HEART SOUNDS BASED MONITORING

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Abhilash Patangay

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: Ahmed Tewfik, PhD

February 2009

Acknowledgements

I would like to thank Dr. Ahmed Tewfik for his support and guidance during this amazing experience. He has guided me in all walks of the process from creating hypothesis to writing research papers. His encouragement and high standards have been a force behind the contributions of this work. I would also like to thank the rest of my committee: Dr. Vladimir Cherkassky, Dr. Emad Ebbini, Dr. Erdman and Dr. Bruce KenKnight for their help, advice and guidance during the PhD process.

I am grateful for the funding and support from Boston Scientific Corporation (BSC) and their Research group at Arden Hills, MN. This thesis would not have been possible without the inspiration and encouragement of the scientists and management at BSC Research. In particular I would like to thank Dr. Julie Thompson for her constant mentoring and advice to help me through the administrative challenges of this collaborative research. I am grateful to Kris Siejko for pioneering the heart sounds research and to Dr. Yi Zhang for championing the ischemia animal study. I am thankful to Dr. Jeffery Von Arx for giving me the confidence to pursue heart sounds research for my doctoral work and Kent Lee for providing the sleep apnea data.

Finally and most importantly I would like to thank my wife for being the inspiration of my life and my parents for their unconditional love and support.

Dedicated to my wife Prashanthi Vemuri

Abstract

Heart failure is the leading cause of mortality worldwide and its complications result in over one million hospitalizations annually in the United States alone. Increasing number of hospitalizations along with escalating health care costs are resulting in a mounting burden on the healthcare system and warrant novel practical approaches to patient management. Remote monitoring of ambulatory patients using wearable and invasive sensors is gaining acceptance as one of the solutions to this problem. Ambulatory sensors that can make physiologic measurements continuously and provide a comprehensive picture of the patient's status over time form the basis of any successful remote monitoring system.

This thesis focuses on one such physiologic sensor, namely heart sounds, for remote ambulatory monitoring. The aim of this research is to develop novel algorithms for accurate measurement and tracking of clinically useful heart sound parameters and the application of these measurements to detect cardiovascular perturbations. This thesis achieves these goals in three main parts.

First, a novel low complexity framework is developed to accurately measure the different heart sound components. For a given heart sound (e.g. S1) a dynamic programming based algorithm is applied to select and track the largest most consistent peak. To establish the value of the proposed framework, it is tested on acute and chronic pre-clinical data collected during heart failure deterioration and compared to a traditional non- tracking algorithm. In all these pre-clinical experiments, the performance of the proposed tracking framework is found to be superior to the traditional non- tracking approach. These results validate that heart sounds measured using the proposed framework contain clinically relevant information about heart failure status that has historically not been available due to the use of a non- tracking approach.

Second, the clinical utility of heart sounds based parameters measured at a non-traditional pectoral location is evaluated in an acute hospitalized setting. In particular the heart sounds ejection time (HSET) which is an indicator of changing LV systolic performance is studied. Heart sound based ejection time measured using our tracking

framework is compared to the stroke volume (SV) measurements recorded during hospitalization. In 20 patients with changes in SV > 10 ml the mean correlation coefficient between HSET and SV is found to be $R = 0.6762$. Also, the HSET is shown to have 70% sensitivity at 80% specificity to detect periods of low stroke volume (SV <50 ml).

Finally, the utility of heart sounds based measurements for the detection of cardiovascular perturbations is evaluated. In particular a system capable of detecting episodes of obstructive sleep apnea (OSA) in heart failure patients with pacemakers and cardiac resynchronization therapy devices is developed. Features for OSA detection are generated by optimally extracting information in the S1 measurements using wavelet decomposition and adaptive dyadic time segmentation. Linear discriminant analysis and support vector machine (SVM) classifiers are trained using different feature selection schemes and tested on an independent test dataset. The tracking based classification is found to consistently outperform non-tracking based classification, emphasizing the importance of tracking. SVM with recursive feature elimination scheme and tracking is shown to have the highest (91.8%) accuracy yielding an improvement of 7% over the non-tracking based approach. The output of the best classifier is used as an OSA severity score which is shown to be correlated significantly ($R = 0.72$, $p<0.05$) with the gold standard apnea hypopnea index.

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 Motivation

Heart failure is the leading cause of mortality worldwide and its complications result in over one million hospitalizations annually in the United States alone [1, 2]. Despite improvements in tests like brain natriuretic peptide (BNP) and echocardiographic measures [3-5] for acute bedside diagnosis of worsening heart failure, issues of variability in measurements and low accuracy warrant the development of novel tools for better patient management. Moreover, the existing tools are not practical for chronic patient monitoring which has gained increased attention due to the escalating health care costs [6, 7]. The traditional health care model has been to schedule periodic clinic visits for preventative check-ups. While this model is successful in the early stage of the disease, it falls short when the disease has progressed to a stage where more frequent follow- ups become necessary. Also, the clinic visits are only snapshots of the patient's health and are often incomplete [8]. The solution to this problem is remote monitoring of ambulatory patients using wearable and invasive sensors. These ambulatory monitoring systems [9] make physiologic measurements continuously and provide a comprehensive picture of the patient's status.

This thesis focuses on the novel application of one such physiologic sensor, namely heart sounds, for chronic ambulatory monitoring. The four heart sounds (S1, S2, S3 and S4) are generated by the beating heart and the resultant blood flow. Traditionally, auscultation exams have been used over the past centuries to look for pathophysiologic sounds using a stethoscope. These exams are snapshots of the

patient's condition and hence suffer from the same pitfalls of traditional periodic physical exams.

To use heart sounds for long term ambulatory monitoring of cardiovascular conditions, it is crucial to detect and track the same heart sound component consistently over time. This tracking is even more important when the heart sound component that is being measured consists of multiple sub- components (e.g. the first heart sound) [10]. This research aims to develop a framework for long term monitoring of heart sounds. The focus is on developing novel algorithms which can facilitate the measurement of these proven clinical parameters and the detection of perturbations in the cardiovascular system based on these measurements. One such important perturbation in the context of heart failure is obstructive sleep apnea (OSA).

OSA is the most prevalent form of sleep disordered breathing where the airways repeatedly close and breathing stops. Also, it has been shown that OSA is a potential risk factor for heart failure patients [11, 12] with rhythm control devices like pacemakers and cardiac resynchronization therapy devices. In patients with such devices, the prevalence of apnea is reported to be much higher than in the general population, with about 50-60% in heart failure patients and over 30% in pacemaker patients [13, 14]. The repetitive interruption of airflow, despite persistent respiratory efforts, is associated with the activation of a number of neural, metabolic and inflammatory disease mechanisms, all of which have been implicated in the pathophysiology of cardiac and vascular disease. Due to abrupt and marked falls in inspiratory intrathoracic pressure the cardiac loading conditions change [15] and OSA can impact the heart through two separate mechanisms. First, during the course of

chronic OSA, multiple episodes of obstructive inspirations during sleep may have deleterious effects on the ventricular function. Second, chronic OSA may lead to sustained hypertension, which has its own effects on the left ventricular loading conditions. Due to these reasons it becomes vital to detect OSA in patients with heart failure. If undiagnosed and untreated, OSA can worsen heart failure eventually leading to hospitalization [12].

The gold standard test for detecting OSA is a polysomnography test [16]. During a polysomnography the patient undergoes a sleep test where numerous physiologic signals are recorded. Due to patient discomfort and the expensive nature of the test, it is impractical for long term monitoring of OSA. In the past several researchers have attempted to solve the problem of detecting OSA non- intrusively using electrocardiogram (ECG) based techniques [17]. While ECG based approaches have been popular and have about 85-92.6% accuracy [17], these methods fail in heart failure patients with implanted cardiac therapy devices which adjust the heart rate automatically. To address these concerns the tracking framework for measuring heart sounds is used to develop an improved hemodynamics based approach for the detection of OSA in patients with rhythm control devices.

1.2 Contributions of this Thesis

The major contributions for this thesis can be divided into three main areas. First, a novel framework for consistently measuring heart sounds is developed in Chapter 3. Heart sounds based parameters are next shown to contain clinically valuable information in Chapter 4. Finally, a novel and practical method for OSA detection in

heart failure patients with rhythm control devices is developed in Chapter 5. These contributions are described in detail below.

1.2.1 Tracking Algorithm

First, a novel low complexity framework for consistently measuring heart sound peaks (S1- S4) is developed. This tracking framework is capable of measuring single and multiple peaks over long periods of time. The value of the proposed framework is established using acute and chronic pre-clinical data collected during heart failure deterioration where the tracking approach is shown to outperform a traditional non-tracking algorithm. Additionally, these results validate that heart sounds measured using the proposed framework contain clinically relevant information regarding heart failure status which has historically not been available due to the use of a non- tracking approach. To our knowledge this is the first research that describes and evaluates a heart sounds monitoring algorithm that can be used to track clinically useful changes in cardiac condition in a chronic setting.

1.2.2 Heart Sounds Ejection Time based Monitoring

Traditionally systolic timing intervals (STIs) have been measured using heart sounds captured at an apical location. Also they have only been used to look for pathophysiologic conditions acutely during clinic visits. Chronic measurements of STIs and their correlation to left ventricular (LV) performance have not yet been evaluated. In this thesis the utility of heart sounds based ejection time (HSET), measured using the tracking framework at a non- traditional pectoral location, as an indicator of changing LV systolic performance is evaluated in an acute hospitalized setting. We show that the

ejection time measured from heart sounds is statistically significantly correlated with stroke volume (SV). Also, evaluation of systolic timing intervals to detect low SV is shown to be promising since timing intervals are robust in the presence of posture and variable transmission effects.

1.2.3 Apnea Detection

The final part of the thesis presents a novel non-snapshot based heart sounds approach for OSA detection using the tracking framework and an adaptive time frequency feature extraction and classification scheme. This approach is shown to provide improved accuracy and reliability over a non- tracking technique. Classification and feature selection schemes are compared to evaluate filter and wrapper based techniques for selecting optimal features. Finally, an algorithm based severity score is developed to determine the intensity of OSA in each patient during the sleep study and shown to significantly correlate with the gold standard apnea hypopnea index (AHI).

1.3 Thesis Outline

The organization of the thesis is as follows. Chapter 2 presents the background that forms the basis of the thesis. Chapter 3 presents the proposed tracking framework for measuring heart sounds over a period of time. Chapter 4 examines the utility of systolic timing intervals measured using the tracking algorithm in hospitalized heart failure patients. Chapter 5 addresses the problem of detecting obstructive sleep apnea using heart sounds. Finally, Chapter 6 describes the technical and clinical value of this research and discusses potential future work.

2 Background

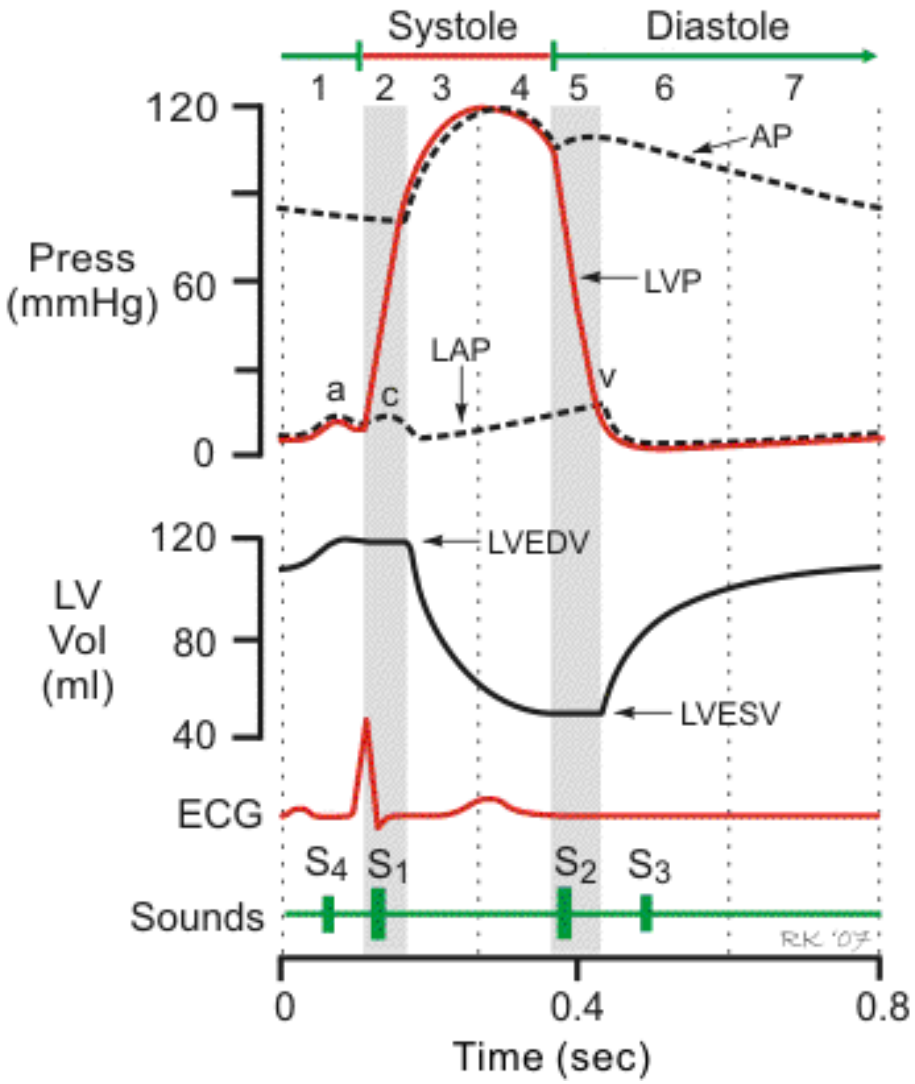
This chapter is a collection of sections that present the information needed for the rest of the thesis. Section 2.12.2 describe the physiology and genesis of heart sounds. 2.3 presents a visualization technique used in this thesis. A survey of the clinical value of heart sounds is presented in Section 2.4. Section 2.5 summarizes the time frequency applications using heart sounds as a signal. This summary along with the physiology and genesis should elucidate the complex nature of the heart sounds waveform. In Section 2.6 a brief survey of target tracking techniques in the context of this thesis is provided. Finally, Section 2.7 sets up the feature generation, selection and classification techniques used in this thesis.

2.1 Heart Sounds Physiology

During rest, the heart beats about 70 times a minute and pumps about 5 liters of blood each minute to the body. While electro-cardiograms (ECG) provide information regarding the electrical activity of the heart, heart sounds provide mechanical information about the cardiovascular system. A cardiac cycle consists of precisely timed electrical and mechanical events that result in rhythmic atrial and ventricular contractions that pump blood into the pulmonary and systemic circulation. As shown in Figure 1, the cardiac cycle is divided into two phases based on the events in the ventricle [18, 19]. The period of ventricular contraction and ejection is called systole, and the period of ventricular relaxation and filling is called diastole. Both the systole and diastole start with periods of isovolumic pressure changes.

The systole begins with isovolumic contraction (phase 2) where the ventricles contract against closed valves and hence no blood is ejected and volume is held constant. This increased ventricular wall tension causes the pressure to increase which eventually opens the aortic and pulmonic valve and the ventricular ejection period occurs (phase 3). Blood ejects through the aorta and the pulmonary artery when the ventricles contract. It is the forced closure against the mitral valve and the sudden rush of blood through the aorta that causes the first heart sound (S1). The total volume of blood ejected by each ventricle is called the stroke volume. The end of systole happens when the aortic valve closes due to reduced flow and pressure (phase 4). This aortic and pulmonary closure, and the movement of blood column after their closure, causes the second heart sound (S2). The S2 consists of two components; the aortic component (A2) and the pulmonic component (P2). Both the S1 and S2 are audible sounds with a bandwidth of 15 Hz to 90 Hz.

During the first part of diastole (phase 5), the ventricles relax and the aortic and pulmonary valves close. The ventricular pressure falls and when it drops below the corresponding atrial pressures the tricuspid and mitral valves open. This is followed by diastolic ventricular filling. The early diastolic filling (phase 6) is passive and in a stiff heart, as in heart failure, results in a third heart sound (S3). During the late diastole (phase 1), when the atrium contracts against a stiff ventricle, a fourth heart sound is heard (S4). When present in adults, the S3 and S4 are indications of pathologic conditions of the heart. Both these sounds are low frequency sounds with a bandwidth of 5 Hz to 25 Hz.



Phases of the cardiac cycle
Phase 1 - Atrial Contraction
Phase 2 - Isovolumetric Contraction
Phase 3 - Rapid Ejection
Phase 4 - Reduced Ejection
Phase 5 - Isovolumetric Relaxation
Phase 6 - Rapid Filling
Phase 7 - Reduced Filling

Acronyms
AP: Aortic Pressure
LVP: Left Ventricular Pressure
LAP: Left Atrial Pressure
LVEDV: LV End Diastolic Volume
LVESV: LV End Systolic Volume

Figure 1: The cardiac cycle showing left sided pressures and events. Reproduced with permission of the author [19].

2.2 Heart Sounds Genesis

Heart sounds have been studied and used for many pathologic conditions of the heart including worsening systolic performance [20-22], murmurs [23-25] and increasing pressures [26, 27]. Although heart sounds have been widely studied, their genesis has been a controversial topic. Figure 2 shows the temporal relationship of various heart sound components with intracardiac pressures. The earliest theory attributed these sounds to the opening and closure of the four cardiac valves [28]. Another theory [29, 30] introduced the notion that vibrations of the cardiac muscle, due to acceleration and deceleration of blood, are responsible for the sounds. It is now commonly believed that the S1 is a combination of vibrations of the LV muscle, blood decelerating against a closed mitral valve (MC) and the accelerating blood after the opening of the aortic valve (AO) [28, 31, 32]. The mitral closure is seen to happen after the end diastolic time and at the start of the isovolumic contraction. The aortic component follows the aortic opening and is followed by a post ejection sound which is lower in frequency. Also, the S2 is now commonly believed to be generated due to the rapid deceleration of blood, after the aortic and pulmonary valves are closed, setting a vibration. The aortic (A2) and pulmonary (P2) components occur coincident with the incisura of the respective pulse pressures indicating a rebound of the blood column to be responsible for the sound generation. These are shown in Figure 2 where both the aortic and pulmonary components are shown to be during the early isovolumic relaxation. In this example a loud P2 is noticed because of pulmonary hypertension.

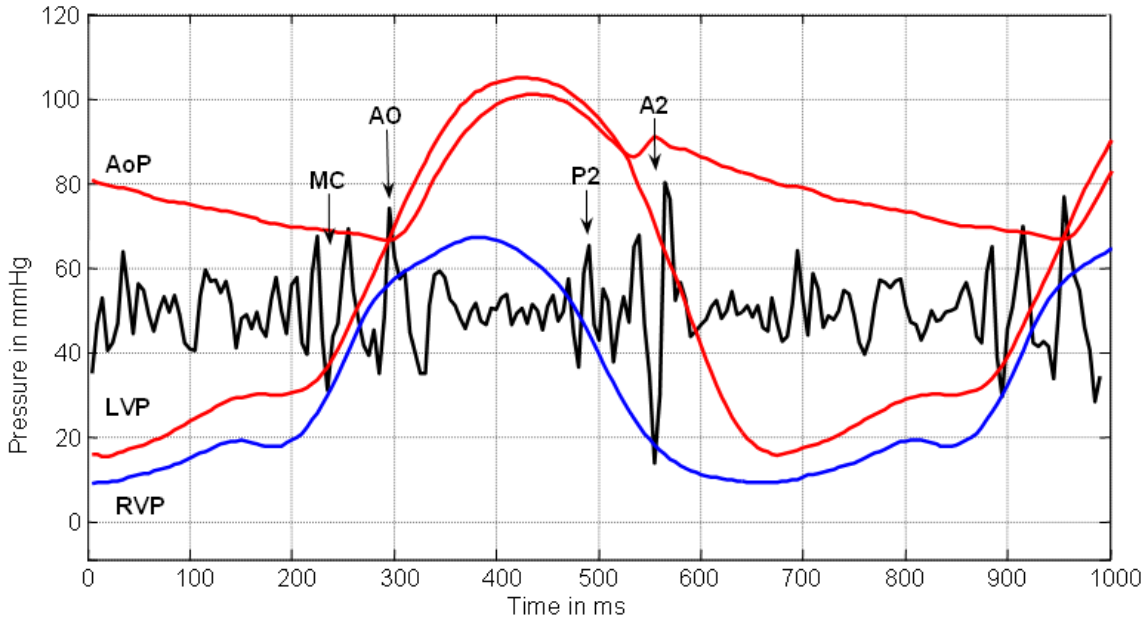


Figure 2: Illustration of the temporal location of various heart sound components relative to intracardiac pressures. The S1 is shown to be composed of the mitral and aortic components. The S2 is composed of the aortic and pulmonic components. AoP: Aortic Pressure, LVP: Left Ventricular Pressure, RVP: Right Ventricular Pressure, MC: Mitral Closure, AO: Aortic Opening, A2 and P2 are the aortic and pulmonary components of S2.

2.3 Phonoimage Based Heart Sound Visualization

Heart sounds have an inherent periodicity to them due to the cardiac activation being periodic in nature. Traditionally heart sounds have been analyzed one cardiac cycle at a time. Conclusions have been made from one beat or an average of several beats. While this is sufficient for certain applications, such analysis lacks the ability to detect or measure changes over time. In [33] for example an algorithm is described which

decomposes the S2 into aortic and pulmonary components using chirplets. While this algorithm can be used to measure the split interval at one instance it is not setup for longitudinal analysis. While split S2's have been shown to be useful for pulmonary hypertension detection [34] their utility is far enhanced if changed in the splitting interval can be measured or visualized over time. For example in Figure 3 snapshots of heart sounds, over a 2.5 hour period, with decreasing splitting intervals are shown. While the multiple plots are useful to see the changes in S2 they cannot be used to visualize longer term changes with increased resolution. A clever representation of this information is shown in Figure 4 and was proposed in [35]. This representation is called a phonoimage. A phonoimage is a plot of a matrix where each row is a heart sound beat and all rows are aligned by the R wave marker (from the ECG). Each column then represents units of time. The colors represent amplitudes, with blue being the low amplitudes and red being the high amplitudes. A colorbar is provided for amplitude reference. Figure 4 can be used to visualize the changes in heart sounds over the 2.5 hour period. These changes include, in this case, a reduction in the S2 splitting interval. Such phonoimages can be used to visualize heart sounds and changes in them over any time duration and any resolution of snapshots in time.

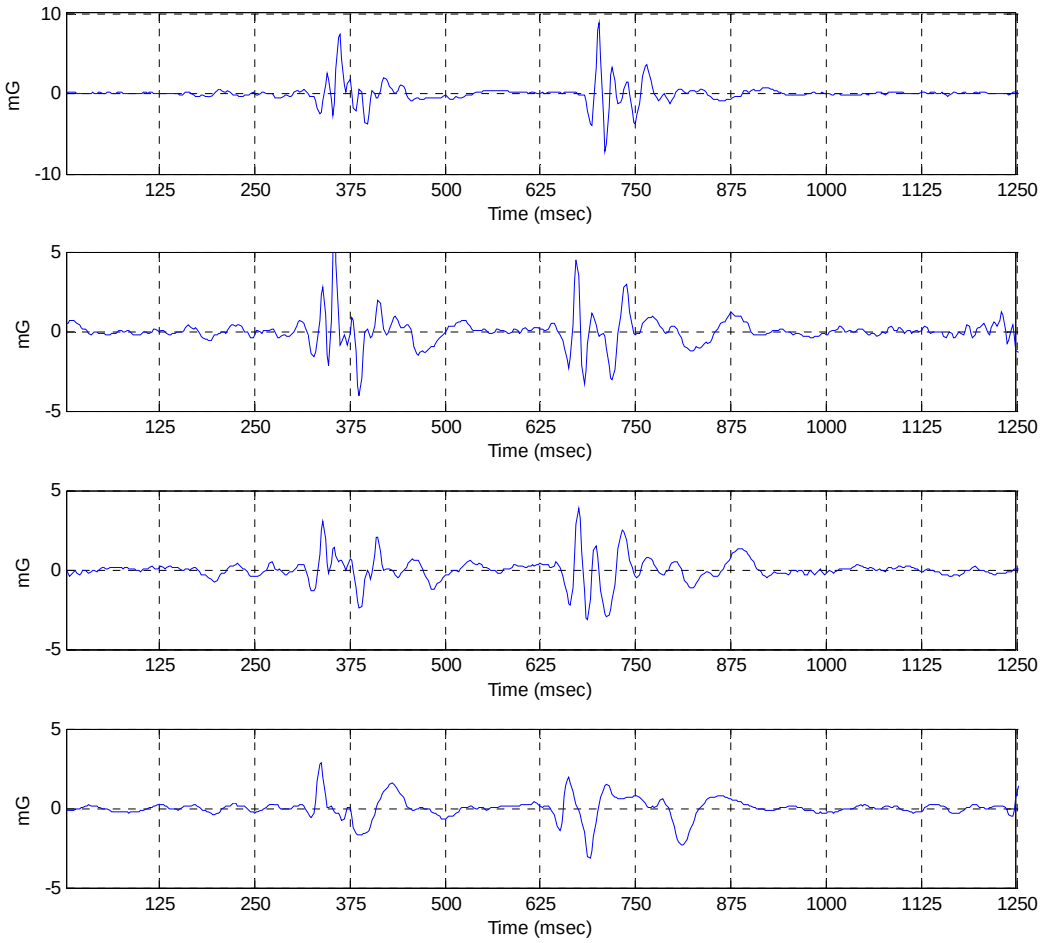


Figure 3: The Split S2 interval decreases over time (2.5 hours) with treatment. Each panel is a snapshot in time during the 2.5 hours of recording (chronologically). The split in S2 is reduced due to treatment during hospitalization.

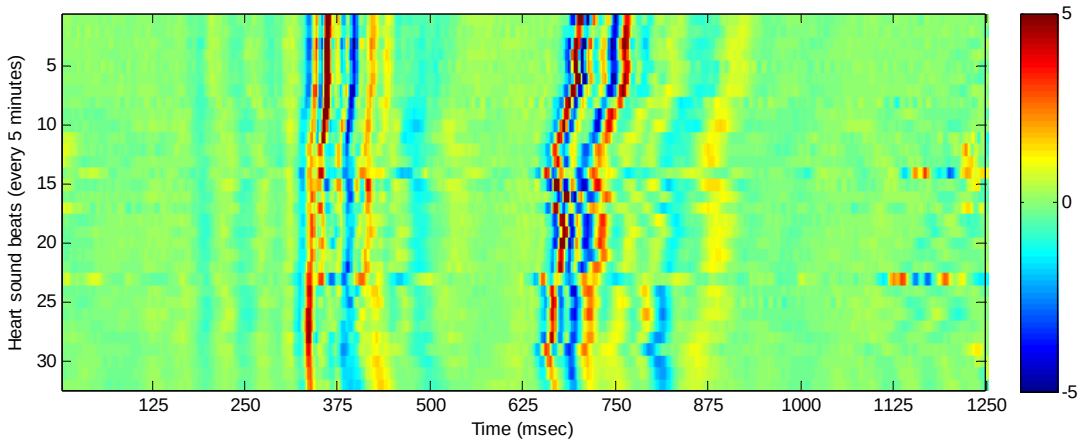


Figure 4: A phonoimage representation of the decreasing splitting interval (over 2.5 hours). Each row of the matrix is a snapshot of heart sounds (taken every 5 minutes). All beats have been aligned by the R wave of the ECG at 250 ms.

2.4 Clinical Utility Review

The S1 heart sound intensity historically has been associated with force of blood ejecting the heart [36]. It has been shown to be related to cardiac contractility [37-39] and changes in S1 frequency have been reported with different contractility states [32]. STIs, like Left Ventricular Ejection Time (LVET) and Pre Ejection Period (PEP), have been used in the past to diagnose patients with compromised left ventricular function [21]. STIs have also been shown to be correlated with stroke volume and cardiac output [20]. Also, estimation of pulmonary artery pressures from the split S2 heart sound has been shown in [34]. In the recent years an automatic S3 and S4 heart sound intensity measure has been developed and was found to be useful in diagnosing left ventricular dysfunction [26, 27]. These papers are a sampling of the research showing the utility of the four main heart sounds for diagnosing patho- physiologic cardiac conditions.

However these papers are limited because information available only in snapshots is used. The goal of this research is to extract clinical information available over long term and incorporate this to better understand the pathophysiology of the heart.

2.5 Time Frequency Representation of Heart Sounds

Time frequency transforms have been thoroughly applied in the context of heart sounds. This section provides an overview of these methods in order to elucidate the complex nature of the heart sound signal. One of the most widely used applications of heart sounds has been murmur detection. Stethoscopes have been commonly used to listen to murmurs, and several researchers [40-43] have attempted to automate both murmur detection and classification. Since cardiac murmurs are high frequency sounds localized to certain parts of the cardiac cycle, time frequency transforms have been used to analyze them. In [44] time frequency techniques are used to improve heart sound signal quality. Time frequency transforms have also been used to understand the genesis of heart sounds. In [39] accelerometers placed on the epicardium were used to describe the complex time frequency characteristics of sounds created during left ventricular ejection. Components with rapidly increasing frequency were observed during early systole which was associated with increasing left ventricular pressure. Time frequency representations were used to describe the multi-component behavior of the S1 heart sound. Unlike the sounds on the epicardium, the S1 recorded on the thorax consisted primarily of stationary or quasi stationary superimposed components. Since then there have been many papers describing the time frequency distributions of heart sounds.

Essentially it is the multi-component nature of heart sounds that require analysis of these signals in both time and frequency dimensions. Figure 5 shows heart sounds recorded synchronously with intracardiac pressures. It is clear from the S1 waveform that the different components have different time frequency localization properties. For example, the early portion of S1, created due to the mitral apparatus is higher in frequency than the later portion which is a result of blood ejecting after aortic opening. This can be seen in the continuous wavelet transform (CWT) of the heart sounds signal (using Haar wavelet) shown in Figure 6. The mitral component is localized around 30-60 Hz where as the aortic component is localized around 15-25 Hz.

While there is rich information in the time frequency domain, regarding the localization and amplitudes of various heart sound components, this information is not consistent across patho-physiologic states. For example in patients with reduced contractility with mitral regurgitation the mitral component is diminished due to leaky valves; in patients with intra ventricular dyssynchrony and increased filling pressures a split S2 is present with a prominent P2 component. Hence each patient has a unique time frequency footprint with patho-physiologic information related to that disease state. This time frequency footprint has information that can be extracted and used for clinical applications. For example in [33, 45] a chirplet decomposition of the second heart sound is used to detect the A2 and the P2 components. The time difference between the two components is used to estimate pulmonary artery pressure.

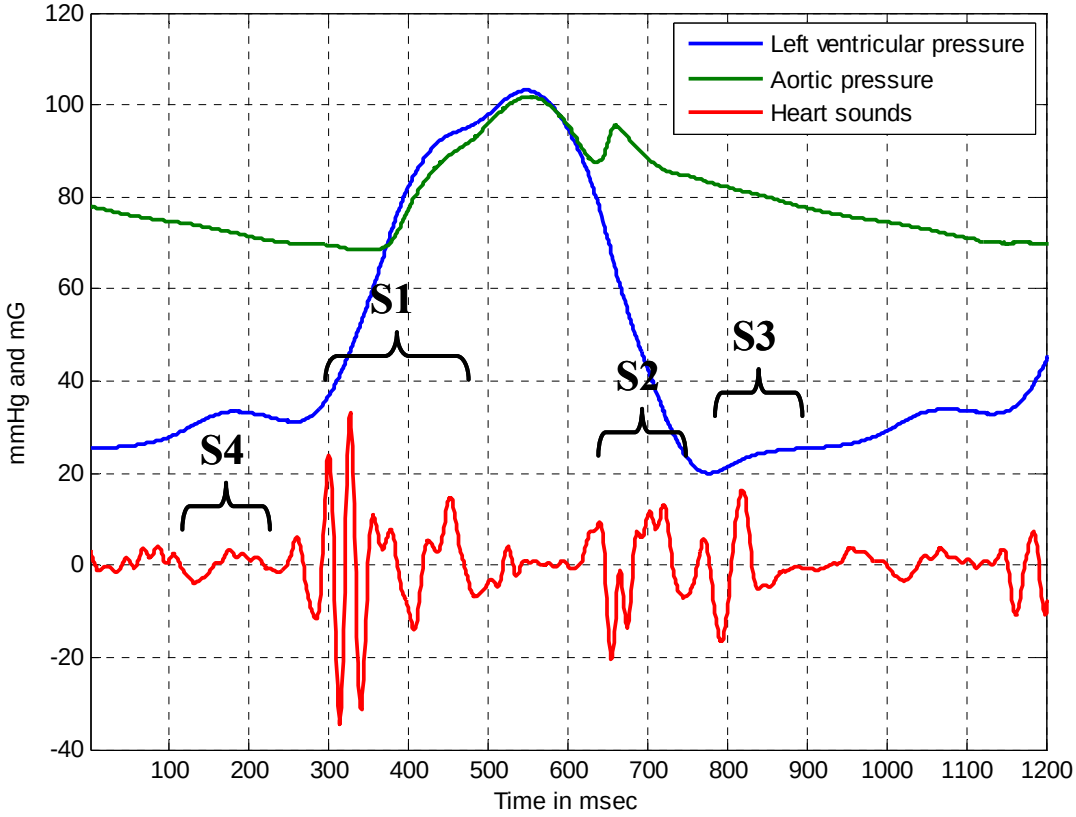


Figure 5: Heart sounds and intracardiac pressures

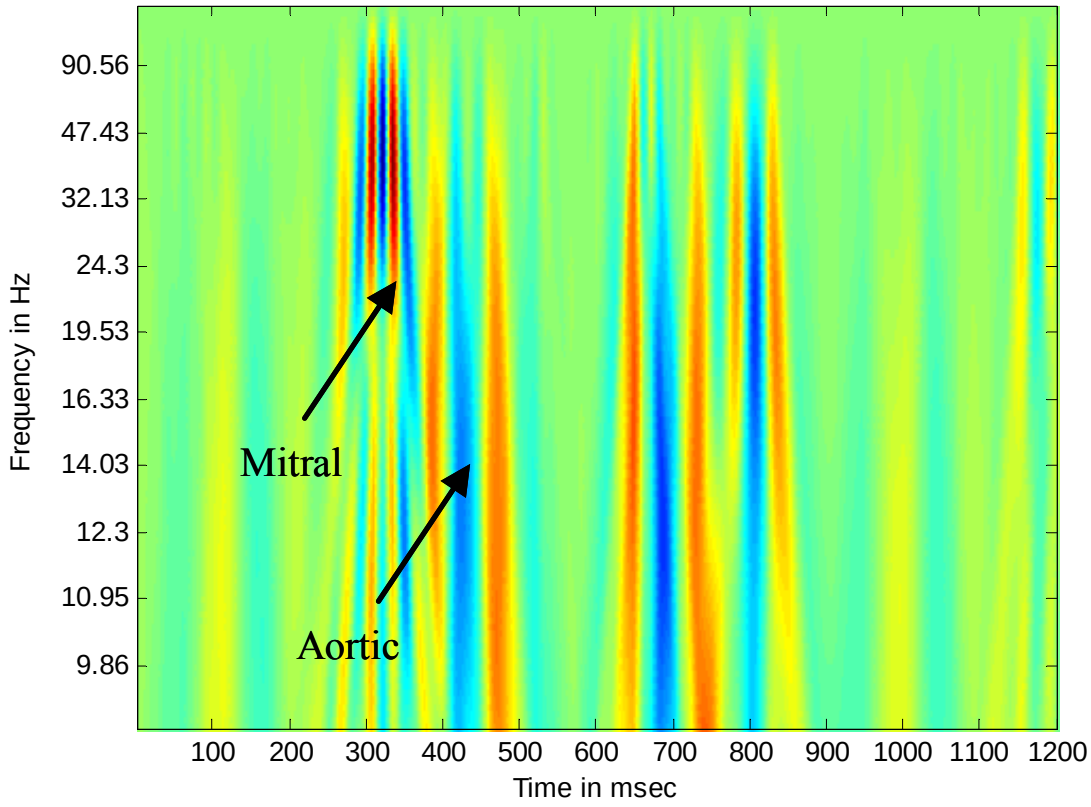


Figure 6: CWT of the heart sounds with time frequency localized components.

Another application of heart sounds has been in the differentiation of normal vs. artificial heart valves. Since artificial valves produce unique valvular sounds, time frequency analysis can be used for their analysis. Most research in time frequency representations of heart sounds has been in the analysis of heart sounds to explain patho-physiologic states. Since the goal of this thesis is to accurately measure heart sound based parameters, these time frequency transforms can be used to extract appropriate physiologic information.

2.6 Target Tracking

The feature tracking problem, in this thesis, has many parallels to target tracking in aviation and defense applications. This section describes target tracking in more detail in order to establish these parallels. Target tracking in defense applications is a well established field of research [46, 47]. Traditional targets have been airborne vehicles in surveillance systems and air traffic controls. Military applications include tracking of a weapon onto its target and tracking of enemy vehicles in radar. Sensor systems like radar, sonar and infrared provide measurements from their respective targets of interest, background noise such as clutter and thermal noise. The goal then is to separate the measurement data into sets of observations or tracks. Typically a host of sensor information and target dynamics and kinematics is used to solve this problem. Information regarding the target, such as target velocity and future position, can be obtained from these tracks.

A typical target tracker is recursive and consists of multiple tracks that it needs to follow. Figure 7 shows a basic target tracker [47] with incoming observations from the sensor and measurement module. These inputs are considered for the update of the existing tracks. Gating tests are performed to determine the possible observation-track pairings, in the candidate generation step, to assign the candidate features to their probable tracks. In the next step observations not assigned to existing tracks can be used to initiate new tracks. A new track can then be reinforced in following recursions. If a certain track does not have any observation it can be discontinued. Finally, the tracks

under consideration are predicted using the filtering and prediction before the new set of observations are obtained.

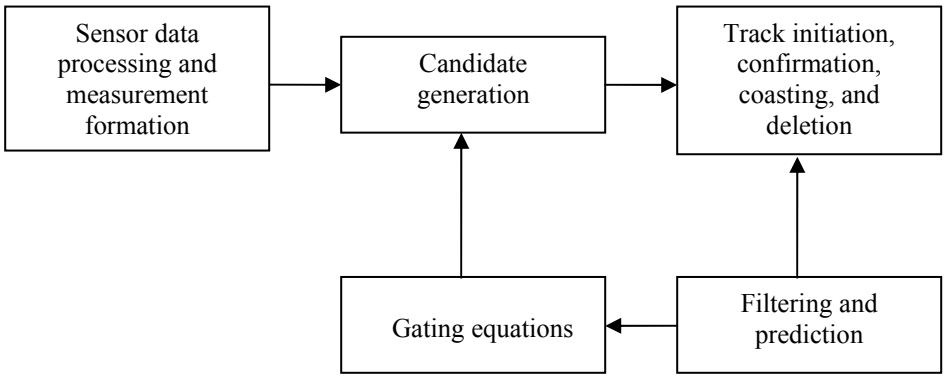


Figure 7: Elements of a multiple target tracker

The sensor preprocessing and measurements are specific to the tracking application. For instance, in radar tracking the processing includes consideration of false alarms, transforming the measured kinematics and target properties into a compatible coordination system and filtering for extraneous noise sources. In this thesis, the sensor preprocessing and measurements could involve measuring the peaks of the heart sounds, developing regression equations based on other physiologic sensor information like heart rate and adapting for known sources of physiologic variation. Information from multiple sources can be used to preprocess the measurements to improve the eventual tracking performance.

Candidate generation is where new observations are separated initially into one of many probable track sets using gating equations. The gating itself could use information from the previous gating equations as well as the sensor specific kinematics. A gate could be a distance metric based threshold to allow only observations which satisfy the distance criteria. The thresholds can be set based on expected or estimated variance of the sensor noise or target dynamics. For example regression equations using sensor information can be used as a gate for heart sound features. Also the gate conditions can be changed based on new track initiation or track deletions.

The track initiation step is used to assign new observations, not assigned to existing tracks, to new tentative tracks. Once a tentative track is formed confirmation logic is usually required to convert it to a track since one single observation cannot constitute a track. A typical rule would be to have received M such confirmations within N recursions [48]. If there is measurement noise or clutter the existing tracks can either be dropped immediately or a coast state can be achieved. In the coast state the recursion skips the current noisy observation and the track still remains even though no valid observation is received. A simple rule could be set that allows for P such coasts within N recursions. If the system coasts more than P coasts the track may be deleted.

One of the approaches to solving the filtering and prediction problem has been using Bayesian inference where the probability density function of the target is estimated recursively. The prediction operation uses a model of target kinematics to predict the probability density function using measurements obtained in realtime. The update phase uses the current sensor measurements to modify the predicted density

function. Bayes theorem is used to update the target density function. In [49] a tracking filter that uses all available data in a posteriori manner is developed. Several suboptimal computationally efficient tracking filters are also developed. A good review of various tracking methods, including the Bayesian approach, can be found in [50].

In [51] an extension of the Viterbi algorithm is proposed for tracking K paths through a trellis. This algorithm can be used to determine the most likely path using an energy cost function which is deterministic. The algorithm finds the set of K paths, no two of which pass through a common state on the trellis, such that the total metric sum for all the K paths is minimum. While the motivation arises from target tracking applications the algorithm can be used when there are no missing detections, no merged measurements and an additive cost function is used.

2.7 Feature Generation and Selection

Measurements obtained from heart sounds using a tracking algorithm contain information regarding cardiovascular status and perturbations. An important aspect of this research is to extract features from the various tracks that are useful for detecting the perturbation. For a given heart sounds recording, there may be multiple heart sound tracks that are measured. These tracks may capture different physiologic changes occurring due to cardiovascular perturbation. In order to find out which heart sound track is useful for the perturbation detection a set of derived features can be calculated. These derived features, for example, can be the energies in different frequency sub-bands of a time frequency decomposition of individual tracks. There are many machine learning algorithms [52, 53] that can be used to either classify the different

perturbations or predict the intensity of perturbation using a regression setting. One example is a SVM based classifier [52, 53]. The two main steps for the optimal performance of a classifier used for cardiac perturbation detection are feature generation and feature selection.

Feature generation: Feature generation is an important task for any predictive learning problem. Poor features, which do not capture the information in the physiologic signals, will lead to poor performance of the prediction algorithm. While the features have to capture the information in the physiologic signal, too many correlated features with increased dimensionality will require a larger training set. There are many feature generation methods for dimensionality reduction like singular value decomposition, time frequency transforms and principal component analysis. These are widely used to capture the information that may be relevant to the problem.

Feature Selection: Feature selection is very important since optimal performance is achieved when the best features are used and noisy features degrade the performance. The feature selection step can be classified into the following two categories:

- **Filter methods:** The filter methods rank features or a subset of features independently of any learning machine (classifier). The ranking is performed by assigning a metric to each feature. For example this metric can be the signal to noise ratio, mutual information, correlation coefficient, odds ratio or accuracy. This metric will provide information about the performance of each feature when used with a learning machine. In a regression problem, where a set of inputs are used to predict an output variable, a correlation coefficient can be

used to estimate the variability in the output due to a given input. A linear or non-linear fit can be used to calculate the correlation coefficient. Another common approach for variable selection is using entropy based (information theoretic) criteria. The estimated densities of discrete variables are used to calculate the mutual information between the input variables and the output to rank them. In a Linear Discriminant Analysis classifier it is common to use Fischer's separability criterion to rank variables to train the classifier. Other techniques use Chi-Square statistics or analysis of variance (ANOVA) type statistical tests to rank features to select them. Other filter based metrics are T-statistics, means separation and Kullback- Leibler distance [54].

- **Wrapper methods:** In supervised learning algorithms the goal is to select a subset of features that when trained on, will provide the best performance on a test dataset [55]. Ranking of candidate features is a common and simple method for selecting among them. One commonly used search strategy is forward selection and backward elimination of features. Others include a beam search where the *k* best paths at each step of feature selection are saved. Features can also be selected based on their weights during the training of a classifier [55, 56]. Other search strategies that are based on the above mentioned methods are also used for improved performance.

SVM Classifier: This section describes the SVM classifier which is used in this thesis. Consider a set of *m* training datasets ($\boldsymbol{x}_k$, y_k) (k=1, 2….*m*), where each training vector $\boldsymbol{x}_k$ has *p* input features and the outcome is $y_k \in$ (-1, 1) where +1 corresponds to one class

(e.g. OSA) and -1 corresponds to the other class (e.g. normal breathing). A linear hyperplane separating both the classes is defined by $f(\boldsymbol{x}) = \boldsymbol{w}\cdot\boldsymbol{x} + w_0$; where $\boldsymbol{w}$ defines the linear boundary. The classification rule is given by $G(\boldsymbol{x}) = \text{sign}(f(\boldsymbol{x}))$, such that a test dataset $\boldsymbol{x}_k$ is declared an OSA epoch if $G(\boldsymbol{x}_k) > 0$ and normal if $G(\boldsymbol{x}_k) \leq 0$. In a SVM formulation, the solution of $\boldsymbol{w}$ satisfies $y_k\cdot(\boldsymbol{w}\cdot\boldsymbol{x}_k + w_0) \geq 0$ subject to the minimization of $\|\boldsymbol{w}\|^2$. In addition there is a soft-margin formulation which allows data that cannot be separated using a decision boundary to improve the generalization of the algorithm. This is done by allowing slack variables (ξ_k) which is the distance from the decision boundary for datasets within the margin or on the wrong side of the decision line. Now the SVM classifier is defined by $y_k\cdot(\boldsymbol{w}\cdot\boldsymbol{x}_k + w_0) \geq 1 - \xi_k$ subject to the minimization $C\sum_{k=1}^{m}\xi_k^2 + \frac{1}{2}\|\boldsymbol{w}\|^2$, where C is the user defined parameter/penalty term that defines the tradeoff between the model complexity and number of separable datasets. The optimal estimation of $\boldsymbol{w}$ and w_0 is a quadratic programming problem whose solution can be found using Lagrange multipliers as shown in [57]. The data points falling on or within the margin or the opposite side of the margin are called support vectors and are the only datasets that are required to define the boundary. Also, the margin is defined by $\gamma = 1/\|\mathbf{w}\|^2$. Figure 8 shows an example of two class dataset with the linear kernel boundary and the support vectors (filled data points).

To achieve better classification, the datasets are mapped into a higher dimensional space by function $\boldsymbol{\varphi}$ and this is commonly known as the "kernel trick". $\boldsymbol{\varphi}$ can be a non-linear transformation that translates the linear boundary into a non-linear

boundary and now the input variables become $\varphi(\boldsymbol{x})$. The most commonly used kernels are linear, polynomial and Gaussian.

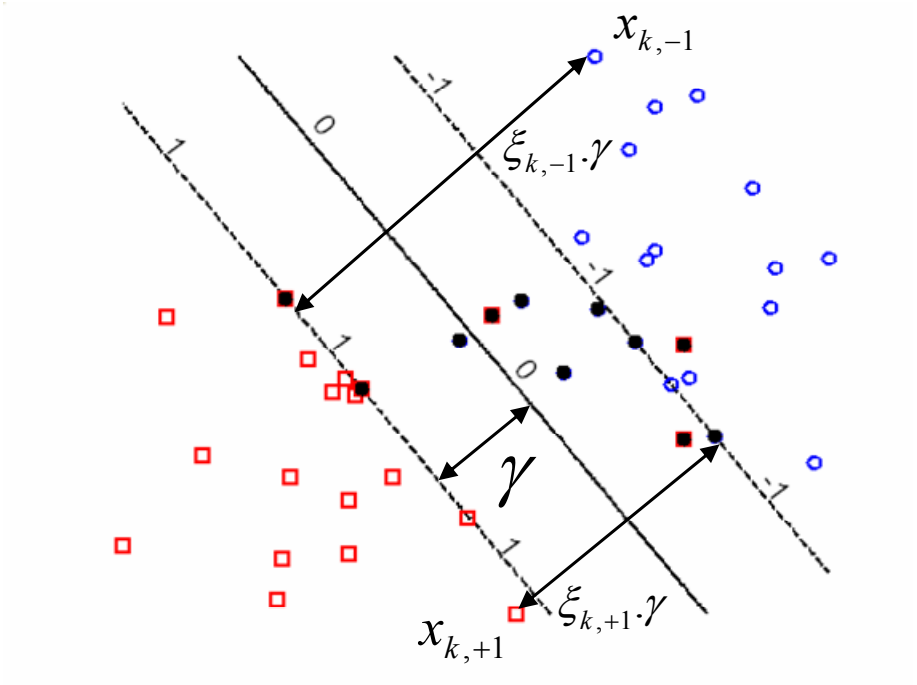


Figure 8: Example of SVM based classification using a linear kernel (*C*=100)

3 Tracking of Heart Sounds

3.1 Abstract

This chapter describes a novel low complexity framework for measuring heart sounds (S1-S4). For a given heart sound component (e.g. S1), candidate peaks are converted into states and assigned appropriate local scores. A low complexity dynamic programming based algorithm is developed to select the largest most consistent peak as the detected heart sound sub- component from these candidate peaks. To establish the value of the proposed framework, we tested it on acute and chronic pre-clinical data collected during heart failure deterioration and compared the tracking approach with a traditional non- tracking algorithm. In two acute animal models we found the S1 amplitude and ejection time to be statistically significantly correlated with the left ventricular (LV) contractility when measured using the tracking algorithm. In another canine model, with a chronic implanted heart sounds sensor, we found the S1 amplitude to be significantly correlated ($R = 0.78$, $p<0.05$) with LV contractility measured over four months in three dogs. As a comparison, measurements made using the non-tracking algorithm did not result in a statistically significant correlation. These results validate that the heart sound measurements using our proposed novel framework contain clinically useful information regarding heart failure status which has historically not been available due to the use of a non- tracking approach.

3.2 Introduction

Algorithms for measuring heart sounds fall in two categories, one which use the electrocardiogram signal as a cue for detections [36, 58, 59] and others which aim to

segment heart sounds without an additional ECG signal [43, 60-62]. These algorithms generally provide a measure of the intensity of the segmented heart sounds in terms of their energies or amplitudes. While they are successful in diagnosing cardiac conditions during the snapshots of recording, they are not setup for tracking chronic changes in cardiac conditions to be useful as a monitoring tool because they make measurements on each cardiac cycle independently.

To use heart sounds for long term ambulatory monitoring of cardiac conditions, it is crucial to detect and track the same heart sound component consistently over time. This tracking is even more difficult when the heart sound component that is being measured consists of multiple sub- components [10]. Also, as described in detail in Section 2.4, the amplitudes and timing intervals of these sub-components have been shown to be useful physiologic measures. For example, in [63] we have shown that the S3 heart sound peak amplitude measured using an implanted heart sound sensor over 6 months, increases significantly with worsening heart failure in a chronic animal study. In this case the S3 was used to track changes in worsening heart failure. There are other changes in the heart sounds which occur due to pathophysiologic changes in the heart which may warrant tracking. An acute myocardial infarction for example may change heart sounds due to cardiac injury. These changes, if captured correctly, can provide valuable information regarding the cardiac condition.

As mentioned earlier, the main challenge for long term monitoring of heart sound components is large variance in amplitude and timing measurements due to noise and physiologic changes. A solution to reduce the variance and misdetections is to use information from other heart sound beats before or after the current beat. Another necessity

for a practical heart sound tracking algorithm is independence from training or user intervention to allow for unbiased testing on newer applications as well as for detection of newer pathologies. Also, the prospective algorithm should have low complexity so that it can be practically implementable in an implanted device [64]. The goal of this chapter is to address these issues and develop a low-complexity framework for long term monitoring using a peak tracking algorithm. Section 3.4 describes the different components of the proposed peak tracking algorithm. As a preliminary test for utility acute and chronic pre-clinical studies with multiple heart failure models (described in Section 3.6) are performed to collect heart sounds data with clinically relevant perturbations. In Section 3.7 measurements using the tracking framework are shown to contain significant information regarding the heart failure status during the perturbations. Also, in all the data analysis the performance of this tracking technique is compared to a non- tracking approach, similar to the literature in this area, where measurements for a given cardiac cycle are made using information from only that cardiac cycle.

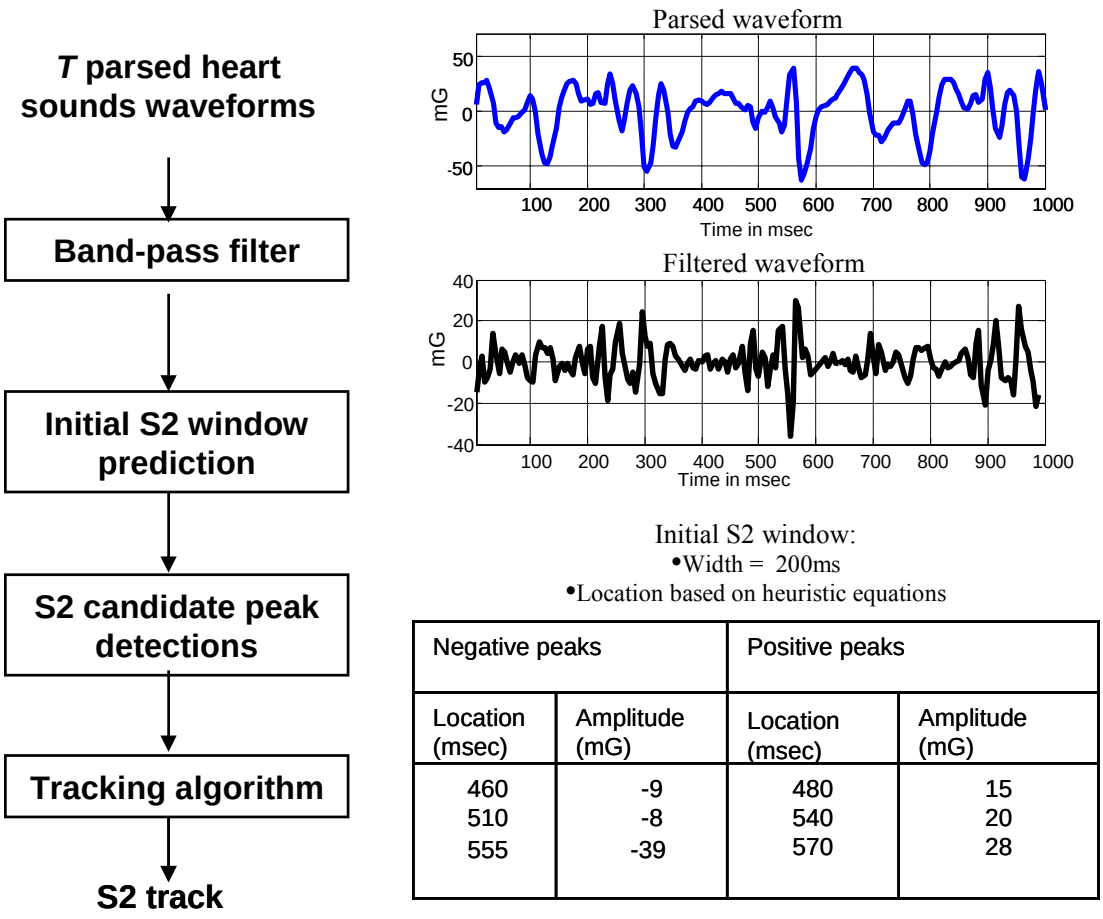
3.3 Non- tracking approach

In the past multiple algorithms for heart sound segmentation and measurement have been developed for datasets that have lasted few minutes at most [23, 34, 36, 43, 58-62]. These techniques have traditionally measured the energy or amplitude of the components, after band- pass or wavelet filtering, independently for each waveform. To compare our tracking framework with a non- tracking based approach the peak envelope of the filtered components in each cardiac cycle is used [58] as the non-tracking measurement. The amplitude and location of this peak envelope for each sound

component are the measurements that are obtained through this non- tracking approach. This envelope based approach extracts the energy of the largest sub-component of a given heart sound. We use this approach, as a comparison to our tracking algorithm, because it has been traditionally used to show the clinical correlates of heart sounds using snapshot based measurements, e.g. [37, 38].

3.4 Tracking Algorithm

The proposed framework uses the peak detections and measurements from individual cardiac cycles to select an appropriate peak track for the entire duration of tracking. Figure 9 shows the various pre-processing steps involved before a tracking algorithm can be used on a set of heart sound waveforms. Raw heart sounds can be recorded using either an accelerometer or a microphone and parsed into individual beats or averaged beats using the ECG marker (e.g. Q wave marker). A set of *T* parsed waveforms can then be filtered using a band pass filter with 15 Hz and 90 Hz filter cutoffs to extract the heart sound components. Based on heuristic rules a window can be setup to detect the possible peak candidates for S2. While our discussion Figure 9 depicts S2 detection as an example; in general it applies to any heart sound component that needs to be tracked. The candidate peaks, for each component, are then converted into states by assigning appropriate local scores to them as described in Section 3.5.1. While the scores can be a function of the energy or frequency characteristics of the peaks for this research we use the amplitude of the peak. These states form the basis for a tracking framework.



Negative peaks		Positive peaks	
Location (msec)	Amplitude (mG)	Location (msec)	Amplitude (mG)
460 510 555	-9 -8 -39	480 540 570	15 20 28

Figure 9: The heart sound detection and tracking algorithm showing the various steps for tracking. S2 is used as an example.

3.5 Specifications for the Tracking Algorithm

The following are the characteristics of a successful heart sound peak tracking algorithm:

- The tracking algorithm should consistently measure the same heart sound component or sub-component for the duration of tracking. During chronic monitoring, heart sounds change morphology due to changing physiology. The

tracking algorithm hence needs to be able to measure the peak corresponding to the same sub-component throughout the duration of measurements. Since the sub-components are expected to have low variability in their location, relative to the ECG marker, consistency can be measured in terms of the locations of the candidate peaks relative to the ECG marker.

- The algorithm should detect the largest peak of the component or sub-component being tracked. It has been shown that detecting the largest peak provides the most information regarding changes in the heart sound sub-components in [37, 38]. We use the amplitude (in mG) of the corresponding candidate peaks as measures to detect the largest peak.
- The peak tracking algorithm should use information from multiple cardiac cycles. This requirement enables the consistent measurement of various heart sounds.
- The peak tracking algorithm should be able to track multiple peaks at the same time. Since heart sounds have multiple sub components, this requirement provides the ability to measure them simultaneously to improve accuracy.
- The peak tracking algorithm should use heuristic rules inherent to the physiologic signal it is to be used for.

Based on these criteria our goal is to select the largest most consistent peak across the set of T parsed waveforms. We define consistency as measuring the same peak corresponding to a given sub-component for the duration of the recording. The performance of our algorithm and its consistency will be determined by comparing with

gold standard LV pressure measurements. Also, consistency is measured in terms of the variability in locations of the selected peaks relative to the ECG marker. Tracking is achieved by converting the candidate peaks into states and assigning a deterministic set of metrics that are maximized using a dynamic programming approach. Since they are part of the dynamic programming framework, the metrics we define use information from the most recent past to make decisions about the present. This agrees with the real-time nature of our application because we ignore any information from the future that could be useful to make decisions about the present. To understand the effects of using information from the future, and to evaluate the performance difference, we run the tracking algorithm in causal and non- causal settings in Section 3.7.4.

3.5.1 Tracking Metric Calculation

Since the algorithm is expected to consistently measure the largest overall sub-component, the metrics that need to be optimized are ones that provide the set of peaks with the largest sum of amplitudes and least variability in terms of location. As defined above, consistency is measured in terms of the variability in locations of the selected peaks relative to the ECG marker. For example in Figure 9 each time instance has 6 possible states (both positive and negative peaks) for the optimal track to follow. Assuming the goal is to find a positive set of peaks, there are 3 states for time k and also assuming that there are 3 states at time $k + 1$; a table of amplitudes and time indices for these potential states (with positive amplitudes) can be formed as shown below.

Amplitude (time)	State *a*	State *b*	State *c*
k	15 (480)	20 (540)	28 (570)
k+1	18 (495)	6 (555)	15 (570)

Table 1: Amplitudes (time indices) of states at time k and k+1

If the criteria is to find the largest most consistent peak then the two potential paths are through state *a* (at time *k* and *k + 1*) and through state *c* (at time *k* and *k + 1*). Of the two potential tracks the one through state *c* has a larger total amplitude and better consistency in terms of location. Hence the choice in this case would be the track through state *c*. Based on this concept of selecting the largest most consistent peak, according to the definition give above, a cost metric of jumping from state x to state x' is given by,

$$J(x',x) = |(I_x - I_{x'})|$$

where, A_x is the amplitude of the peak at state x, I_x is its time index. Also a local score of each state is simply given by,

$$d(x_t, x) = GA_x .$$

where *G* is a tunable gain constant. The cost of jumping to a state with larger amplitude is lower than jumping to a state with smaller amplitude. On the other hand the cost of jumping to a state that is farther is higher than jumping to a state that is closer. Also the cost of jumping away a certain fixed distance from the current state in either direction is the same. This concept is next captured in the single peak tracking algorithm.

3.5.2 Single Peak Tracker

To implement the single peak tracking solution a trellis based representation of heart sounds peaks is first described. A trellis diagram is a representation of the possible state transitions for the tracker over time. Consider the trellis diagram shown in Figure 10a, which depicts the different possible peak tracks through various heart sound peaks. Each stage of the trellis corresponds to a set of heart sounds peaks from the heart sound waveform at a given time. The trellis shows the 3 states at time k and time $k+1$. Assuming there are 4 possible states at time $k+2$ there are 36 possible tracks from state k to state $k+2$. The goal of the single peak tracker is to choose the track which provides the largest most consistent peak. Each state transition in the trellis has a cost associated with it. This cost is the peak state transition metric described in Section 3.5.1. Choosing $G = 1$, the various metric of state transitions from the states at time k to states at time $k+1$ are tabulated below.

(J, d)	a, $k+1$	b, $k+1$	c, $k+1$
a, k	(15, 18)	(75, 6)	(90, 15)
b, k	(45, 18)	(15, 6)	(30, 15)
c, k	(75, 18)	(15, 6)	(0, 15)

Table 2: Cost of state transitions from time *k* to time *k+1*

Again, the transition through state c provides a reasonably large amplitude peak with least cost in terms of variability. To find the track that provides the least cost across N

states the dynamic programming forward backward technique is used. The following definitions are first provided:

- Each path between state x' and x has a jump penalty given by $J(x',x)$. This jump penalty is related to the time indices of states x' and x.
- The set of all possible predecessors of x is given by $M(x)$
- The local score at point x is given by $d(x_t,x)$. This local score is a function of the energy or amplitude of the peak that corresponds to state x. The local score can be defined based on physiologic expectation of a particular peak.
- The total score of a track at time t at a state x is given by $D(t,x)$.
- The best predecessor for a given state x at time t is given by $B(t,x)$.

Based on these definitions a forward recursive pass for calculating the total score is given by,

$$D(t,x) = \max_{x' \in \mathrm{M}(x)} \{D(t-1,x') - J(x',x)\} + d(X_t,x).$$

The total score of a state at time t depends on the energy of the peak at that state and the least jump penalty from the previous states at time $t-1$. Also at each time t the best predecessor is stored as a backpointer given by,

$$B(t,x) = \arg\max_{x' \in \mathrm{M}(x)} \{D(t-1,x') - J(x',x)\}.$$

In the backward pass these backpointers are used to trace back the track with the largest score starting with the state with the largest D at time T. Table 2 shows the values of the

metrics J and D for various paths between time k and time $k+1$. If we use the cost function described above, then the track that has the least cost is the track that has state c in it.

When there are consistently more than three heart sound peaks which have comparable amplitudes, as in an S1 complex, the single peak tracker will not pick the appropriate peak track. This is because the track that has the largest score may not always be the best, for physiologic correlations, when there are multiple peaks. Also, the single peak tracker may jump tracks due to morphology changes which happen frequently with physiologic variation. To overcome these challenges a multiple peak tracker is developed in the next section.

3.5.3 Multiple Peak Tracker

A multiple peak tracker aims to detect and measure more than one peak track simultaneously. The goal is to maximize the total probability that the K tracks are the correct heart sound peaks. The tracker does not find the best set of K tracks by finding the best track, then the next best track and so on sequentially. This solution would work in the case where a segment of a track can be a part of another track. In the case of heart sounds if a peak, for a given beat, is a part of a track then it cannot be a part of another track through that beat. Since the tracks cannot share the same peak, the K tracks are no longer independent and cannot be selected sequentially. Also, in the case of heart sounds tracking, the K tracks must be parallel and non-intersecting. That is any two tracks cannot be overlapping in their peak detections.

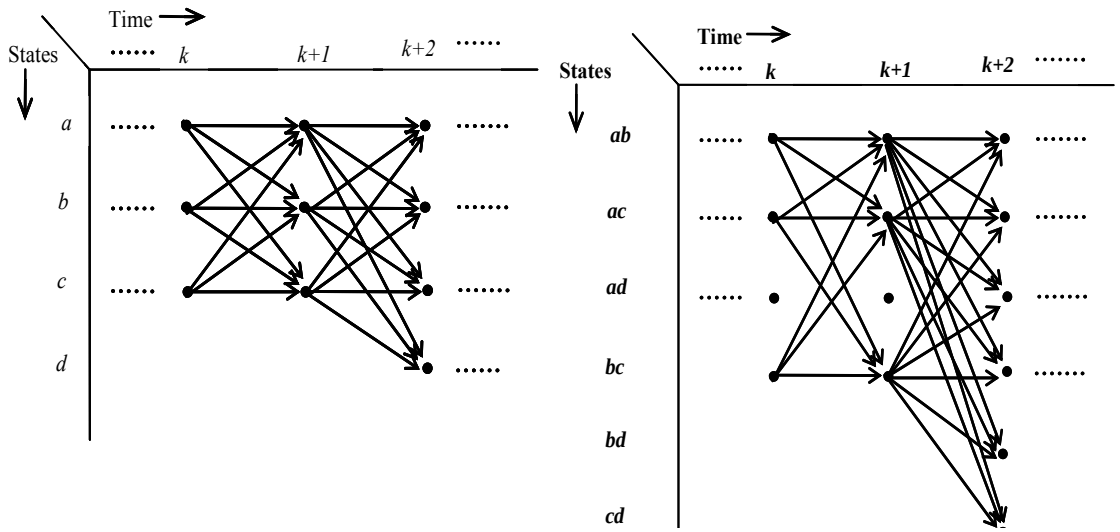


Figure 10 a and b: Trellis diagram representations of heart sound peaks for single and multiple peak trackers.

The multiple peak tracker problem can be solved by converting it into a single peak tracking problem. Consider the trellis in Figure 10 a where there are 3 states at time *k* and *k + 1* and 4 four states at time *k + 2*. To generate two peak tracks through this trellis we would have to select 2 peaks at each time instance such that the two tracks have the largest overall score and yet do not overlap. To select these two tracks we first generate an expanded trellis as shown in Figure 10b.

The states of the expanded trellis are a *K* fold product of the states from the original trellis. If there are N_k states at a given time *k* the number of states in the expanded trellis is given by N_k / K. The branches between the expanded trellis states have new values for amplitudes and locations. The amplitudes for the new states are the sum of amplitudes of the original states; these are tabulated in Table 3 along with the respective time indices. Based on this concept of selecting the largest most consistent

peak jump penalty metric, similar to the one for single peak tracker, associated with jumping from state x to state x' is given by,

$$J(x_1'x_2'...x_K', x_1x_2...x_K) = \sum_{i=1}^{K} |(I_{x_i} - I_{x_i'})|.$$

Also a local score of each expanded state is given by,

$$d(x_t, x_1x_2...x_K) = G\sum_{i=1}^{K} A_{x_i}.$$

where, A_{x_i} is the amplitude of the peak at state x_i of the original trellis I_{x_i} is its time index and *G* is a tunable gain constant.

	State *ab*	State *ac*	State *bc*
k	35 (480, 540)	48 (480, 570)	43 (540, 570)
k+1	24 (495, 555)	33 (495, 570)	21 (555, 570)

Table 3: Amplitudes (time indices) of expanded states

Using the cost metrics defined above the multiple peak tracker has been converted to a single peak tracker. The best single track, which maximizes the overall score, in the expanded trellis can hence be calculated. This single best track through the expanded trellis is actually a *K* fold set of multiple states chosen simultaneously to maximize the overall score. For example in Table 3, the expanded states *ac* have the most consistency (15 msec shift) and largest sum amplitudes. Hence choosing states *ac* in the expanded trellis will provide the largest score (for times *k* and *k + 1*). However,

choosing states *ac* is the same as choosing two tracks through states *a* and *c* in the original trellis.

An illustration of the peak trackers applied to heart sounds collected over long term is shown in Figure 11 where the three sub-components of S1 and the peak S2 are detected using multiple and single peak tracker respectively. The top panel is a unique representation of parsed and averaged heart sound beats [35]. Consecutive averaged waveforms are aligned using the ECG marker such that the corresponding marker is at 250 ms for each waveform. The 600 waveforms correspond to the 10 hours of recording and the color represents the amplitude of each waveform in mG. As seen in the figure there are 3 positive peaks detected for S1 which are the three sub-components. For the S2 a peak negative peak is tracked using a single peak tracker. Over the 10 hour period these sub- components change due to change in patient status which can be captured using the track measurements.

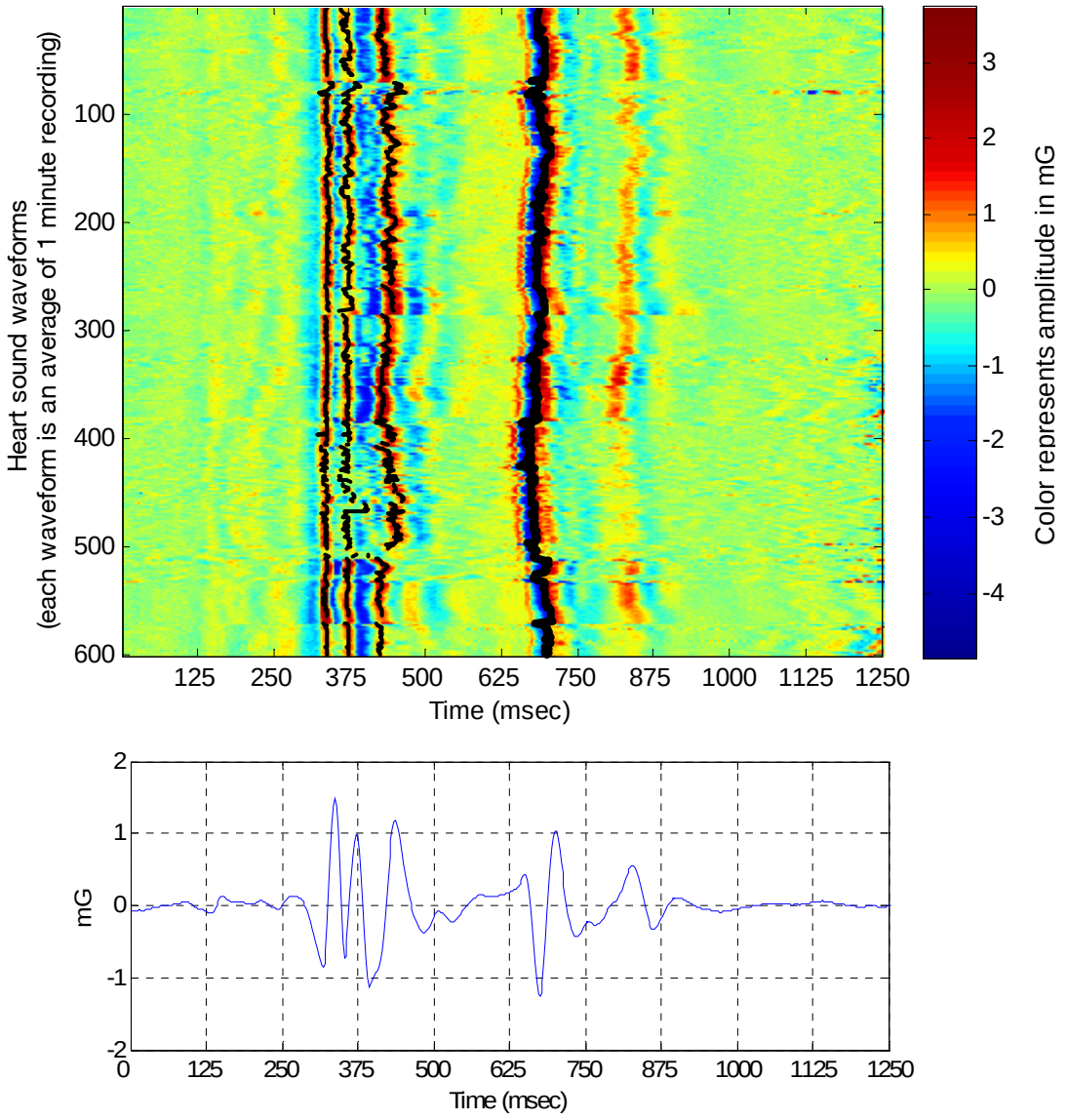


Figure 11: Illustration of a multiple and single peak trackers detecting the three dominant peaks of the S1 and peak S2 respectively during a 10 hour recording.

3.5.4 Initiation, Coasting and Merging- Splitting

In order to start the single or multiple peak tracker the number of tracks needs to be initialized first. The number of tracks depends on the kind of heart sound being

detected. For example the S3 or S4 heart sounds have a single component/ peak that needs to be detected. In the case of S2 there may be multiple sub-components/ peaks if a split S2 is expected. For the S1 there are usually 2-3 peak tracks that need to be detected simultaneously. Hence to find the number of potential tracks a peak detection algorithm is applied to the average waveform (for the period of the tracking). For example the average waveform (panel 2) of Figure 11 is used to determine that, on average, there are three positive peaks, representing the three S1 sub-components, which need to be tracked.

During the dynamic programming peak track generation some heart sound waveforms may provide noisy candidates which can be skipped. This step of skipping over one or many heart sound waveforms, since they are noisy, is called coasting. Coasting can be performed if: the number of candidates for a give waveform is less than *K,* the candidates have a low local score *d,* the candidates have a high jump penalty *J* from the previous state, the waveform does not satisfy some minimum noise threshold. Coasting can be performed if one or a set of *P* (e.g. 10 beats) beats are noisy. After the noisy beats are coasted over; the scores and backpointers, for the following beats, are continued from the waveforms before the coasting. If more than *P* beats are noisy, and need to be coasted over, the system will break the track. If the tracking stops it may have to be manually restarted or the dataset may be labeled corrupt. Another option is to restart the tracking with new total scores *D* and backpointers *B*. This restart is useful when there has been a long gap in the data collection and the previous *D* and *B* are not representative of the current data. For example, when the physiology of heart sound changes, leading to morphology change which result in tracks merging or splitting, the

algorithm will coast due to changes in the tracking metrics. Similar to breaking the track and reinitiating due to noisy beats, the merging- splitting happens when physiology behaves as a source of noise. The number of beats P is selected to tradeoff the number of measurements required vs. noise immunity.

In our analysis, we have found that selecting $K = 3$ for the S1 measurement and $K = 1$ for S2 provides the best performance in terms of detection accuracy and noise immunity. Additionally, we choose $G = 1$ and $P=10$ in our analysis since it provides the best performance.

3.5.5 Tracking complexity

The parameters that influence the complexity of the dynamic programming tracking algorithm are the number of tracks (K) and the depth of the trellis (N). Assuming the number of candidate peaks for each heart sound waveform is fixed at M, the complexity [65] of the tracking algorithm is of order $O(W)$, where W is given by,

$$W = \left(\frac{M}{K} \right)^2 K^3 N \ .$$

It should be noted that the complexity of the algorithm is linear with N. This is because the forward backward dynamic programming algorithm is used as the solution for the search. Also for every pair of subsets of size K from time t to time $t+1$, *a* K by K assignment problem must be solved which has complexity K^3. Note that while the complexity is cubic with K the application of heart sounds tracking anticipates large N (e.g. N = 1000) and small K (e.g. K = 1, 2). Hence the fact that the tracking algorithm has complexity linear in N makes it extremely attractive for implementation.

3.6 Pre-clinical Experiments

The goal of the pre-clinical studies was to understand the changes in heart sounds with worsening cardiac systolic performance and to evaluate the utility of the tracking algorithm to detect these pathophysiologic changes. Two acute and one chronic animal model, described below, were used to create heart failure in swine and canine. For the acute studies, heart sounds were captured using a surface accelerometer synchronously with ECG and intra-cardiac pressures (Millar Instruments) sampled at 1000 Hz. A chronic canine study was performed next, to evaluate the tracking algorithm in a long term setting, with an accelerometer based implanted heart sounds sensor and an implanted LV pressure monitor for gold- standard measurements; both sampled at 200 Hz. It is to be noted that 200 Hz sampling was chosen for complexity reasons with the knowledge that the heart sounds of interest (S1 and S2) are known to have most of their energy in the 15 Hz to 100 Hz bandwidth [36, 39, 66, 67]. The pre-clinical studies complied with all animal use regulations as set forth in the Animal Welfare Act, Title 9, CFR, Chapter 1, Subchapter A and adhered to the principles outlined in the "Guide for the Care and Use of Animals," National Institutes of Health Publication.

3.6.1 Acute Animal Model I

Ischemic heart failure was created using a coronary occlusion protocol. The swine model was selected because of the similarity to the anatomical geometry, scale, and extent of arterial collateral branching to the human. Other large animal models, such as canine, have different collateral branching from the human, which impact the temporal profile of the signal changes during ischemia. Since phenylephrine, epinephrine and

dobutamine impact the degree of ischemia, their usage was maintained at a minimal level during ischemia and acute myocardial infarction induction and they were administered under the consensus of the facility veterinarian. The animals were sedated, intubated and placed on a volume-controlled ventilator. Oxygen and isoflurane levels were set as determined appropriate by the facility veterinarian and adjusted as needed to maintain a surgical plane of anesthesia. The animals were also given heparin and anti-arrhythmic drugs as per the protocol. After baseline data collection, closed-chest ischemia induction was performed using the balloon occlusion and reperfusion method. In brief, after introducing a guide wire in the pre-specified artery, an angioplasty balloon of appropriate size and length for the specified artery was positioned in the pre-specified location. The balloon was inflated for approximately 60 minutes. Up to two of the following three coronary arteries were occluded in one study: LAD, LCX and RCA.

3.6.2 Acute Animal Model II

This acute heart failure model included inducing mitral regurgitation by chordae damage and increasing afterload using vaso-constrictive pharmaceuticals. This procedure required the use of a specialized steerable guide catheter and biopsy catheter introduced through the left or right common carotid artery, down the ascending aorta, through the aortic valve, and into the left ventricle where it is was used to capture the chordae associated with the mitral valve and then retracted to cause damage to the chordae. This process was repeated until the mitral valve had been weakened sufficiently to cause mitral regurgitation leading to elevated left atrial pressures. Either additional mitral valve weakening or vaso-constrictive pharmaceuticals (Phenylephrine) were used to further elevate the left atrial pressures. Left atrial and left ventricular

pressures were monitored to determine the effectiveness of the induction method. Phenylephrine was used to increase and control the left sided pressures once sufficient mitral regurgitation was developed.

3.6.3 Chronic Animal Model

Three adult mongrel canines were implanted with an implantable cardioverter defibrillator (ICD) modified for heart sounds waveform capture capability and a chronic LV pressure monitor. The ICD was placed on the left side, ventrally at the 3rd and 4th intercostals. The heart failure model involved microsphere injection in the Left Circumflex to target a 10 – 20% infarct followed by high rate pacing (200 bpm). Clinical treatment for heart failure was based on pulmonary symptoms (dyspnea, crackles) and lethargy. Twice a week pacing was interrupted for one hour at the end of which heart sounds, ECG and LV pressure were acquired from the implanted devices using inductive telemetry for at least 15 minutes.

3.7 Results

This section describes the comparisons of heart sound measurements, made using the tracking and non- tracking algorithm, with gold-standard measurements from the intra-cardiac pressures. Filtering and averaging were performed to mitigate any muscle noise and motion artifacts. A band pass filter with 15 to 90 Hz cutoffs was used for both the acute and chronic study data. For the acute study, 10 second average heart sounds waveforms were computed using the ECG- R wave as the reference to parse individual beats. For the chronic implanted study 1 minute averages were computed. The tracking algorithm was then used to consistently detect the same peak amplitude peak of the sub-

components in the average heart sounds waveforms collected during the course of the study. These measurements, made using both the tracking and non- tracking approach (Section 3.3), were compared with gold- standard measurements to evaluate the utility of the tracking algorithm.

3.7.1 Ischemia Induction

Multiple heart sounds parameters, including S1 intensity and heart sounds ejection time, were extracted using the tracking algorithm described in Section III. These parameters were analyzed for their correlation to the aortic pulse pressure (PP) and maximum left ventricular dP/dt, which are measures of the cardiac systolic performance, in the context of ischemia. An example of the changes in these parameters, in one animal, is shown in

Figure 12 where the S1 intensity and ejection time normalized by cardiac cycle length (HSET) change after the balloon occlusion. It is interesting to note that the cardiac contractility has a reflex increase/ decrease in the minutes following the occlusion, in this animal, which is also seen in the S1 intensity and HSET interval parameters. The S1, measured using the tracking algorithm, intensity for the duration of the experiment was correlated with PP and max dp/dt with an average correlation coefficient of 0.48 ($p<0.01$) and 0.51 ($p<0.01$) across six animals. The HSET interval, measured using the tracking algorithm, was also correlated with PP and max dP/dt with an average correlation coefficient of 0.54 ($p<0.01$) and 0.62 ($p<0.01$) across six animals. As a comparison to the tracking algorithm results, measurements made using the non-tracking algorithm were also compared to LV dp/dt. The non- tracking based S1 had a correlation coefficient of 0.32 ($p<0.05$) and 0.36 ($p<0.05$) with LV dp/dt and PP

respectively. On the other hand, the non- tracking based HSET, due to higher noise, did not have a statistically significant correlation with either LV dp/dt or PP. These results are also summarized in Table 4.

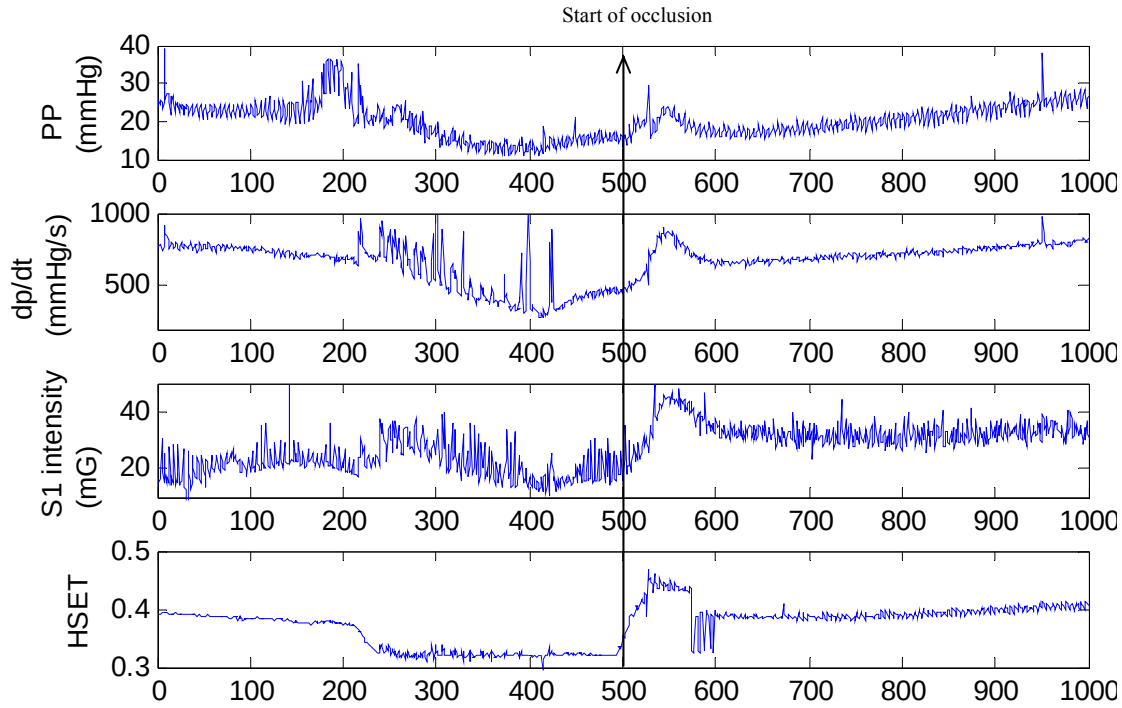


Figure 12: Example of changes in systolic performance and heart sound parameters due to coronary artery occlusion. The correlation coefficient, in this example, between S1 amplitude and LV dP/dt is 0.54.

3.7.2 MR Induction and Afterload Increase

This acute study protocol was attempted in 3 dogs of which 2 dogs developed moderate to high grade mitral regurgitation post cordae damage as determined by the attending

veterinarian using a stethoscope. This resulted in more than 50 % reduction in left ventricular contractility as measured by LV dP/dt. In one of the 3 animals due to challenges in pappilary and cordae anatomy MR induction was not achieved. In all the three animals there were significant changes in contractility due to phenylepherine infusions. As in the case of the Ischemia protocol the S1 amplitude, measured using the tracking algorithm, was statistically significantly correlated with max LV dP/dt and PP with an average correlation coefficient of 0.6 and 0.58 respectively. The HSET, measured using the tracking algorithm, on the other hand had an average correlation coefficient of 0.52 and 0.47 with max LV dP/dt and PP respectively. The correlation coefficient, in general, was higher in this protocol, compared to the ischemia model, due to the larger changes in max LV dP/dt. In order to assess the performance of heart sounds to predict changes in contractility (max LV dP/dt), percent changes from baselines were compared for S1 amplitude and HSET to max LV dP/dt. As a comparison to the tracking algorithm results, measurements made using the non-tracking algorithm were also evaluated for hemodynamic correlations. The non-tracking based S1 had a correlation coefficient of 0.39 ($p<0.05$) and 0.32 ($p<0.05$) with LV dp/dt and PP respectively. Similar to the ischemia protocol results the non- tracking based HSET did not have a statistically significant correlation with either LV dp/dt or PP as shown in Table 4. Figure 13 shows the two measurements at instances where the contractility was 25% higher or lower than the baseline values for the three animals. Baseline contractility was defined as the mean of the 2 minutes of data recording preceding any drug or surgical intervention.

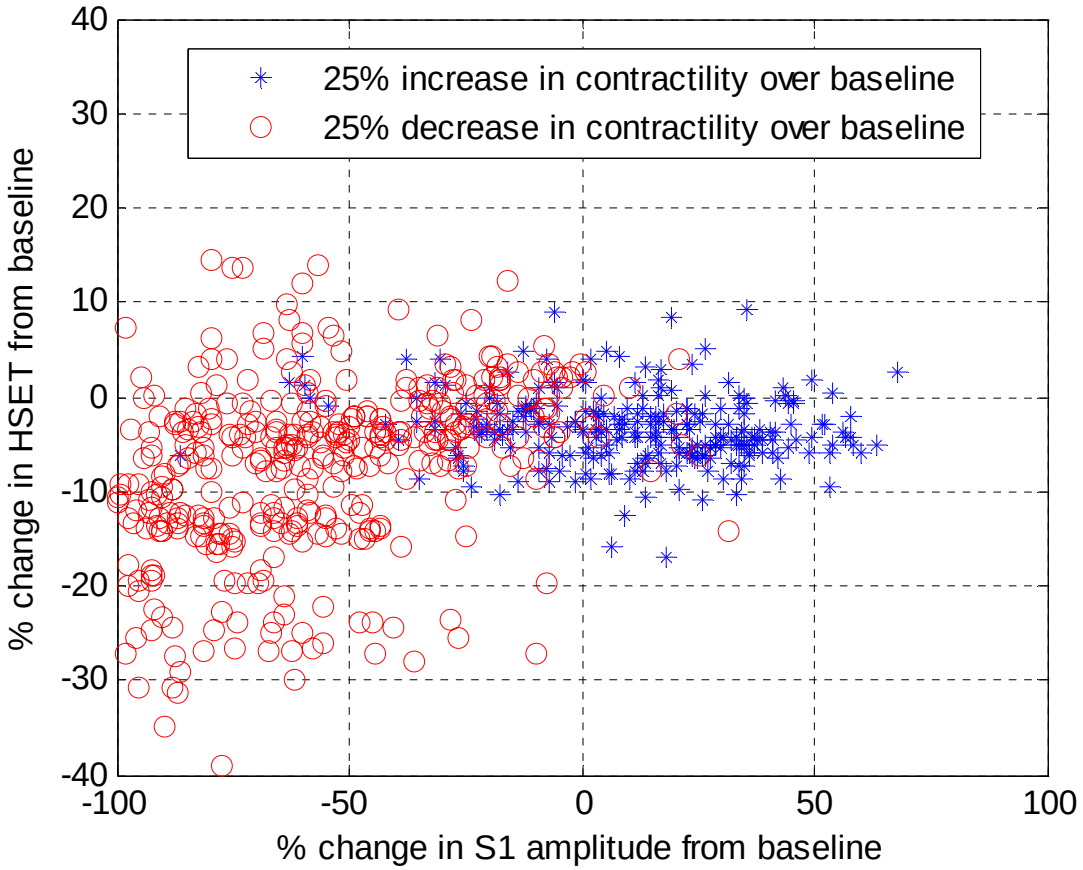


Figure 13: S1 amplitude and HSET at increased and decreased contractility states.

When we used the S1 amplitude as a threshold to detect time instances with greater than 25% decrease in contractility from the baseline, it had a receiver operating characteristic with 85.2% sensitivity at 80% specificity (accuracy = 82.6%). The HSET had a receiver operating characteristic with 50.6% sensitivity at 80% specificity to detect instances with greater than 25% decrease in contractility from the baseline (accuracy = 60.3%). These curves are shown in Figure 14.

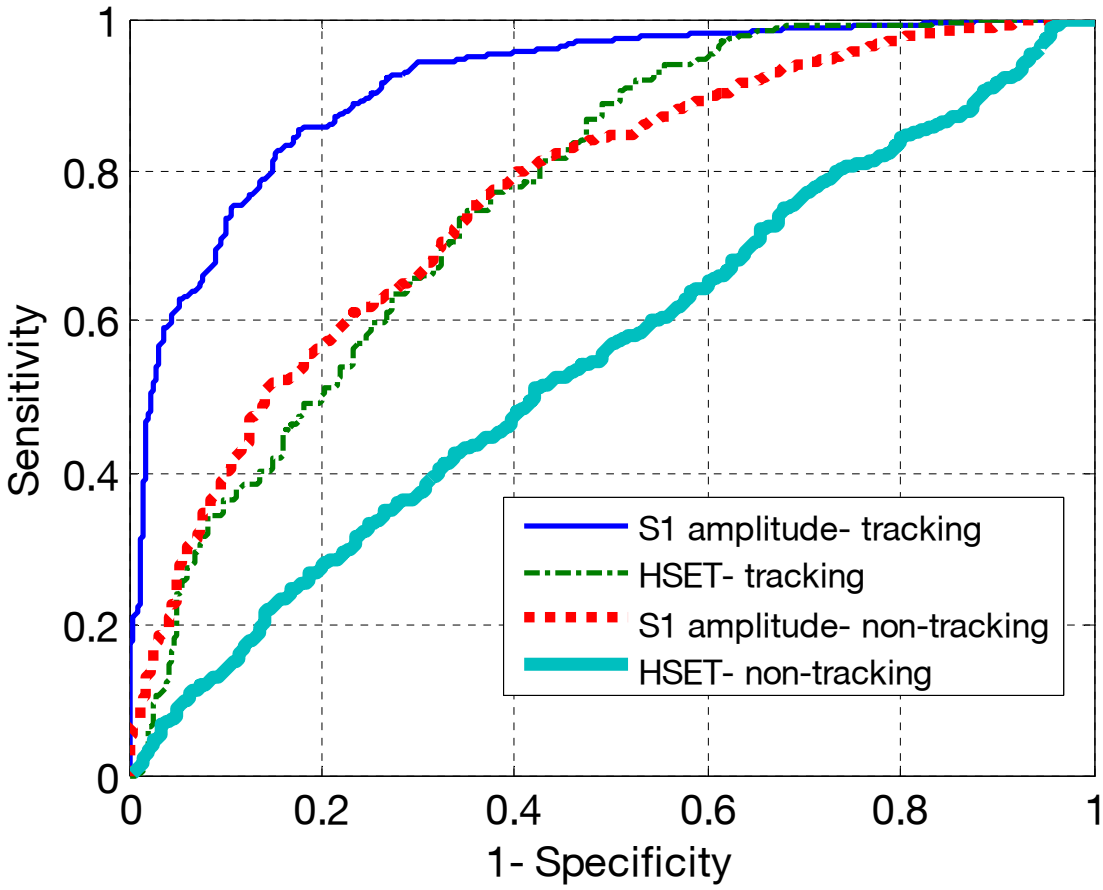


Figure 14: Receiver operating curves for using S1 and HSET to detect periods with greater than 25% decrease in contractility from the baseline.

The poor performance of HSET, compared to the ischemia model, we think is associated with the competing effect of afterload and heart rate due to vaso- constrictive drugs, in this model, which could not be compensated for by normalization. Also it is important to note that theses measurements are averages of 10 second time durations considered independently. The predictive ability of these trends will improve with longer averaging and a change detection algorithm. In comparison, as shown in Table 1, the S1 measured using the non- tracking algorithm had an accuracy of 68.5% (56.8%

sensitivity at 80% specificity) to detect time instances with greater than 25% decrease in contractility from the baseline. On the other hand the HSET without tracking had no-predictive power to detect periods of reduced contractility, as reflected in the non-significant correlation coefficient values.

3.7.3 Chronic Heart Failure

After heart failure induction, the left ventricular shortening fraction, a measure of cardiac stroke, decreased by 45.63 %. Data in three dogs was collected for an average duration of 116 days. As described in the animal model heart sounds, ECG and LVP were inductively acquired twice a week. Data were processed using the tracking algorithm and measurements were compared with gold standard max LV dP/dt. Since the aortic pressure was not acquired, PP was not measured. Figure 15 shows waveforms of heart sounds before and after heart failure was induced using the rapid pacing protocol in one animal. As expected the S1 amplitude was reduced after heart failure induction with reduced contractility. In the same animal Figure 16 shows trends of the sparsely sampled S1 amplitude and max LV dP/dt over the 160 day data collection duration. Immediately after implant the max LV dP/dt is high due to the rapid pacing protocol. As the heart failure sets in the LV contractility as measured by max dP/dt is reduced which is also reflected in the S1 amplitude trends.

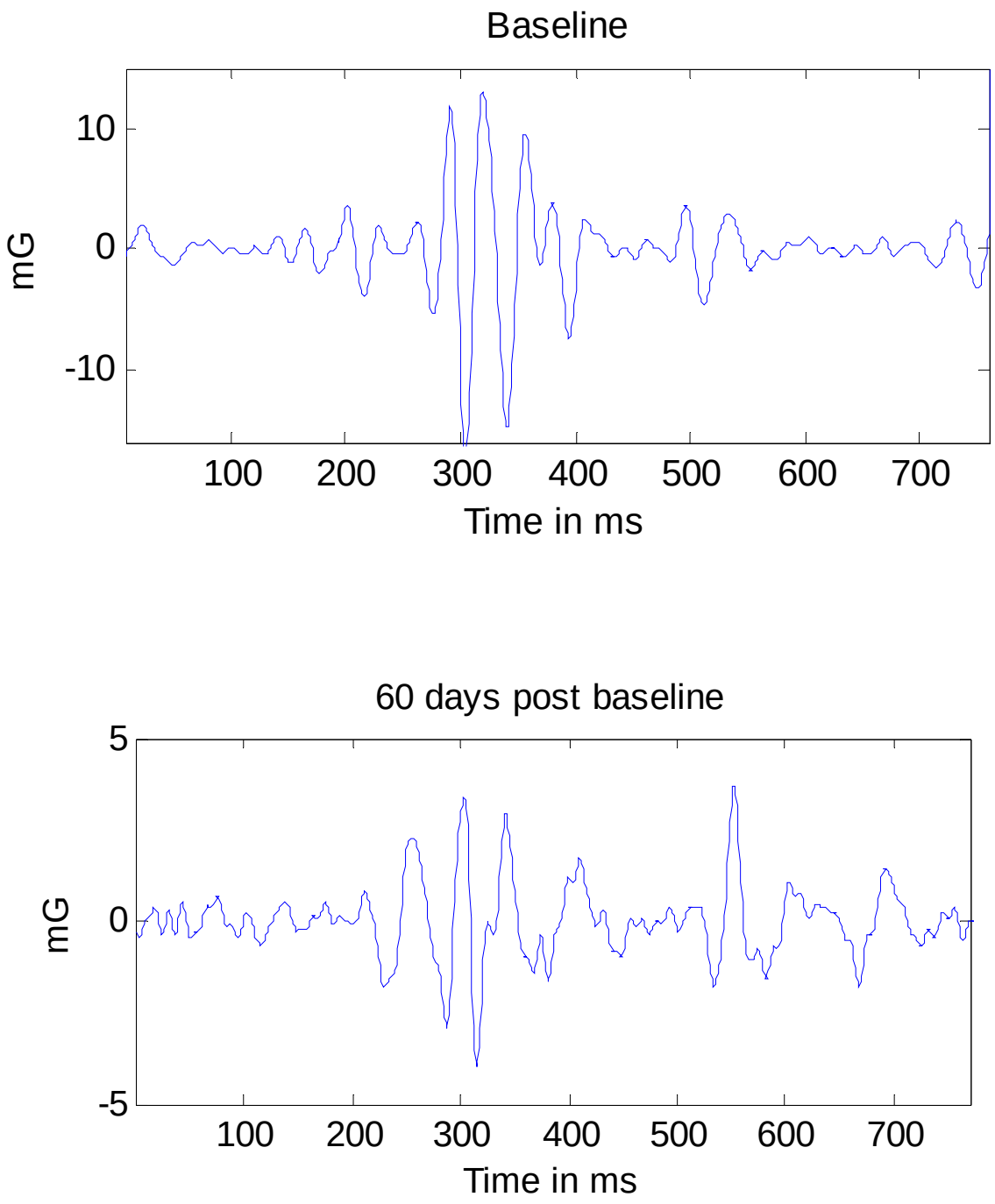


Figure 15: Heart sounds waveforms before and after heart failure induction showing decrease in S1 amplitude with reduced contractility.

Across three animals the S1 intensity was correlated to the max LV dP/dt with a correlation coefficient of 0.78 ($p<0.05$), evaluated over an average duration of 4 months in 3 dogs. Figure 17 shows the scatter plot between max LV dP/dt and S1 amplitude in 3 dogs. When we used the S1 amplitude as a threshold to detect a reduction in contractility from baseline the resulting receiver operating characteristic with maximum accuracy of 84% (Sensitivity = 80% and Specificity = 89%) is shown in Figure 18. When measured using the non- tracking algorithm the S1 did not have a statistically significant correlation with LV dp/dt and had a poor accuracy of 60% (40% sensitivity at 80% specificity) to detect decreases in contractility from baseline, as shown in Figure 18

3.7.4 Tracking vs. Non- Causal Tracking

The tracking algorithm described in Section 3.4 assumes that optimal track up to time *t* does not depend on the observations after time *t*, i.e. the future. To evaluate the validity of this assumption, data from the acute and chronic animal models is used. S1 heart sound measurements are obtained both with a causal tracker, where the measurements are made chronologically, and the non- causal tracker, where measurements are made in a reverse-chronological order. Measurements from the two scenarios are compared with each other to determine the differences in information available. Also, measurements in the two scenarios were compared with LV dp/dt to determine differences in their correlation coefficients with gold standard information. In the acute animal studies 0.26% of the peak measurements were different in the two scenarios. These differences were due to 12 instances of coasting in 3 animals. The rest of the 6 datasets did not have any instances of coasting or differences in peak measurements.

Because these differences in measurements were in less than 1% of all measurements made, there were no differences in the correlation coefficients with LV dp/dt in the two scenarios. In the chronic animal data, again there were no differences in measurements made using the causal and non- causal approach, which resulted in the same correlation coefficient value for the two scenarios. This analysis supports the assumption that using the dynamic programming framework for making measurements and not using future observations does not result in a degradation of the clinical information extracted from heart sounds.

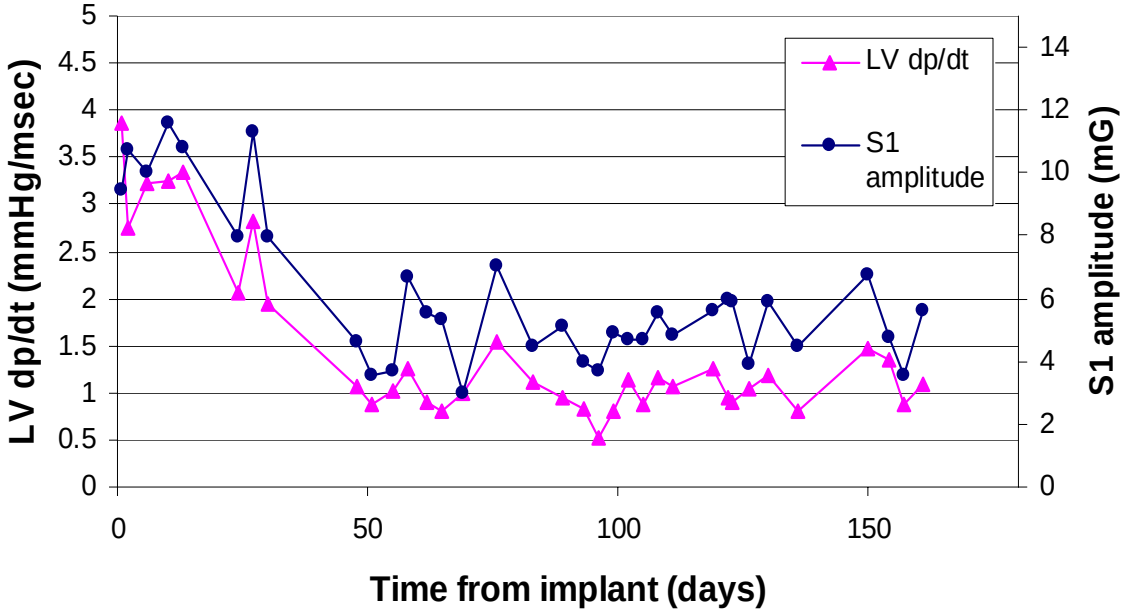


Figure 16: Trends of S1 amplitude and LV dP/dt shown over the duration of data collection.

Animal Model	Metric	Tracking	Non-tracking
Ischemia Model	Correlation coefficient between S1 and dp/dt	0.48	0.32
	Correlation coefficient between S1 and PP	0.51	0.36
	Correlation coefficient between HSET and dp/dt	0.54	0.11 n.s.
	Correlation coefficient between HSET and PP	0.62	-0.19 n.s.
MR-induction	Correlation coefficient between S1 and dp/dt	0.6	0.39
	Correlation coefficient between S1 and PP	0.58	0.32
	Correlation coefficient between HSET and dp/dt	0.52	0.21 n.s.
	Correlation coefficient between HSET and PP	0.47	0.04 n.s.
	Accuracy of S1 to detect time instances with greater than 25% drop in contractility	82.6%	68.5%
Chronic	Correlation coefficient between S1 and dp/dt	0.78	0.34 n.s.
	Accuracy to detect decreases in contractility from baseline	84%	59%

Table 4: Comparison of the hemodynamic correlates of heart sounds measured using the tracking and non- tracking framework for the various experiments. Results that are statistically not significant are labeled n.s.

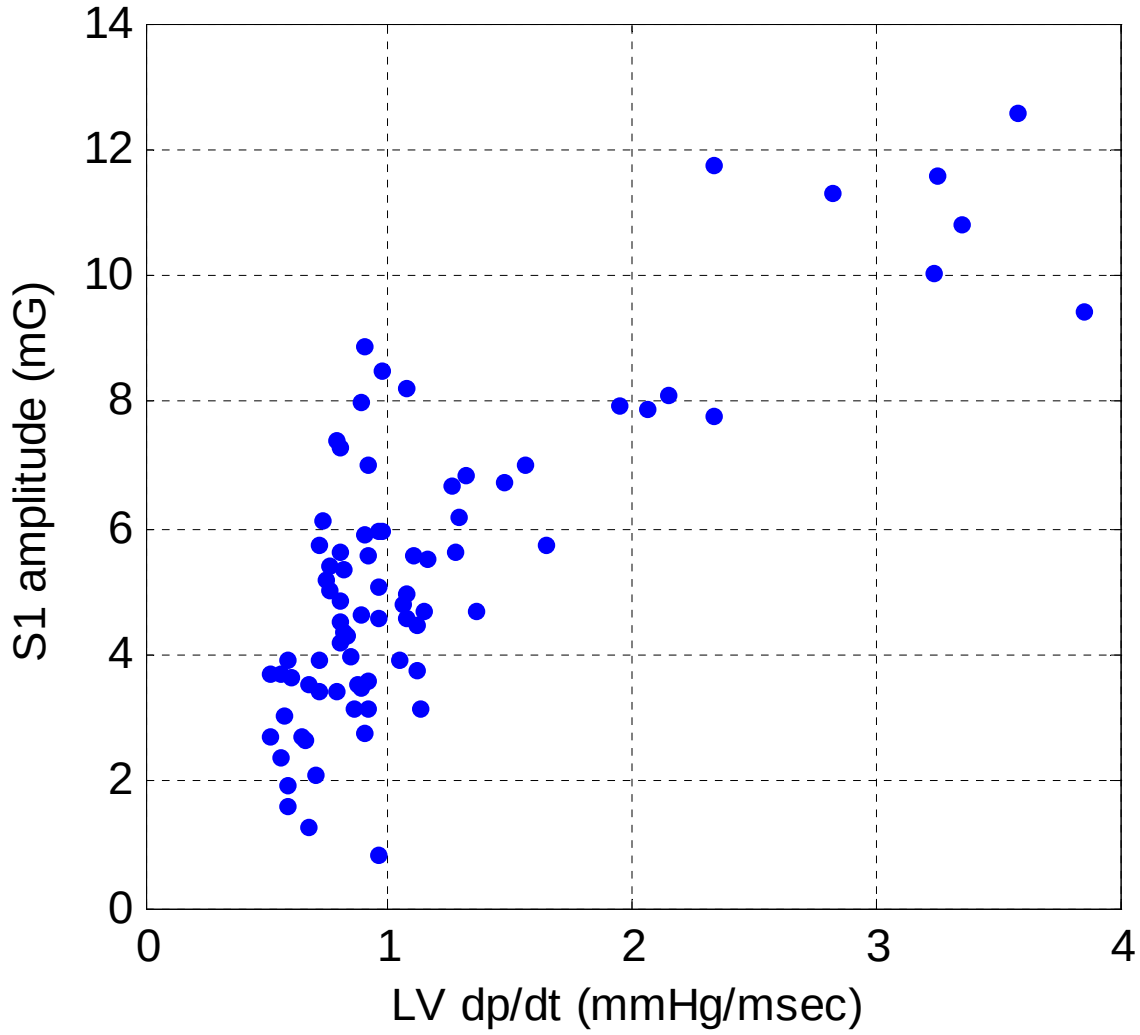


Figure 17: Scatterplot showing the relationship between S1 amplitude and LV dP/dt measured in 3 dogs for an average duration of 4 months.

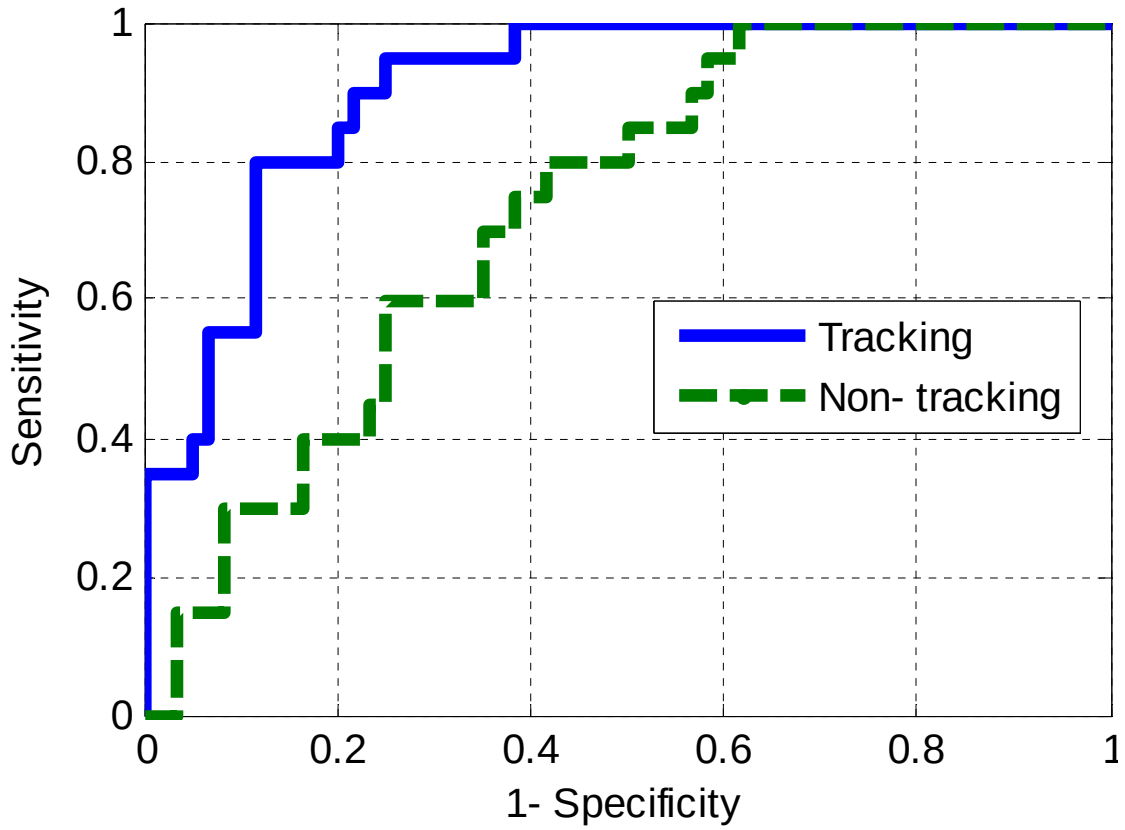


Figure 18: Receiver operating curve for using S1 to detect periods with a decrease in contractility from the baseline.

3.8 Discussion and Conclusions

Historically the four main heart sounds (S1, S2, S3 and S4) have been shown to be clinically useful for diagnosing cardiac conditions during clinic visits but, unlike ECG, have never been used to track patient condition changes over time. In this research a novel framework for long term monitoring of these heart sounds is presented. The duration of tracking is dependent on the application and can vary from a few minutes in

the case of acute event detections (e.g. ischemia) to months in case of chronic monitoring (e.g. heart failure monitoring).

There are multiple noise sources like speech, muscle movement and other physiologic variations that make heart sound monitoring challenging. While traditionally microphones have been used to listen to heart sounds, one can use an accelerometer based sensor to reduce some effects of these noise sources. Also averaging multiple beats reduces these effects. One disadvantage of using an accelerometer vs. a microphone is the high noise level during motion. This may limit the application of such a system to monitoring during resting states and cannot be applied for example during ECG stress tests. The coasting and break lock features, discussed in Section 3.5.4, are used to skip such periods of noise while maintaining the track information. The frequency and duration of these algorithm resets can be changed to trade-off data quality with the number of available measurements. Also, it is important to note that our system requires an external or implanted electrical reference marker (e.g. ECG) for beat parsing and averaging. For example the ECG- Q wave in the acute animal studies and the sensed or paced ventricular pulse in the chronic implanted data are used as the reference marker. This approach, while requiring additional processing, provides the noise immunity needed for ambulatory monitoring. Also, most current implanted devices have a readily available electrical marker measured using implanted leads. Another important note is that the dynamic programming framework used in this research assumes that optimal track up to time *t-1* does not depend on the measurements at time *t*. While the motivation for this causal assumption is realtime

processing and low complexity nature of our application, it is shown in Section 3.7.4 to perform just as well as a non- causal framework.

The core of this research is the development of a dynamic programming based heart sounds tracking algorithm which is tested on acute and chronic pre-clinical data collected during heart failure deterioration. The S1 amplitude and ejection time (HSET) are shown to be statistically significantly correlated with left ventricular contractility measured by dP/dt and aortic pulse pressure. Also, the S1 amplitude and HSET are shown to change with worsening heart failure in all the three animal models. These animal models were chosen to encompass a variety of etiologies for heart failure deterioration. The ischemia model signifies the onset of heart failure during an acute myocardial infarction. The MR induction model attempts to emulate reduction in systolic performance due to increasing filling pressures and left ventricular anatomy changes. It should be remembered that heart sounds are not a surrogate for invasive LV pressure parameters. Instead the goal of this research was to show that the measured heart sounds contain information regarding the cardiac contractility and when measured over long term may provide clinical value. Finally, the chronic animal model with an implanted heart sound sensor is used to validate the applicability of the tracking algorithm for long term monitoring of cardiac mechanical performance. These results, even though they were derived from a limited number of animals, reflect the possible clinical utility of the heart sound tracking algorithm. In all these experiments the utility of the tracking framework is highlighted through comparisons of the performance of non- tracking algorithm to the tracking algorithm. The S1 amplitude measured using non- tracking framework is shown to have poor predictive power compared to S1

amplitude measured using the tracking framework. Also, measurements of timing intervals using the non- tracking framework are shown to have no clinical value due to the lack of statistically significant correlations. These comparisons along with hemodynamic correlates of the measurements made using the tracking framework show the promise of our novel approach for long term monitoring of cardiac conditions.

4 Heart Sounds Ejection Time based Monitoring

4.1 Abstract

Background: Systolic timing intervals (STIs) have been shown to be useful diagnostic tools in multiple cardiac conditions. Traditionally STIs are measured using heart sounds measured at an apical location. Also they have only been used to detect patho-physiologic conditions acutely during clinic visits. These exams are snapshots of the patient's condition and chronic measurements of STIs and their correlation to LV performance have not yet been evaluated. In this study the utility of heart sounds based ejection time (HSET), measured at a non- traditional pectoral location, as an indicator of changing LV systolic performance was evaluated in an acute hospitalized setting. **Methods:** Heart sounds were recorded at pectoral implant locations in patients who were hospitalized for worsening heart failure. Patients were enrolled in the study for a maximum of 72 hours and their stroke volume (SV) measurements were made using a Swan- Ganz catheter. The tracking algorithm described in Chapter 3 was used to analyze the large volume of continuous heart sounds signal data that was recorded for each patient. Computed heart sound based ejection time was compared to the SV measurements recorded in the case report forms. This study enrolled 20 patients who had least three SV measurements during their hospitalization. **Results:** In patients with changes in SV> 10 ml the mean correlation coefficient was 0.6762 ($p<0.001$). To understand the utility of heart sound parameters to detect poor systolic performance, receiver operating curve analysis was performed for each heart sound parameter to detect periods of low stroke volume (SV < 50ml). The HSET had a moderate 70% sensitivity at

80% specificity to detect periods of low stroke volume. As a comparison, CI when used to detect low SV performed with 84% sensitivity at 89% specificity. **Conclusions:** We found that measuring and tracking systolic timing intervals, during the recovery period of a heart failure hospitalization, using heart sounds measured at the non- traditional pectoral region may be clinically useful. This is supported by our findings that the ejection time measured from heart sounds is statistically significantly correlated with SV and can be applied to detect periods of poor systolic performance.

4.2 Introduction

Systolic timing intervals (STIs) are time periods between events of a cardiac cycle; the most popular being ejection time and pre- ejection period. As an indicator of cardiac systolic performance, STIs have been proven to be useful diagnostic tools in multiple cardiac conditions including cardiovascular disease [68], congestive heart failure, valvular disease [69], coronary artery disease [70], ischemia [71] and recently in the optimization of cardiac rhythm management devices [72]. Also, STIs have shown to be correlated to measures of cardiac systolic function such as left ventricular end diastolic volume [73], stroke volume [68], cardiac output [74] and left ventricular ejection fraction [73].

Traditionally STIs have been used to look for patho-physiologic conditions acutely during clinic visits. These exams are snapshots of the patient's condition and chronic measurements of STIs and their correlation to LV performance have not yet been evaluated. Also, traditional STIs have been measured using heart sounds at an

apical location. In this study the utility of heart sounds based systolic timing intervals as measures of changing LV systolic performance and as tools for diagnosis and management of heart failure was evaluated in an acute hospitalized setting. Heart sounds were measured at the pectoral locations which are prospective implant sites for cardiac resynchronization therapy devices. The goal of this study was to understand the relationship between heart sounds measured at implant site locations and invasive hemodynamics measured during the course of recovery in patients hospitalized for worsening heart failure.

4.3 Methods

4.3.1 Subjects and Data

Patients who were hospitalized for heart failure hospitalization with an ejection fraction (EF) < 40% or a myocardial infarction (MI) and were being monitored using a Swan-Ganz catheter were enrolled in this study. They were enrolled in the study for a maximum of 72 hours and there was no required patient follow-up. Exclusion criteria included patients whose primary diagnosis was not heart failure or myocardial infarction; patients with any of the following: chronic obstructive pulmonary disease or chronic lung disease, sepsis, shock, significant septal defects, left-to-right shunts, intra-aortic balloon pumps, end-stage renal disease requiring dialysis or any wound, skin condition or physical barrier that prohibited taping accelerometer sensors to the chest at the appropriate sites for a duration of 48 hours; post cardiac surgery patients; patients whose pulmonary capillary wedge pressure (PCWP) was less than 20 mmHg before enrollment.

After patient consent, the physician (or designee) indicated both the presence and loudness of heart (and lung) sounds on a patient enrollment worksheet. Any other clinical signs of heart failure hospitalization that the patients exhibited were also noted. As shown in Figure 19, an accelerometer was affixed to the chest using surgical adhesive tape. A weighted (80 g) accelerometer shaped approximately as an implantable defibrillator was positioned at the typical implantable cardioverter defibrillator (ICD) left pectoral implant site. However, if patients had implanted devices, this implant site accelerometer was affixed to the right side to avoid potential for mechanical interactions. Additionally three ECG leads were attached to the subject in any acceptable 3-lead configuration. Data were recorded for at least 24 hours, and not more than 72 hours. Once the heart sounds recording had started, periodic Swan-Ganz and cuff pressure measurements were recorded approximately every 2 hours as specified in the study protocol. The data recorder clock time was logged on the patients chart along with the pressure measurements, which enabled later synchronization to the heart sounds data.

4.3.2 Data Processing

The tracking algorithm described in Section 3 was used to analyze the large volume of continuous heart sounds signal data that was recorded for each patient. The HSET was measured as the time interval between the S2 and S1 heart sounds of a given beat. Computed heart sound features were compared to the clinical reference measurements recorded in the case report forms. Since the heart sound measurements were effectively sampled at 1-minute intervals relative to the 2-hour PA pressure measurements, the

heart sounds measurements were averaged in time with a moving average filter to minimize data registration errors and reduce the measurement variation noise.

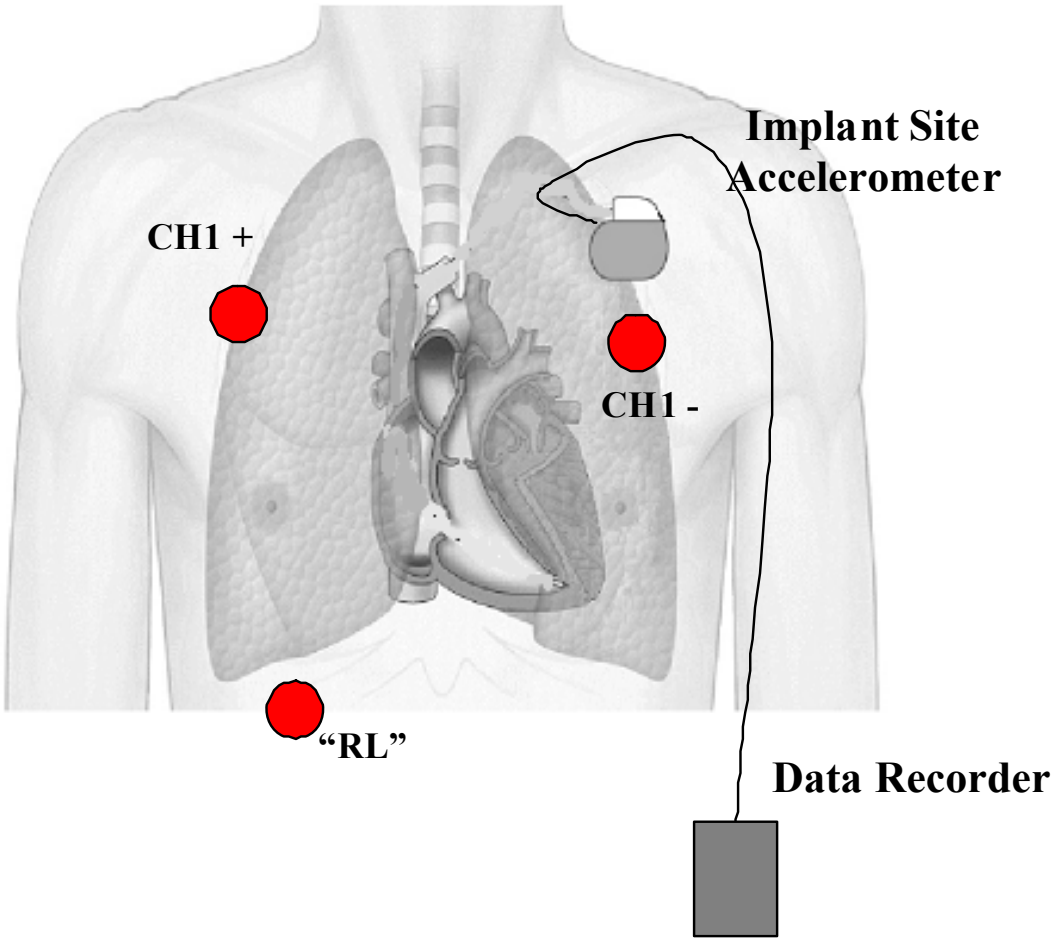


Figure 19: Protocol recommended placement of accelerometer sensor and ECG electrodes.

4.3.3 Statistical Analysis

The heart sound measurements were compared to the clinical reference standard data using statistical tests. Each patient's heart sound dataset was aligned to be synchronous with the clinical variables. The key clinical variable, used as the measure of systolic performance, for this analysis was SV. The SV was preferred over cardiac index (CI) or cardiac output, when available, to look at systolic performance changes independent of heart rate. A 5 minute average of the HSET was calculated around each SV measurement for a Pearson correlation analysis between heart sounds and SV. A correlation coefficient value was calculated for each patient to understand the similarity of HSET to SV. Additionally, an analysis scheme that looks at intra-patient correlations was performed to evaluate the utility of the tracked features across the population. The intent of this analysis was to determine if a detection scheme could be employed without use of trending or change detection to estimate sensitivity and specificity of detecting low SV due to reduced cardiac function.

4.4 Results

4.4.1 Patient Population

This study enrolled 40 patients of which 20 patients had at least three SV measurements. For the purposes of this research this subset of 20 patients was analyzed. Demographical summary is shown in Table 5. The patients enrolled in the study were relatively young (mean age of 52), predominantly New York Heart Association (NYHA) class 4, male (70%), and with a high prevalence of implanted devices (75%)

which accounted for a large number of right-sided implant site accelerometer placements. S3 was audible with a stethoscope in 65%, and accompanied dyspnea in 50% of patients. Also, 75 % of patients presented with edema, with only 3 patients (15%) absent of any overt signs. Most of the patients improved in their systolic performance due to pump improvement during hospitalization.

Demographic	Number (%) or Mean+/- SD
Male	14 (70%)
Age	52 +/- 12
Weight	100 +/- 32
Ischemic etiology	12 (60%)
NYHA <= 3	3 (15)%
NYHA = 4	17 (85%)
Implanted PG	15 (75 %)
Clinical status upon enrollment	
Dyspnea	15 (75 %)
Edema	
Pulmonary	13 (65%)
Peripheral	15 (75%)
Both	10 (50%)
Neither	3 (15%)
First Swan-Ganz measurements	
SV (ml)	51 +/- 17
Cardiac Index ($l/m^2/min$)	1.97 +/- .73
Last Swan-Ganz measurements	
SV (ml)	57 +/- 22
CI ($l/m^2/min$)	2.44 +/- .72

Table 5: Demographics for the 20 patients analyzed in this research

4.4.2 Example Patient Data

To illustrate the data analysis methods, a case study for a single patient first is presented. This patient had an implanted device on the left side, necessitating placement of weighted heart sounds sensor on the right pectoral region. The patient was a NYHA class III, ischemic, and upon enrollment presented with and no overt signs of peripheral or pulmonary edema at the initial physical exam. The first and last hemodynamic measurements (from the case report forms) are shown in Table 6.

	First Measurement	**Last Measurement**
Systolic BP (mmHg)	118	100
Diastolic BP (mmHg)	80	61
Pulmonary Artery Systolic Pressure (PAS) (mmHg)	68	43
Pulmonary Artery Diastolic Pressure (PAD) mmHg)	30	15
Pulmonary Capillary Wedge Pressure (PCWP) (mmHg)	32	10
SV (ml)	66	123
CI (l/min/m2)	1.43	3.12

Table 6: The first and the last hemodynamic measurements recorded

It can be concluded that the patient entered the study with poor systolic function and exited with improved perfusion. In order to visualize the relative timing of various heart sound components, with respect to hemodynamic events, an echocardiogram recording is shown in Figure 20.

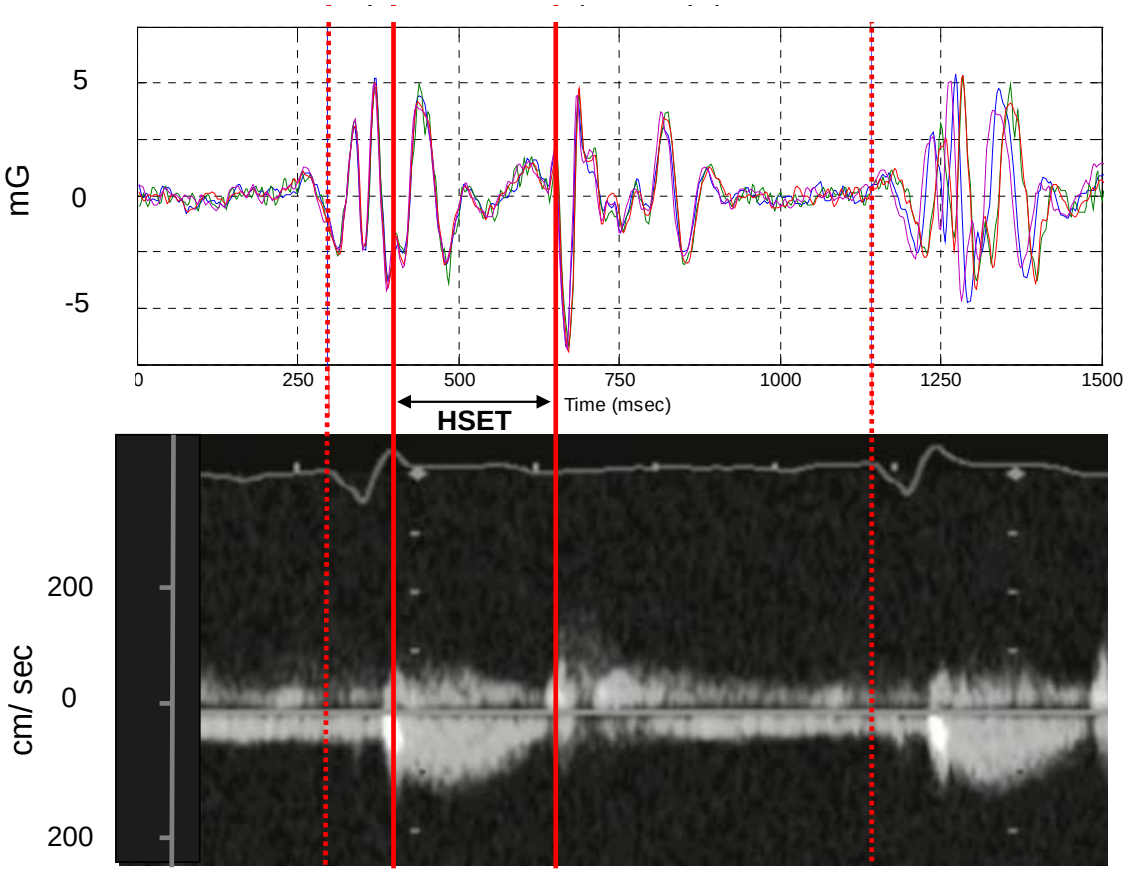


Figure 20: Heart sounds waveforms shown temporally aligned with aortic outflow echocardiogram. The bars show alignment with ECG R wave markers.

The echo is temporally aligned; the vertical bars were used as manual alignment references and signify the ECG- Q wave locations. The echocardiogram shows the velocity of the blood ejecting through the aortic valve during systole. Five averaged heart sounds waveforms (1 minute averages) collected during the echocardiogram recording are also shown.

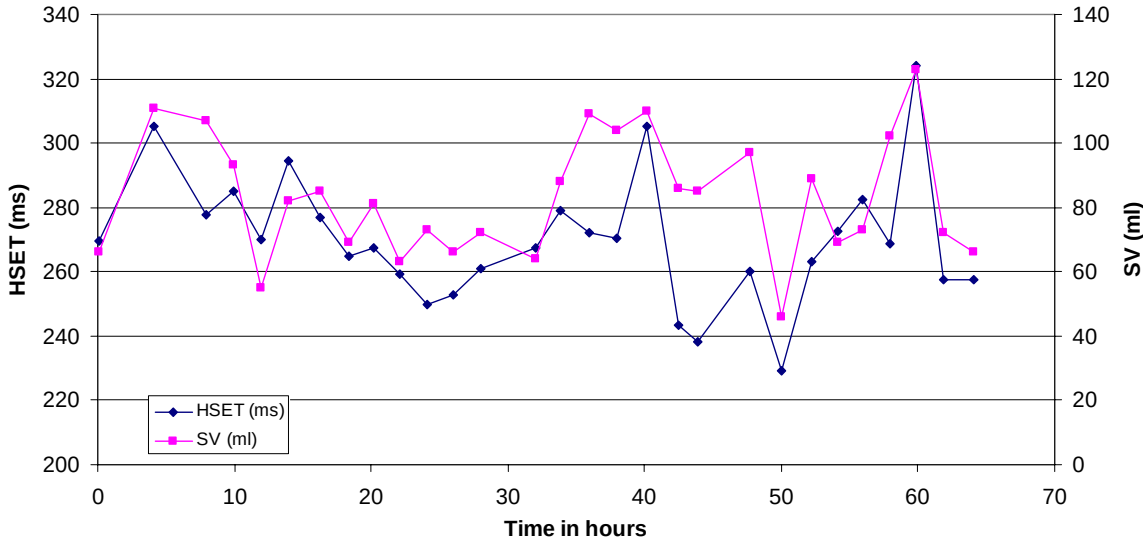


Figure 21: SV and HSET trends in an example dataset

As expected the S1 and S2 heart sounds occur during the start and end of ejection. Also other heart sounds are shown. Note that the amplitude scale on the heart sounds signal is in mG because an accelerometer is used as the sensor. Since the goal of this study was to understand the correlation of heart sound components to hemodynamics, SV was measured every 2 hours and recorded on the case report forms along with other Swan-Ganz and standard measurement. Figure 21 shows this measured SV over a course of 64

hours. As described in Section 4.3.2, the raw heart sounds were first processed to extract clinically useful parameters. Based on physiologic understanding of hemodynamics and heart sounds genesis it was hypothesized that the HSET would be correlated with SV. This relationship is shown in Figure 21, where the trends in SV are tracked by the HSET. The HSET also has a correlation coefficient of R = 0.65 ($p<0.05$) with SV. This suggests that the HSET can be used for the estimation of a patient's systolic performance. This hypothesis is tested across the population in the following sections.

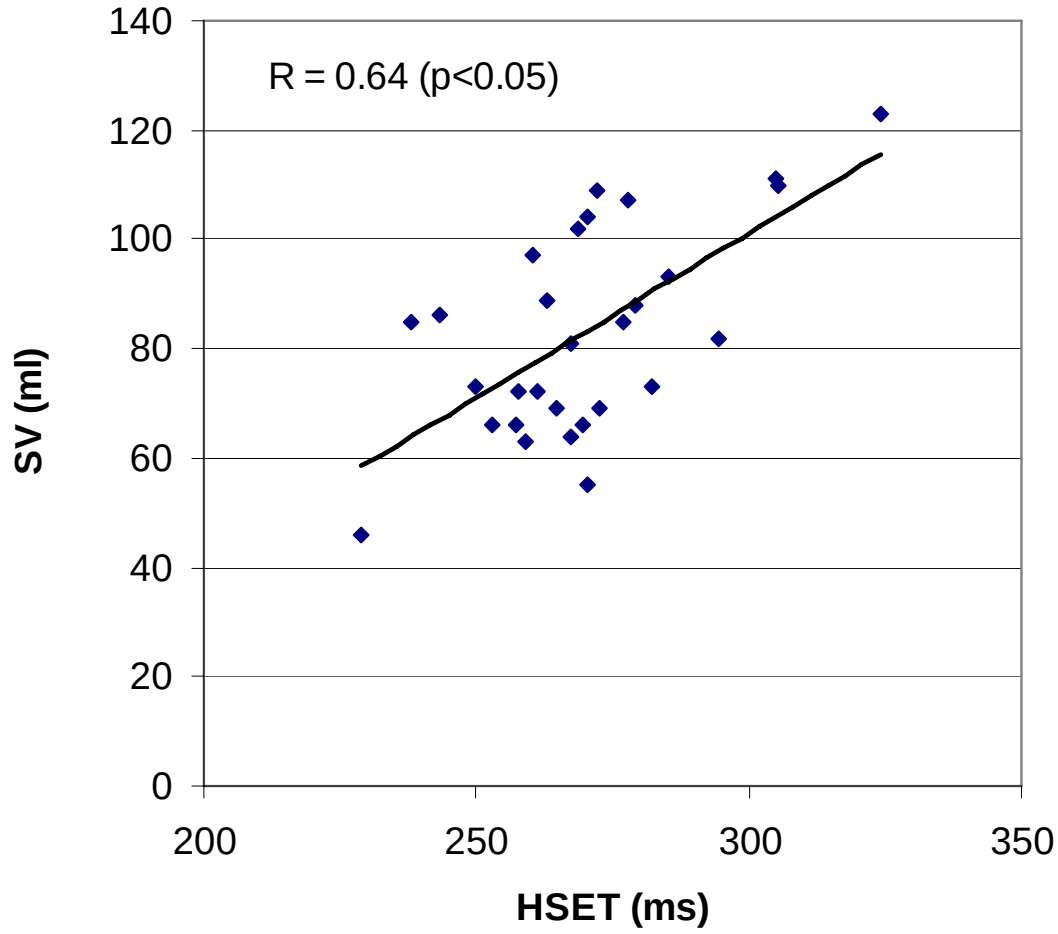


Figure 22: Scatterplot between HSET and SV in an example patient

4.4.3 Intra- Patient Trend Correlation to Stroke Volume

The example patient analysis described in the previous section was performed on all the 20 patients to understand the extent of correlation of heart sounds parameters to SV. During our analysis we found that there were four patients who did not have significant correlations between SV and CI or SV and HSET. The case report form data from these patients were further analyzed to understand the lack of correlations. There were no relationships between clinical data (e.g. medications, etiology etc) in these patients and their poor correlations. The only objective relationship found was that the changes in SV, in these patients, were less than 10 ml. Also since the CI is expected to be highly correlated with SV, this suggested that there was significant measurement noise in the Swan- Ganz technique. In the patients with changes in SV> 10 ml the mean correlation coefficient between SV and HSET was 0.6762. This sub-group statistic is reported since it is expected that with small changes in SV the changes in HSET will be within the noise floor.

4.4.4 Population Based SV Correlates

While the previous sections considered trends for each patient separately, this section examines the measurements from all patients aggregated into a single reference frame. The goal is to evaluate the potential of acute heart sounds measurements to detect adverse hemodynamic conditions such as low cardiac output without chronic trending. This capability could have a great potential in risk stratification and for assessment of overall disease status or severity relative to population. Figure 24 shows the scatter plots

between HSET and SV with data from all patients. In this analysis the four patients who had less than 10 ml change in SV are not excluded because as a part of the overall population they were not found to be any different in their hospitalization condition (e.g. drugs administered, discharge time etc). The HSET has a statistically significant correlation coefficient of $R = 0.51$ ($p < 0.05$) across patients. The scatter between CI and SV is also provided for relative comparison of the two measures.

To understand the utility of heart sound parameters for the detection of poor systolic performance, receiver operating curve analysis is performed for each heart sound parameter to detect periods of low stroke volume (SV < 50ml). The HSET had a moderate sensitivity of 70% at 80% specificity to detect periods of low stroke volume. As a comparison, CI when used to detect low SV performs with 84% sensitivity at 89% specificity.

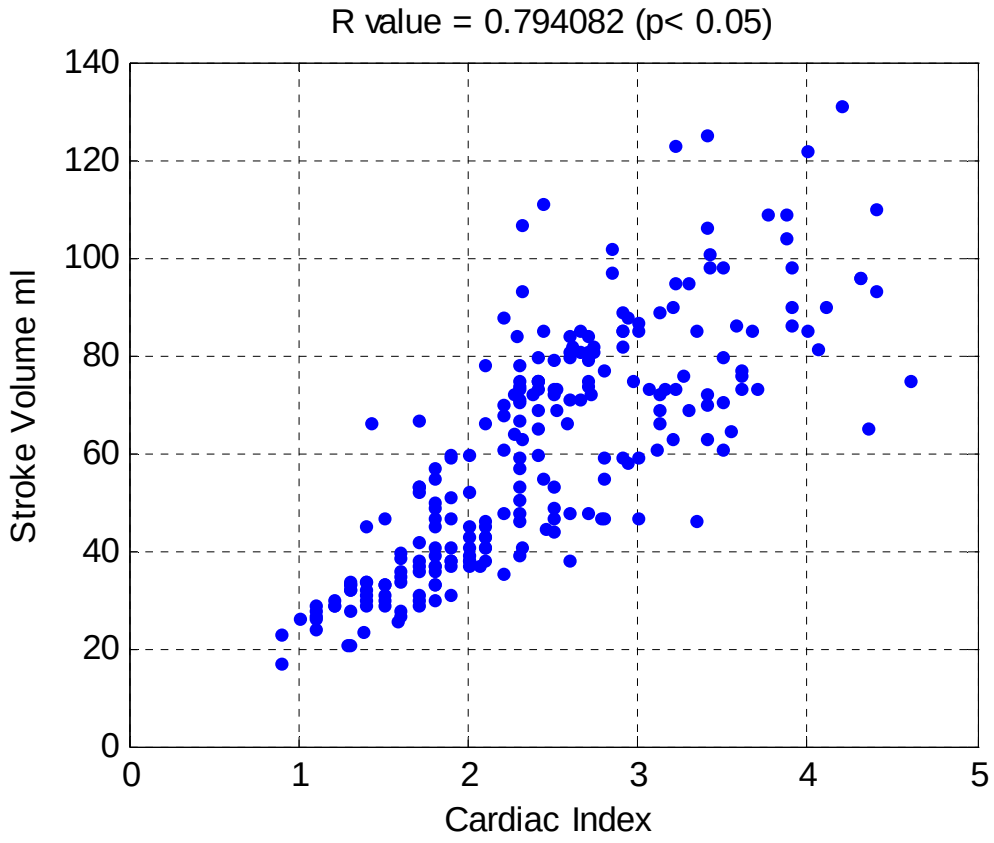


Figure 23: Scatter plots between CI and SV in 20 patients.

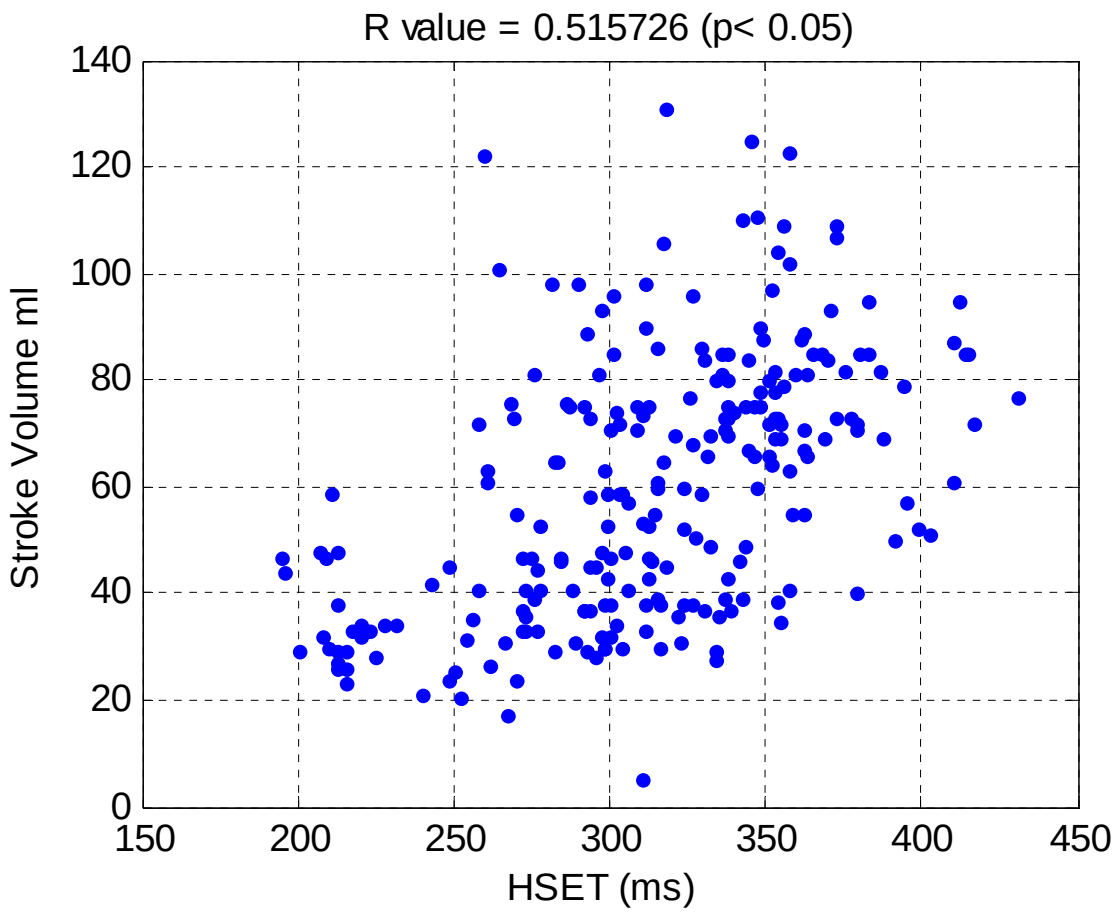


Figure 24: Scatter plots between HSET and SV in 20 patients.

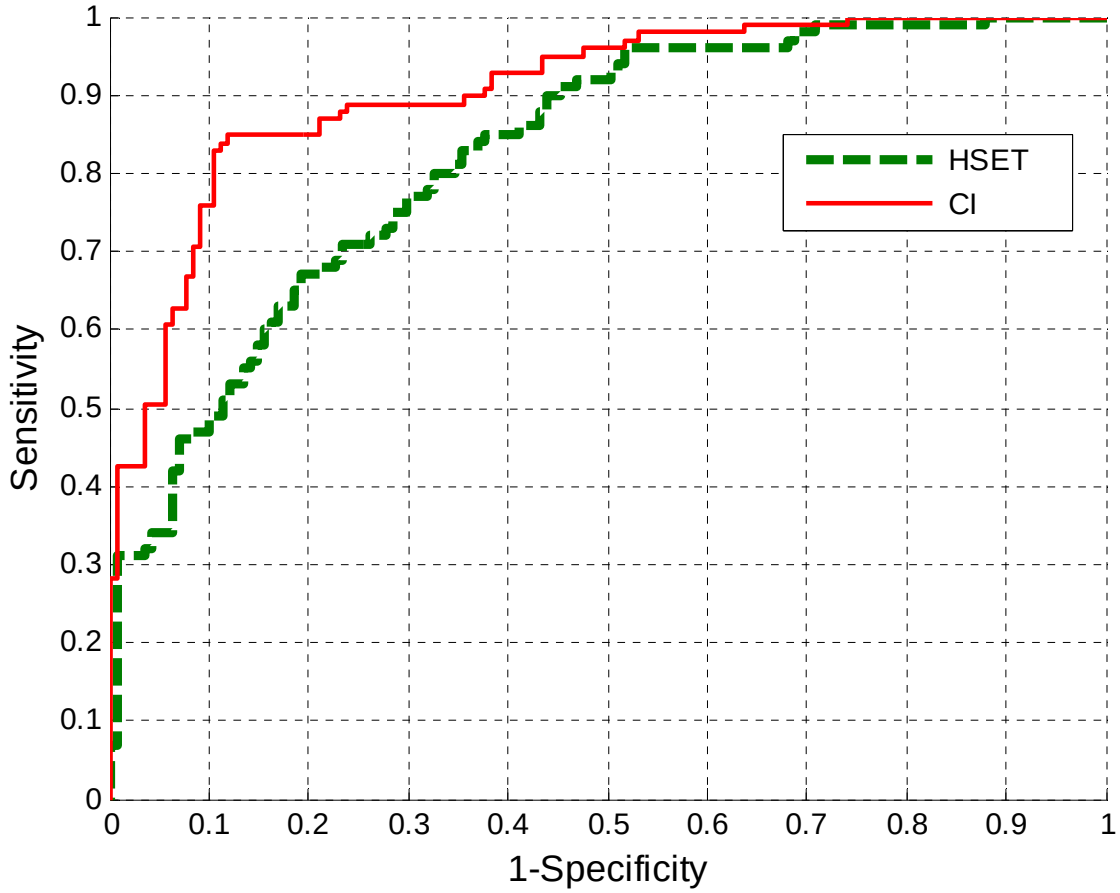


Figure 25: ROC curves to detect SV < 50ml.

4.5 Discussion and Conclusion

The goal of this analysis was to understand the value of measuring systolic timing intervals over long term in heart failure patients. Historically, STIs and other heart sound parameters have been measured during clinic visits which are snapshots of the patient's cardiac condition. While STIs have been shown to be correlated to LV systolic performance, longitudinal measurements of STIs and their chronic relationship to hemodynamics has not been evaluated before. In this study heart sounds were measured

for over 48 hours in hospitalized heart failure patients with simultaneous Swan-Ganz measurements. This study showed that the ejection time measured from heart sounds is significantly correlated with SV.

Traditionally systolic timing intervals have been normalized for improved hemodynamic correlation [20] using the ratio of LVET and PEP which has been shown to be significantly correlated to SV. In our analysis such ratios which employ normalization by the pre-ejection period or its equivalent R-S1 interval only serve to worsen the correlation, likely due to the high variability observed in this surrogate. Note that trending of HSET may obviate the need for normalization when only changes in SV are desired. Absolute estimates of SV will require more precise estimates of PEP and HSET which are not easily obtained from heart sounds and ECG alone [74]. Evaluation of systolic timing intervals to detect low SV was more promising since timing intervals are robust in the presence of posture and variable transmission effects. The HSET had a sensitivity of 70% and a specificity of 80% for a SV threshold of 50 ml.

In Chapter 3 and Chapter 5 we have shown that the S1 amplitude, when measured using the tracking algorithm also contains information regarding cardiac hemodynamics. In this clinical study the S1 amplitude did not perform similarly. When tested using the intra-patient correlation analysis the S1 amplitude was statistically significantly correlated with SV in only 6 of 20 patients. Also, the population analysis resulted in a statistically non- significant correlation between S1 amplitude and SV indicating that the S1 amplitude had poor predictive power to detect periods of low SV. We hypothesize that the poor/ inconsistent coupling between the accelerometer sensor and chest during the 48 hour

recording along with posture related noise was responsible for these findings. Additionally, the variable sensor placement (right vs. left), due to a large percentage patients enrolled with devices, may have an impact on the transmission properties of heart sound amplitudes. Since the ejection time is a measure of time duration, in contrast to an amplitude measure, it may be less sensitive to coupling and transmission related noise. Also, these coupling and variable transmission effects will be reduced when an implanted device is used for measurement of heart sounds.

Since a majority of patients had implanted PG devices, this required the placement of the external weighted accelerometer on the right side. Although the large presence of implanted devices was beneficial from a heart failure etiology standpoint, suboptimal placement of the ICD implant site accelerometer on the right side may have served to underestimate the performance. The impact of this sensor placement will be studied in future research.

This study was designed to understand correlations of HS to hemodynamics during the recovery from hospitalization. It was anticipated that there would be a steady increase in cardiac performance after admission into the hospital. This was rarely the case in the data collected; instead many small fluctuations in the data were found. These small fluctuations, often influenced by intravenous medications, may not be a representative of longer term shifts in stroke volume that evolve days and weeks before a heart failure hospitalization event.

In conclusion, this study showed that the heart sounds based ejection time measurement was correlated to change in stroke volume during recovery from heart failure hospitalization. While this study looked at this relationship during the recovery, further research is required to evaluate the utility of STIs for general heart failure monitoring.

5 OSA Detection using Heart Sound Monitoring

5.1 Abstract

In this chapter we develop a novel accurate and practical heart sounds based framework for detecting periods of obstructive sleep apnea (OSA) in heart failure patients with pacemakers and cardiac resynchronization therapy devices. S1 heart sounds are measured using the dynamic programming based tracking algorithm which has linear complexity with the monitoring duration. Information in the S1 measurements is then optimally extracted using wavelet decomposition and an adaptive dyadic time segmentation to generate features for apnea detection. As a comparison to the tracking approach, the traditional non- tracking algorithm is also used for feature generation. Two classifiers: linear discriminant analysis (LDA) and support vector machine (SVM) are trained using different feature selection schemes and tested on an independent test dataset. We found that the tracking based classification consistently outperformed non-tracking based classification, emphasizing the importance of tracking. SVM with recursive feature elimination scheme and tracking had the highest (91.8%) accuracy yielding an improvement of 7% over the non-tracking based approach. Finally, the output of the best classifier was used as an OSA severity score and was found to correlate significantly ($r^2 = 0.52$, $p<0.05$) with the gold standard apnea hypopnea index.

5.2 Introduction

OSA is the most prevalent form of sleep disordered breathing where the airways repeatedly close and breathing stops. Also, it has been shown that OSA is a potential

risk factor for heart failure patients [11, 12] with rhythm control devices like pacemakers and cardiac resynchronization therapy devices. In patients with such devices, the prevalence of apnea is reported to be much higher than in the general population, with about 50-60% in heart failure patients and over 30% in pacemaker patients [13, 14]. If undiagnosed and untreated, OSA can worsen heart failure eventually leading to hospitalization [12].

In the past several researchers have attempted to solve the problem of detecting OSA non- intrusively using ECG based techniques [17]. While ECG based approaches have been popular and have about 85-92.6% accuracy [17], these methods fail in heart failure patients with implanted cardiac therapy devices which adjust the heart rate automatically. To address these concerns the tracking framework for measuring heart sounds is used to develop an improved hemodynamics based approach for the detection of OSA in patients with rhythm control devices. The S1 heart sound is first consistently measured over the sleep duration in 17 patients undergoing a sleep study. Wavelet decomposition and time segmentation is used to generate features for the LDA and SVM classifiers which are then optimized using feature selection schemes and used to detect periods of OSA. Finally, an algorithm based severity score is developed to determine the intensity of OSA for each patient during the sleep study and shown to significantly correlate with the gold standard apnea hypopnea index (AHI).

5.3 Physiology

A cardiac cycle consists of precisely timed electrical and mechanical events that result in rhythmic atrial and ventricular contractions that pump the blood into the pulmonary

and systemic circulation respectively. The forced closure against the mitral valve and the sudden rush of blood through the aorta causes the first heart sound (S1) [10]. The S1 amplitude has been shown to be correlated with the intensity of ventricular contraction measured by the maximum change in left ventricular pressure at the beginning of systole (maximum LV dP/dt) [37, 38]. The S1 is an audible sound with most of its energies in the 15 Hz to 100 Hz bandwidth [39, 66, 67]. During OSA there is a marked decrease in intrathoracic pressure which results in an increasingly stronger contraction of the left ventricle [15]. As a result the S1 heart sound generated during the OSA period has a crescendo like change in amplitudes as shown in the shaded regions of Figure 26. On the other hand, during periods of normal breathing the left ventricular pressure is modulated by respiration at the same frequency as the respiratory drive. Hence, during normal breathing the S1 amplitude is modulated at the respiration frequency. This is also reflected in the figure during periods of respiratory flow.

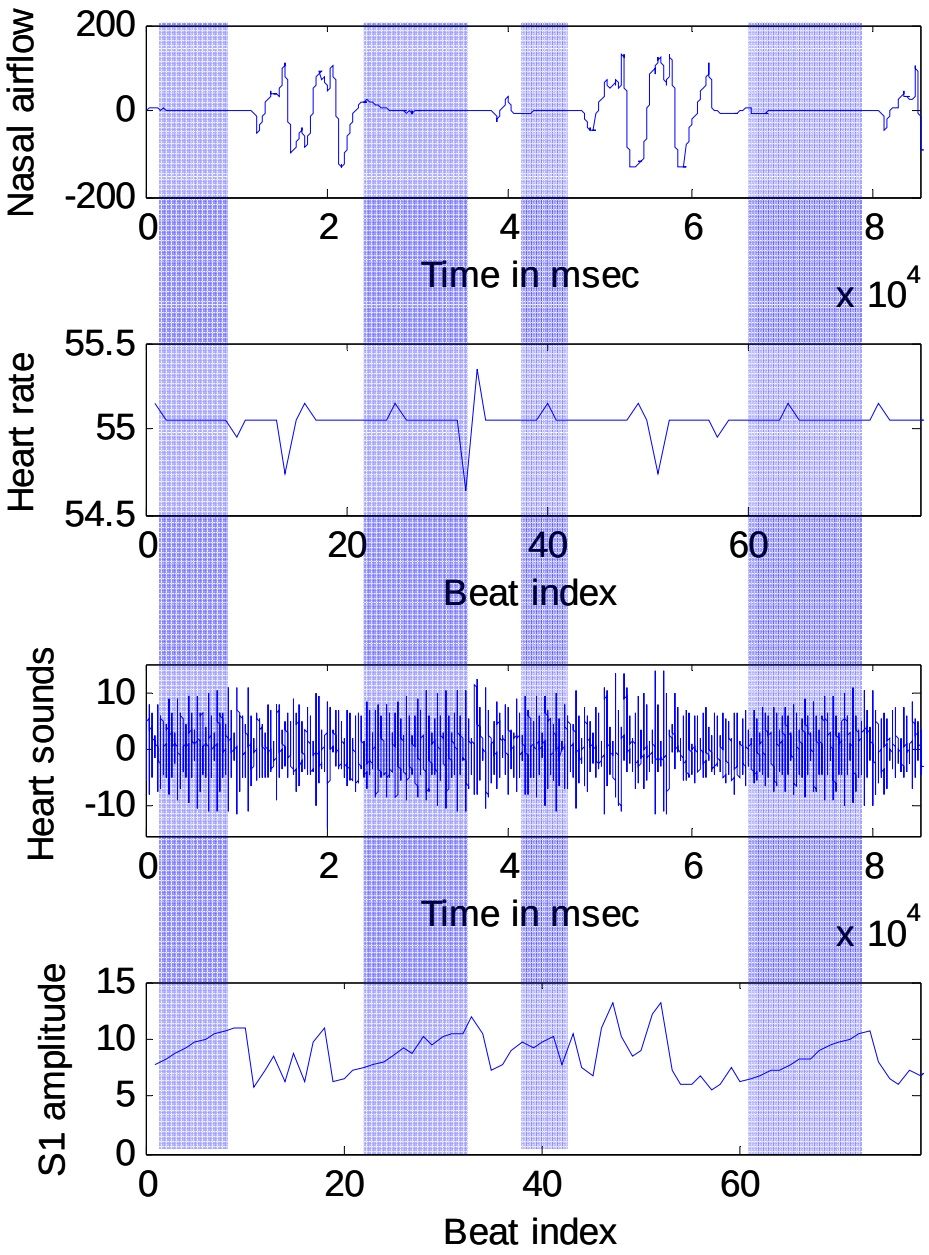


Figure 26: Panel 1 shows periods of OSA with loss of nasal airflow. Periods of apnea are shaded. Panel 2 shows the heart rate kept relatively constant due to pacing. Panel 3 shows the heart sounds signal with a crescendo like increase in amplitudes during apnea. Panel 4 shows the increasing S1 heart sound due to reduced intrathoracic pressures.

5.3.1 Data

The data used in this research was from seventeen patients who were prescribed polysomnography tests. Demographics for each of the 17 patients are shown in Table 7. The mean age of the patients was 63 years with 12/ 17 being male. Most patients had some form of heart disease. Many were overweight (mean BMI = 29.3) with high apnea hypopnea index (mean AHI = 54.5). OSA was defined as an absence of airflow for more than 10 seconds. We parsed data from each of the patients into two minute epochs for analysis. Each two minute epoch is composed of ECG, heart sounds, nasal flow and the annotation from the trained clinician who was blinded to the goals of this research. Annotations include the number of OSA events, central sleep apnea (CSA) events and hypopnea events. Also based on these, the AHI for each sleep test is calculated. The ECG and heart sounds are sampled at 1000 Hz; nasal flow at 25 Hz. In total 600 two minute epochs of OSA and 600 two minute epochs of normal breathing were extracted from the data. Data from half the patients was randomly selected for training (parameter optimization and feature selection) and the other half for testing.

ID	Age	Sex	BMI	AHI	Cardiac history	EKG info
01	63	M	28.4	-	CAD	SSS
02	72	M	24.8	-	Ablation, CAD	SSS- brady-tachy
03	63	M	27.5	40.7	Ischemic, Hypertension	2° heart block Möbitz
04	63	M	25.1	17.9	CAD, Hypertension	2° heart block Möbitz
05	49	F	37.9	17.8	Hypertension	SSS-SA arrest
06	61	F	33.8	27.0	Hypertension	2° heart block Möbitz
07	58	F	33.7	42.7	Unspecified	2° heart block Möbitz
08	65	M	34.6	48.8	CAD	2° heart block
09	65	M	32.7	84.6	CAD	SSS
10	63	M	28.4	85.2	Unspecified	SSS
11	56	F	24.2	32.7	Unspecified	2° heart block Möbitz
12	63	M	27.2	56.9	CAD	SSS
13	72	M	28.4	73.6	CAD	SSS
14	55	M	35.4	64.6	Valvular	SSS
15	61	F	22.1	96.4	Surgical	SSS
16	70	M	27.7	98.1	Hypertension	2° heart block Möbitz
17	68	M	26.1	30.3	Ablation, Hypertension	Complete Heart Block
Mean	63		29.29	54.5		

Table 7: Patient demographics. SSS: Sick Sinus Syndrome

5.4 Methods

Figure 27 shows a block diagram of the steps involved in extracting an apnea severity score from the two minute epochs. The heart sounds in each two minute epoch are first used to measure the S1 amplitude for each epoch. This is done using the dynamic programming based tracking algorithm, described in Section 5.4.1, for improved consistency and accuracy of the measurement. Following the S1 measurements, wavelet decomposition is used to generate features that are predictive of OSA in Section 5.4.2. Features generated, using the wavelet decomposition, are used as inputs to LDA and SVM classifiers which are described in Section 5.4.3. Features that have the most predictive content are sub- selected using filter and wrapper based techniques described in Section 5.4.4.

5.4.1 S1 Measurement

The dynamic programming tracking algorithm, as described in detail in [75] and Section 3.4, uses the peak detections of the S1 heart sound across multiple cardiac cycles to detect the largest peak consistently for the duration of the two minute epoch. This tracking is important because the S1 heart sound has multiple sub-components which need to be consistently measured for improved accuracy. Heart sounds are parsed into individual beats with the ECG reference markers detected using [76]. These parsed beats are band pass filtered with 15 Hz and 90 Hz filter cutoffs. A window is setup around the S1 waveform and candidate peaks are measured. The candidate peaks over multiple cardiac cycles are used to detect the largest most consistent peak using the dynamic programming tracker.

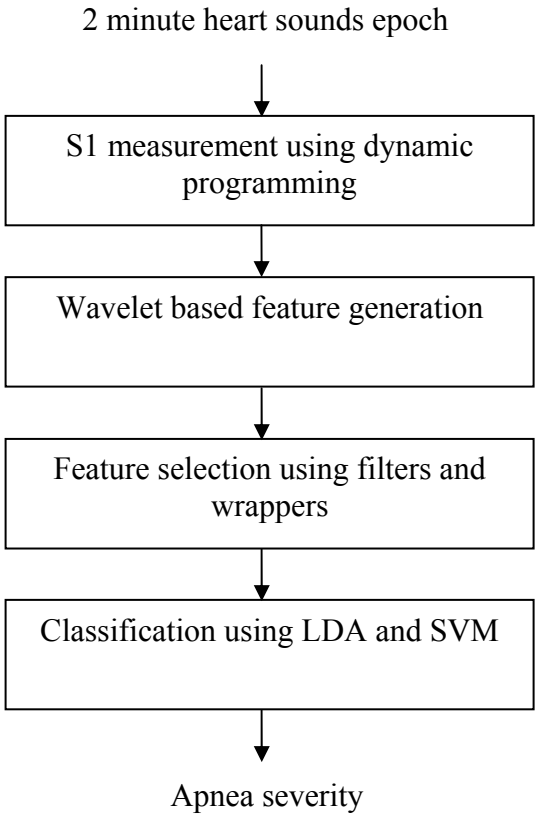


Figure 27: Various steps involved in extracting the apnea severity score from each 2 minute epoch of recorded heart sounds.

5.4.2 Wavelet Based Feature Generation

The S1 amplitude measured using the tracking algorithm is next used to generate features using a wavelet transform based sub-band decomposition followed by time segmentation. We use this approach to extract predictive features because the S1 amplitude modulation due to apnea is localized in both frequency and time within each epoch. As described in Section 5.3 the dominant frequency component in the S1 heart

sound amplitude during normal breathing is due to the modulation caused by changes in left ventricular preload. This modulation due to the respiratory drive is also present during periods of OSA. But, this respiratory modulation is not the dominant component since OSA causes the S1 heart sound to increase in a crescendo like signal. This modulation of the S1 amplitude was shown in Figure 26. To generate predictive features, we decompose the mean subtracted S1 amplitude measured during each epoch into frequency sub- bands using dyadic stationary wavelet transform (SWT). The SWT eliminates the down-sampling operation and the final transform is shift invariant which is important in our machine learning framework. We use a 7 level SWT decomposition using a 'db5' wavelet. Sub-bands of level 1 through 3 are found to correspond to decompositions with respiratory information (greater than 5 breaths/ minute) and sub-bands from levels 4 through 7 are found to contain information about apnea events (less than 5 breaths/ minute). These findings are consistent with the physiology of OSA generation and its modulation of the cardiac preload [15]. Figure 28 shows the SWT decomposition (Level 1), of the mean subtracted S1 amplitude measured during the apnea epoch of Figure 28. The approximation signal captures the crescendo like increases in S1 amplitude during an apnea event. Since the detail signal corresponds to frequencies greater than 0.33 Hz, it has information about the respiratory drive. Periods of apnea in the detail signal have smaller amplitudes because of the loss of breathing during those times.

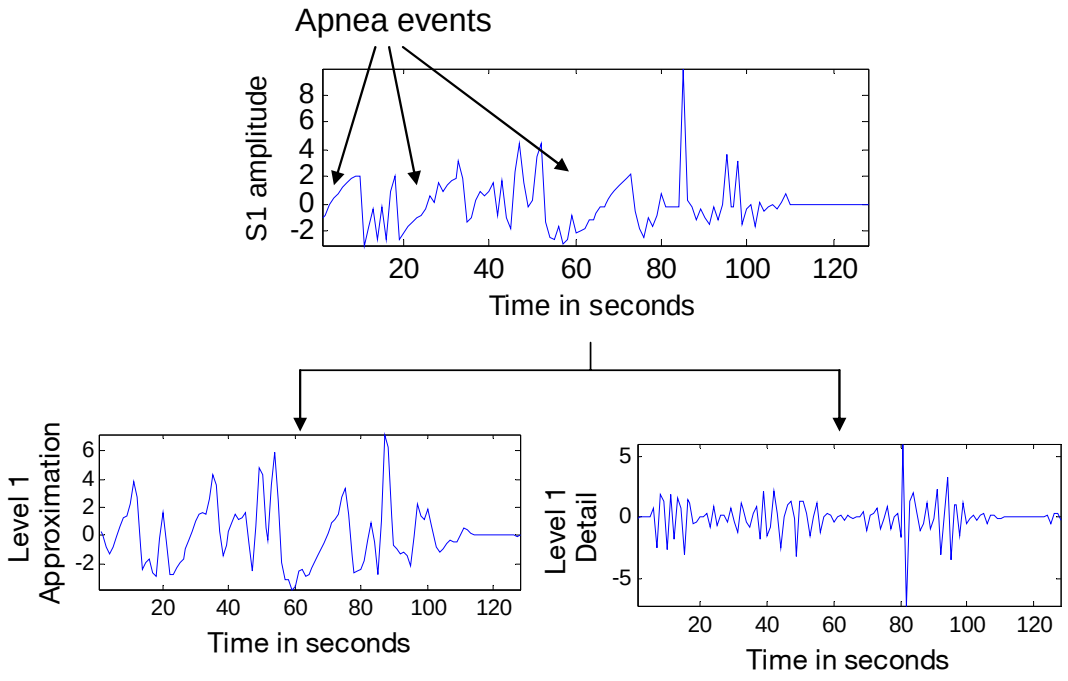


Figure 28: Level 1 SWT decomposition of the example apnea epoch.

Since apnea events are localized for a short duration (10 seconds to a minute) within each epoch; we use a second tree structure by adaptively segmenting in time the output of each sub-band filter to extract the time varying characteristics of these signals. We use a 3 level dyadic time segmentation to decompose each sub- band in time. This corresponds to using time segments of 128 (initial length), 64, 32 and 16 seconds at each level respectively. An example of this time segmentation of the Level 1 detail is shown in Figure 29. Energies in each of these time segments are used to generate predictive features. Since apneas will be localized within one or a combination of these time segments, the mean and max energies in time segments of the same length (for example the 8 segments with 16 seconds duration) are used in the feature vector.

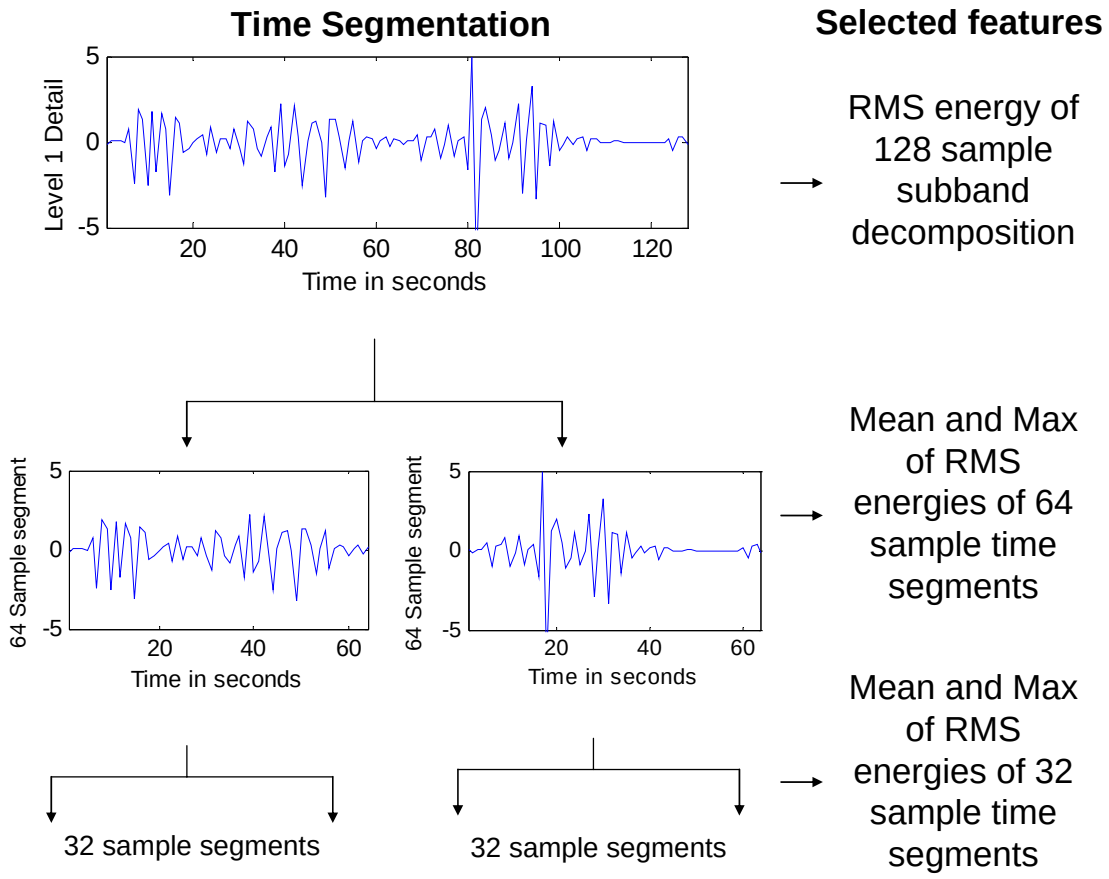


Figure 29: Time segmentation and feature generation from the Level 1 Detail obtained using an SWT.

Hence, a mean and max energy across the different time segments is obtained. Along with the total energy in the initial epoch, the feature vector for each sub-band now has 7 energy values corresponding to the different time segments. We use these energy values for all the 14 sub-bands to obtain a feature vector with 98 features for the LDA and SVM classifiers

5.4.3 Classification

S1 heart sound measurements from each two minute epoch are used to calculate the RMS energies in the various time segments of the wavelet sub-bands. These form the features $\boldsymbol{r}_k$, corresponding to that epoch, which are used for classification. Consider a set of m such datasets ($\boldsymbol{r}_k$, y_k) (k=1, 2….m), where each feature vector $\boldsymbol{r}_k$ has p = 98 input features and the outcome is $y_k \in$ (-1, 1) where +1 corresponds to OSA and -1 corresponds to normal breathing. A linear hyperplane separating both the classes is defined by $f(\boldsymbol{r}) = \boldsymbol{w}\cdot\boldsymbol{r} + w_0$ where $\boldsymbol{w}$ defines the linear boundary. We use two classifiers with subsequent feature selection schemes (described in Section **Error! Reference source not found.**.5.4.4). Since our goal is to compare the tracking based measurements to a non- tracking approach, the two classifiers are used to understand the relative performance gain due to the classification schemes. We use the commonly used LDA classifier because of its ease of implementation and popularity. Also, we used the SVM classifier because of its improved performance due to optimization of the margin and wrapper based feature selection schemes.

Linear Discriminant Analysis: An LDA is a method to identify the best discriminating hyperplane in a p dimensional space. The weight vector w used in LDA is calculated as

$$\boldsymbol{w} = (\Sigma_{Normal} + \Sigma_{OSA})^{-1}(\mu_{Normal} - \mu_{OSA})$$

where, Σ and μ are the covariance matrix and mean of the two classes respectively. A detailed description is provided in [53].

Support Vector Machine: The SVM classifier is defined by $y_k \cdot (\boldsymbol{w} \cdot \boldsymbol{x}_k + w_0) \geq 1 - \xi_k$ subject to the minimization

$$C\sum_{k=1}^{m} \xi_k^2 + \frac{1}{2}\|\boldsymbol{w}\|^2,$$

where C is the user defined parameter/penalty term that defines the tradeoff between the model complexity and number of separable datasets. The C that provides the maximum accuracy during the five fold cross validation is chosen. The margin is defined by $\gamma = 1/\|\mathbf{w}\|^2$ and the optimal estimation of $\boldsymbol{w}$ and w_0 is a quadratic programming problem whose solution can be found using Lagrange multipliers as shown in [52]. We use this margin to select optimal features in the wrapper based scheme which is described in Section 5.4.4. Also, while there are many kernel choices available, we use a linear kernel as it is found to have the best performance during the five fold cross validation.

5.4.4 Feature Selection

Feature selection is used to enhance the generalization performance of the LDA and SVM classifiers by eliminating irrelevant variables and hence reducing complexity. Traditionally, feature selection schemes have been classified into two categories: filter methods and wrapper methods. Filters select the best subset of features using predefined evaluation criteria that are independent of the classifier. In this research we use Fisher's criterion, which is a popular filter method as the low complexity filter method for both the LDA and SVM classifier. On the other hand, wrapper methods use information from the SVM classifier to select the best subset of features. The wrapper based feature selection criteria employed in this research is based on the popular SVM- recursive feature

elimination technique described in [77]. Since this recursive technique can be computationally intensive, we also use a simple SVM weight based wrapper technique as a comparison.

Fisher Criterion (FC): The FC estimates the correlation of each feature to the two classes, OSA and normal breathing. For each feature the following *FC* score is calculated, where μ and σ are the mean and standard deviation of the feature it belongs to.

$$FC = \frac{\mu_{OSA} - \mu_{Normal}}{\sigma_{OSA}^2 - \sigma_{Normal}^2}$$

Features are selected with the largest Fisher criterion score, which corresponds to the ones with the most discrimination power.

Recursive Feature Elimination (RFE): The RFE technique, first proposed in [77], is based on the concept of maximizing the SVM margin. For each feature, its importance is measured by the contribution it has on the margin of the trained SVM. In the RFE method the feature with the least influence on the margin of the trained SVM is eliminated based on backward feature elimination technique. A ranking criterion based on the value of $(w_i)^2$ is used to determine the feature that is eliminated.

Weight Threshold (WT) Method: As a comparison to the RFE technique the weight based feature selection used in [56] is also applied. In this technique features are ranked based on their importance, which is inferred from the absolute value of their weights. The criterion is to choose features whose weights are greater than an adjustable lower threshold.

5.4.5 Results

The tracking algorithm described in Section 5.4.1 is used to measure the S1 amplitude for each epoch. As a comparison the non- tracking approach of taking the maximum amplitude in a predefined S1 window [56] is also used to generate non-tracking S1 amplitude. These two measurements are next used to independently train the LDA and SVM based classifiers for OSA detection. The results for training using cross validation and selection of the best features are described next.

5.4.6 Training and feature selection

For the SVM we select C=100 parameter value and a linear kernel using five fold cross validation without any feature selection. The two classifiers and three feature selection techniques are next compared, using five fold cross validation, with and without the tracking framework. Figure 30 shows the accuracy with different number of features used for the eight scenarios. Accuracy is defined as the mean of sensitivity and specificity. Based on cross validation the number of features which gives the highest accuracy for each of the eight scenarios is selected. The number of features selected for the FC, WT and RFE based schemes is 55%, 40% and 20% respectively. It is interesting to note that the filter based FC approach, irrespective of the classifier used, has much worse accuracy compared to the other two approaches. On the other hand, the LDA with FC has comparable performance to the SVM with FC. This suggests that choice of the classification scheme does not affect the performance significantly.

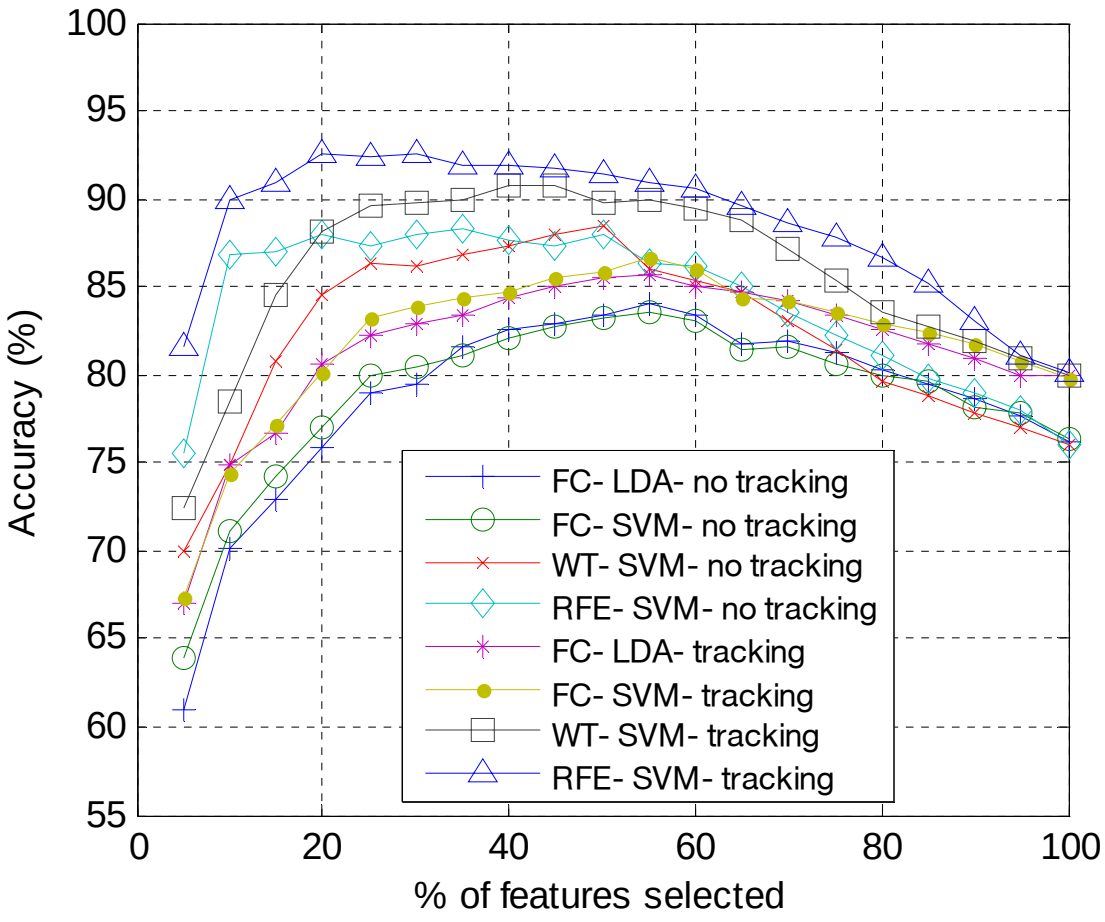


Figure 30: Accuracy of the eight scenarios during training using five fold cross validation.

Also, while the WT wrapper technique is better than the FC filter method, iteratively selecting features using the RFE method achieves improved accuracy. More importantly, using the tracking algorithm has improved performance over using a non-tracking technique in all the scenarios.

5.4.7 Testing

The accuracy on the test dataset for the eight scenarios using the reduced feature set for each of feature selections and S1 measurement technique combinations is shown in Table 8.

Accuracy (%)	No- Tracking	Tracking
FC- LDA	77.3 %	85.6 %
FC- SVM	78.8 %	85.9 %
WT- SVM	82.4 %	87.4 %
RFE- SVM	84.8 %	91.8 %

Table 8: Accuracy for the eight scenarios on the test set.

The RFE based feature selection with tracking outperforms other techniques with far fewer features. Also, the performance of the tracking based measurements is consistently better than the non- tracking based measurements. These results, on the test dataset are consistent with the ones on the training dataset obtained during cross validation. This overall performance is better than that of existing ECG based techniques described in [17] which have about 85- 92.6% accuracy. Note that the ECG results in [17] were on a different dataset. Since, our dataset is based on patients whose cardiac rhythms are controlled using rhythm control devices like pacemakers; the ECG based techniques cannot be applied.

The RFE based feature selection criteria in both the tracked and non-tracked frameworks performed better than the simple weight based technique because the RFE

re-trains the SVM at each feature elimination stage which leads to a more predictive set of features being selected as a group. The weight based technique selects the top individually predictive features which may not be the most predictive as a group. This analysis also highlights the importance of using tracking framework for monitoring. The non-tracking algorithm because of noisy measurements and no rejection criteria had worse performance than the tracking algorithms. Tracking of heart sounds improved the performance of the algorithm by over 7% over non- tracking at the same number of features.

5.4.8 Selected Features

The RFE feature selection scheme for the SVM classifier using the tracking algorithm based S1 measurements performs best with 20 features. Figure 31 shows the different features that were selected. Rows correspond to the different sub-bands and columns to the time segments within each sub-band. The selected features in the respiratory frequency bands are shaded black and selected features in the apnea frequency bands are shaded gray. As expected the most predictive sub-bands correspond to the apnea frequencies (below 0.08 Hz) and respiration frequencies (0.16 Hz to 0.5 Hz). These findings are consistent with our physiologic understanding of the changes in S1 modulation during OSA events [15]. It is interesting to note that time segments shorter than 32 seconds do not yield useful features. This suggests that a substantial number of apnea events in our dataset are longer than 16 seconds in duration. This finding is again consistent with the definition of apneas as the loss of breathing for more than 10 seconds.

Time Segments (TS) / Sub-bands	Energy of 128 sec TS	Max energy of 64 sec TS	Mean energy of 64 sec TS	Max energy of 32 sec TS	Mean energy of 32 sec TS	Max energy of 16 sec TS	Mean energy of 16 sec TS
A1 (0- 0.33 Hz)							
A2 (0- 0.16 Hz)							
A3 (0- 0.08 Hz)							
A4 (0- 0.04 Hz)							
A5 (0- 0.02 Hz)							
A6 (0- 0.01 Hz)							
A7 (0- 0.005 Hz)							
D1 (0.33- 0.5 Hz)							
D2 (0.16- 0.33 Hz)							
D3 (0.08- 0.16 Hz)							
D4 (0.04- 0.08 Hz)							
D5 (0.02- 0.04 Hz)							
D6 (0.01- 0.02 Hz)							
D7 (0.005- 0.01 Hz)							

Figure 31: Features selected for the SVM classifier using RFE criteria are shown. Sub-bands and time segments (TS) corresponding to normal respiratory drive are shown in black and sub-bands corresponding to apnea frequencies are shaded gray.

5.4.9 CSA episodes

It is important to note that our system is not affected by CSA episodes. During such events the physiology of cardiovascular changes in heart sounds is different. In particular during CSA episodes there is a cessation of airflow which results in a cessation of the S1 amplitude modulation. Figure 32 shows an example of the S1

amplitude during a CSA episode. As described the respiratory modulation is reduced and there is no crescendo like change.

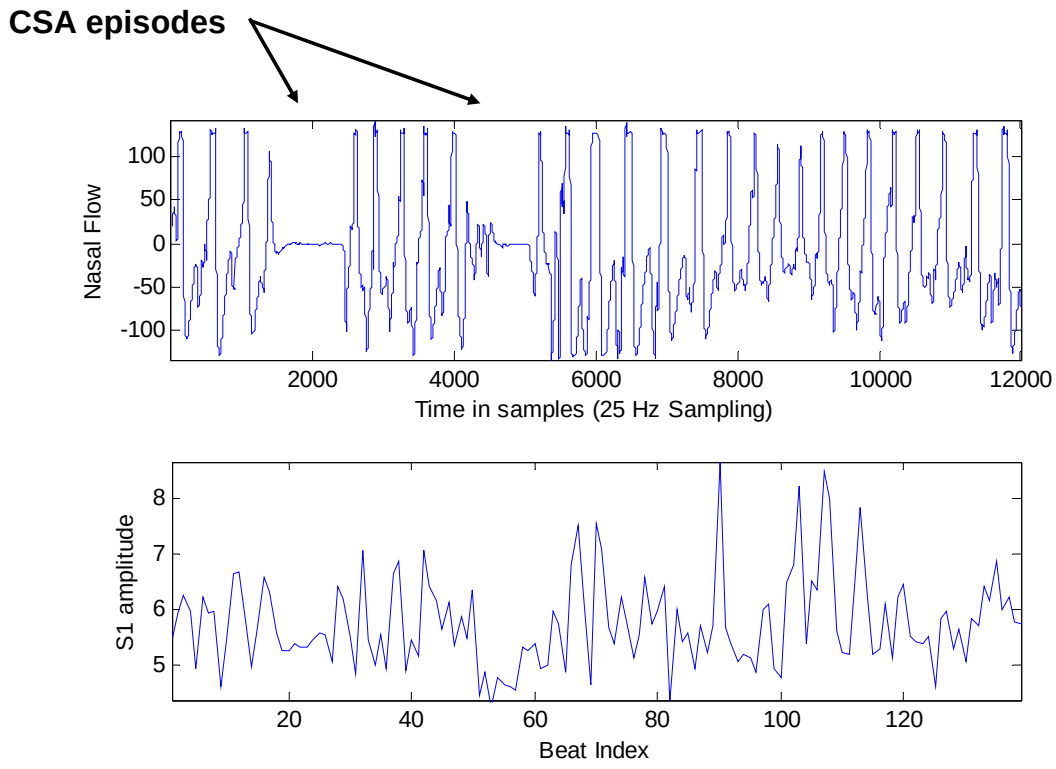


Figure 32: Example of the S1 amplitude measured during a CSA episode. The nasal flow and S1 amplitude are time aligned.

Since the features used for OSA detection aim to capture the crescendo like modulation in the S1 amplitude during periods of CSA will be treated as periods of normal breathing. To test this hypothesis the 22 epochs corresponding to 57 episodes of CSA were analyzed along with 22 randomly chosen epochs of OSA and 22 randomly chosen epochs of normal breathing. The SVM algorithm with features selected using the RFE technique was used to classify these epochs. All CSA epochs were classified as

epochs of normal breathing, supporting the hypothesis that CSA episodes will not affect the proposed algorithm.

5.4.10 Apnea Severity Score

The tracking algorithm measurements, features selected using the RFE technique and the SVM trained on the training set are next used to calculate the number of epochs per hour that are detected for each patient. This metric provides an indication of the severity of OSA in the patients in our dataset. This severity metric is compared with the traditional apnea hypopnea index (AHI), which is defined as the number of apnea or hypopnea episodes in an hour. Since any given epoch in our dataset can have multiple apnea episodes, our severity metric, while being correlated to AHI is not a direct estimate of the traditional AHI. Figure 33 shows the number of apnea epochs detected per hour vs. the AHI for the 17 patients during their sleep tests. Training and testing patients are plotted as separate scatters. Note that the same classifier optimized using the training patients is used to measure this OSA metric in both the training and the testing patients. The algorithm based OSA severity has a statistically significant correlation (R = 0.72, $p<0.05$) with AHI for the test set patients. This result is comparable with other more direct AHI surrogates that have been developed for apnea monitoring [78].

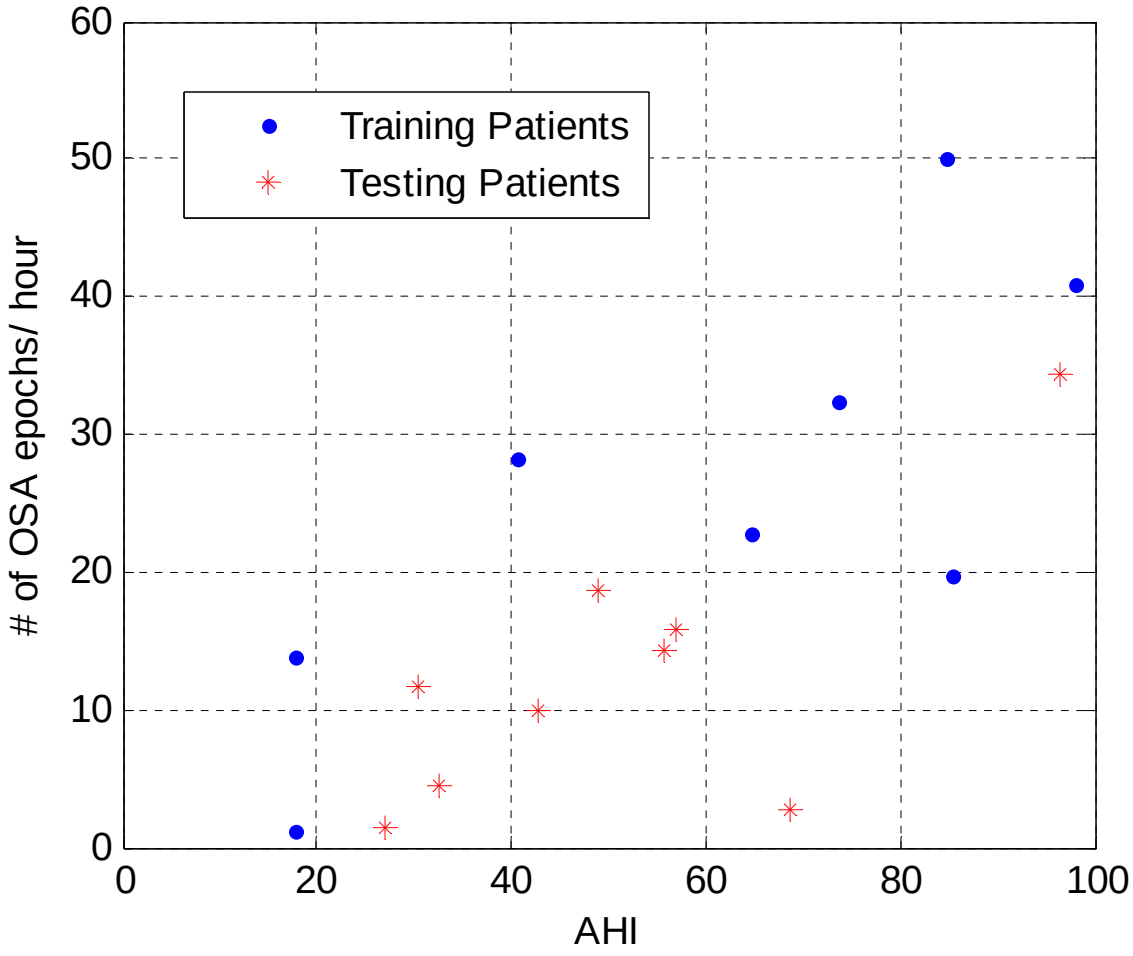


Figure 33: Algorithm based apnea severity vs. the gold standard AHI score for the training and testing patients.

Table 9 shows the clinical observations for each patient, during their sleep test, along with the algorithm results. The dataset is separated into data used for training and testing. Averages for each parameter for the training and testing data are shown to summarize the differences in the datasets. Note that patients were randomly chosen to be part of the two groups. While, it seems that the training dataset has more OSA

episodes and AHI, the test dataset has high AHI and frequent OSA episodes relative to normal patients.

	Sleep Time (ST in min)	REM (%)	# OSA	# CSA	# Hypopnea	AHI	SaO2 mean	SaO2<90 (% of ST)	# Epochs detected	# Epochs detected/ hour
Train Pt 1	357	7.8	204	0	303	85.2	93	3.8	59	19.8
Train Pt 2	330	0.0	199	0	206	73.6	93	2.5	89	32.4
Train Pt 3	315.5	26.0	91	0	249	64.7	92	5.6	60	22.8
Train Pt 4	296.5	17.6	281	0	204	98.1	91	10.1	101	40.9
Train Pt 5	268	8.9	303	21	54	84.6	90	35.7	112	50.1
Train Pt 6	336	12.5	63	0	165	40.7	91	11.8	79	28.2
Train Pt 7	288	25.3	16	19	51	17.9	93	0.6	33	13.8
Train Pt 8	322	16.4	9	2	85	17.9	93	1.6	3	1.1
Train mean	**314.1**	**14.3**	**145.8**	**5.3**	**164.6**	**60.3**	**92**	**9.0**	**67.0**	**26.1**
Test Pt 1	242	21.1	1	3	105	27.0	93	0.4	3	1.5
Test Pt 2	292	25.5	34	0	174	42.7	93	2.9	24	9.9
Test Pt 3	231	0.0	60	0	128	48.8	92	3.5	36	18.7
Test Pt 4	302.5	29.9	15	0	150	32.7	95	0.1	11	4.4
Test Pt 5	402.5	32.5	162	0	220	56.9	94	0.0	53	15.8
Test Pt 6	255	19.0	129	1	280	96.5	95	0.0	73	34.4
Test Pt 7	368	28.6	41	11	134	30.3	93	0.2	36	11.7
Test Pt 8	208.5	16.0	15	0	179	55.8	94	0.3	25	14.4
Test Pt 9	409.5	12.8	12	0	456	68.6	91	4.1	9	2.6
Test mean	**301.2**	**20.6**	**52.1**	**1.7**	**202.9**	**51.1**	**93.3**	**1.3**	**30.0**	**12.6**

Table 9: Performance of the trained algorithm compared to the gold standard measurements.

It is also interesting that a majority of patients had and AHI > 30, which is the clinically significant value for AHI. This suggests that the patients in this dataset had advanced apnea. Also, Table 9 shows that the patients in this dataset had relatively few CSA

episodes (96.3 OSA events/ hour vs. 3.3 CSA events/ hour on average). Hence the confounding effects of CSA will be minimal in this dataset. Also, important is the consistent relationship between the number of OSA episodes and value of AHI. The number of hypopnea episodes, while being high, does not affect the relationship between the number of OSA and value of AHI. This lets us compare the algorithm based OSA severity with the absolute AHI values across the patient population. As a comparison, the SaO2 mean and the duration of sleep with duration < 90%, which has been suggested as a metric for high AHI detection [79], are also shown for each patient. The SaO2 metric does not have a statistically significant correlation with AHI in this dataset.

5.5 Discussion and conclusions

The goal of this research was to develop a system for detecting OSA in high risk patients with rhythm control devices like pacemakers and cardiac resynchronization therapy devices. We use a S1 heart sound based tracking framework with an adaptive time frequency feature extraction and classification scheme for the OSA detection system. The data used in this research was collected during a standard polysomnography test. While the number of patients is limited, patients with a range of AHI were enrolled to be able to evaluate the feasibility of this approach. Also, the number of CSA episodes during the sleep tests was minimal and hence the confounding effects of CSA on our algorithm will be minimal. It is important to note that our system is not affected by CSA episodes. In Section 5.4.9 it is shown that during such events the physiology of cardiovascular changes in heart sounds is different. and that these

episodes will not affect the OSA severity metric developed in this thesis. Also, it is important to note that traditionally apnea severity is measured in events per hour. To estimate the number of events per hour, we parse time into epochs which are classified into OSA epoch or normal epoch. The number of epochs in each category is then used to measure the OSA severity.

With these specifications in mind we trained LDA and SVM based classifiers that can detect relatively short durations of obstructive sleep apnea using the novel hemodynamics based approach. To reduce complexity and improve generalization performance feature selection methods are employed and compared. The RFE criterion with SVM classifier in both the tracked and non-tracked frameworks is shown to perform better than the other techniques. This is because the RFE re-trains the SVM at each feature elimination stage which leads to a more predictive set of features being selected as a group. The Fischer criterion and the weight based techniques select the top individually predictive features which may not be the most predictive as a group. Also, while we find that the SVM classification accuracy is improved with feature selection schemes, its performance with the FC is similar to the LDA performance. This suggests that the classification scheme, in our case, is less of importance than the use of the tracking algorithm and feature selection schemes to generate optimal features.

This research also highlights the importance of consistent tracking of heart sounds, which is shown to improve the performance by 7% over not tracking using RFE feature selection. The tracking algorithm provides more consistent and less noisy S1 amplitude measurements which add to the predictive power of the features generated. This tracking becomes even more important when patients are monitored over longer

durations (days or months) because of larger variation in physiology. The tracking framework with features selected using the RFE criteria has a 91.8% accuracy when applied to the test dataset. While our dataset is different from the one used for the ECG algorithm testing in [17], this result shows the favorable performance of our approach compared to other popular approaches.

In summary this research presents a novel tracking based approach, using non-snapshot monitoring of heart sounds, to detect perturbations in the cardiovascular system due to OSA. The main contributions of this research are the development of the novel tracking framework with a classification scheme. This technique is shown to perform favorably compared to literature and more importantly the low complexity tracking framework is shown to significantly improve its accuracy over a non- tracking approach. Filter and wrapper based feature selection schemes are compared and evaluated for selecting optimal features. Additionally, an algorithm based severity score is shown to be correlated with the gold standard apnea hypopnea index measure.

6 Conclusions

6.1 Technical Contributions

The goal of this thesis was to develop algorithms to facilitate accurate measurement and tracking of clinically useful heart sound parameters and apply them for the detection of cardiovascular perturbations. Two technical contributions made it possible for us to achieve this goal:

- A dynamic programming based heart sounds tracking algorithm was developed. This algorithm has linear complexity in the number of heart sounds beats which makes it attractive for implementation in an implanted device. When tested on acute and chronic pre-clinical heart failure models, we found that the tracking framework outperformed the traditional non- tracking algorithm. The performance gain along with the implementation ease could be useful for long term monitoring of patients using heart sound parameters.

- A non-snapshot based heart sounds approach for OSA detection using the tracking framework was developed for patients with rhythm control devices. This framework was based on a novel adaptive time frequency feature extraction and classification scheme. The classification and feature selection schemes were optimized to give the best performance. The tracking framework developed when compared with a traditional non- tracking method was found to provide improved accuracy and reliability. While the proposed framework in this thesis was shown to detect OSA, the application of these techniques can be extended to other perturbations in the cardiovascular system. The general idea of feature

generation using wavelet transforms followed by time segmentation can be applied to many other physiologic perturbations where the changes in physiology are localized in both time and frequency.

6.2 Clinical Findings

The development of algorithms to make heart sound measurements continuously will facilitate continuous tracking of the patient's status. Listed here are some of the clinical correlations and potential utility of the heart sound based measurements we found using the algorithms developed in this thesis:

- The S1 amplitude and ejection time (HSET) measured using the algorithms developed in Chapter 3 were significantly correlated with left ventricular contractility measured by dP/dt and aortic pulse pressure. Also, the S1 amplitude and HSET showed change with worsening heart failure in the three animal models. The chronic animal model with an implanted heart sound sensor validated the applicability of the tracking algorithm for long term monitoring of cardiac mechanical performance. These comparisons along with hemodynamic correlates of the measurements made using the tracking framework illustrate the potential of our novel approach for long term monitoring of cardiac conditions.

- The clinical study data analysis in Chapter 4 revealed that in patients with changes in SV> 10 ml the HSET and SV was significantly correlated. Also, this analysis showed that HSET has a moderate sensitivity of 70% at 80% specificity to detect periods of low stroke volume (SV< 50ml). We found that the ejection time measured from heart sounds contains clinically useful information regarding the

cardiac systolic performance and also the value of measuring systolic timing intervals in hospitalized heart failure patients.

- In Chapter 5, a system for detecting OSA in high risk patients with rhythm control devices like pacemakers and cardiac resynchronization therapy devices was developed. This algorithm based severity score was used determine the intensity of OSA for each patient during the sleep study and was shown to significantly correlate with the gold standard apnea hypopnea index (AHI).

6.3 Future Work

The algorithms developed in this thesis can be directly applied to patient monitoring using heart sounds. While they have been studied here in the context of heart failure preclinical and clinical studies and apnea detection, their application to other cardiac disease states (e.g. ischemia detection, pulmonary hypertension) needs to be further explored. These applications will require additional pre-processing steps like time-frequency transformation of the heart sounds to extract optimal signal content specific to the disease state. These steps can be easily added to the perturbation detection framework shown in Figure 27. Also, while we use an SVM based classifier for OSA detection, other applications may require exploration of other classification algorithms. Depending on the number of features and their content, the detection algorithm may range from a simple threshold detector to a sophisticated neural network. These application specific areas need to be explored further.

Appendix- Glossary

Cardiac valves: the heart valves are valves in the heart that maintain the unidirectional flow of blood by opening and closing depending on the difference in pressure on each side. There are four main valves of the heart: mitral, tricuspid, aortic and pulmonary.

Circadian: A circadian rhythm is a roughly-24-hour cycle in the physiological processes of living beings, including plants, animals, fungi and cyanobacteria. This rhythm causes a cyclic variation in many physiologic variables like heart rate, respiration rate, etc.

Contractility: Myocardial contractility is a term used in physiology to describe the performance of cardiac muscle. It is often defined as: the intrinsic ability of a cardiac muscle fiber to contract at a given fiber length.

Epicardium: Epicardium describes the outer layer of heart. Its largest constituent is connective tissue and functions as a protective layer.

Hemodynamics: Hemodynamics (literally, "blood dynamics"), is the study of the properties and flow of blood.

Mitral regurgitation: Mitral regurgitation (MR), also known as mitral insufficiency, is the abnormal leaking of blood through the mitral valve, from the left ventricle into the left atrium of the heart.

Pacing: Pacing is the mechanism of sending current pulses to selected areas of the heart in order to activate that region.

Pulse pressure: Pulse pressure is the change in blood pressure seen during a contraction of the heart.

Stroke volume: Stroke volume is the amount of blood pumped by the left ventricle of the heart in one contraction.

Systolic performance: The systolic performance of the heart is a measure of its ability to pump blood during ejection. A weak heart has poor systolic performance.

References

[1] W. Rosamond, K. Flegal, K. Furie, A. Go, K. Greenlund, N. Haase, S. M. Hailpern, M. Ho, V. Howard, B. Kissela, S. Kittner, D. Lloyd-Jones, M. McDermott, J. Meigs, C. Moy, G. Nichol, C. O'Donnell, V. Roger, P. Sorlie, J. Steinberger, T. Thom, M. Wilson, and Y. Hong, "Heart disease and stroke statistics--2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee," *Circulation*, vol. 117, pp. e25-146, 2008.

[2] J. Butler and A. Kalogeropoulos, "Worsening heart failure hospitalization epidemic we do not know how to prevent and we do not know how to treat!," *J Am Coll Cardiol*, vol. 52, pp. 435-7, 2008.

[3] C. W. Yancy, M. Lopatin, L. W. Stevenson, T. De Marco, and G. C. Fonarow, "Clinical presentation, management, and in-hospital outcomes of patients admitted with acute decompensated heart failure with preserved systolic function: a report from the Acute Decompensated Heart Failure National Registry (ADHERE) Database," *J Am Coll Cardiol*, vol. 47, pp. 76-84, 2006.

[4] J. Remes, H. Miettinen, A. Reunanen, and K. Pyorala, "Validity of clinical diagnosis of heart failure in primary health care," *Eur Heart J*, vol. 12, pp. 315-21, 1991.

[5] N. J. Fortuin, W. P. Hood, Jr., and E. Craige, "Evaluation of left ventricular function by echocardiography," *Circulation*, vol. 46, pp. 26-35, 1972.

[6] C. Berry, D. Murdoch, and J. McMurray, "Economics of chronic heart failure," *Eur. J. Heart Fail*, vol. 3, pp. 283- 291, 2001.

[7] P. Heidenreich, C. Ruggerio, and B.Massie, "Effect of a home monitoring system on hospitalization and resource use for patients with heart failure," *Am. Heart J.*, vol. 138, pp. 633- 640, 2000.

[8] W. H. Wu, A. A. T. Bui, M. A. Batalin, D. Liu, and W. J. Kaiser, "Incremental Diagnosis Method for Intelligent Wearable Sensor Systems," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 11, pp. 553-562, 2007.

[9] U. Anliker, J. A. Ward, P. Lukowicz, G. Troster, F. Dolveck, M. Baer, F. Keita, E. B. Schenker, F. Catarsi, L. Coluccini, A. Belardinelli, D. Shklarski, M. Alon, E. Hirt, R. Schmid, and M. Vuskovic, "AMON: a wearable multiparameter medical monitoring and alert system," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 8, pp. 415-427, 2004.

[10] J. Abrams, "Current concepts of the genesis of heart sounds. I. First and second sounds," *Jama*, vol. 239, pp. 2787-9, 1978.

[11] E. A. Phillipson, "Sleep apnea--a major public health problem," *N Engl J Med*, vol. 328, pp. 1271-3, 1993.

[12] E. Shahar, C. W. Whitney, and S. Redline, "Sleep disordered breathing and cardiovascular disease: cross sectional results of the Sleep Heart Health Study," *Am J Respir Crit Care Med*, vol. 163, pp. 19- 25, 2001.

[13] S. Javaheri, T. J. Parker, J. D. Liming, W. S. Corbett, H. Nishiyama, L. Wexler, and G. A. Roselle, "Sleep apnea in 81 ambulatory male patients with stable heart

failure. Types and their prevalences, consequences, and presentations," *Circulation*, vol. 97, pp. 2154-9, 1998.

[14] S. Garrigue, J. L. Pepin, P. Defaye, F. Murgatroyd, Y. Poezevara, J. Clementy, and P. Levy, "High prevalence of sleep apnea syndrome in patients with long-term pacing: the European Multicenter Polysomnographic Study," *Circulation*, vol. 115, pp. 1703-9, 2007.

[15] J. D. Parker, "Acute and chronic effects of airway obstruction on canine left ventricular performance," *Am J Respir Crit Care Med*, vol. 160, pp. 1888- 1896, 1999.

[16] *Systematic Review of the Literature Regarding the Diagnosis of Sleep Apnea: Summary, Evidence Report/ Technology Assessment Number 1 Agency for Health Care Policy and Research (AHCPR), Department of Health and Hyman Services, U.S. Public Health Service*, vol. E002, 1999.

[17] G. B. Moody, R. G. Mark, A. Goldberger, and T. Penzel, "Stimulating rapid research advances via focused competition: the Computers in Cardiology Challenge 2000," presented at Computers in Cardiology 2000, 2000.

[18] A. Vander, Sherman, J. and Luciano, D., *Human Physiology*: McGraw-Hill, New York, 2005.

[19] R. E. Klabunde: www.cvphysiology.com, 2007.

[20] A. M. Weissler, "Interpreting systolic time intervals in man," *J Am Coll Cardiol*, vol. 2, pp. 1019-20, 1983.

[21] R. S. Stack, Y. H. Sohn, and A. M. Weissler, "Accuracy of systolic time intervals in detecting abnormal left ventricular performance in coronary artery disease," *Am J Cardiol*, vol. 47, pp. 603-9, 1981.

[22] B. J. Northover, "Left ventricular systolic time intervals in patients with acute myocardial infarction," *Br Heart J*, vol. 43, pp. 506-13, 1980.

[23] M. El-Segaier, O. Lilja, S. Lukkarinen, L. Sornmo, R. Sepponen, and E. Pesonen, "Computer-based detection and analysis of heart sound and murmur," *Ann Biomed Eng*, vol. 33, pp. 937-42, 2005.

[24] R. J. Kulangara, W. B. Strong, and M. D. Miller, "Evaluation of heart sounds and murmurs in children," *Postgrad Med*, vol. 72, pp. 151-3, 156-60, 1982.

[25] R. M. Rangayyan and R. J. Lehner, "Phonocardiogram signal analysis: a review," *Crit Rev Biomed Eng*, vol. 15, pp. 211-36, 1987.

[26] M. Shapiro, B. Moyers, G. M. Marcus, I. L. Gerber, B. H. McKeown, J. C. Vessey, M. V. Jordan, M. Huddleston, E. Foster, K. Chatterjee, and A. D. Michaels, "Diagnostic characteristics of combining phonocardiographic third heart sound and systolic time intervals for the prediction of left ventricular dysfunction," *J Card Fail*, vol. 13, pp. 18-24, 2007.

[27] G. M. Marcus, I. L. Gerber, B. H. McKeown, J. C. Vessey, M. V. Jordan, M. Huddleston, C. E. McCulloch, E. Foster, K. Chatterjee, and A. D. Michaels, "Association between phonocardiographic third and fourth heart sounds and objective measures of left ventricular function," *Jama*, vol. 293, pp. 2238-44, 2005.

[28] D. Chen, L. G. Durand, and H. C. Lee, "Time-frequency analysis of the first heart sound. Part 1: Simulation and analysis," *Med Biol Eng Comput*, vol. 35, pp. 306-10, 1997.
[29] A. A. Luisada and F. Portaluppi, "The main heart sounds as vibrations of the cardiohemic system: old controversy and new facts," *Am J Cardiol*, vol. 52, pp. 1133-6, 1983.
[30] A. A. Luisada, "The first heart sound in normal and pathological conditions," *Jpn Heart J*, vol. 28, pp. 143-56, 1987.
[31] D. Chen, L. G. Durand, Z. Guo, and H. C. Lee, "Time-frequency analysis of the first heart sound. Part 2: An appropriate time-frequency representation technique," *Med Biol Eng Comput*, vol. 35, pp. 311-7, 1997.
[32] D. Chen, L. G. Durand, H. C. Lee, and D. W. Wieting, "Time-frequency analysis of the first heart sound: Part 3: Application to dogs with varying cardiac contractility and to patients with mitral mechanical prosthetic heart valves," *Med Biol Eng Comput*, vol. 35, pp. 455-61, 1997.
[33] X. Jingping, L. G. Durand, and P. Pibarot, "Extraction of the aortic and pulmonary components of the second heart sound using a nonlinear transient chirp signal model," *Biomedical Engineering, IEEE Transactions on*, vol. 48, pp. 277-283, 2001.
[34] D. Chen, P. Pibarot, G. Honos, and L. G. Durand, "Estimation of pulmonary artery pressure by spectral analysis of the second heart sound," *Am J Cardiol*, vol. 78, pp. 785-9, 1996.
[35] K. Z. Siejko, "Phonocardiographic image-based atrioventricular delay optimization," U.S. Patent US7123962 B2, 2006.
[36] L. G. Durand and P. Pibarot, "Digital signal processing of the phonocardiogram: review of the most recent advancements," *Crit Rev Biomed Eng*, vol. 23, pp. 163-219, 1995.
[37] J. Genest, Jr. and L. G. Durand, "Relationship of the left ventricular and apical first sounds to the left ventricular pressure derivative," *Med Biol Eng Comput*, vol. 23, pp. 95-8, 1985.
[38] T. Sakamoto, R. Kusukawa, D. M. Maccanon, and A. A. Luisada, "Hemodynamic Determinants of the Amplitude of the First Heart Sound," *Circ Res*, vol. 16, pp. 45-57, 1965.
[39] J. C. Wood, A. J. Buda, and D. T. Barry, "Time-frequency transforms: a new approach to first heart sound frequency dynamics," *IEEE Trans Biomed Eng*, vol. 39, pp. 730-40, 1992.
[40] Z. M. Zin, S. H. Salleh, S. Daliman, and M. D. Sulaiman, "Analysis of heart sounds based on continuous wavelet transform," 2003.
[41] S. M. Debbal and F. Bereksi-Reguig, "Analysis of the second heart sound using continuous wavelet transform," *J Med Eng Technol*, vol. 28, pp. 151-6, 2004.
[42] L. Khadra, M. Matalgah, B. el-Asir, and S. Mawagdeh, "The wavelet transform and its applications to phonocardiogram signal analysis," *Med Inform (Lond)*, vol. 16, pp. 271-7, 1991.

[43] H. Sava, P. Pibarot, and L. G. Durand, "Application of the matching pursuit method for structural decomposition and averaging of phonocardiographic signals," *Med Biol Eng Comput*, vol. 36, pp. 302-8, 1998.
[44] Z. Xuan, L. Durand, L. Senhadji, H. C. Lee, and J. L. Coatrieux, "Analysis-synthesis of the phonocardiogram based on the matching pursuit method," *Biomedical Engineering, IEEE Transactions on*, vol. 45, pp. 962-971, 1998.
[45] X. Jingping, L. Durand, and P. Pibarot, "Nonlinear transient chirp signal modeling of the aortic and pulmonary components of the second heart sound," *Biomedical Engineering, IEEE Transactions on*, vol. 47, pp. 1328-1335, 2000.
[46] N. Gordon and D. Salmond, "Aspects of target tracking: problems and techniques," presented at IEE Colloquium on Target Tracking and Data Fusion 1998.
[47] S. S. Blackmann, *Multiple-target tracking with radar applications*: Artech House, Inc., Dedham, MA, 1986.
[48] G. V. Trunk, "Survey of Radat ADT," *Microwave Journal*, vol. 26, pp. 77-88, 1983.
[49] R. Singer, R. Sea, and K. Housewright, "Derivation and evaluation of improved tracking filter for use in dense multitarget environments," *Information Theory, IEEE Transactions on*, vol. 20, pp. 423-432, 1974.
[50] Y. Bar-Shalom, "Tracking methods in a multitarget environment," *Automatic Control, IEEE Transactions on*, vol. 23, pp. 618-626, 1978.
[51] J. K. Wolf, A. M. Viterbi, and G. S. Dixon, "Finding the best set of K paths through a trellis with application to multitarget tracking," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 25, pp. 287-296, 1989.
[52] R. T. T. Hastie, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*: New York: Springer-Verlag, 2001.
[53] P. E. H. R. O. Duda, and D. G. Stork, *Pattern Classification*: 2 ed: Wiley Interscience, 2000.
[54] I. Guyon, *Feature Extraction: foundations and applications*: Springer, 2006.
[55] I. Guyon, "An introduction to variable and feature selection " *J. Mach. Learn. Res.* , vol. 3, pp. 1157-1182 2003
[56] A. Patangay, P. Vemuri, and A. Tewfik, "Monitoring of Obstructive Sleep Apnea in Heart Failure Patients," presented at Engineering in Medicine and Biology Society, EMBS '07. 29th Annual International Conference of the IEEE, 2007.
[57] T. Hastie and R. Tibshirani, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.*: New York, Springer-Verlag, 2006.
[58] H. Liang, S. Lukkarinen, and I. Hartimo, "Heart sound segmentation algorithm based on heart sound envelogram," presented at 24th Computers in Cardiology, 1997.
[59] M. B. Malarvili, I. Kamarulafizam, S. Hussain, and D. Helmi, "Heart sound segmentation algorithm based on instantaneous energy of electrocardiogram," presented at 31st Computers in Cardiology, 2003.

[60] S. Rajan, E. Budd, M. Stevenson, and R. Doraiswami, "Unsupervised and Uncued Segmentation of the Fundamental Heart Sounds in Phonocardiograms Using a Time-Scale Representation," 2006.
[61] H. Liang and I. Hartimo, "A heart sound feature extraction algorithm based on wavelet decomposition and reconstruction," 1998.
[62] H. Liang, L. Sakari, and H. Iiro, "A heart sound segmentation algorithm using wavelet decomposition and reconstruction," 1997.
[63] A. Patangay, K. Siejko, D. Da Cunha, A. Pedraza, and R. Hamlin, "S3 intensity measured chronically from an implanted ICD is correlated to hemodynamics during HF decompensation," *Journal of Cardiac Failure*, vol. 14, pp. S111, 2008.
[64] A. Patangay, J. Stahmann, and R. Sweeney, "Monitoring of Heart Sounds," U.S. Patent Application US20080119750 A1, 2007.
[65] D. A. Castanon, "Efficient algorithms for finding the K best paths through a trellis," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 26, pp. 405-410, 1990.
[66] C. Danmin, L. G. Durand, and H. C. Lee, "Application of the cone-kernel distribution to study the effect of myocardial contractility in the genesis of the first heart sound in dog," 1995.
[67] J. C. Wood and D. T. Barry, "Quantification of first heart sound frequency dynamics across the human chest wall," *Med Biol Eng Comput*, vol. 32, pp. S71-8, 1994.
[68] A. M. Weissler, W. S. Harris, and C. D. Schoenfeld, "Systolic time intervals in heart failure in man," *Circulation*, vol. 37, pp. 149-59, 1968.
[69] R. L. Moskowitz and B. M. Wechsler, "Left Ventricular Ejection Time in Aortic and Mitral Valve Disease," *Am J Cardiol*, vol. 15, pp. 809-14, 1965.
[70] W. S. Aronow, A. F. Bowyer, and M. A. Kaplan, "External isovolumic contraction times and left ventricular ejection time-external isovolumic contraction time ratios at rest and after exercise in coronary heart disease," *Circulation*, vol. 43, pp. 59-65, 1971.
[71] P. Hamosh, J. N. Cohn, K. Engelman, and M. F. E. Broder, "Systolic time intervals and left ventricular function in acute myocardial infarction," *Circulation*, vol. 45, pp. 375-81, 1972.
[72] M. Zuber, S. Toggweiler, L. Quinn-Tate, L. Brown, A. Amkieh, and P. Erne, "A comparison of acoustic cardiography and echocardiography for optimizing pacemaker settings in cardiac resynchronization therapy," *Pacing Clin Electrophysiol*, vol. 31, pp. 802-11, 2008.
[73] C. L. Garrard, Jr., A. M. Weissler, and H. T. Dodge, "The relationship of alterations in systolic time intervals to ejection fraction in patients with cardiac disease," *Circulation*, vol. 42, pp. 455-62, 1970.
[74] A. M. Weissler, W. S. Harris, and C. D. Schoenfeld, "Bedside technics for the evaluation of ventricular function in man," *Am J Cardiol*, vol. 23, pp. 577-83, 1969.

[75] A. Patangay and A. Tewfik, "Low Complexity Tracking for Long Term Monitoring of Heart Sounds," presented at Engineering in Medicine and Biology Society, EMBS '08. 30th Annual International Conference of the IEEE, 2008.

[76] Q. Zhang, "Matlab package for robust and efficient location of T- wave ends in ECG and its evaluation with PhysioNet data," *http://www.irisa.fr/sosso/zhang/biomedical/*, 2005.

[77] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *J. Machine Learning Res*, vol. 3, pp. 1439-1461, 2003.

[78] Shalaby A, "Feasibility of automated detection of advanced sleep disordered breathing utilizing an implantable pacemaker ventilation sensor.," *Pacing Clin Electrophysiol*, vol. 29, pp. 1036-43, 2006.

[79] S. Gyulay, L. G. Olson, M. J. Hensley, M. T. King, K. M. Allen, and N. A. Saunders, "A comparison of clinical assessment and home oximetry in the diagnosis of obstructive sleep apnea," *Am Rev Respir Dis*, vol. 147, pp. 50-3, 1993.