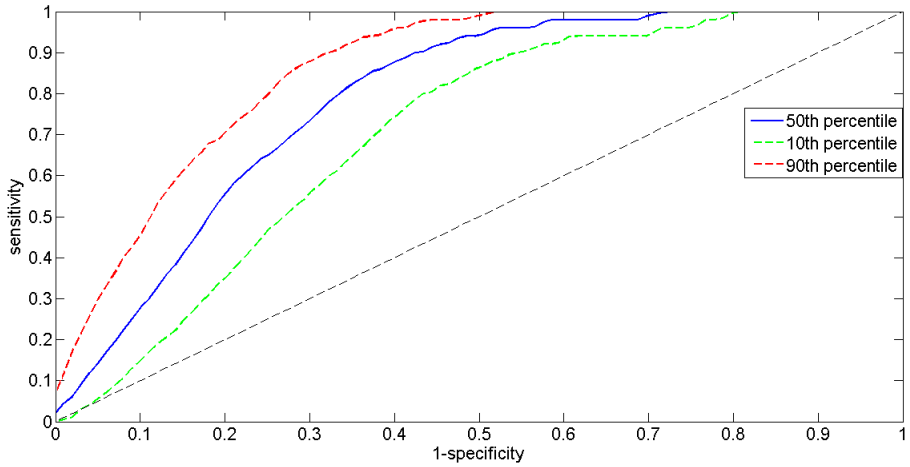


Figure 8.9: ROC curve to illustrate the performance of novelty detection in separating healthy children (excluding children with asthma) from children with infection in the Walsgrave database using a model of normality trained using the "Expanded Normal" database. The 10^{th}, 50^{th} and 90^{th} percentile ROCs are plotted. The performance of a random classifier with 50% accuracy is also shown for reference.

Comparison	Sensitivity (%)	Specificity (%)	AUC
No infection excluding asthma patients (36) vs rest of the database (527)	68 (60-73)	78 (75-84)	0.785 (0.723-0.842)
No infection including asthma patients (41) vs rest of the database (527)	65 (59-70)	71 (66-78)	0.705 (0.635-0.769)

Table 8.10: Identification of children with no infection in the Walsgrave database using novelty detection with model of normality built using data from the "Nottingham" database and the Oxford Paediatric Triage study. Median (10-90%) sensitivity and specificity of the point closest to (0,1) on the ROC curve are reported

8.5.4 Combining rule-in and rule-out to aid clinical decisions

The novelty detection classification system can be employed with an upper threshold (to rule in patients suspected of serious infection) and a lower threshold (to rule out healthy patients). The choice of the appropriate threshold in each case depends on the requirements, the available resources and the clinical setting. For illustration, figure 8.10 (a) shows the results obtained on the Walsgrave dataset after selecting thresholds based on selecting the operating point closest to (0,1) on the ROC curve. The results suggest that the lower threshold is too high since there are 22% of

Age (years)	Temperature	Heart Rate	SpO_2	Respiratory rate	Diagnosis
11.9	36.6	80	100	18	appendicitis
11.0	36.7	94	98	20	appendicitis
1.0	36.9	139	98	29	pneumonia
7.3	36.6	99	98	18	bacteraemia

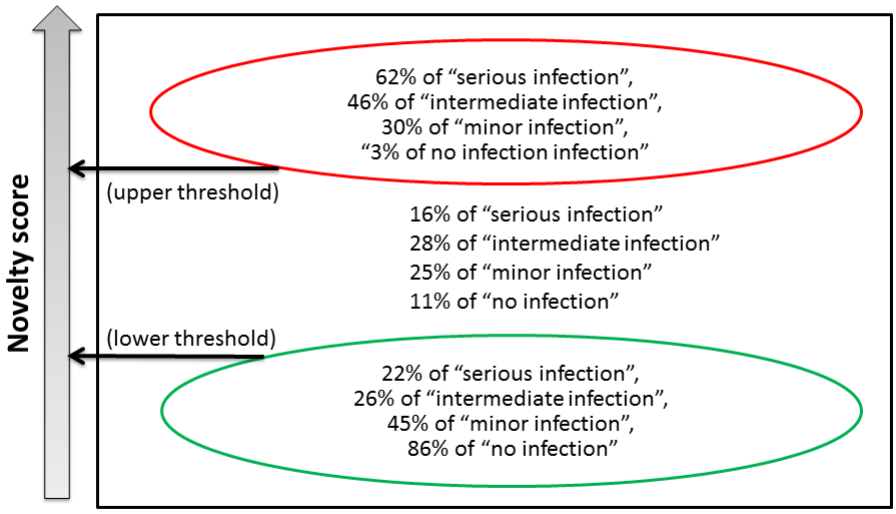
Table 8.11: Vital signs, age and final diagnosis for the 4% patients in the "serious infection" category wrongly classified as healthy.

patients with "serious infection" below the threshold, even though 86% of patients with no infection are also ruled out. This threshold can be reduced (as shown in figure 8.10 (b)) since it is more important not to miss any child with a serious infection than it is not to rule out a child who is healthy. The vital sign values, age and the final diagnosis for the 4% patients with serious infection still below the lower threshold in figure 8.10 (b) are shown in Table 8.11. These 4 patients have a low novelty score because their vital signs appear to be normal for their age. It is to be hoped that the presenting symptoms (e.g. stomach pain for appendicitis) would identify them as having a serious infection.

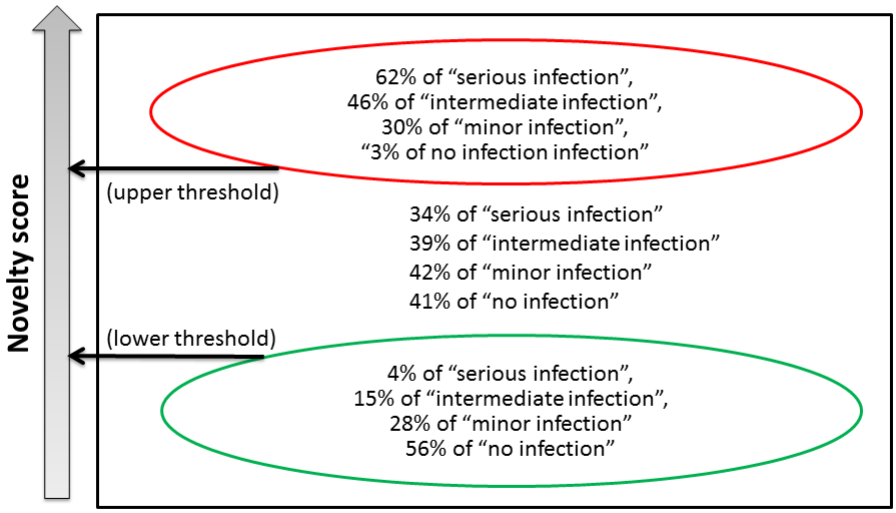
8.6 Data fusion using independent test set (Pinderfields)

Most of the patients recruited in the Pinderfields study had all the four vital signs available, however the SpO_2 measurement was found to be missing in a significant minority of patients (see Table 8.12). It was therefore decided to classify the test data using a 4-vital sign model of normality as before in this chapter (Walsgrave dataset) and then using a 3-vital sign model of normality (without SpO_2 in the model).

Table 8.13 provides the classification results using the 4-dimensional model of normality trained on the "Expanded Normal" group using, as before, data from the Oxford Paediatric Triage study and the Nottingham database. Subsequently, a 3-dimensional model of normality was constructed using the same (Expanded Normal) dataset but without SpO_2, with the classification results being given in Table 8.14.



(a) based on selecting the threshold closest to (0,1) for both the rule-in and rule-out cases



(b) based on a reduced "lower threshold" to lower the percentage of children with "serious infection" being classified as healthy. The vital sign values, age and final diagnosis for the 4% of children with "serious infection" in the green category are shown in Table 8.11)

Figure 8.10: Classification results of the Walsgrave dataset using "Nottingham + Oxford" model of normality, combining rule-in and rule-out

Vital Signs present	Control cases	Pneumonia	UTI	Sepsis
Pulse, Temperature, Respiratory rate, SpO_2	461	72	16	72
Pulse, Temperature, Respiratory rate	781	78	31	94

Table 8.12: Summary of the number of patients in the various groups in the Pinderfields database

Control cases (461) vs.	Sensitivity (%)	Specificity (%)	AUC
Pnuemonia (72)	74	75	0.802
Sepsis (72)	68	61	0.679
UTI (16)	69	50	0.553
Pneumonia/Sepsis (144)	63	74	0.741
Pneumonia/Sepsis/UTI (160)	69	64	0.723

Table 8.13: Identification of children with Pneumonia/Sepsis/UTI in the Pinderfields database using novelty detection with a 4-D model of normality built using data from the "Nottingham" database and the Oxford Paediatric Triage study. Sensitivity and specificity of the point closest to (0,1) on the ROC curve are reported

Control cases (781) vs.	Sensitivity (%)	Specificity (%)	AUC
Pnuemonia (78)	74	70	0.785
Sepsis (94)	64	67	0.689
UTI (31)	68	54	0.599
Pneumonia/Sepsis (172)	70	67	0.733
Pneumonia/Sepsis/UTI (203)	66	67	0.712

Table 8.14: Identification of children with Pneumonia/Sepsis/UTI in the Pinderfields database using novelty detection with a 3-D model of normality built using data from the "Nottingham" database and the Oxford Paediatric Triage study. Sensitivity and specificity of the point closest to (0,1) on the ROC curve are reported

Comparing the results in Table 8.13 with those in Table 8.14, it can be seen that the addition of SpO_2 does not lead to any significant change in results. This could possibly be due to correlation of SpO_2 with other vital signs. For example, a high respiratory rate might be indicative of breathing difficulties which are likely to lower SpO_2.

8.7 Comparison of results with previous work

The only comparable method to the novelty detection classification scheme presented in this thesis is the clinical decision rule developed by Thompson et al. [140] and reviewed in section 8.3.1. Here, the classification performance of both methods on the Pinderfields data set as an independent data set are compared.

From Table 8.15, it can be seen that the novelty score based classifier performs consistently better across the main childhood illnesses than the clinical decision rule developed by Matthew et al. [140]. For similar sensitivity values, the specificity is between 8% to 18% better, except for sepsis where the improvement is only 4%.

Prediction (versus the case control group)	Clinical decision rule (sensitivity/specificity)	novelty score based classification (sensitivity/specificity)
Pneumonia	81% / 49%	81% / 60%
Sepsis	75% / 49%	76% / 53%
UTI	59% / 49%	61% / 57%
Pneumonia/Sepsis/UTI	75% / 49%	75% / 67%

Table 8.15: Comparison of results on the Pinderfields dataset

8.8 Conclusion

This chapter has shown that it is possible to use the novelty detection framework (developed in the previous chapter) to improve the triage process. Two completely independent datasets (Walsgrave and Pinderfields) were used to test novelty score based classification. Although individual vital signs are disturbed to a varying degree by a large number of illnesses (both self-limiting illnesses and serious infection), the novelty score combines the information from multiple vital signs to predict with reasonable accuracy the presence and severity of an infection. In particular, an upper threshold and a lower threshold on the novelty score can be selected to rule in patients with serious infection and rule out patients unlikely to have an infection.

With the upper and lower thresholds chosen for the Walsgrave dataset, 62% of the children with a serious infection but only 3% of healthy children had a novelty score above the upper threshold.

At the same time, 56% of children with no infection and just 4% of children with serious infection had a novelty score below the lower threshold.

Results using a case-controlled study showed that the novelty score can be used to distinguish patients who are seriously ill (e.g. pneumonia, sepsis, UTI) from children with self-limiting infections (e.g. tonsillitis, viral infection, URTI) admitted to hospital. Best results were achieved with distinguishing patients with pneumonia from the control group (AUC=0.80) followed by children with sepsis (AUC=0.68) and least good for patients with UTI (AUC=0.55).

Comparison with previous work showed that the novelty score based classification provided slightly better results than the simple clinical decision rules developed in [140].

Chapter 9

Conclusion

The work in this thesis has provided evidence for a monitoring and data fusion system that can be used in different clinical settings to improve the diagnosis of children during their initial assessment based on four vital signs. The major topics covered in the thesis were:

- Review of the physiological basis for the respiratory rate information in the photoplethysmogram (PPG) and of the existing algorithms for the estimation of respiratory rate from the PPG (Chapter 2)

- Acquisition of vital sign data, during routine clinical assessment in a clinical trial from 272 children presenting to the Emergency Department at the John Radcliffe Hospital, Oxford, UK (Chapter 3)

- Development and validation of time-domain and frequency-domain techniques for the estimation of respiratory rate from the PPG, with a focus on a novel approach based on autroregressive modelling (Chapters 4 and 5)

- Development of adaptive template matching for signal quality assessment and investigation of methods for incorporating signal quality information to improve the accuracy of the respiratory rate estimation algorithms (Chapter 6)

- Investigation of the extent of vital sign disturbance caused by various illnesses using novelty detection and ranking of diagnoses based on a novelty score, using data from over 1300

unwell children presenting to the ED (Chapter 7)

- Assessment of the diagnostic accuracy of a data fusion system based on novelty detection and applied to paediatric vital sign data from independent datasets, the aim being to distinguish children with serious infection from children who are healthy and/or with a self-limiting infection (Chapter 8)

9.1 Overview

The initial investigations (Chapter 4) focused on respiration-induced amplitude modulation (AM) of the PPG. Peak detection methods and autoregressive-model methods were developed for respiratory rate estimation. Various approaches were developed to address some of the limitations of AR-modelling including multiple AR modelling, envelope normalization and peak detection in the frequency spectrum. The techniques developed were subsequently applied to the analysis of the Respiratory Sinus Arrhythmia (RSA) component in the PPG. Respiratory rate estimation based on RSA analysis consistently gave better results.

In Chapter 6, an algorithm for deriving a signal quality index (SQI) based on adaptive template matching was developed. The SQI was used to identify the highest quality PPG segments for estimating the respiratory rate. The accuracy of the respiratory rate estimates was improved by fusing the time-domain and frequency-domain approaches after identifying the limitation of the latter at high respiratory rates. Further improvement was possible after fusing the AM-based and the RSA-based algorithms using pulse rate. The final algorithm developed had a mean absolute error of 6.6 breaths/minute, with 119 patients (54%) having a respiratory rate estimate within 5 breaths per minute. A source of error was identified to be the inherent physiological variability in respiratory rate, which was quantified using PPG recordings from 86 patients with high-quality data.

In the final part of the thesis, the effect of childhood illnesses on vital signs was investigated, using a dataset comprising the data from the 1,386 children presenting to an ED (Nottingham). A model of normality was constructed using vital sign data from a group of normal children and the

mean novelty scores for various diagnosis groups were then determined and ranked. This work provided insights into which vital signs are affected by a given diagnosis group and to which extent this occurs. The use of the normal children from the "Oxford paediatric study dataset" enabled us to test the model of normality with an independent test set.

In chapter 8, the novelty detection framework was extended to classify patients based on their novelty score. It was shown that the novelty score combines the information from multiple vital signs to predict with reasonable accuracy the presence and severity of infection. Results on an independent test set (Walsgrave) showed that appropriate thresholds on the novelty score could be determined to rule in patients with a serious infection and rule out patients unlikely to have any infection. The same novelty detection framework applied to data from another independent dataset, a case-controlled study (Pinderfields), showed that the novelty score could be used to distinguish patients with a severe infection (e.g. pneumonia, vital infection, UTI) from those with self-limiting infections (e.g. tonsillitis, URTI).

Figure 9.1 summarises the major contributions of this thesis (shaded in green) to the problems encountered in the context of paediatric triage (shaded in blue).

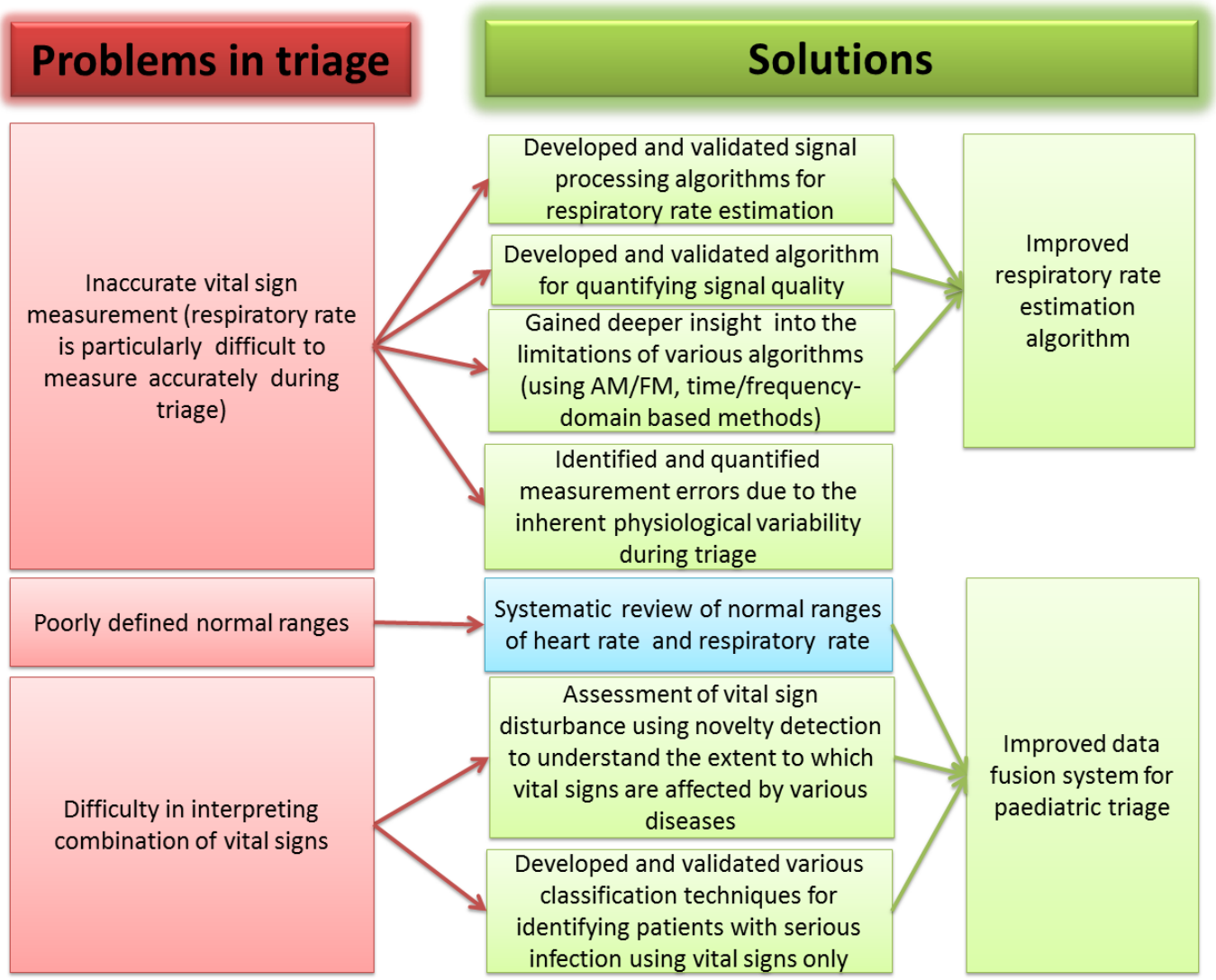


Figure 9.1: Solutions to problems in paediatric triage addressed in this thesis

9.2 Future Work

Future work is divided into two main parts: that which will be done with the available data, already analysed in this thesis; and work that will require a new clinical study.

9.2.1 Use of available datasets

Respiratory rate estimation algorithm:

(i) The template creation stage used in deriving a Signal Quality Index (SQI) can be improved

by using a modelling approach similar to the work done in [88], by using a sum of two Gaussians to represent a single PPG cycle for example. Two Gaussians will be a reasonable choice, with one to model the systolic peak of the PPG and the second to model the dicrotic notch found in some PPG waveforms. The algorithm used in this thesis creates a template using PPG beat segments that are within certain thresholds determined from what is considered to correspond to a physiologically plausible heart rate. However, it is still possible that certain PPG beat segments will be corrupted even if they are within the thresholds applied. The modelling approach proposed will allow us to use prior information (based on what a single PPG cycle should look like) and discard any PPG beat segments before creating a template.

(ii) It might be possible to improve the signal reconstruction stage described in section 6.7 by additional pre-processing before applying the AR model to the reconstructed signal. The concatenation of different sections of high-quality PPG segments can introduce discontinuities which will affect the performance of the AR algorithm. In this thesis, this effect was partly mitigated by using small sliding window (15 seconds) with large overlaps (14 seconds) and partly by ensuring that the different PPG segments were joined together using a common salient point.

Data fusion for paediatric triage: In this thesis, novelty detection was based on a Parzen windows model of normality. While it is clear that multiclass classification is not likely to be useful in the context of paediatric triage, it might be possible to improve the model of normality by using a different technique (e.g. one-class support vector machines, Gaussian mixture models etc.).

9.2.2 Future work requiring a new clinical trial

The ED study provided evidence for the feasibility of applying the proposed monitoring system to paediatric triage and showed that respiratory rate could be accurately estimated from PPG waveform data (within 5 breaths per minute of the reference rate) in over half the children.

Event marker switch interfaced with the software: The evaluation of the algorithm is likely to have been limited by the accuracy of the manual method of counting chest wall movement and the inherent physiological variability of respiratory rate during the two-minute monitoring period. A future study should, first of all, eliminate the error introduced by the physiological variability by

ensuring that the timing of the nurse's respiratory rate estimation is known. This could be achieved by using a simple wireless switch interfaced to the software (event marker). The nurses could be instructed to press the button as soon as they started counting chest wall movements and to press again as soon as they stopped. This will also record the counting period used by the nurse and will enable a further assessment of the limitations of manual estimation of respiratory rate.

Effect of illness on respiratory rate estimation: With a larger number of children with clearly recorded diagnoses, it would be useful to investigate if certain illnesses affect the accuracy of respiratory rate estimation. This would enable us to find out if certain illnesses are more likely to reduce the accuracy of AM and/or RSA-based techniques (e.g. children with URTI versus children with lower respiratory tract infections).

Potential biomarkers in heart rate and respiratory variability: A major advantage of having an automated monitoring device such as that developed in this thesis is the ability to record and save continuous PPG data. Even though current practice relies on a single measurement, the evidence in this thesis has shown that there is variability in both heart rate and respiratory rate during a two-minute monitoring period. It would be interesting to investigate the variability patterns in vital sign data to identify whether there may be biomarkers of various illnesses.

Combining monitoring and data fusion: It was not possible to apply the novelty detection algorithms to the data acquired during the ED clinical trial due to the small number of patients in the trial. A future study should aim not only to provide PPG waveform data for respiratory rate estimation but also to recruit a large number of children with a clearly-defined final diagnosis. The novelty detection algorithm could then be tested, both with the respiratory rate estimated by the triage nurse and the respiratory rate estimated by the automated algorithm. Such an investigation would then also allow us to quantify the improvement, if any, of using respiration rate estimated from the PPG rather than by counting chest wall movement. This would also allow us to find out how much error in respiratory rate can be tolerated before the error becomes clinically significant.

Bibliography

[1] PS Addison and JN Watson. Secondary wavelet feature decoupling (swfd) and its use in detecting patient respiration from the photoplethysmogram. In *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, volume 3, pages 2602–2605. IEEE, 2003.

[2] P.S. Addison and J.N. Watson. Secondary transform decoupling of shifted nonstationary signal modulation components: application to photoplethysmography. *International Journal of Wavelets, Multiresolution and Information Processing*, 2(1):43–57, 2004.

[3] M. Akre, M. Finkelstein, M. Erickson, M. Liu, L. Vanderbilt, and G. Billman. Sensitivity of the pediatric early warning score to identify patient deterioration. *Pediatrics*, 125(4):e763, 2010.

[4] S.T. Anderson. *Advanced electrocardiography*. SpaceLabs, 1992.

[5] H.H. Asada, H.H. Jiang, and P. Gibbs. Active noise cancellation using mems accelerometers for motion-tolerant wearable bio-sensors. In *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, volume 1, pages 2157–2160. IEEE, 2004.

[6] L. Bernardi, D. Hayoz, R. Wenzel, C. Passino, A. Calciati, R. Weber, and G. Noll. Synchronous and baroceptor-sensitive oscillations in skin microcirculation: evidence for central autonomic control. *American Journal of Physiology-Heart and Circulatory Physiology*, 273(4):H1867, 1997.

[7] R. Beveridge et al. Caep issues. the canadian triage and acuity scale: a new and critical element in health care reform. canadian association of emergency physicians. *The Journal of emergency medicine*, 16(3):507, 1998.

[8] S. Bhal, V. Tygai, N. Kumar, V. Sreenivas, JM Puliyel, et al. Signs of inflammation in children that can kill (sick score): Preliminary prospective validation of a new non-invasive measure of severity-of-illness. *Journal of postgraduate medicine*, 52(2):102, 2006.

[9] A. Bhattacharyya. On a measure of divergence between two multinomial populations. *Sankhyā: The Indian Journal of Statistics (1933-1960)*, 7(4):401–406, 1946.

[10] C.M. Bishop and SpringerLink (Service en ligne). *Pattern recognition and machine learning*, volume 4. springer New York, 2006.

[11] SE Bleeker, G. Derksen-Lubsen, DE Grobbee, ART Donders, KGM Moons, and HA Moll. Validating and updating a prediction rule for serious bacterial infection in patients with fever without source. *Acta Paediatrica*, 96(1):100–104, 2007.

[12] SE Bleeker, HA Moll, EW Steyerberg, ART Donders, G. Derksen-Lubsen, DE Grobbee, and KGM Moons. External validation is necessary in prediction research:: A clinical example. *Journal of clinical epidemiology*, 56(9):826–832, 2003.

[13] M. Bolanos, H. Nazeran, and E. Haltiwanger. Comparison of heart rate variability signal features derived from electrocardiography and photoplethysmography in healthy individuals. In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, pages 4289–4294. IEEE, 2006.

[14] A.A. Bove and W.P. Santamore. Ventricular interdependence. *Progress in cardiovascular diseases*, 23(5):365–388, 1981.

[15] A.P. Bradley. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern recognition*, 30(7):1145–1159, 1997.

[16] A.J. Brent, M. Lakhanpaul, M. Thompson, J. Collier, S. Ray, N. Ninis, M. Levin, and R. MacFaul. Risk score to stratify children with suspected serious bacterial infection: observational cohort study. *Archives of disease in childhood*, 96(4):361–367, 2011.

[17] R.J. Brilli, R. Gibson, J.W. Luria, T.A. Wheeler, J. Shaw, M. Linam, J. Kheir, P. McLain, T. Lingsch, A. Hall-Haering, et al. Implementation of a medical emergency team in a large pediatric teaching hospital prevents respiratory and cardiopulmonary arrests outside the intensive care unit*. *Pediatric Critical Care Medicine*, 8(3):236, 2007.

[18] J. Bryce, C. Boschi-Pinto, K. Shibuya, and R.E. Black. Who estimates of the causes of death in children. *The Lancet*, 365(9465):1147–1152, 2005.

[19] S. Cazares, M. Moulden, C.W.G. Redman, and L. Tarassenko. Tracking poles with an autoregressive model: a confidence index for the analysis of the intrapartum cardiotocogram. *Medical engineering & physics*, 23(9):603–614, 2001.

[20] J.M. Chamberlain, K.M. Patel, U.E. Ruttimann, and M.M. Pollack. Pediatric risk of admission (prisa): a measure of severity of illness for assessing the risk of hospitalization from the emergency department. *Annals of emergency medicine*, 32(2):161–169, 1998.

[21] KW Chan and YT Zhang. Adaptive reduction of motion artifact from photoplethysmographic recordings using a variable step-size lms filter. In *Sensors, 2002. Proceedings of IEEE*, volume 2, pages 1343–1346. IEEE, 2002.

[22] S.M. Chapman, M.P.W. Grocott, and L.S. Franck. Systematic review of paediatric alert criteria for identifying hospitalised children at risk of critical deterioration. *Intensive care medicine*, 36(4):600–611, 2010.

[23] K.H. Chon, S. Dash, and K. Ju. Estimation of respiratory rate from photoplethysmogram data using time–frequency spectral estimation. *Biomedical Engineering, IEEE Transactions on*, 56(8):2054–2063, 2009.

[24] D. Clifton, J.G. Douglas, P.S. Addison, and J.N. Watson. Measurement of respiratory rate from the photoplethysmogram in chest clinic patients. *Journal of clinical monitoring and computing*, 21(1):55–61, 2007.

[25] F.M. Coetzee and Z. Elghazzawi. Noise-resistant pulse oximetry using a synthetic reference signal. *Biomedical Engineering, IEEE Transactions on*, 47(8):1018–1026, 2000.

[26] JG Cronin. The introduction of the manchester triage scale to an emergency department in the republic of ireland. *Accident and emergency nursing*, 11(2):121–125, 2003.

[27] D. Cysarz, R. Zerm, H. Bettermann, M. Frühwirth, M. Moser, and M. Kröz. Comparison of respiratory rates derived from heart rate variability, ecg amplitude, and nasal/oral airflow. *Annals of biomedical engineering*, 36(12):2085–2094, 2008.

[28] A.V. Deshmane, R.G. Mark, L.J. Kessler, et al. *False arrhythmia alarm suppression using ecg, abp, and photoplethysmogram*. PhD thesis, Massachusetts Institute of Technology, 2009.

[29] JC Dorlas and JA Nijboer. Photo-electric plethysmography as a monitoring device in anaesthesia. *British journal of anaesthesia*, 57(5):524, 1985.

[30] H. Duncan, J. Hutchison, and C.S. Parshuram. The pediatric early warning system score: a severity of illness score to predict urgent medical need in hospitalized children. *Journal of critical care*, 21(3):271–278, 2006.

[31] E.D. Edwards, CVE Powell, B.W. Mason, and A. Oliver. Prospective cohort study to test the predictability of the cardiff and vale paediatric early warning system. *Archives of disease in childhood*, 94(8):602–606, 2009.

[32] P. Egdell, L. Finlay, and DK Pedley. The paws score: validation of an early warning scoring system for the initial assessment of children in the emergency department. *Emergency Medicine Journal*, 25(11):745–749, 2008.

[33] AS El-Radhi and W. Barry. Thermometry in paediatric practice. *Archives of disease in childhood*, 91(4):351–356, 2006.

[34] S. Fleming. *Measurement and fusion of non-invasive vital signs for routine triage of acute paediatric illness*. PhD thesis, University of Oxford, 2010.

[35] S. Fleming, L. Tarassenko, M. Thompson, and D. Mant. Non-invasive measurement of respiratory rate in children using the photoplethysmogram. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 1886–1889. IEEE, 2008.

[36] S. Fleming, M. Thompson, R. Stevens, C. Heneghan, A. Pluddemann, I. Maconochie, L. Tarassenko, and D. Mant. Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies. *The Lancet*, 2011.

[37] S.G. Fleming and L. Tarassenko. A comparison of signal processing techniques for the extraction of breathing rate from the photoplethysmogram. *International Journal of Biological and Medical Sciences*, 2(4), 2007.

[38] M. Folke, L. Cernerud, M. Ekström, and B. Hök. Critical review of non-invasive respiratory monitoring in medical care. *Medical and Biological Engineering and Computing*, 41(4):377–383, 2003.

[39] J.Y.A. Foo and S.J. Wilson. Estimation of breathing interval from the photoplethysmographic signals in children. *Physiological measurement*, 26:1049, 2005.

[40] J.E. Gentle, W. Härdle, and Y. Mori. *Handbook of computational statistics: concepts and methods*. Springer Verlag, 2004.

[41] P. Gibbs and H.H. Asada. Reducing motion artifact in wearable bio-sensors using mems accelerometers for active noise cancellation. In *American Control Conference, 2005. Proceedings of the 2005*, pages 1581–1586. IEEE, 2005.

[42] E. Gil, J. María Vergara, and P. Laguna. Detection of decreases in the amplitude fluctuation of pulse photoplethysmography signal as indication of obstructive sleep apnea syndrome in children. *Biomedical Signal Processing and Control*, 3(3):267–277, 2008.

[43] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, 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):e215–e220, 2000.

[44] J.M. Goldman, M.T. Petterson, R.J. Kopotic, and S.J. Barker. Masimo signal extraction pulse oximetry. *Journal of clinical monitoring and computing*, 16(7):475–483, 2000.

[45] S. Gouin, J. Gravel, D.K. Amre, and S. Bergeron. Evaluation of the paediatric canadian triage and acuity scale in a pediatric ed. *The American journal of emergency medicine*, 23(3):243–247, 2005.

[46] P. Grossman, FH Wilhelm, and M. Spoerle. Respiratory sinus arrhythmia, cardiac vagal control, and daily activity. *American Journal of Physiology-Heart and Circulatory Physiology*, 287(2):H728, 2004.

[47] C. Haines, M. Perrott, and P. Weir. Promoting care for acutely ill children.development and evaluation of a paediatric early warning tool. *Intensive and Critical Care Nursing*, 22(2):73–81, 2006.

[48] P.S. Hamilton and W.J. Tompkins. Quantitative investigation of qrs detection rules using the mit/bih arrhythmia database. *Biomedical Engineering, IEEE Transactions on*, (12):1157–1165, 1986.

[49] M.J. Hayes and P.R. Smith. A new method for pulse oximetry possessing inherent insensitivity to artifact. *Biomedical Engineering, IEEE Transactions on*, 48(4):452–461, 2001.

[50] B. Hjorth. Eeg analysis based on time domain properties. *Electroencephalography and Clinical Neurophysiology*, 29(3):306–310, 1970.

[51] C.W. Hsieh, C.W. Mao, M.S. Young, T.L. Yeh, and S.J. Yeh. Respiratory effect on the pulse spectrum. *Journal of medical engineering & technology*, 27(2):77–84, 2003.

[52] JA Innes, SC De Cort, W. Kox, and A. Guz. Within-breath modulation of left ventricular function during normal breathing and positive-pressure ventilation in man. *The Journal of Physiology*, 460(1):487, 1993.

[53] A. Johansson. Neural network for photoplethysmographic respiratory rate monitoring. *Medical and Biological Engineering and Computing*, 41(3):242–248, 2003.

[54] A. Johansson, P.Å. Öberg, and G. Sedin. Monitoring of heart and respiratory rates in newborn infants using a new photoplethysmographic technique. *Journal of clinical monitoring and computing*, 15(7):461–467, 1999.

[55] A. Johansson and T. Strömberg. Influence of tidal volume and thoraco-abdominal separation on the respiratory induced variation of the photoplethysmogram. *Journal of clinical monitoring and computing*, 16(8):575–581, 2000.

[56] WS Johnston and Y. Mendelson. Extracting breathing rate information from a wearable reflectance pulse oximeter sensor. In *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, volume 2, pages 5388–5391. IEEE, 2004.

[57] J. Kaiser and R. Schafer. On the use of the i0-sinh window for spectrum analysis. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 28(1):105–107, 1980.

[58] S.M. Kay and S.L. Marple Jr. Spectrum analysis: a modern perspective. *Proceedings of the IEEE*, 69(11):1380–1419, 1981.

[59] B.S. Kim and S.K. Yoo. Motion artifact reduction in photoplethysmography using independent component analysis. *Biomedical Engineering, IEEE Transactions on*, 53(3):566–568, 2006.

[60] B.R. Kirkwood and J.A.C. Sterne. *Essential medical statistics*. Wiley-Blackwell, 2003.

[61] H. Kocoglu, S. Goksu, M. Isik, Z. Akturk, and Y.A. Bayazit. Infrared tympanic thermometer can accurately measure the body temperature in children in an emergency room setting. *International journal of pediatric otorhinolaryngology*, 65(1):39–43, 2002.

[62] B. Krieger, D. Feinerman, A. Zaron, and F. Bizousky. Continuous noninvasive monitoring of respiratory rate in critically ill patients. *Chest*, 90(5):632–634, 1986.

[63] R. Krishnan, B. Natarajan, and S. Warren. Analysis and detection of motion artifact in photoplethysmographic data using higher order statistics. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 613–616. IEEE, 2008.

[64] S. Kullback and R.A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[65] CM Lee and YT Zhang. Reduction of motion artifacts from photoplethysmographic recordings using a wavelet denoising approach. In *Biomedical Engineering, 2003. IEEE EMBS Asian-Pacific Conference on*, pages 194–195. IEEE, 2003.

[66] J. Lee and K. Chon. Time-varying autoregressive model-based multiple modes particle filtering algorithm for respiratory rate extraction from pulse oximeter. *Biomedical Engineering, IEEE Transactions on*, (99):1–1, 2011.

[67] J. Lee and K.H. Chon. An autoregressive model-based particle filtering algorithms for extraction of respiratory rates as high as 90 breaths per minute from pulse oximeter. *Biomedical Engineering, IEEE Transactions on*, 57(9):2158–2167, 2010.

[68] J. Lee and KH Chon. Respiratory rate extraction via an autoregressive model using the optimal parameter search criterion. *Annals of biomedical engineering*, 38(10):3218–3225, 2010.

[69] J. Lee, W. Jung, I.T. Kang, Y. Kim, and G. Lee. Design of filter to reject motion artifact of pulse oximetry. *Computer Standards & Interfaces*, 26(3):241–249, 2004.

[70] P. Leonard, N.R. Grubb, P.S. Addison, D. Clifton, and J.N. Watson. An algorithm for the detection of individual breaths from the pulse oximeter waveform. *Journal of clinical monitoring and computing*, 18(5):309–312, 2004.

[71] P.A. Leonard, D. Clifton, P.S. Addison, J.N. Watson, and T. Beattie. An automated algorithm for determining respiratory rate by photoplethysmogram in children. *Acta Pædiatrica*, 95(9):1124–1128, 2006.

[72] Q. Li and G. D. Clifford. Dynamic time warping and machine learning for signal quality assessment of pulsatile signals. *Physiological measurement*, in submission, 2012.

[73] L.G. Lindberg, H. Ugnell, and PÅ Öberg. Monitoring of respiratory and heart rates using a fibre-optic sensor. *Medical and Biological Engineering and Computing*, 30(5):533–537, 1992.

[74] G. Lorenzi-Filho, H.R. DAJANI, R.S.T. Leung, J.S. Floras, and T.D. Bradley. Entrainment of blood pressure and heart rate oscillations by periodic breathing. *American journal of respiratory and critical care medicine*, 159(4):1147, 1999.

[75] P.B. Lovett, J.M. Buchwald, K. Stürmann, and P. Bijur. The vexatious vital: neither clinical measurements by nurses nor an electronic monitor provides accurate measurements of respiratory rate in triage. *Annals of emergency medicine*, 45(1):68–76, 2005.

[76] D. Lowe and M.E. Tipping. Neuroscale: Novel topographic feature extraction using rbf networks. 1997.

[77] S. Lu, H. Zhao, K. Ju, K. Shin, M. Lee, K. Shelley, and K.H. Chon. Can photoplethysmography variability serve as an alternative approach to obtain heart rate variability information? *Journal of clinical monitoring and computing*, 22(1):23–29, 2008.

[78] K. Mackway-Jones. *Advanced paediatric life support: the practical approach*. BMJ books, 2005.

[79] K. Mackway-Jones. *Advanced paediatric life support: the practical approach*. BMJ books, 2005.

[80] K.V. Madhav, M.R. Ram, E.H. Krishna, K.N. Reddy, and K.A. Reddy. A robust signal processing method for extraction of respiratory activity from artifact corrupted ppg signal. In *Recent Advances in Intelligent Computational Systems (RAICS), 2011 IEEE*, pages 451–456. IEEE, 2011.

[81] T. Maldonado and J.R. Avner. Triage of the pediatric patient in the emergency department: are we all in agreement? *Pediatrics*, 114(2):356–360, 2004.

[82] P.A. Maningas, D.A. Hime, and D.E. Parker. The use of the soterion rapid triage system in children presenting to the emergency department. *The Journal of emergency medicine*, 31(4):353–359, 2006.

[83] P.A. Maningas, D.A. Hime, D.E. Parker, and T.A. McMurry. The soterion rapid triage system: evaluation of inter-rater reliability and validity. *The Journal of emergency medicine*, 30(4):461–469, 2006.

[84] P. Margolis and A. Gadomski. Does this infant have pneumonia? *JAMA: the journal of the American Medical Association*, 279(4):308, 1998.

[85] M. Markou and S. Singh. Novelty detection: a review–part 1: statistical approaches. *Signal Processing*, 83(12):2481–2497, 2003.

[86] M. Markou and S. Singh. Novelty detection: a review part 2:: neural network based approaches. *Signal Processing*, 83(12):2499–2521, 2003.

[87] MK Marks, M. South, and JB Carlin. Reference ranges for respiratory rate measured by thermistry (12-84 months). *Archives of disease in childhood*, 69(5):569, 1993.

[88] P.E. McSharry, G.D. Clifford, L. Tarassenko, and L.A. Smith. A dynamical model for generating synthetic electrocardiogram signals. *Biomedical Engineering, IEEE Transactions on*, 50(3):289–294, 2003.

[89] L.J. Mengelkoch, D. Martin, and J. Lawler. A review of the principles of pulse oximetry and accuracy of pulse oximeter estimates during exercise. *Physical Therapy*, 74(1):40–49, 1994.

[90] DJ Meredith, D. Clifton, P. Charlton, J. Brooks, CW Pugh, and L. Tarassenko. Photoplethysmographic derivation of respiratory rate: a review of relevant physiology. *Journal of Medical Engineering & Technology*, 36(1):1–7, 2012.

[91] A. Monaghan. Detecting and managing deterioration in children. *Paediatr Nurs*, 17(1):32–35, 2005.

[92] M. Moser, M. Lehofer, A. Sedminek, M. Lux, H.G. Zapotoczky, T. Kenner, and A. Noordergraaf. Heart rate variability as a prognostic tool in cardiology. a contribution to the problem from a theoretical point of view. *Circulation*, 90(2):1078–1082, 1994.

[93] B.K. Muma, D.J. Treloar, K. Wurmlinger, E. Peterson, and A. Vitae. Comparison of rectal, axillary, and tympanic membrane temperatures in infants and young children. *Annals of emergency medicine*, 20(1):41–44, 1991.

[94] 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.

[95] MR Neuman and N. Wang. Motion artifact in pulse oximetry. In *Engineering in Medicine and Biology Society, 1990., Proceedings of the Twelfth Annual International Conference of the IEEE*, pages 2007–2008. IEEE, 1990.

[96] L. Nilsson, A. Johansson, and S. Kalman. Macrocirculation is not the sole determinant of respiratory induced variations in the reflection mode photoplethysmographic signal. *Physiological measurement*, 24:925, 2003.

[97] L. Nilsson, A. Johansson, and S. Kalman. Respiratory variations in the reflection mode photoplethysmographic signal. relationships to peripheral venous pressure. *Medical and Biological Engineering and Computing*, 41(3):249–254, 2003.

[98] L. Nilsson, A. Johansson, and S. Kalman. Respiration can be monitored by photoplethysmography with high sensitivity and specificity regardless of anaesthesia and ventilatory mode. *Acta anaesthesiologica scandinavica*, 49(8):1157–1162, 2005.

[99] M. Nitzan, I. Faib, and H. Friedman. Respiration-induced changes in tissue blood volume distal to occluded artery, measured by photoplethysmography. *Journal of biomedical optics*, 11:040506, 2006.

[100] A. Nocera and A. Garner. An australian mass casualty incident triage system for the future based upon triage mistakes of the past: the homebush triage standard. *Australian and New Zealand journal of surgery*, 69(8):603–608, 1999.

[101] A.V. Oppenheim and R.W. Schafer. *Discrete-time signal processing*, volume 3. Pearson Prentice Hall Upper sadle River, NJ, 2010.

[102] A.V. Oppenheim, R.W. Schafer, J.R. Buck, et al. *Discrete-time signal processing*, volume 1999. Prentice hall Englewood Cliffs, NJ:, 1989.

[103] S.J. Orfanidis. *Optimum signal processing: An introduction, 2nd Edition*. Prentice-Hall, Englewood Cliffs, NJ, 1996.

[104] J. Pardey, S. Roberts, and L. Tarassenko. A review of parametric modelling techniques for eeg analysis. *Medical engineering & physics*, 18(1):2–11, 1996.

[105] C.S. Parshuram, J. Hutchison, K. Middaugh, et al. Development and initial validation of the bedside paediatric early warning system score. *Crit Care*, 13(4):R135, 2009.

[106] G. Pearson and H. Duncan. Early warning systems for identifying sick children. *Paediatrics and Child Health*, 21(5):230–233, 2011.

[107] G.A. Pearson and C.E. into Maternal. *Why Children Die: A Pilot Study 2006 England;(South West, North East and West Midlands), Wales and Northern Ireland*. CEMACH, 2008.

[108] M.M. Pollack, K.M. Patel, U.E. Ruttimann, et al. The pediatric risk of mortality iii–acute physiology score (prism iii-aps): a method of assessing physiologic instability for pediatric intensive care unit patients. *The Journal of pediatrics*, 131(4):575–581, 1997.

[109] J.G. Proakis and D.G. Manolakis. *Digital signal processing: principles, algorithms, and applications*, volume 3. Prentice Hall Upper Saddle River, NJ:, 1996.

[110] U. Rajendra Acharya, K. Paul Joseph, N. Kannathal, C.M. Lim, and J.S. Suri. Heart rate variability: a review. *Medical and Biological Engineering and Computing*, 44(12):1031–1051, 2006.

[111] K.A. Reddy, B. George, and V.J. Kumar. Use of fourier series analysis for motion artifact reduction and data compression of photoplethysmographic signals. *Instrumentation and Measurement, IEEE Transactions on*, 58(5):1706–1711, 2009.

[112] K.A. Reddy and V.J. Kumar. Motion artifact reduction in photoplethysmographic signals using singular value decomposition. In *Instrumentation and Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE*, pages 1–4. IEEE, 2007.

[113] AR Relente and LG Sison. Characterization and adaptive filtering of motion artifacts in pulse oximetry using accelerometers. In *Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference, 2002. Proceedings of the Second Joint*, volume 2, pages 1769–1770. IEEE, 2002.

[114] O.D. Richard, E.H. Peter, and G.S. David. Pattern classification. *A Wiley-Interscience Public*, pages 373–378, 2001.

[115] M. Richardson and M. Lakhanpaul. Nice guidelines: Assessment and initial management of feverish illness in children younger than 5 years: summary of nice guidance. *BMJ: British Medical Journal*, 334(7604):1163, 2007.

[116] J. Roukema, EW Steyerberg, A. Van Meurs, M. Ruige, J. Van Der Lei, and HA Moll. Validity of the manchester triage system in paediatric emergency care. *Emergency medicine journal*, 23(12):906–910, 2006.

[117] TL Rusch, R. Sankar, and JE Scharf. Signal processing methods for pulse oximetry. *Computers in biology and medicine*, 26(2):143–159, 1996.

[118] F. Rusconi, M. Castagneto, N. Porta, L. Gagliardi, G. Leo, A. Pellegatta, S. Razon, and M. Braga. Reference values for respiratory rate in the first 3 years of life. *Pediatrics*, 94(3):350–355, 1994.

[119] J.W. Sammon Jr. A nonlinear mapping for data structure analysis. *Computers, IEEE Transactions on*, 100(5):401–409, 1969.

[120] SS Sastry. *Introductory methods of numerical analysis*. PHI Learning Pvt. Ltd., 2005.

[121] S. Saxena, A. Majeed, and M. Jones. Socioeconomic differences in childhood consultation rates in general practice in england and wales: prospective cohort study. *BMJ*, 318(7184):642–646, 1999.

[122] A. Schafer and K.W. Kratky. Estimation of breathing rate from respiratory sinus arrhythmia: Comparison of various methods. *Annals of Biomedical Engineering*, 36(3):476–485, 2008.

[123] M. Scoble. Implementing triage in a children's assessment unit. *Nursing standard*, 18(34):41–44, 2004.

[124] N. Selvaraj, A. Jaryal, J. Santhosh, KK Deepak, and S. Anand. Assessment of heart rate variability derived from finger-tip photoplethysmography as compared to electrocardiography. *Journal of medical engineering & technology*, 32(6):479–484, 2008.

[125] K.H. Shelley, A.A. Awad, R.G. Stout, and D.G. Silverman. The use of joint time frequency analysis to quantify the effect of ventilation on the pulse oximeter waveform. *Journal of clinical monitoring and computing*, 20(2):81–87, 2006.

[126] R. Shiavi. *Introduction to applied statistical signal analysis: guide to biomedical and electrical engineering applications*. 2007.

[127] B.W. Silverman. *Density estimation for statistics and data analysis*, volume 26. Chapman & Hall/CRC, 1986.

[128] EA Simoes, R. Roark, S. Berman, LL Esler, and J. Murphy. Respiratory rate: measurement of variability over time and accuracy at different counting periods. *Archives of disease in childhood*, 66(10):1199, 1991.

[129] J.E. Sinex. Pulse oximetry: principles and limitations. *The American journal of emergency medicine*, 17(1):59–66, 1999.

[130] P.K. Stein, R.E. Kleiger, and J.N. Rottman. Differing effects of age on heart rate variability in men and women. *The American journal of cardiology*, 80(3):302–305, 1997.

[131] M. Stewart, U. Werneke, R. MacFaul, J. Taylor-Meek, HE Smith, and IJ Smith. Medical and social factors associated with the admission and discharge of acutely ill children. *Archives of Disease in Childhood*, 79(3):219–224, 1998.

[132] CP Subbe, M. Kruger, P. Rutherford, and L. Gemmel. Validation of a modified early warning score in medical admissions. *Qjm*, 94(10):521–526, 2001.

[133] J.A. Sukor, SJ Redmond, and NH Lovell. Signal quality measures for pulse oximetry through waveform morphology analysis. *Physiological Measurement*, 32:369, 2011.

[134] L. Tarassenko, A. Hann, A. Patterson, E. Braithwaite, K. Davidson, V. Barber, and D. Young. Biosign: multi-parameter monitoring for early warning of patient deterioration. In *Medical Applications of Signal Processing, 2005. The 3rd IEE International Seminar on (Ref. No. 2005-1119)*, pages 71–76. IET, 2005.

[135] L. Tarassenko, A. Hann, and D. Young. Integrated monitoring and analysis for early warning of patient deterioration. *British journal of anaesthesia*, 97(1):64–68, 2006.

[136] L. Tarassenko, P. Hayton, N. Cerneaz, and M. Brady. Novelty detection for the identification of masses in mammograms. In *Artificial Neural Networks, 1995., Fourth International Conference on*, pages 442–447. IET, 1995.

[137] L. Tarassenko, A. Nairac, N. Townsend, I. Buxton, and P. Cowley. Novelty detection for the identification of abnormalities. *International Journal of Systems Science*, 31(11):1427–1439, 2000.

[138] J.F. Thayer, J.J. Sollers III, E. Ruiz-Padial, and J. Vila. Estimating respiratory frequency from autoregressive spectral analysis of heart period. *Engineering in Medicine and Biology Magazine, IEEE*, 21(4):41–45, 2002.

[139] S. Theodoridis and K. Koutroumbas. *pattern recognition*, volume 4. Academic Pr, 2009.

[140] M. Thompson, N. Coad, A. Harnden, R. Mayon-White, R. Perera, and D. Mant. How well do vital signs identify children with serious infections in paediatric emergency care? *Archives of disease in childhood*, 94(11):888–893, 2009.

[141] M. Thompson, R. Mayon-White, A. Harnden, R. Perera, D. McLeod, and D. Mant. Using vital signs to assess children with acute infections: a survey of current practice. *The British Journal of General Practice*, 58(549):236, 2008.

[142] M.J. Thompson, N. Ninis, R. Perera, R. Mayon-White, C. Phillips, L. Bailey, A. Harnden, D. Mant, and M. Levin. Clinical recognition of meningococcal disease in children and adolescents. *The lancet*, 367(9508):397–403, 2006.

[143] J. Tibballs and S. Kinney. Reduction of hospital mortality and of preventable cardiac arrest and death on introduction of a pediatric medical emergency team*. *Pediatric Critical Care Medicine*, 10(3):306, 2009.

[144] J. Tibballs, S. Kinney, T. Duke, E. Oakley, and M. Hennessy. Reduction of paediatric in-patient cardiac arrest and death with a medical emergency team: preliminary results. *Archives of disease in childhood*, 90(11):1148–1152, 2005.

[145] K.K. Tremper and S.J. Barker. Pulse oximetry. *Anesthesiology*, 70(1):98, 1989.

[146] M.P. Tulppo, T.H. Mäkikallio, T. Seppänen, R.T. Laukkanen, and H.V. Huikuri. Vagal modulation of heart rate during exercise: effects of age and physical fitness. *American Journal of Physiology-Heart and Circulatory Physiology*, 274(2):H424, 1998.

[147] A. Van den Bruel, B. Aertgeerts, R. Bruyninckx, M. Aerts, and F. Buntinx. Signs and symptoms for diagnosis of serious infections in children: a prospective study in primary care. *The British Journal of General Practice*, 57(540):538, 2007.

[148] A. Van den Bruel, R. Bruyninckx, E. Vermeire, P. Aerssens, B. Aertgeerts, and F. Buntinx. Signs and symptoms in children with a serious infection: a qualitative study. *BMC family practice*, 6(1):36, 2005.

[149] A. Van den Bruel, T. Haj-Hassan, M. Thompson, F. Buntinx, D. Mant, et al. Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review. *The Lancet*, 375(9717):834–845, 2010.

[150] C. van Ravenswaaij-Arts, L.A.A. Kollee, J.C.W. Hopman, G. Stoelinga, and H.P. van Geijn. Heart rate variability. *Annals of Internal Medicine*, 118(6):436, 1993.

[151] M. Van Veen, E.W. Steyerberg, M. Ruige, A.H.J. van Meurs, J. Roukema, J. van der Lei, and H.A. Moll. Manchester triage system in paediatric emergency care: prospective observational study. *BMJ: British Medical Journal*, 337, 2008.

[152] H. Wang, K. Siu, K. Ju, and K.H. Chon. A high resolution approach to estimating time-frequency spectra and their amplitudes. *Annals of biomedical engineering*, 34(2):326–338, 2006.

[153] J. Weng, Z. Ye, and J. Weng. An improved pre-processing approach for photoplethysmographic signal. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 41–44. IEEE, 2006.

[154] M.W. Wukitsch, M.T. Petterson, D.R. Tobler, and J.A. Pologe. Pulse oximetry: analysis of theory, technology, and practice. *Journal of Clinical Monitoring and Computing*, 4(4):290–301, 1987.

[155] F. Yasuma and J. Hayano. Respiratory sinus arrhythmia*. *Chest*, 125(2):683–690, 2004.

[156] T.S. Yeh, M.M. Pollack, U.E. Ruttimann, P.R. Holbrook, and A.I. Fields. Validation of a physiologic stability index for use in critically ill infants and children. *Pediatric research*, 18(5):445–451, 1984.

[157] Y. Zhou, Y. Zheng, C. Wang, and J. Yuan. Extraction of respiratory activity from photoplethysmographic signals based on an independent component analysis technique: Preliminary report. *Instrumentation Science and Technology*, 34(5):537–545, 2006.

[158] W. Zong, T. Heldt, GB Moody, and RG Mark. An open-source algorithm to detect onset of arterial blood pressure pulses. In *Computers in Cardiology, 2003*, pages 259–262. IEEE, 2003.

Appendix A

Relationship between pole location and frequency response

The development in this appendix follows closely to the discussion given in [101, 109].

A general linear constant-coefficient difference equation is represented by equation A.1,

$$\sum_{k=0}^{P} a_k y(n-k) = \sum_{k=0}^{M} b_k x(n-k) \tag{A.1}$$

which can be represented as: $y(n) = \frac{1}{a_0}(\sum_{k=0}^{M} b_k x(n-k) - \sum_{k=1}^{P} a_k y(n-k))$

where $y(n)$ is the current output of the time series ($t = n$) , $y(n-k)$ is the time series output at $t = n-k$, a_k are the P coefficients which quantifies the linear relationship between the current output (i.e. at $t = n$) and the outputs at $t = (n-k)$, $x(n)$ is the current input of the time series ($t = n$) , $x(n-k)$ is the time series input at $t = n-k$ and b_k are the $M+1$ coefficients which quantifies the linear relationship between the current output (i.e. at $t = n$) and the inputs at $t = (n-k)$. Equation A.1 reduces to the time-domain AR model equation given by equation 4.1 when $M = 0$ and $b_0 x(n) = e(n)$.

Applying the z-transform to equation A.1 to get the system transfer function, $H(z)$,

$$\sum_{k=0}^{P} a_k Y(z) z^{-k} = \sum_{k=0}^{M} b_k X(z) z^{-k}$$

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^{M} b_k z^{-k}}{\sum_{k=0}^{P} a_k z^{-k}}$$

Factorising the polynomials in the numerator and the denominator gives

$$H(z) = \frac{Y(z)}{X(z)} = \frac{b_0}{a_0} \frac{\prod_{k=1}^{M}(1 - c_k z^{-1})}{\prod_{k=1}^{P}(1 - d_k z^{-1})} \tag{A.2}$$

In equation A.2, each of the factor in the numerator $(1-c_k)$ contributes a zero at $z=c_k$ and a pole at $z=0$ and each factor in the denominator contributes a pole at $z=d_k$ and a zero at $z=0$. Evaluating $H(z)$ at the unit circle (i.e. when $z=e^{j\omega}$ implying $|z|=1$) will give us the frequency response of the system.

$$H(e^{j\omega}) = \frac{b_0}{a_0}\frac{\prod_{k=1}^{M}(1-c_k e^{-j\omega})}{\prod_{k=1}^{P}(1-d_k e^{-j\omega})} \tag{A.3}$$

and since the magnitude , $|H(e^{jw})|$, of the above equation is equal to the product of all the terms in the equation,

$$|H(e^{j\omega})| = \frac{|b_0|}{|a_0|}\frac{|1-c_1e^{-j\omega}|...|1-c_ke^{-j\omega}|...|1-c_Me^{-j\omega}|}{|1-d_1e^{-j\omega}|...|1-d_ke^{-j\omega}|...|1-d_Pe^{-j\omega}|}$$

$$|H(e^{j\omega})| = \frac{|b_0|}{|a_0|}\frac{|e^{jw}-c_1|...|e^{jw}-c_k|...|e^{jw}-c_M|}{|e^{jw}-d_1|...|e^{jw}-d_k|...|e^{jw}-d_P|} \tag{A.4}$$

To aid in the geometric interpretation of equation A.4, lets assume a simple situation with a single zero at $z=c_k$ and a single pole at $z=d_k$ as shown in figure A.1. Lets also assume that the frequency response $H(e^{j\omega})$ at a specific frequency, w, represented by a vector of unit length $e^{j\omega}$ making an angle w with the real axis as shown by the vector OC in the figure needs to be evaluated. Finding the frequency response of $H(e^{j\omega})$ at frequency w is equivalent to evaluating the z-transform at the point C of the complex plane. From the figure, it can be seen that:

$$OC = OA + AC$$

$$OC = OB + BC$$

Furthermore, $OC = e^{jw}$, $OA = d_k$ and $OB = c_k$. Therefore,

$$AC = e^{jw} - d_k \tag{A.5}$$

$$BC = e^{jw} - c_k \tag{A.6}$$

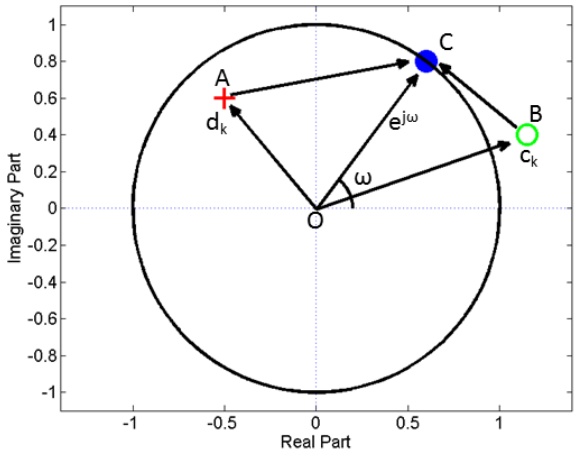


Figure A.1: Geometric interpretation of the contribution of a pole and a zero to the frequency magnitude spectrum, a pole shown in red at A, a zero shown in green at B and the magnitude evaluated on the unit circle at point C (modified from [109])

Relating equations A.5 and A.6 to equation A.4, it can be seen that each factor in the numerator, $|e^{jw} - c_k|$, in equation A.4 corresponds to the distance of the k^{th} zero to the point C (where the frequency magnitude is to be evaluated), and each factor, $|e^{jw} - d_k|$, corresponds to the distance of the k^{th} pole to the point C. Thus it can be concluded that the magnitude response of a system, $H(z)$, is the product of the lengths of zero vectors,$|Z_k|$ (corresponding to the distance from each zero to the point on the unit circle where $H(z)$ is evaluated) divided by the product of the lengths of the pole vectors,$|V_k|$ (corresponding to the distance from each pole to the point on the unit circle where $H(z)$ is evaluated) i.e.

$$|H(e^{j\omega}) = \frac{\prod_{k=1}^{M} |Z_k|}{\prod_{k=1}^{P} |V_k|}$$

Using the concepts developed, how the presence of poles affect the frequency spectrum using a simple example of a system function given by equation A.7 having two poles (which will be conjugate, a condition imposed by having real signals) can now be found.

$$H(z) = \frac{1}{(1 - re^{j\theta}z)(1 - re^{-j\theta}z))} \tag{A.7}$$

Such a case is illustrated in figure A.2 with the two poles occurring at P and Q shown by red

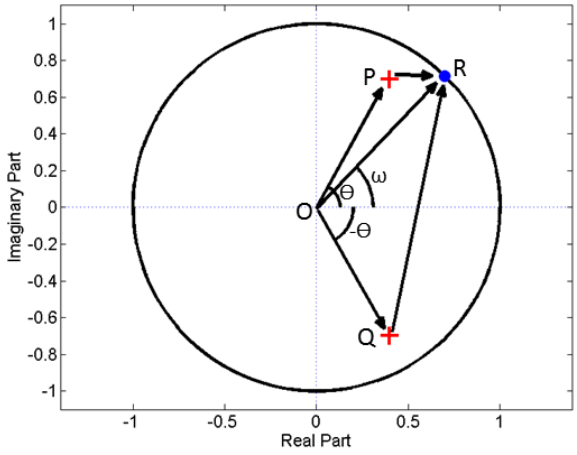


Figure A.2: Pole-zero plot for a system with two complex conjugate poles (modified from [101])

marks. Relating equation (A.7) with the figure, it can be seen that $|OP| = |OQ| = |r|$. Furthermore, since $H(z)$ is an all-pole model, therefore no zeros occur away from the origin and the two zeros due to the two terms in the denominator occur at the origin, O. The frequency magnitude response at ω, evaluated at a point on the unit circle, at R is:

$$|H(e^{jw})| = \frac{|OR|^2}{|PR||QR|} = \frac{1}{|PR||QR|}$$

since $|OR| = 1$. When $\omega \approx \theta$, $|PR|$ is small and changes significantly as ω varies around θ while |QR| changes only slightly. The magnitude response in this region is ,thus, dominated by the pole located at P. By symmetry, the frequency magnitude response is dominated by the pole located at Q around $\omega \approx -\theta$. In general, a pole located at an angle θ dominates the frequency magnitude response around $\omega \approx \theta$.

It can also be seen from the figure that the closer a pole is to the unit circle, the smaller will be the magnitude of the pole vector and the higher will be the magnitude of the frequency response, $|H(e^{j\omega})|$. Lastly, using geometry, the shortest distance from a pole located at an angle θ inside the unit circle to the unit circle occurs at $\omega = \theta$. Thus, assuming there are no closely located poles, the frequency magnitude response will be highest at $\omega = \theta_h$ where θ_h corresponds to the angle of the pole with the highest magnitude (equivalent to the pole that is closest to the unit circle).